



**Job Search Assistance Programs in Europe:
Evaluation Methods and Recent Empirical
Findings**

Stephan L. Thomsen

FEMM Working Paper No. 18, May 2009

F E M M

Faculty of Economics and Management Magdeburg

Working Paper Series

JOB SEARCH ASSISTANCE PROGRAMS IN EUROPE: EVALUATION METHODS AND RECENT EMPIRICAL FINDINGS

Stephan L. Thomsen*
University of Magdeburg & ZEW, Mannheim

This version: June 2, 2009

Abstract

Job search assistance programs are part of active labor market policy in many countries. The main characteristics of these activities are an intensified counseling and a job search monitoring; in addition, several countries integrate courses teaching further skills into the programs. Job search assistance programs should help to increase the employment chances and to reduce the unemployment duration of the job seekers. In this paper, recent empirical findings from evaluation studies for 9 European countries are reviewed and implications with regard to the effectiveness of the activities are derived. To make the findings of various studies evaluating the different programs comparable, the methodological issues of the empirical approaches applied to estimate the causal effects of the programs are discussed in detail. In addition, relevant characteristics of the unemployment insurance systems, the assignment process, and the content of programs are presented to derive meaningful implications. The comparison of the programs takes account of individual effects and, if available, cost benefit considerations. The results show that job search assistance programs tend to provide an effective means to reduce individual unemployment, particularly if provided as combinations of intensive counseling and short-term training courses.

Keywords: Job search assistance programs, active labor market policy, evaluation methods, Europe

JEL Classification: J68, C31

*Stephan L. Thomsen is Assistant Professor of Labor Economics at Otto-von-Guericke-University Magdeburg and Research Associate at Centre for European Economic Research (ZEW) Mannheim.
Address: PO Box 4120, D-39016 Magdeburg, e-mail: stephan.thomsen@ovgu.de, phone: +49 391 6718431, fax: +49 391 6711700.

1 Introduction

In many European countries, programs to activate the unemployed (active labor market policy, ALMP) play an important role within labor market policy. Designed as selective policies that are (more or less) targeted to well-defined groups of the labor force, programs aim at reducing individual unemployment by offering a variety of activities such as training courses, employment subsidies, or public employment services. The core argument for offering ALMP is the economic reasoning that promoting employment (or activities directly aiming at employment) is better than compensating unemployment (by generous unemployment benefits). In that sense, the main purposes of ALMP are not only to improve employment outcomes, to lower benefit rates and the benefit dependency rates, but also to contribute to the quality and productivity of jobs and to strengthen social cohesion (European Commission, 2006). However, high and persistent levels of unemployment and a growing share of long-term unemployment lead to budgetary pressure and to doubts about the effectiveness of ALMP in many countries. Accompanied by demographic change (ageing societies) and new economic challenges (e.g. globalization) this has forced politicians, economists, and society to critically evaluate the efforts in labor market policy programs.

Therefore, the number of studies providing empirical evaluations of the effects of various ALMP programs in European countries has risen steadily over the last decade. This development was accompanied and/or in part enabled by a growing availability of better data (particularly, micro register data), overcoming limitations of earlier household or labor force surveys, as well as improvements in evaluation methodology to estimate treatment effects from non-experimental data leading to more reliable estimation results. Nevertheless, European Commission (2006) still argued that an ‘evaluation culture’ for labor market policies is still missing in many European countries and contrasts markedly with the practice in other countries like the US or Canada. However, the situation improves as evaluation studies become more common using state-of-the-art evaluation methods to identify the causal effects of programs and better data.

The overviews by Martin and Grubb (2001), Calmfors, Forslund, and Hemström (2001) and Kluge (2006) show that the doubts of the public about the effectiveness of ALMP programs correlate positively with the findings from the scientific evaluations: The majority of programs, in particular employment subsidies and longer-term vocational training programs, tend to be ineffective in reducing unemployment and bringing people back to employment, or result in modest positive improvements of the labor market situation of the participants. The budgetary pressure together with this evidence at hand led to a shift of the single activities within ALMP in recent years. Programs for identification of needs, job search assistance, guidance, and training as part of a personalized action plan have become more important within the mixture of ALMP in a number of countries. The main reasons are the large degree of flexibility, the lower expenditure relative to more traditional and longer lasting programs, and the possibility to commit

job seekers to report actions more frequently. Recently, these job search assistance programs in Europe have been analyzed in a number of empirical studies. However, the evidence is dispersed as studies differ in a number of respects. First, the programs offered are heterogeneous across countries with regard to the design and the duration, i.e. some programs provide intensified job search assistance via counseling and monitoring only whereas others comprise job training courses in addition. Second, some programs are addressed to all job seekers whereas others are explicitly targeted to certain groups, e.g. young unemployed or long-term unemployed. In the latter case, programs may prove effective for the groups in focus but generalizing the findings to other groups of job seekers could be problematic. Third, a number of different estimation methods are applied to evaluate programs' effects in the available empirical studies. All these differences have to be regarded when comparing the results.

The aim of this survey is to condense the findings from recent evaluation studies of European job search assistance programs. Two questions lay at the core of this paper. The first is: How effective are job search assistance programs in European countries in bringing people back to work or shortening the duration of unemployment? And according to that, the second question asks: What are the reasons and determinants of effective - or respectively - less effective programs? By and large, nine countries could be considered in this survey: Austria, Denmark, Finland, France, Germany, Hungary, the Netherlands, Portugal, and the United Kingdom. For all of these countries, at least one empirical study evaluating the effects of the job search assistance program is available. For some countries, programs have been analyzed by several authors. As the focus of this paper is the cross-country comparison of European programs, I have selected a number of representative studies with regard to the empirical findings and the estimation method applied in these cases.¹

Moreover, I do not consider programs conducted in the US. A comprehensive overview on these programs similar to the kind considered here is provided by Meyer (1995). From his summary of five different job search assistance programs he concludes that the activities provided are an effective means to shorten unemployment duration and could be cost-efficient. However, with regard to the reasons he could not disentangle whether positive effects are due to a stricter enforcement of job seekers or due to a frequent verification of the job search behavior. Whereas the first refers to an improvement of the job search abilities, the latter coincides with a stricter monitoring and sanctioning of benefits if job seekers do not cooperate. In particular the latter aspect of programs exhibits some threat effect of participation to job seekers. Fortunately, Ashenfelter, Ashmore, and Deschênes (2005) complement the findings of Meyer (1995) by analyzing a job search program containing only the monitoring component of the program. The empirical results clearly indicate that solely increasing the monitoring of the job seekers provides no means to improve labor market outcomes. With regard to the differences in the design of the European programs reviewed here, this evidence could be important. Therefore, I will consider the evidence of the US in the discussion.

To make results comparable across countries, a particular emphasis is spent on the three aspects mentioned above. To do so, I will first review the main characteristics of the empirical methods including the empirically non-testable identifying assumptions (section 2). This review should help to understand heterogeneity in the estimates and the limitations of the different approaches. After that, a summary of important characteristics of the countries' labor market systems and the programs considered is provided (section 3). Based on this preparation, in section 4 the main findings of the studies are compared. Although job search assistance programs in Europe tend to exhibit positive effects in terms of increased employment and a reduced unemployment duration, the picture is mixed and not all programs provided are successful. The reasons for these differences will be explored and discussed in section 5. A short conclusion is given in the final section.

2 Evaluation methods

The crucial task of program evaluation is to identify the causal effect of participation in the program in study on an outcome variable, e.g. the duration of unemployment, the probability of employment, or the earnings. On the micro level, the fundamental problem of this task lies in the missing counterfactual situation of participation. In the simplest case, persons could only choose (whether on themselves, selected by a caseworker, or upon some specific rule makes no difference at this point) to participate in a program or not. Clearly, no person is observable in both states at the same point of time. To solve this problem, experimental and non-experimental approaches are suggested in the literature that both require non-testable assumptions to construct the missing counterfactual. Without collecting data specifically designed to test these assumptions (see Heckman, Ichimura, Smith, and Todd, 1996) no validation of the unbiasedness of the estimates is possible. With regard to the evaluation of job search assistance programs in Europe, three approaches (experimental estimators, propensity score matching estimators, and multivariate duration models) have been used and will be discussed below. To start with I will first review the framework for the evaluation of treatment effects.

2.1 Potential outcome approach and treatment effect

The standard model in the microeconomic evaluation literature is the so-called *potential outcome approach* which has been variously attributed to Neyman (1923; 1935), Fisher (1935), Roy (1951), Quandt (1972; 1988) or Rubin (1974). In the simple form, the model considers two possible states of the world.² An individual i is imagined to either participate in a program or not. Let Y_i^1 and Y_i^0 denote the potential outcomes corresponding to the states, where 1 denotes treatment and 0 non-treatment. According to this definition, the individual causal effect of

treatment is defined as the difference of the two potential outcomes, i.e.

$$\Delta_i = Y_i^1 - Y_i^0. \quad (1)$$

However, since the individual cannot be in both states of the world at the same time, the actual observed outcome for each individual i can be written as:

$$Y_i = Y_i^1 \cdot D_i + (1 - D_i) \cdot Y_i^0, \quad (2)$$

where $D_i \in \{0, 1\}$ is a binary treatment indicator, with $D = 1$ denoting participation and $D = 0$ denoting non-participation.³ To complete the notation, let X denote variables that are unaffected by treatment – the so-called *attributes* by Holland (1986). Eq. (2) makes clear that one of the outcomes is unobservable for each individual, i.e. only Y_i^1 or Y_i^0 is observable. For that reason, there is no opportunity to calculate individual effects directly from the data, and Δ_i is not observed for anyone.

Since direct estimation of the individual effect of treatment in eq. (1) is impossible, evaluation has to focus on population averages of gains from treatment. The most common parameter of interest in the empirical literature is the average effect of treatment on the treated (ATT).⁴ The ATT is defined as

$$\begin{aligned} \Delta^{ATT} = E(\Delta|D = 1) &= E(Y^1 - Y^0|D = 1) \\ &= E(Y^1|D = 1) - E(Y^0|D = 1), \end{aligned} \quad (3)$$

which is the difference of the expected outcomes with and without treatment for participants. As it focuses directly on the actual participants, it determines the realized gross gain for this group (Heckman, LaLonde, and Smith, 1999). Thus, its importance for policy makers becomes obvious as programs are generally targeted to certain groups; and by comparing the program effect with its costs, the ATT is a reasonable approach to measure the performance of the program, i.e. deciding whether the program is a success or not (see Heckman and Robb, 1985b, and Heckman, Ichimura, and Todd, 1997).

To render the model useful for causal analysis, one has to invoke the stable unit treatment value assumption (SUTVA, see e.g. Rubin, 1986). SUTVA rules out any cross-effects or general equilibrium effects that may occur among potential program participants because of their participation decision (Lechner, 2001). In other words, the potential outcomes of an individual depend on the individual's participation decision only and are not affected by the treatment status of other individuals. Furthermore, whether an individual participates or not does not depend on the participation decision of other individuals. This additional feature excludes peer-effects (Sianesi, 2004). If one is willing to estimate the effect of the program for a person drawn randomly from the participants sample, those effects are negligible and SUTVA could be assumed to be fulfilled.

In addition, Heckman, LaLonde, and Smith (1999) point out that microeconomic evaluation concentrates on direct effects only. A full evaluation of the program of interest would require an enumeration of all outcomes of interest for every person, both in the actual state of the world as well as in the alternative state(s). In the most general view, almost everyone in a modern economy participates in each social program either directly or indirectly. Direct effects affect the situation of only those persons enrolled to the programs. Effects that do not result from participation directly are defined as indirect effects. The indirect effects could occur for participants and non-participants. For example, participants may pay taxes or unemployment insurance contributions to support the program just as persons who do not participate. Furthermore, indirect effects occur for persons with whom the participants compete in the labor market and for the firms that hire the participants. The problem of the indirect effects is ignored in the econometric and statistical evaluation literature, and treatment outcomes are equated with the direct outcome Y^1 in the program state; no treatment outcomes are equated with the direct outcome Y^0 in the no-program state. However, this is a crucial assumption in the traditional evaluation literature (Heckman and Smith, 1998). This should be kept in mind when discussing the results of the job search assistance programs below.

2.2 Experimental estimators

It becomes obvious from eq. (3) that the second term on the right-hand side is unobservable. The term describes the hypothetical outcome of the participants if they had not participated in the program. In an experimental evaluation where participants are assigned randomly to the program or a control group, the missing counterfactual data of the treatment can be derived by using information from that control group. The hypothetical outcome of participants if they had not participated would not differ from the expected outcome of the non-participants, i.e.

$$E(Y^0|D = 1) = E(Y^0|D = 0). \quad (4)$$

Therefore, it is possible to approximate the counterfactual outcome of the participants by the non-participants' outcomes. The ATT can easily be computed since the data on program participants identify the mean outcome in the treated state, $E(Y^1|D = 1)$, and the randomized-out comparison provides the direct estimate for $E(Y^0|D = 1)$ (Smith and Todd, 2005).⁵

Social experiments have been seen as the ideal way to evaluate the impacts of programs in particular in the US (Smith, 2000b). In his survey, Smith (2000a) notes a set of advantages of social experiments over standard non-experimental methods. First, they are easy to explain to policy makers because most educated persons understand the issue of random assignment. Second, as becomes obvious by eq. (4), social experiments produce consistent estimates of the impact of treatment on the treated and they are less controversial than non-experimental methods. Third, for conductors of experiments it is more difficult to cheat, i.e. to produce the impact they want,

because the evaluators could not choose from a set of estimators. Fourth, experiments provide an opportunity to examine the efficacy of different alternative non-experimental estimators.⁶

However, social experiments also have some important drawbacks. First, they cannot address many questions of interest to researchers or policy makers. As they generate choice-based, endogenously stratified samples that are difficult to use in any other economic question, they only allow the evaluation of the impact of treatment on the treated for one program with one set of participants and eligibility rules (Heckman, LaLonde, and Smith, 1999). Second, social experiments may be hard to accomplish as they entail high costs and ethical issues concerning the use of random assignment. Further practical problems with the implementation of social experiments are mentioned in the literature: the problem of non-compliance, the problem of substitution bias, and of randomization bias. Non-compliance occurs if persons assigned to the treatment group do not participate or if members assigned to the control group participate in the program. Selective non-compliance may lead to biased estimates of the program effects (see Bijwaard and Ridder, 2000). Randomization bias describes the phenomenon if persons selected for the program differ from persons who would participate in the program under usual conditions, i.e. the effects of the program are estimated for an unrepresentative sample. Closely related to that, social experiments may also lead to non-representative estimates due to a Hawthorne effect. In that case, people are aware of the experimental situation and change their behavior. Finally, a substitution bias could occur if members of the control group participate in similar programs to the experimental treatment (see Heckman and Smith, 1995).

2.3 Non-experimental estimators

Whereas in the experimental situation the randomized-out control group provides a direct estimate of the non-treated outcome of the treated, there is no such group available in non-experimental data. Therefore, in a non-experimental evaluation, analysts must replace the missing data with data on non-participants along with additional assumptions invoked when using the method of social experiments since no direct estimate for this counterfactual mean is available and eq. (4) will usually not hold, i.e. $E(Y^0|D = 1) \neq E(Y^0|D = 0)$. Using the unadjusted outcomes of the non-participants to approximate the missing counterfactual in the ATT will lead to selection or evaluation bias:

$$\begin{aligned} \Delta^{ATT} &= E(Y^1|D = 1) - E(Y^0|D = 0) \\ &= E(Y^1 - Y^0|D = 1) + \underbrace{\{E(Y^0|D = 1) - E(Y^0|D = 0)\}}_{= B}. \end{aligned} \quad (5)$$

The term in the curly brackets is the selection bias, B , i.e. the difference between the hypothetical and actual outcomes after non-participation. The reason why this selection bias could not be assumed to be zero with non-experimental data is that participants and non-participants would also have had different non-treatment outcomes even in the absence of the program.

An important share of the non-experimental evaluation literature deals with providing estimators for average treatment effects of receiving or not receiving a binary treatment under the assumption that the treatment satisfies some kind of exogeneity. This assumption, variously referred to as selection on observables by Barnow, Cain, and Goldberger (1980), unconfoundedness by Rosenbaum and Rubin (1983), or conditional independence assumption by Lechner (1998) denotes that the receipt of treatment is independent of the potential outcomes with and without treatment if certain observable attributes are held constant. In his review on non-parametric estimators that are based on this exogeneity assumption, Imbens (2004) distinguishes five classes of estimators that comprise regression, matching on covariates, methods based on the propensity score, combinations of these approaches, and Bayesian methods.

However, in particular the matching estimator has become a popular approach to estimate causal treatment effects. The main reasons for the popularity of the matching estimator are its underlying idea as well as the simplicity of explanation (Heckman, LaLonde, and Smith, 1999). Therefore, matching estimators are frequently used for program evaluation and in consulting business.

Propensity score matching

The basic idea of the matching approach is to find in a large group of non-participants those individuals who are similar to the participants in all relevant pre-treatment characteristics X ('statistical twins'). For that reason, the method appeals to the intuitive principle that it is possible to 'adjust away' differences between participants and non-participants using the available regressors (Heckman, LaLonde, and Smith, 1999). Originated in the statistical literature⁷, matching thus generates a comparison group that resembles an experimental control group in one key respect: conditional on X , the distribution of the counterfactual outcome, Y^0 , of the participants is the same as the observed distribution of the outcome Y^0 of the comparison group (Heckman, LaLonde, and Smith, 1999). In the method of matching, the construction of the correct sample counterpart (for the missing information on the treated outcomes had they not participated) consists in pairing each program participant with one or more members of a comparison group (Blundell and Costa Dias, 2002). Therefore, the matching approach allows to compare treated and non-treated outcomes directly without having to impose structure on the problem. This is the analogy to random assignment in a (social) experiment.

An advantage of the method of matching is its generality due to the non-parametric nature of the approach, i.e. no particular distribution has to be assumed. Furthermore, matching is flexible and may be combined with other methods to produce more accurate estimates of the treatment effects allowing for less restrictive assumptions. One example is the so-called conditional difference-in-differences (DiD) suggested by Heckman, Ichimura, and Todd (1997) that combines matching and the DiD estimator. For the evaluation of the job search assistance

programs, this approach is used by Centeno, Centeno, and Novo (2006) for Portugal and by Blundell, Costa Dias, Meghir, and van Reenen (2004) for the UK. However, since matching methods concern themselves solely with selection of observable variables to solve the selection problem, they require very rich data in order to make the estimates of the treatment effects credible (Smith, 2000a).

The key assumption in the statistical matching literature for the construction of a valid comparison group is that conditional on all relevant pre-treatment characteristics, X , the potential outcomes, Y^1, Y^0 , are independent of the treatment assignment, D (see Rubin, 1977). If the sole parameter of interest is the average effect of treatment on the treated and not the impact on the distribution, for identification of the parameter the conditional mean independence assumption has to be invoked:

$$E(Y^0|X, D = 1) = E(Y^0|X, D = 0) = E(Y^0|X), \quad (6)$$

where \perp denotes independence, and X are covariates that are unaffected by the treatment. As a consequence of eq. 6, the distribution of non-treatment outcomes

$$F(Y^0|X, D = 1) = F(Y^0|X, D = 0) = F(Y^0|X) \quad (7)$$

is independent of the treatment assignment (Heckman, LaLonde, and Smith, 1999). Furthermore, to guarantee that a participant for each non-participant is found, a common support condition is required:

$$Pr(D = 1|X) < 1. \quad (8)$$

Eq. 8 implies that there is an overlap in the distribution of X between the treated and the non-treated group.⁸ Furthermore, it prevents X to be a perfect predictor for treatment or non-treatment respectively. Failure to the common support assumption would lead to biased estimates of the treatment impact as it cannot be identified for all values of X (Heckman, LaLonde, and Smith, 1999). In that case, matching can only be performed within the common support of treated and non-treated individuals. In consequence, the estimated ATT has then to be re-defined for those treated falling in the common support (Blundell, Dearden, and Sianesi, 2004).

If eq. 6 and 8 hold, the ATT in eq. (3) can be rewritten for the matching estimator following Smith and Todd (2005) as

$$\begin{aligned} \Delta_{MAT}^{ATT} &= E(Y^1 - Y^0|D = 1) \\ &= E(Y^1|D = 1) - E_{X|D=1}\{E_Y(Y^0|X, D = 1)\} \\ &= E(Y^1|D = 1) - E_{X|D=1}\{E_Y(Y^0|X, D = 0)\}. \end{aligned} \quad (9)$$

The first term, $E(Y^1|D = 1)$, can be estimated from the observed outcomes of the treated individuals; the second term, $E_{X|D=1}\{E_Y(Y^0|X, D = 0)\}$, can be estimated from the observed outcomes of the (conditional on the X for the treated) matched non-treated.⁹

Matching on X can become hazardous when X is of high dimension (‘curse of dimensionality’, see e.g. Pagan and Ullah, 1999). To deal with this dimensionality problem, Rosenbaum and Rubin (1983) suggest to use balancing scores, $b(X)$. Balancing scores are functions of the relevant covariates X , such that the conditional distribution of X given $b(X)$ is independent of assignment to treatment, i.e. the same for the treated and the non-treated individuals

$$X \perp\!\!\!\perp D | b(X). \quad (10)$$

This means that for treated and non-treated individuals with the same balancing score the distributions of the covariates X are balanced across the treated and the non-treated group. One possible balancing score is the probability of participating in a program, i.e. the propensity score $p(X) = E(D = 1|X)$ that summarizes the information of the relevant covariates X into a scalar index function. Therefore, all biases due to observable covariates can be removed by conditioning solely on the propensity score.

When the propensity score, $p(X)$, is known, the curse of dimensionality for the X can be eliminated; and solving the fundamental evaluation problem requires only to pair treated and non-treated individuals who have the same $p(X)$ as this balances the distributions of X across groups. When the propensity score is unknown, it could be estimated by parametric, semi-parametric or non-parametric methods. However, non-parametric estimation is not preferable since the curse of dimensionality will reappear in the estimation of the propensity score. Therefore, much of the empirical literature uses probit or logit models.

As noted above, the idea of the matching estimator is to find for each treated individual i comparable persons j from the comparison group. Let N_1 denote the number of treated individuals ($D = 1$) and N_0 the number of comparison individuals ($D = 0$). Matches are constructed based on a neighborhood $C(p(X_i))$, where $p(X_i)$ is the propensity score for individual i . Possible matches (neighbors) to treated person i are persons j in the comparison sample whose propensity scores are in the neighborhood $C(p(X_i))$, i.e. $p(X_j) \in C(p(X_i))$. The persons matched to individual i are those in the set A_i , where $A_i = \{j \in D = 0 | p(X_j) \in C(p(X_i))\}$ (Smith and Todd, 2005). With $0 \leq W(i, j) \leq 1$ defining the weight placed on the non-treated observation j for forming a comparison with observation i , the general form of the matching estimator for the ATT is given by

$$\Delta_{MAT}^{ATT} = \frac{1}{N_1} \sum_{i \in \{D=1\}} \left[Y_i^1 - \sum_{j \in \{D=0\}} W(i, j) Y_j^0 \right]. \quad (11)$$

The weights always satisfy $\sum_{i \in \{D=0\}} W(i, j) = 1 \forall i$, i.e. the total weight of all comparisons sums up to one for each treated individual. The literature provides a variety of alternative matching schemes (e.g. nearest-neighbor, kernel density, caliper) to estimate the treatment effects. The different matching estimators vary in the weights attached to the members of the comparison group (Heckman, Ichimura, Smith, and Todd, 1998).¹⁰

Whether the identifying assumption holds or not has raised some discussion recently. The critical question is that the optimizing behavior of the decision makers, e.g. the individual or the caseworkers, precludes their choices being independent of the potential outcomes. Imbens (2004) presents three arguments concerning the reliability of the assumptions. These arguments comprise statistical, data-descriptive, and empirical questions as well as the occurrence of selection on unobservable factors. First, as the natural starting point for any evaluation is the comparison of average outcomes for treated and non-treated individuals, the quality of the comparison may be enhanced by adjusting away any difference in outcomes for differences in exogenous attributes, where attributes are exogenous in the sense that they are not affected by treatment. Although this may not lead to the final word on efficacy of the treatment, its absence would seem difficult to rationalize if one seriously attempts to understand the evidence regarding the impact of the treatment (Imbens, 2004). Second, the empirical question of the evaluation asks which individuals should be compared. Therefore, economic theory on the decision process of treatment may provide some guidance in choosing the variables that need to be adjusted for versus those that do not need to be adjusted for. The conditional mean independence assumption is fulfilled if the researcher observes all variables that need to be adjusted for (relevant covariates). However, if variables that are needed to be adjusted for are not observed, strong assumptions will be required for the identification of the effects of interest. Third, even when agents optimally choose their treatment, two agents with the same values for observed characteristics may differ in their treatment choices. The unconfoundedness assumption must not be invalidated in this case if the differences in the choice are driven by unobserved factors that are themselves unrelated to the outcomes of interest. This may be the case if the objective of the potential participant to participate is distinct from the outcome that is of interest for the evaluator. This third argument is in line with the discussion of Heckman, LaLonde, and Smith (1999) about the validation of the outcomes. Since different persons may value the same state of the world differently even if they experience the same ‘objective’ outcomes, this must be considered in the economic model. A good example is a program that is in part due to paternalistic or altruistic preferences. In that case, allocation of individuals may be guided by equity concerns, whereas evaluation may focus on program efficiency. While the efficiency criterion focusses on maximizing the social return to a public program investment, i.e. it concentrates on groups for whom the impact is largest, the equity criterion aims at groups who are most in ‘need of services’.

Multivariate duration models

An alternative approach of program evaluation that has been applied frequently in the context of job search assistance programs are duration models. In contrast to propensity score matching, longitudinal information is used to identify the treatment effects. By imposing more structure to solve the selection problem, consideration of observable and unobservable influences becomes

possible. However, to take account of possibly endogenous participation, the recent literature suggests the use of multivariate models where the interdependence of the transitions of interest (transition into program and transition into employment) are estimated simultaneously.

A useful multivariate duration framework for the evaluation of treatment effects is introduced by Abbring and van den Berg (2003) based on the population of inflows into unemployment. To illustrate the approach some further notation is useful. The duration until the individual enters employment (T_e) and the duration until he/she joins a program (T_p) are measured from the point of time an individual enters unemployment. T_e and T_p are assumed to be non-negative and continuous random variables with realizations denoted as t_e and t_p . The durations T_e and T_p are assumed to vary with time-invariant observable characteristics (x) and time-invariant unobservable characteristics (v_e, v_p). For the observable characteristics, no exclusion restrictions are necessary, i.e. the observable characteristics could be the same for both durations. In addition, unobserved variables may affect the realizations of the event. Therefore, let v_e capture the unobserved heterogeneity of T_e and v_p the corresponding unobserved heterogeneity of T_p .

The fundamental assumption of the model is that any dependence between T_e and T_p conditional on x and (v_e, v_p) stems from the causal effect of T_p on T_e . Then, the joint distribution $T_e, T_p | x, v$ is the product of the conditional distributions $T_e | T_p, x, v$ and $T_p | x, v$. Assuming further that $T_e, T_p | x, v$ is absolutely continuous the conditional distributions in terms of their hazard rates could be specified (Abbring and van den Berg, 2004). It is common, to specify both hazard rates as mixed proportional hazard (MPH) models,

$$\theta_e(t | t_p, x, v_e) = \lambda_e(t) \exp(x' \beta_e) v_e \mu(t - t_p, x, v)^{I(t > t_p)}, \quad (12)$$

$$\theta_p(t | x, v_p) = \lambda_p(t) \exp(x' \beta_p) v_p. \quad (13)$$

The hazard rate for the transition into employment (eq. 12) at time t consists of a baseline hazard $\lambda_e(t)$, a systematic part $\exp(x' \beta_e)$, the unobserved heterogeneity term v_e , and the treatment effect $\mu(t - t_p, x, v)^{I(t > t_p)}$. The underlying assumption of the MPH specification is that duration dependence and individual heterogeneity enter the hazard multiplicatively (see Lancaster, 1979). The duration dependence, i.e. the shape of the hazard over time, is represented by the baseline hazard. Individual heterogeneity is regarded by the systematic part and the unobserved heterogeneity term. It is common to MPH models to specify the systematic part such that $\theta_e(t | t_p, x, v_e)$ and $\theta_p(t | x, v_p)$ are multiplicative in each element of x . The transition rate from unemployment into program (eq. 13) is specified analogously with baseline hazard $\lambda_p(t)$, systematic part $\exp(x' \beta_p)$ and unobserved heterogeneity term v_p . In contrast to the method of matching, assuming a MPH specification imposes more structure on the problem of potential selection bias and causal effects of treatment. In particular, assuming proportionality of the baseline hazard rates for all individuals and that the single factors affect the transition rate multiplicatively has to be discussed carefully in each case.

The treatment effect $\mu(t - t_p, x, v)^{I(t > t_p)}$ represents the causal effect of t_p on the hazard rate $\theta_e(t|t_p, x, v_e)$, where $I(t > t_p)$ is an indicator function taking the value 1 if $t > t_p$. The treatment effect can be interpreted as a shift of the hazard rate by $\mu(t - t_p, x, v)$ that is directly associated with the expected remaining unemployment duration. In that sense, a positive treatment effect will shorten the expected remaining unemployment duration. Hence, in the most general specification, the treatment effect is allowed to depend on the time since treatment has started ($t - t_p$), on observable characteristics x , and on unobservable factors v . For this reason, the model is very flexible with regard to consideration of different types of heterogeneity, but computation is burdensome. Hence, many applications consider simpler cases of the treatment effect only (depending on the time since treatment or observable characteristics only).

The basic assumption of the model is that any selectivity relates to observable and unobservable factors. Technically, selectivity means that those individuals observed to receive a treatment at t_p are a non-random subset with respect to t_e . Whereas any selectivity conditional on observable characteristics is captured by the systematic part in eq. (12), possible selection on unobservable factors is captured by a dependence of v_e and v_p . Generally, it could be assumed that (v_e, v_p) is a random vector with distribution function $G(v_e, v_p)$ independent of x . If selectivity cannot be fully captured by the observable characteristics, a dependence of the unobserved heterogeneity terms will be observed. Then, the indicator function for the treatment effect appears as an endogenous time-varying regressor.

A further important aspect of the model is the consideration of the information on the timing of the treatment within the unemployment spell. As Abbring and van den Berg (2003) demonstrate, this additional information conveys useful information on the treatment effect in the presence of selectivity. It enables a distinction to be made between time-invariant selection effects embodied by observable and unobservable characteristics and a causal treatment effect that becomes effective at the moment the treatment starts. A positive causal treatment effect leads to a pattern where a transition into employment is typically realized very quickly after a transition into treatment, no matter how long the elapsed duration of unemployment is. In contrast, in the case of a selection effect there would be a correlation between the points in time of the transitions into employment and the program. For example, a positive selection effect results in a pattern where a quick transition into the program is followed by a quick transition into employment, i.e. both transitions occur very rapidly after the unemployment spell has started. Thus, the main difference between a treatment and a selectivity effect is that the treatment effect affects the transition rate into employment only after treatment has been realized whereas selectivity affects the transition rate everywhere.

Identification of the treatment effect requires that individuals do not anticipate future treatments. Anticipatory effects are present if, for example, those individuals who are informed about a future program reduce their search activity in order to wait for the program. In this case, the hazard rate at t of an individual who anticipates a future treatment at time t_p , will

be different from the hazard rate of an individual who obtains an alternative treatment at time t_p^* for $t \leq \min\{t_p, t_p^*\}$.¹¹ Due to the anticipatory effect, the information on the timing of the event would not be sufficient for identification since a causal change of the hazard occurs at the moment the information shock of the treatment arrives. However, information on the timing when people start to anticipate future participation is usually not available. Therefore, convincing argumentation in favor of the no-anticipation assumption to hold in the specific case is crucial. Nevertheless, it should be noted that the assumption of no anticipatory effects does not rule out that individuals act on the determinants of T_p . In other words, individuals are allowed to adjust their optimal behavior to the determinants of the treatment process, but not to the realizations of t_p . Abbring and van den Berg (2003) prove that with assumptions similar to those made in standard univariate MPH models, the bivariate model in eqs. (12) and (13) and the treatment effect in particular are identified. The identification is nonparametric, since no parametric assumptions with respect to the baseline hazard and the unobserved heterogeneity distribution are required.

Having outlined the empirical methods used to evaluate the treatment effects of job search assistance programs, I will now turn to the characteristics of the programs and discuss the results of the empirical studies thereafter.

3 Set-up and Eligibility

3.1 Registration, job search requirements, and participation

In recent years, many European economies have adopted policy measures that aim at increasing the intensity and efficiency of job search and at mitigating the disincentive effects of unemployment compensation on labor supply. By adoption of these activities, a number of changes within the mixture of labor market policy programs have been established: most European countries have strengthened the counseling and job brokerage activities and have introduced measures to check compliance with eligibility conditions and job search requirements. These requirements usually cover a frequent reporting of the job search activities. If the requirements are not met, sanctions by benefit revocation to enforce job search obligations and/or acceptance of suitable job offers could be imposed. ALMP programs should force job seekers to search more actively for work. In addition, most European countries have developed activation strategies to coordinate benefit administration for public unemployment insurance and spending on ALMPs. Economic theory and empirical results suggest significant interactions between unemployment benefits and activation policies, meaning that disincentive effects associated with over-generous unemployment benefits (high and long-lasting) can be, at least partially, counteracted by the adoption of well-designed ALMPs.

However, although over-generous payments could lead to a potential abuse of the unemployment

insurance system by claimants if they decide to reduce search efforts, this argumentation is not without doubt. Meyer (1995) has raised this issue, an empirical investigation is provided by Ashenfelter, Ashmore, and Deschênes (2005). The results provide no support that failure of the benefit claimants to actively seek for work has been caused by overpayments in the unemployment insurance system. Despite that evidence, increasing the frequency of reporting job search efforts is part of the job search assistance programs in all countries considered in this review.

Include Table 1 about here

In order to make job search programs comparable across countries, the characteristics of the unemployment insurance systems as well as the specific characteristics of the programs have to be regarded. Table 1 summarizes selected characteristics of the unemployment systems and the process of assignment to job search programs useful for the analysis.

A first obvious thing to note is that European countries' labor market policy do not differ much (anymore); an important reason for that may be seen in the collective agreements within the European Union of the past years. In the majority of countries reviewed, benefit entitlement starts with registration at the employment office (Austria, Denmark, Germany, Hungary, and Portugal) or even before (France, the Netherlands, the UK). The only exception is Finland where benefit entitlement starts after registration. In addition, most countries have adopted strategies to intervene early in the unemployment spell. As becomes obvious from Table 1, compulsory interviews during the first month of unemployment are common in most of the countries, and typically a profiling is conducted and/or an individual action plan covering the requirements of job search for the unemployed persons is set up. Moreover, unemployed persons are forced to report the job search efforts frequently; in intervals ranging from every two weeks (UK) to about every two months (Germany), but variable requirements are possible (Portugal) as well.

With regard to the assignment rules of participation in ALMP programs there are some smaller differences between countries. Although participation is compulsory in all countries if persons are assigned to the program by the Public Employment Service, only Denmark, the Netherlands, and the UK have compulsory participation after some specific duration of unemployment. Of those three, the activation policy of the Netherlands postulates participation after one month of unemployment already; the requirements in the UK are more relaxed with up to 22 months of unemployment prior compulsory participation (for adults). Moreover, there are heterogeneous requirements of participants to continue job search during participation. Continuing job search is only required for participants in Denmark, Finland, and Germany; the other countries do not impose comparable rules.

By and large, despite these smaller differences mentioned labor market policy is quite comparable across countries. All countries favor the use of close monitoring of the job seekers comprising

a frequently reporting of the activities. Moreover, if job seekers are assigned to ALMP programs, participation is compulsory in the majority of cases. The requirement to continue job search during participation imposed by some countries may speed up exit to employment. In contrast, in countries that do not postulate continued job search during participation, locking-in effects may be possible. Locking-in effects occur if participants reduce their job search efforts during participation leading to a smaller transition probability of leaving unemployment (or the program) for employment. However, in both cases the quality of jobs and the matching quality has to be regarded. In sum, given the similarity of the European labor market systems, large differences in impacts of the job search assistance programs due to the general framing of the labor market policy could not be expected.

3.2 The programs

Although the job search programs in the different countries aim at shortening the duration of unemployment by increasing the job chances of the job seekers (or the job match probability), they differ in a number of respects. Some countries provide intensified counseling schemes with frequent interviews over the unemployment spells only (Hungary, the Netherlands, and RESTART program in the UK), whereas in other countries additional activation programs or even qualification programs are an integral part of the job search assistance programs. A key difference to more traditional ALMP programs (e.g., employment subsidies or vocational training) is the modular set-up of the programs. For example, Denmark offers at first a job search program for up to two weeks that could be extended by an activation program with a possible duration of up to three months if the persons remains unemployed. In a similar way, job search programs in Germany consist of three modules that could be arranged variably covering job search assistance, intensified counseling and short-term training programs providing specific skills and techniques. This set-up is comparable to the New Deal for Young People in the UK, where a combination of individual job search assistance followed by subsidized courses is offered. Similar combinations of programs are available in Austria, Finland, France, and Portugal.

Include Table 2 about here

Programs differ across countries with respect to the nature (whether programs are mandatory or voluntary) and the duration. Most programs are mandatory for the participants, i.e. employment officers are allowed (or even forced) to sanction benefit entitlements of the job seekers if they do not comply with participation. The only exception is Finland where programs are arranged with voluntary entitlement of participants. With regard to the duration of programs the differences are quite large. Depending on the type of program (or module), programs' durations range from repeated half-an-hour interviews over some pre-defined interval of unemployment (e.g. RESTART in the UK, or Counseling & Monitoring in the Netherlands) over one day profiling courses (e.g. skill assessment in France) to on average two to four weeks programs teaching

job search skills and further specific skills and techniques (e.g. in Germany, Finland, Denmark, and France). In addition, due to the modular set-up of many programs lasting durations for up to several months are possible. However, these longer program durations usually do not refer to full-time courses but usually to repeated frequent meetings of job seekers and caseworkers (e.g. in Hungary, the Netherlands, Portugal, or France). Nevertheless, in contrast to traditional further training programs that usually last for a couple of months or even several years, the job search assistance programs are all considerably shorter.

4 Programs' effects

Two questions lay at the core of this paper. The first is: How effective are job search assistance programs in European countries in bringing people back to work or shortening the duration of unemployment? According the answer of the first question, the second question asks: What determines effective - or respectively - less effective programs? To answer both questions, a number of complications have to be considered. First, as shown above, programs differ with regard to eligibility criteria and the contents provided across countries. Therefore, there exists no 'the' job search assistance program. However, by taking account of the single building blocks of the measures provided, comparison could potentially identify more and less effective modules and tools. A second complication arises from differences in the evaluation method used to estimate program effects. As discussed in section 2, the estimators applied to solve the selection problem - experimental and non-experimental estimators - have different underlying assumptions that are empirically not testable. Thus, different estimators may lead to different parameter estimates of treatment effects for the same program. Finally, the studies selected for this review consider different types of outcome variables. Besides others, the spectrum comprises transition rates to employment, unemployment rates, matching probabilities, life satisfaction etc.

Despite these differences, consideration of evaluation studies from 9 different countries, namely Austria, Denmark, Finland, France, Germany, Hungary, the Netherlands, Portugal, and the United Kingdom, allows to identify patterns in effects and effectiveness of programs. A summary of the evaluation studies by country is provided in Table 3. However, comparison of programs country by country may be afflicted by the problems of the different evaluation methods applied. For that reason, I will organize the presentation of the empirical results according to the different estimation methods.

Include Table 3 about here

4.1 Results from experimental data

Social experiments to evaluate job search assistance programs have been conducted in 7 of 9 European countries reviewed here for which studies are available. Given the aloofness of European countries against social experiments that existed for a long time, this number provides interesting evidence that the situation has changed in recent years. The picture of the empirical results of the impacts of the programs is mixed. With the exception of Gorter and Kalb (1996) who report negative effects on finding a permanent job for formerly temporary employed job seekers of a job search assistance program in the Netherlands, all experimental studies establish no effects or small positive effects of job search assistance programs. What determines these positive effects? To answer this question, the results are ambiguous to some extent. Good examples are Denmark and the UK. Although Graversen and van Ours (2008) find positive effects on the transition rate to employment (conditional on time upon program entry) of the job search and job activation program, the authors emphasize that not the activities provided are responsible for the increased transition probability but the threat and punishment job seekers fear. In a similar way, Dolton and O'Neill (2002) who evaluate the effects of the UK RESTART program (targeted to long-term unemployed persons) report positive effects for men in terms of reduced unemployment rates but no effects for women. Here, the authors distinguish short-run and long-run effects; whereas services provided in programs are responsible for the long-run results, short-run effects are mainly determined by the associated threat components of the program.

Other studies reporting positive after program effects do not follow that interpretation. Centeno, Centeno, and Novo (2006) for Portugal establish positive effects of a program for long-term unemployed (REAGE) in terms of a shortened duration of unemployment. Positive effects of a job search assistance program on the reemployment probability and on life satisfaction of the participants are also found by Vuori and Silvonen (2005) for Finland (Tyhönön Job Search Program). Based on a social experiment in Germany, Büttner (2007) shows that participation increases the employment rate for young unemployed persons only. Finally, for Hungary where Micklewright and Nagy (2005) analyze the effects of a job search program on the transition rate to employment, positive evidence is given for women aged over 30 only.

In contrast to that, for other groups of participants in job search assistance programs no effects are established. Hämäläinen, Uusitalo, and Vuori (2007) who analyze the effects from two experiments in Finland (one is the same as that analyzed by Vuori and Silvonen, 2005) cannot establish any significant effects from the experimental estimators, and, therefore, partly contradict the results of Vuori and Silvonen (2005). Also Büttner (2007) for Germany reports zero results for job seekers who participate in internships within the job search assistance program. Moreover, in contrast to his positive results for young unemployed persons the evidence for Hungary (Micklewright and Nagy, 2005) establishes insignificant effects for this group. A

similar finding also holds for the so-called INSERJOVEM program in Portugal that is targeted to young unemployed persons but the analysis of Centeno, Centeno, and Novo (2006) does not obtain positive results for that group.

In the Netherlands, evidence on a social experiment conducted in 1989/90 is available from Gorter and Kalb (1996); in addition, another social experiment conducted in 1998/99 has been evaluated by van den Berg and van der Klaauw (2006). Gorter and Kalb (1996) find no or weak positive effects on the transition rate to employment for persons with a permanent job before; the results of the program analyzed by van den Berg and van der Klaauw (2006) are similar with at best small positive effects on the exit rate to work.

4.2 Results from non-experimental data

In addition to social experiments it is quite common to evaluate treatment effects of programs based on non-experimental data available for already implemented ALMP programs. To solve the problem of potential selectivity bias in the estimates, two approaches are mainly used throughout the empirical studies: propensity score matching estimators and multivariate duration models. As described above (section 2) both methods differ with regard to data requirements and identifying assumptions of the treatment effect. To start with, I will first review the results obtained from propensity score matching estimators (and related methods) and discuss the results of duration models thereafter.

Propensity score matching and variants

Job search assistance programs in Germany have been empirically evaluated by Biewen, Fitzenberger, Osikominu, and Waller (2007) and Lechner and Wunsch (2008). Both studies apply a cross-section propensity score matching estimator on administrative data of the Federal Employment Office. By using the same comprehensive data base, program heterogeneity is considered in detail in both studies. However, whereas Lechner and Wunsch (2008) compare participants to non-participants (over the whole period of analysis), Biewen, Fitzenberger, Osikominu, and Waller (2007) apply an approach suggested by Sianesi (2004) that defines participation dynamically over the unemployment spell. Here, unemployed persons are defined as non-participants as long as they do not participate in a program or leave unemployment for work. Therefore, participation and non-participation depends on the timing of comparison and effects are compared between today's participants and today's non-participants, who are potentially tomorrow's participants. A further difference of both studies is that Lechner and Wunsch (2008) concentrate on West Germany only but distinguish different sub-programs of job search assistance (short-term training measures, short-combined measures, and job search assistance).

The overall picture revealed by Lechner and Wunsch (2008) is more disappointing than the experimental evidence provided by Büttner (2007) (for Germany). On average, none of the

sub-programs considered reveals positive effects in terms of reducing the individual unemployment rate and job search assistance programs tend to have negative effects on the employment chances of the participating individuals, too. Nevertheless, some heterogeneity in the effects is obtained for selected groups of the unemployed. People with a vocational education, long-term unemployed persons and low-qualified persons experience an increase in the employment chances due to participation in a combined program that aims at removing specific skill deficits or in a short-term training program. In contrast to that, the results of Biewen, Fitzenberger, Osikominu, and Waller (2007) are more positive. Here, contradicting the result by Lechner and Wunsch (2008) programs provided in West Germany lead to clear positive effects on the employment rate. In addition, East German programs are not as effective and participants experience no or only a small increase in the employment chances due to participation.

Since both studies use the same data, a similar estimation approach, and investigate the identical program the explanation for the different results obtained is given by the different definition of participation and non-participation. Whereas Lechner and Wunsch (2008) compare participation to a no-program state, interpretation of the program effect is straightforward. However, in a comprehensive system of labor market policy programs, participation at different points of time in the unemployment spell is possible and, therefore, using the no-program state as the comparison requires conditioning on future outcomes at the time of treatment start. Hence, the no-program group may be particularly selective with respect to good job chances (i.e. no activation is required) or bad job chances (i.e. activation is denied due to cream-skimming). In that sense, the estimated treatment effects may be biased. If one assumes the first of the two possible directions of bias, the estimates provided are downward biased and would explain the disappointing findings. To mitigate this problem, Lechner and Wunsch (2008) conduct a pre-match of participants and non-participants with respect to the distribution of starting dates. The probability distribution of starting dates in the participants' group is estimated and is matched based on observable characteristics to the non-participants' group. Non-participants are kept in the analysis based on these estimated potential starting dates. However, this approach still includes a conditioning on future outcomes since it depends on the time horizon of the analysis.

In contrast to that, Biewen, Fitzenberger, Osikominu, and Waller (2007) do not have to condition on future outcomes in the analysis as they define participation and non-participation dynamically. However, this complicates interpretation of the estimates for two reasons. First, estimated effects could only be interpreted conditional on the timing of treatment start in the unemployment spell, i.e. the effect of a program that has started in the first quarter of unemployment is different from the effect of a program that has started in the third quarter of unemployment since the timing of the treatment start in the unemployment spell is interpreted as an integral part of the program effect. Second, since the comparison requires non-participation up to the point in time the participants start the program only, the individuals of the compar-

ison group are likely to participate later in time. Particularly if matching is successful, persons with comparable participation probabilities will be matched. Hence, program effects may be compared to postponed entry into programs only. As programs are not counted as employment, this leads to increased estimates of the treatment effect on average.

For the evaluation of the New Deal for Young People in the UK, Blundell, Costa Dias, Meghir, and van Reenen (2004) use a combination of a difference-in-differences estimator and the matching estimator as suggested by Heckman, Ichimura, Smith, and Todd (1998). Here, matching is used to rule out selectivity in the cross-section; in addition, participants and non-participants may be affected differently by macroeconomic events. The difference-in-differences estimator takes account of these time effects that are independent of the program.¹² The empirical estimates are positive and complement the picture revealed experimentally by Dolton and O'Neill (2002) of the RESTART program. With regard to the two outcome variables considered (share of persons leaving unemployment, outflow into employment) the New Deal program is effective for men and women. However, it has to be noted that the program is intended to promote young unemployed persons, whereas the programs in Germany are addressed to all job seekers. Hence, generalizing the findings of the UK to the whole group of job seekers would require strong further assumptions.

Multivariate duration models

Multivariate duration models have been applied to evaluate the effects of job search assistance programs in Austria, France, and Germany. The studies for Austria (Weber and Hofer, 2004a; 2004b) and Germany (Hujer, Thomsen, and Zeiss, 2006) analyze the effects of programs on the transition to employment; Crépon, Dejemeppe, and Gurgand (2005) for France further take account of recurrence of unemployment. Similar to Germany, the job search assistance program in Austria is not targeted to certain groups of unemployed persons. In addition to job counseling it contains a job application course that should be provided to all job seekers before completing the first four months of unemployment. As noted above, the job search assistance program in Germany offers optional short-term training courses to the job seekers as well. The program used in France has a modular set-up comparable to Germany and includes project support activities targeted to low-ability unemployed persons.

In contrast to the experimental studies and non-experimental studies using propensity score matching methods all studies applying multivariate duration models independently of the country exhibit positive effects of programs on the outcome variables in consideration (except for one group in Austria). For France, Crépon, Dejemeppe, and Gurgand (2005) find positive effects in terms of a reduced average unemployment duration as well as a lower probability of unemployment recurrence after participation. The results for Germany by Hujer, Thomsen, and Zeiss (2006) indicate an increase in the transition rate to employment after participation. In

addition, individual heterogeneity in the effects shows that job search programs are particularly effective for low-skilled unemployed men and women. The treatment effect on the hazard rate is modeled in a flexible way, i.e. it depends not only on the timing of treatment in the unemployment spell but is allowed to change with time after the start of the program. For this reason, it is possible to analyze the half-life of treatment effects. The results show that programs affect the transition rate only in the short- to medium-run after participation, i.e. participants' hazard rates increase directly after participation but the effect are diminishing after a couple of months. Finally, after about one year after the start of program no differences between participants and non-participants remain. This finding is interpreted as a depreciation of the program's content over time; the abilities taught in the courses as well as the value of intensified counseling start to decrease after participation.

A similar effect is found by Weber and Hofer (2004b) for the Austrian program. Weber and Hofer (2004a) evaluate the average treatment effect with regard of the timing of treatment in the unemployment spell, but do not present different program effects conditional on the preceding unemployment duration. The results indicate a strong positive effect on the hazard rate to employment. In contrast, Weber and Hofer (2004b) extend the approach by estimating treatment effects for pre-defined unemployment durations. The detailed analysis shows that only programs provided during the first year of unemployment have positive effects on shortening the duration of unemployment; for long-term unemployed persons (with a duration of more than one year), the effects are reversed and participation prolongs unemployment duration on average.

4.3 Cost benefit considerations

In contrast to empirical studies of US programs, in evaluation studies of the European programs cost benefit considerations are scarce. Only three of the studies considered in this review provide cost benefit analyses, and these are quite short or more or less back-of-the-envelope calculations. These analyses are limited to costs for the UI system and do not consider social costs and benefits of programs in addition. Therefore, all cost benefit analyses provided depend strongly on the estimated individual treatment effects of the programs. For the UK, Dolton and O'Neill (2002) analyze the cost-benefit relationship of RESTART. Associated with the positive estimates of the program in terms of reduced unemployment rates, the result shows clear positive effect for participating men in the long-run (with an estimated gain of 494 British pounds). In contrast to that, the average short-run gain (for about one year) is negative but not significant (about -24 British pounds), or slightly positive (86 British pounds) but the latter refers to an unrestricted sample, i.e. job placement in the sample is not only due the RESTART activities. Therefore, the program seems to be cost-effective for males and tend to be cost-effective for women. Compared to the results summarized by Meyer (1995), benefits are larger than in the US. In a similar way, van den Berg and van der Klaauw (2006) provide a cost benefit analysis on the individual level.

Programs tend to be cost-effective by saving about 903 Euro on average (but insignificant) if constant treatment effects over time are assumed. However, the cost analysis provided depends strongly on the assumptions imposed on the development of the treatment effect over time. In a different scenario, where a fading out of the treatment effect until 6 months after the start of the program is assumed, the return shrinks to 56 Euro (and is again insignificant).

In contrast to that, Lechner and Wunsch (2008) for Germany find a negative cost benefit relationship of participation. Compared to remaining on unemployment benefits, the extra costs of the program (course costs, tuition, etc.) together with the estimated prolongation of the unemployment duration of the participants result in costs exceeding the benefits by about 2,000 to 2,500 Euro per participant. Hence, Lechner and Wunsch (2008) rate German job search assistance programs inefficient from the perspective of the UI system. Nevertheless, compared to other programs in Germany, these estimates are clearly smaller. However, social costs and benefits as well as the expected future contributions to social security are not regarded in the static analysis.

5 Discussion

Job search assistance programs are provided in many countries to improve the reemployment prospects of unemployed persons and to reduce the budgetary pressure on the unemployment insurance systems. The empirical evidence for the 9 countries reviewed for which meaningful empirical studies are available clarifies that several different combinations of services are subsumed in the job search assistance programs. The main differences refer to the type of job finding services provided, whether or not additional short-term training courses and work-tests are an integral part of the program, and to the reporting requirements of the job seeker to the responsible caseworker. Nevertheless, despite this variety quite a number of different approaches seem to be successful in reaching the intended goals. In addition, since job search assistance programs improve the matching of workers and jobs, they are unlikely to encourage job seekers to find a job quickly at the expense of others who are displaced by those in contrast to more traditional programs.

To allow for a generalization of the empirical effects, it is useful to identify the different effects these programs exhibit within a theoretical framework. A formal discussion within a job search model is provided by Hujer, Thomsen, and Zeiss (2006); for the case at hand, I will refer to the main implications of that model. On the one hand, job search assistance programs attempt to improve the job placement on side of the employment agency as well as the self-contained job search of the participants. On the other hand, programs contain training elements that are used to adjust the qualification of the job seekers to the demand of the market. By simplifying the main characteristics of job search assistance programs that way, two channels of how programs affect the job search could be deduced. With regard to the first channel, one can expect an effect

on the search behavior of the participants by increasing the intensity as well as the efficiency of the job search efforts. Hence, job search could be assumed to be more productive leading to a reduction of unemployment duration. The second channel that teaches job relevant skills may improve the job opportunities of the participants by allowing them to apply for jobs which are on average associated with higher wages.

However, both channels do not work in the same direction. Within a search model framework (see Mortensen, 1986, and van den Berg, 1994) the first channel of providing job search assistance will have a direct positive effect on the job search efficiency on the one hand. However, an increased job search efficiency could be translated into a higher job offer arrival rate that corresponds to an increase in the reservation wage. This increased reservation wage may exhibit an indirect negative effect on the transition rate to employment on the other hand. Hence, increasing the job search efficiency directly lowers the unemployment duration of the individual, but makes the workers more selective with respect to the wage offers at the same time. The second channel improves the job-relevant skills of the participants, and, therefore, increases their job opportunities. This is equivalent to a raised productivity. Therefore, participants could apply for jobs which are on average associated with higher wages. An increased mean of the wage offer distribution increases the hazard rate of leaving unemployment for employment since the reservation wage increases by less than the mean of the wage offer distribution. Hence, for the given higher mean the workers are less selective with respect to the wage offers. Given this pattern of the two channels of the effects, knowing the impact of a job search assistance program with regard to employment is impossible *ex ante*. It depends on the relative importance of the two channels in the case analyzed.

By and large, the empirical evidence from social experiments for Europe shows that job search assistance programs could lead to positive effects in terms of increased employment rates and reduced unemployment, but not every program is successful. The positive findings are in line with the experiences of the US (see Meyer, 1995) and support the use of these activities to activate the unemployed. However, even in cases where no short-term effect of the program is observable participants' employment chances seem not to be harmed by participation. Explaining the success of the programs requires to investigate the relative importance of increased services (including counseling and the job search programs) and work search requirements in the determination of the labor market outcomes. Here, the conclusions differ for the US and the European evidence. For the US, Ashenfelter, Ashmore, and Deschênes (2005) note that a stricter verification of the search efforts does not contribute to a decrease in the duration of unemployment, i.e. the threat of enforcing job search seems to be ineffective. Hence, they argue that subsidized job search assistance plays the major role for the empirical findings of job search assistance programs summarized by Meyer (1995). In contrast, for Europe Dolton and O'Neill (2002) and Graversen and van Ours (2008) relate the positive effects of the programs analyzed to the threat component in particular.

The value of experimentally obtained estimates may be affected by the problems described in section 2. In addition, Meyer (1995) notes that the caseworker at the program sites may not be representative of the entire pool of service providers. If they differ substantially in their experience and ability, the difference in outcomes may be biased by this personnel ability. Nevertheless, since the results from the non-experimental studies complement the findings of the social experiments, this bias seems to be of minor importance for the empirical results of job search assistance programs reviewed here.

The picture of the effects of job search assistance programs revealed from the studies applying variants of propensity score matching estimators is mixed. Whereas some studies establish clear positive effects, others do find positive effects for selected sub-groups only, or even negative effects. The non-experimental evidence is complemented by the studies applying multivariate duration models. The cross-country comparison shows that intensifying counseling activities and providing job application abilities to job seekers clearly improve the chances of leaving unemployment for employment. However, the skills provided during short-time courses do not last very long. Effects of programs therefore diminish over time after participation. In addition, offering short-courses teaching specific skills and abilities (like in France or Germany) tend to help low-skilled unemployed persons and long-term unemployed persons in particular. In contrast, purely supporting job search activities (like in Austria) is a means for short-term unemployed persons with a high employability only; long-term unemployed persons could experience a prolonged unemployment duration if no further training of abilities is provided.

How to manage the use of job search assistance programs? An important study that analyzes this question is provided by Wunsch (2007). She simulates the optimal use of active and passive labor market policy calibrated to parameters for West Germany taken from Lechner and Wunsch (2008). A particular focus is laid on the trade-off between the aims of the programs, i.e. ALMP programs should increase the exit rates to employment by improving job search effectiveness or skills, and that programs are costly, i.e. course costs usually exceed unemployment benefits and participation may reduce the time available to actively search for work (locking-in effects). With regard to the job search assistance programs, the results of the simulation show that programs should be used early in the unemployment spell as well as after depreciation of the search effectiveness. In that case, despite the fact that no positive effects could be established by Lechner and Wunsch (2008), the simulation results suggest that a more-targeted use could lead to an increase of the exit rate of unemployment of about 2 to 10 percentage points per half-month of unemployment.

Moreover, she distinguishes certain levels of job search effectiveness of the unemployed. With regard to that the results of the simulation indicate further that job search assistance programs should not be used for people with the highest level of search effectiveness; this result, however, is contradictory to the empirical finding of Weber and Hofer (2004b) for Austria. In addition, Wunsch (2007) could not recommend the use of job search assistance programs for the lowest

qualified job seekers since returns to work are too low for that group. Although this seems contradictory to the finding of Hujer, Thomsen, and Zeiss (2006) who report particularly strong effects for low-qualified men and women, it has to be noted that the study does not consider the cost benefit relationship of the program.

Taking account of the cost benefit relationship of job search assistance programs the simulations of Wunsch (2007) show that a shortening of the programs to a duration of between one half to two months may be optimal; moreover, the accumulation rate of job search effectiveness must exceed 3% per half-month for the job search assistance program to be successful. If that could be reached budget savings of 3 to 8 percent could be realized compared to an optimal labor market policy without job search assistance programs. Clearly, the results of Wunsch (2007) are calibrated to job search assistance programs in West Germany; generalizing the findings to other countries may reveal different figures.

6 Conclusion

The aim of this paper was to condense the findings of the empirical effects of job search assistance programs in Europe. These programs are part of ALMP in many countries and aim at increasing the employment chances of the unemployed persons by providing assistance with the job search activities, like an intensified counseling and a frequent reporting of the job search efforts, and short subsidized courses to enhance the quality of applications but also the skills of the individuals. Although the programs differ across countries with respect to the spectrum of services provided and the relative weight associated to the single services, the empirical evidence for the different programs tends to be quite positive. These positive effects seem to be due to the improved matching of job seekers and jobs, but also due to the threat component of programs. This threat component comprises the possibility of the job seeker to be sanctioned by benefit revocation in case of non-compliance. However, not every program is effective; the empirical evidence shows that particularly programs that combine a number of different services, e.g. intensified counseling and job application training, are effective, whereas programs that rely on a sole intensified counseling only show no or only modest positive effects.

The robustness of the findings is proved by the variety of empirical methods applied. On the one hand, experimental estimates may lead to consistent estimates of programs' impacts if random assignment of participants is conducted carefully. However, social experiments may suffer from a number of effects related to the experimental situation. Therefore, generalizing the findings may require strong assumptions. Nevertheless, the majority of the experimental evidence shows zero or positive effects of job search assistance programs in terms of a reduced unemployment duration and increased employment rates. This evidence is complemented by a bulk of studies using non-experimental data; here, variants of propensity score matching methods and multivariate duration models are used frequently to evaluate the effects of job search

assistance programs. Framing effects that may apply to social experiments could be excluded but estimates may suffer from violations of the non-testable identifying assumptions (the latter would hold for social experiments as well, however, it is more likely that the identifying assumption holds if randomization is successful). Despite the different methods applied the non-experimental evidence complements the picture revealed experimentally. Independently of the country the majority of the job search assistance programs exhibits positive effects on both, the decrease in the unemployment duration and the increase in the employment rates.

Acknowledgements

I would like to thank the the *Stifterverband für die Deutsche Wissenschaft* (Claussen-Simon-Stiftung) for financial support. The usual disclaimer applies.

Notes

¹For Germany for example, there exist a number of recent further studies evaluating the effects of job search assistance programs, e.g., Wolff and Jozwiak (2007), Osikominu (2008) and Fitzenberger, Orlyanskaya, Osikominu, and Waller (2008). However, since they apply similar empirical methods and do not differ much in the results they establish I refrain from taking account of these.

²The model has been extended for the case of multiple treatments by Lechner (2001) and Imbens (2000).

³Alternatively, in the case of J mutually exclusive treatments (e.g. for the case of evaluating different ALMP programs), D could be an indicator for the $J+1$ possible states the individual faces. D could also be $\mathbb{R}_+ := [0, \infty)$, representing a continuum of doses of some medication, for example (see Abbring, 2003).

⁴Heckman, LaLonde, and Smith (1999) discuss further parameters that may be of interest: for example, the average effect of treatment (ATE) defined as:

$$\Delta^{ATE} = E(\Delta) = E(Y^1 - Y^0) = E(Y^1) - E(Y^0).$$

The ATE computes the difference of the expected outcomes after participation and non-participation. It answers the question what the impact of treatment would be if individuals are randomly assigned to treatment. However, for policy implications it is only of minor relevance as persons are included for whom the program was never intended (Heckman, 1997). Further parameters of interest may be the proportion of people taking the program who benefit from it, or the increase in the proportion of outcomes above a certain threshold outcome value due to a policy.

⁵See Heckman, Ichimura, and Todd (1997), for a description of how randomization solves the evaluation problem.

⁶The interested reader is referred to the paper of LaLonde (1986) and the responses and extensions by Dehejia and Wahba (1999; 2002) and Smith and Todd (2005).

⁷See for example Rubin (1974; 1977; 1979; 1991), Rosenbaum and Rubin (1983; 1985), and the overview by Rosenbaum (2002). However, the idea of matching is not new. Heckman, Ichimura, Smith, and Todd (1998) note that the method of matching was first used by Fechner (1860).

⁸It has to be noted that the conditional mean independence assumption and the common support condition are not specific to the matching estimator, but apply to all non-experimental evaluation estimators that condition on exogenous covariates.

⁹The idea of conditioning on X to eliminate selection bias may also justify linear regression. However, two drawbacks of this method relative to matching have to be noted. First, matching is a non-parametric method and therefore does not require any parametric assumption, like the linearity implicit in linear regression. Second, matching emphasizes the common support problem, whereas in analyses that estimate impacts simply by running regressions on X , the issue is rarely even investigated (Smith, 2000a).

¹⁰For a further detailed discussion the interested reader is referred to the overviews by Heckman, LaLonde, and Smith (1999) and Imbens (2004). Furthermore, see Heckman, Ichimura, and Todd (1997;1998) for some additional estimators and their properties respectively, e.g. kernel matching or local polynomial matching, and Smith and Todd (2005).

¹¹The alternative treatment at t_p^* includes the non-treatment case, see Abbring and van den Berg (2003).

¹²Participation and non-participation is defined in the classical way similar to Lechner and Wunsch (2008).

References

- ABBRING, J. H. (2003): “Dynamic Econometric Program Evaluation,” Discussion Paper No. 804, IZA, Paper prepared for the H. Theil Memorial Conference Amsterdam 2002.
- ABBRING, J. H., AND G. J. VAN DEN BERG (2003): “The Nonparametric Identification of Treatment Effects in Duration Models,” *Econometrica*, 71(5), 1491–1517.
- (2004): “Analyzing the Effect of Dynamically Assigned Treatments Using Duration Models, Binary Treatment Models, and Panel Data Models,” *Empirical Economics*, 29(1), 5–20.
- ASHENFELTER, O., D. ASHMORE, AND O. DESCHÊNES (2005): “Do Unemployment Insurance Recipients Actively Seek Work? Evidence from Randomized Trials in Four U.S. States,” *Journal of Econometrics*, 125(1+2), 53–75.
- BARNOW, B., G. CAIN, AND A. GOLDBERGER (1980): “Issues in the Analysis of Selectivity Bias,” in *Evaluation Studies*, ed. by E. Stromsdorfer, and G. Farkas, vol. 5, pp. 290–317. Sage Publications, Beverly Hills.
- BIEWEN, M., B. FITZENBERGER, A. OSIKOMINU, AND M. WALLER (2007): “Which Program for Whom? Evidence on the Comparative Effectiveness of Public Sponsored Training Programs in Germany,” Discussion Paper No. 07-42, ZEW.
- BIJWAARD, G., AND G. RIDDER (2004): “Correcting for Selective Compliance in a Re-employment Bonus Experiment,” *Journal of Econometrics*, forthcoming.

- BLUNDELL, R., AND M. COSTA DIAS (2002): “Alternative Approaches to Evaluation in Empirical Microeconomics,” *Portuguese Economic Journal*, 1, 91–115.
- BLUNDELL, R., M. COSTA DIAS, C. MEGHIR, AND J. VAN REENEN (2004): “Evaluating the Employment Impact of a Mandatory Job Search Assistance Program,” *Journal of the European Economic Association*, 2(4).
- BLUNDELL, R., L. DEARDEN, AND B. SIANESI (2004): “Evaluating the Impact of Education on Earnings in the UK: Models, Methods and Results from the NCDS,” Working Paper No. 03/20, The Institute for Fiscal Studies.
- BÜTTNER, T. (2007): “Ankündigungseffekt oder Maßnahmewirkung? Eine Evaluation von Trainingsmaßnahmen zur Überprüfung der Verfügbarkeit,” Discussion Paper No. 25/2007, IAB.
- CALMFORS, L., A. FORSLUND, AND M. HEMSTRÖM (2001): “Does Active Labour Market Policy Work? Lessons From The Swedish Experience,” *Swedish Economic Policy Review*, 85, 61–124.
- CENTENO, L., M. CENTENO, AND A. NOVO (2006): “Evaluating Job Search Programs for Old and Young Individuals: Heterogeneous Impact on Unemployment Duration,” Working Paper, Banco des Portugal.
- CRÉPON, B., M. DEJEMEPPE, AND M. GURGAND (2005): “Counseling the Unemployed: Does It Lower Unemployment Duration and Recurrence,” Discussion Paper No. 1796, IZA, Bonn.
- DEHEJIA, R. H., AND S. WAHBA (1999): “Causal Effects in Nonexperimental Studies: Reevaluating the Evaluation of Training Programs,” *Journal of the American Statistical Association*, 94(448), 1053–1062.
- (2002): “Propensity Score Matching Methods for Nonexperimental Causal Studies,” *The Review of Economics and Statistics*, 84(1), 151–161.
- DOLTON, P., AND D. O’NEILL (2002): “The Long-Run Effects of Unemployment Monitoring and Work-Search Programs: Experimental Evidence from the United Kingdom,” *Journal of Labor Economics*, 20(2), 381–403.
- European Commission (2006): “Employment in Europe 2006,” Brussels.
- FECHNER, G. (1860): *Elemente der Psychophysik*. Breitkopf und Härtel, Leipzig.
- FISHER, R. (1935): *Design of Experiments*. Hafner, New York.
- FITZENBERGER, B., O. ORLYANSKAYA, A. OSIKOMINU, AND M. WALLER (2008): “Déjà Vu? Short-Term Training in Germany 1980-1992 and 2000-2003,” Discussion paper no. 3540, IZA Bonn.

- GORTER, C., AND G. R. J. KALB (1996): “Estimating the Effect of Counseling and Monitoring the Unemployed Using a Job Search Model,” *Journal of Human Resources*, 31(3), 590–610.
- GRAVERSEN, B. K., AND J. C. VAN OURS (2008): “How to Help Unemployed Find Jobs Quickly: Experimental Evidence from a Mandatory Activation Program,” *Journal of Public Economics*, (92), 2020–2035.
- HÄMÄLÄI NEN, K., R. UUSITALO, AND J. VUORI (2007): “Long-term effects of job search training: Evidence from two randomized field experiments,” Working Paper, Government Institute for Economic Research.
- HECKMAN, J. (1997): “Instrumental Variables: A Study of the Implicit Behavioral Assumptions Used in Making Program Evaluations,” *The Journal of Human Resources*, 32(3), 441–461.
- HECKMAN, J. J., H. ICHIMURA, J. A. SMITH, AND P. TODD (1996): “Sources of Selection Bias in Evaluating Social Programs: An Interpretation of Conventional Measures and Evidence on the Effectiveness of Matching as a Program Evaluation Method,” *Proceedings of the National Academy of Sciences of the United States of America*, 93, 13416–13420.
- (1998): “Characterizing Selection Bias Using Experimental Data,” *Econometrica*, 66, 1017–1098.
- HECKMAN, J. J., H. ICHIMURA, AND P. TODD (1997): “Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme,” *Review of Economic Studies*, 64, 605–654.
- HECKMAN, J. J., R. J. LALONDE, AND J. A. SMITH (1999): “The Economics and Econometrics of Active Labor Market Programs,” in *Handbook of Labor Economics*, ed. by O. Ashenfelter, and D. Card, vol. 3A, chap. 31, pp. 1865–2097. Elsevier, Amsterdam.
- HECKMAN, J. J., AND R. ROBB (1985): “Alternative Methods for Evaluating the Impact of Interventions - An Overview,” *Journal of Econometrics*, 30, 239–267.
- HECKMAN, J. J., AND J. A. SMITH (1995): “Assessing the Case for Social Experiments,” *Journal of Economic Perspectives*, 9, 85–110.
- (1998): “Evaluating the Welfare State,” Working Paper No. 6542, NBER.
- HOLLAND, P. W. (1986): “Statistics and Causal Inference,” *Journal of the American Statistical Association*, 81, 945–960.
- HUJER, R., S. L. THOMSEN, AND C. ZEISS (2006): “The Effects of Short-Term Training Measures on Individual Unemployment Duration in Western Germany,” Discussion Paper No. 06-065, ZEW, Mannheim.

- IMBENS, G. W. (2000): “The Role of the Propensity Score in Estimating Dose-Response Functions,” *Biometrika*, 87(3), 706–710.
- (2004): “Nonparametric Estimation of Average Treatment Effects under Exogeneity: A Review,” *The Review of Economics and Statistics*, 86(1), 4–29.
- KLUVE, J. (2006): “The Effectiveness of European Active Labor Market Policy,” Discussion Paper No. 2018, IZA.
- LALONDE, R. J. (1986): “Evaluating the Econometric Evaluations of Training Programs with Experimental Data,” *The American Economic Review*, 76, 604–620.
- LANCASTER, T. (1979): “Econometric Methods for the Duration of Unemployment,” *Econometrica*, 47, 939–956.
- LECHNER, M. (1998): *Training the East German Labour Force. Microeconomic Evaluations of Continuous Vocational Training after Unification*. Physica-Verlag, Heidelberg, New York.
- (2001): “Identification and Estimation of Causal Effects of Multiple Treatments under the Conditional Independence Assumption,” in *Econometric Evaluation of Labour Market Policies*, ed. by M. Lechner, and F. Pfeiffer, vol. 13 of *ZEW Economic Studies*, pp. 43–58. Physica-Verlag, Heidelberg.
- LECHNER, M., AND C. WUNSCH (2008): “What Did All the Money Do? On the General Ineffectiveness of Recent West German Labour Market Programmes,” *Kyklos*, 61(1), 134–174.
- MARTIN, P., AND D. GRUBB (2001): “What works and for whom: A review of OECD countries experiences with active labour market policies,” *Swedish Economic Policy Review*, 8, 9–56.
- MEYER, B. (1995): “Lessons from the U.S. Unemployment Insurance Experiments,” *Journal of Economic Literature*, 33(1), 91–131.
- MICKLEWRIGHT, J., AND G. NAGY (2005): “Job Search Monitoring and Unemployment Duration in Hungary: Evidence from a Randomised Control Trial,” Budapest Working Papers on the Labour Market 2005/09, University Budapest.
- MORTENSEN, D. (1986): “Job Search and Labor Market Analysis,” in *Handbook of Labor Economics Vol.II*, ed. by O. Ashenfelter, and D. Card, pp. 849–919. Elsevier, Amsterdam.
- NEYMAN, J. S. (1923): “On the Application of Probability Theory to Agricultural Experiments. Essays on Principles. Chapter 9,” *Statistical Science*, 5(4), 463–480, Reprint in English from 1990.
- (1935): “Statistical Problems in Agricultural Experiments,” *The Journal of the Royal Statistical Society*, 2, 107–180.

- OECD (2007): “OECD Employment Outlook,” Paris.
- OSIKOMINU, A. (2008): “Is Short Training Short-Lived and Long Training Long-Lasting? A Multi-State Duration Analysis of the Dynamic Effects of Training Schemes for the Unemployed,” mimeo, Albert-Ludwigs-University Freiburg.
- PAGAN, A., AND A. ULLAH (1999): *Nonparametric Econometrics*. Cambridge University Press, Cambridge.
- QUANDT, R. (1972): “Methods for Estimating Switching Regressions,” *Journal of the American Statistical Association*, 67(338), 306–310.
- (1988): *The Economics of Disequilibrium*. Basil Blackwell, Oxford.
- ROSENBAUM, P. R. (2002): *Observational Studies*. Springer, New York.
- ROSENBAUM, P. R., AND D. B. RUBIN (1983): “The Central Role of the Propensity Score in Observational Studies for Causal Effects,” *Biometrika*, 70(1), 41–50.
- (1985): “Constructing a Control Group Using Multivariate Matched Sampling Methods that Incorporate the Propensity Score,” *The American Statistician*, 39(1), 33–38.
- ROY, A. D. (1951): “Some Thoughts on the Distribution of Earnings,” *Oxford Economic Papers*, 3, 135–145.
- RUBIN, D. B. (1974): “Estimating Causal Effects to Treatments in Randomised and Nonrandomised Studies,” *Journal of Educational Psychology*, 66, 688–701.
- (1977): “Assignment to Treatment Group on the Basis of a Covariate,” *Journal of Educational Studies*, 2, 1–26.
- (1979): “Using Multivariate Matched Sampling and Regression Adjustment to Control Bias in Observational Studies,” *Journal of the American Statistical Association*, 74, 318–328.
- (1986): “Statistics and Causal Inference: Comment. Which Ifs Have Causal Answers,” *Journal of the American Statistical Association*, 81(396), 961–962.
- (1991): “Practical Implication of Modes of Statistical Inference for Causal Effects and the Critical Role of the Assignment Mechanism,” *Biometrics*, 47, 1213–1234.
- SIANESI, B. (2004): “An Evaluation of the Active Labour Market Programmes in Sweden,” *The Review of Economics and Statistics*, 86(1), 133–155.
- SMITH, J. A. (2000a): “A Critical Survey of Empirical Methods for Evaluating Active Labor Market Policies,” *Schweizerische Zeitschrift für Volkswirtschaft und Statistik*, 136(3), 1–22.

- (2000b): “Evaluation aktiver Arbeitsmarktpolitik: Erfahrungen aus Nordamerika,” *Mitteilungen aus der Arbeitsmarkt und Berufsforschung*, 3, 345–356.
- SMITH, J. A., AND P. TODD (2005): “Does Matching Overcome LaLonde’s Critique of Nonexperimental Estimators?,” *Journal of Econometrics*, 125(1/2), 305–354.
- VAN DEN BERG, G., AND B. VAN DER KLAUW (2006): “Counseling and Monitoring of Unemployed Workers: Theory and Evidence from a Controlled Social Experiment,” *International Economic Review*, 47(3), 895–936.
- VAN DEN BERG, G. J. (1994): “The Effects of Changes of the Job Offer Arrival Rate on the Duration of Unemployment,” *Journal of Labor Economics*, 12(3), 478–498.
- VUORI, J., AND J. SILVONEN (2005): “The benefits of a preventive job search program on re-employment and mental health at 2-year follow-up,” *Journal of Occupational and Organizational Psychology*, 78(1), 43–52.
- WEBER, A., AND H. HOFER (2004a): “Are Job Search Programmes a Promising Tool? A Microeconomic Evaluation for Austria,” Discussion Paper No. 1075, IZA.
- (2004b): “Employment Effects of Early Interventions on Job Search Programmes,” Discussion Paper No. 1076, IZA.
- WOLFF, J., AND E. JOZWIAK (2007): “Does short-term training activate means-tested unemployment benefit recipients in Germany?,” Discussion paper 29/07, IAB.
- WUNSCH, C. (2007): “Optimal Use of Labour Market Policies,” Discussion Paper, SIAW, St. Gallen.

Tables

Table 1: Registration, job-search requirements and ALMP participation^a

country	benefit entitlement start ^b	timing of first intensive interview ^c	frequency at which unemployed are obliged to report job-search activities	compulsory ALMP participation after some set unemployment duration ^d	continuing job-search requirement during participation in ALMP
Austria	simultaneously	often at first interview, with IAP	once a month	no	no
Denmark	simultaneously	within a month, with profiling	at least once every three weeks	yes, six months for unemployed under 30 and over 60 years; nine months else	required
Finland	after	within a month, with profiling	from one week to one month	no	required
France	before	within five (until 2007: eight) days, with profiling and IAP	once a month (after fourth month)	no	no
Germany	simultaneously	usually within 10 days, with profiling and IAP	depends on profiling category: on average six times a year	no	required
Hungary	simultaneously	“as soon as possible”, with profiling and IAP	monthly for regular benefit recipients	no	no
Netherlands	before (retrospective)	at registration, with profiling	every four weeks	yes, one month/six months (reintegration trajectory)	–
Portugal	simultaneously	at registration, with profiling	variable requirements (until 2006: not specified)	no	no
United Kingdom	before	within one to three weeks, with IAP if profiled is expected to exhaust UB	every two weeks	yes, ten months for youth, 22 months for adults (25-49 years)	no

^a Source: OECD (2007), Tables 5.1, 5.2 and 5.5.

^b before, simultaneously, after = benefit entitlement starts before/ simultaneously with/ after registration for placement.

^c IAP=individual action plan.

^d In all countries participation in ALMP is compulsory if unemployed is referred to program by Public Employment Services.

Table 2: Description of the Job Search Assistance Programs

Country	Name/Paraphrase	Contents	Duration	Nature
Austria	Job search programs/ job-coaching	<ul style="list-style-type: none"> - not focused on target groups - include job application practices (writing application letters, behavior in job talks) - aim: every new entrant into unemployment should be enrolled in a course before completing the first four months of unemployment - implementation faced administrative problems 	up to 7 weeks	mandatory
Denmark	Special Job Search Program at 2 sites: Storstrøm County and South Jutland County	<ul style="list-style-type: none"> a) job search program b) activation program (if unemployed for 4 months, for people below 30 or older than 60 years: after 6 months) c) longer meeting with case worker to work out a new job plan (if unemployed for 6 to 7 months) 	<ul style="list-style-type: none"> a) up to 2 weeks b) at least 3 months 	mandatory
Finland	Tyhönön Job Search Program	<ul style="list-style-type: none"> - two randomized experiments - group sessions (6 to 17 members, Exp. 1; 8 to 20, Exp. 2) - role plays, discussions, group problem solving in order to recognize one's marketable skills, social networks, contacting employers, drawing up job applications, and preparation of job interviews 	<ul style="list-style-type: none"> Exp. 1: up to one week; Exp 2: up to 10 days 	voluntary
France	Four counseling schemes within Plan d'Aide au Retour d'Emploi	<ul style="list-style-type: none"> modular set-up of four different types of services 1) skill assessment: evaluation of the individual's skills 2) project assessment: evaluation of the individual's skills for people with a professional experience 3) job-search support: personal advisor teaches on-the-job search methods, defines course of actions and provides logistic support 4) project support: for low-ability workers, similar to project assessment 	<ul style="list-style-type: none"> 1) 1 day 2) up to 42 days 3+4) up to 3 months 	mandatory

continued on next page

Table 2: Description of the Job Search Assistance Programs (continued)

Country	Name/Paraphrase	Features	Duration	Nature
Germany	Training measures and job-seeker assessment (Trainingsmaßnahmen und Maßnahmen zur Eignungsfeststellung)	<ul style="list-style-type: none"> - modular set-up of three types of interventions used separately or in combination: <ul style="list-style-type: none"> a) aptitude tests assessing job seeker's availability, willingness and ability to search (jsa), b) short combined measures removing specific minor skill deficits for job applications and interviews (scm), c) short term training providing necessary skills and techniques (st) 	<ul style="list-style-type: none"> a) up to 4 weeks, b) up to 2 weeks, c) up to 8 weeks; combinations of a), b) and c) up to 12 weeks	<ul style="list-style-type: none"> mandatory by assignment
Hungary	Job search monitoring in six counties	<ul style="list-style-type: none"> - intensified contact to employment office (every 3 weeks) - report of job search activities 	4.5 months	mandatory
Netherlands	Counseling & Monitoring at 7 sites	<ul style="list-style-type: none"> - intensive counseling by caseworker on applications, potentially suitable jobs, advice on the direction of search - more time spent on checking information provided by the unemployed - decision on activities, e.g. advice on writing applications, behavior in a job interview, participation in application-training course - only for people younger than 57.5 years and applying for permanent positions 	<ul style="list-style-type: none"> every four weeks for one year 	mandatory

continued on next page

Table 2: Description of the Job Search Assistance Programs (continued)

Country	Name/Paraphrase	Features	Duration	Nature
	Counseling & Monitoring at 2 sites	<ul style="list-style-type: none"> - monthly meetings with caseworker for a period of 6 months after registration - first: intake meeting with discussion of job applications, information about job opportunities, stimulation of workers to contact PES - in the following months: follow-up meetings to evaluate past progress and plan future activities - caseworkers could sanction unemployed persons if they do not cooperate properly 	every four weeks for six months	mandatory
Portugal	REAGE program for long-term unemployed and IN-SERJOVEM for young unemployed persons	<ul style="list-style-type: none"> - intensive job search assistance - small basic skills training - each individual is enrolled in a number of interviews to help her improve job-search skills - elaboration of a "Personal Employment Plan" including unemployed individual's search efforts 	6 months	mandatory
United Kingdom	RESTART	<ul style="list-style-type: none"> - combination of tighter monitoring and increased job-search assistance targeted to long-term unemployed persons - compulsory interview 	interview of 20 to 25 minutes, repeated every six months, possible further programs	mandatory
	New Deal for Young People	<ul style="list-style-type: none"> - combination of individual job search assistance followed by subsidized options, e.g. wage subsidies, temporary government jobs, full time education, training; - target group: 18 to 24 year old claiming UI for six months 	up to 4 months	mandatory

Table 3: Evaluation Studies on Job Search Assistance Programs in Europe

Country	Name/Paraphrase	Authors	Program Period	Design	Method	Outcome(s)	Results	Cost-Benefit-Analysis
Austria	Job search programs/job-coaching	Weber and Hofer (2004a)	1999 to 2001	non-experimental	bivariate mixed proportional hazards model	transition to employment	+ positive effects of programs on the hazard rate to employment; participation reduces unemployment duration significantly	no cba
		Weber and Hofer (2004b)	1999 to 2001	non-experimental	bivariate mixed proportional hazards model	transition to employment, effects conditional on time upon program entry	+ positive effects of program for entries during the first year of unemployment, - negative effects after more than one year of unemployment	no cba
Denmark	Special Job Search Program at 2 sites: Storstrøm County and South Jutland County	Graversen and van Ours (2008)	2005 to 2006	experimental, control group individuals start a job search program after one year of unemployment	proportional hazard rate models, mixed proportional hazard rate models	job finding rate	+ positive effect on job finding rate, authors conclude that threat and punishment are responsible rather than help	no cba
Finland	Tyhönön Job Search Program	Hämäläinen, Uusitalo, and Vuori (2007)	1996/97 and 1999	experimental	experimental estimators and non-experimental matching-estimators	employment, unemployment, labor market training, out of labor force	0 no effects for any of the programs from experimental estimates, + positive long-term effects from non-experimental matching estimates	no cba
		Vuori and Siltonen (2005)	1996/97	experimental	experimental	re-employment, job satisfaction, self-esteem, depressive symptoms	+ positive effects on re-employment and job satisfaction based on one-tailed test, 0 no effects based on two-tailed tests	no cba

continued on next page

Table 3: Evaluation Studies on Job Search Assistance Programs in Europe (continued)

Country	Name/Paraphrase	Authors	Program Period	Design	Method	Outcome(s)	Results	Cost-Benefit-Analysis
France	Four counseling schemes within Plan d'Aide au Retour d'Emploi	Crépon, Djemeppe, and Gurgand (2005)	2001 to 2004	non-experimental	multivariate mixed proportional hazards model	duration of unemployment, unemployment recurrence	+ positive effects on reduced unemployment recurrence, + positive in terms of reduced unemployment duration	no cba
Germany	Training measures and job-seeker assessment (Trainingsmaßnahmen und Maßnahmen zur Eignungsfeststellung)	Hujer, Thomssen, and Zeiss (2006)	2000 to 2003	non-experimental	multivariate mixed proportional hazards model	transition to employment	no distinction between modules, West Germany: + positive effects for men and women on the transition rate to employment, + strong positive effects for low-qualified men, + positive effects for low-qualified women	no cba
		Lechner and Wunsch (2008)	2000 to 2002	non-experimental	propensity score matching	employment rate, unemployment rate, further participation	West Germany: 0 no effects for scm and st, - negative effects for jsa, + positive effects of st for people with vocational education, + positive findings of scm for persons with more than 5 months of unemployment, + positive for low-qualified unemployed persons	- negative cost efficiency due to prolongation of unemployment
		Biewen, Fitzenberger, Osikominu, and Waller (2007)	2002 to 2002	non-experimental	propensity score matching	employment rate	no distinction between modules: + positive effects in West Germany, 0/(+) insignificant/marginally significant effects in East Germany	no cba
		Büttner (2007)	2005	experimental	propensity score matching	threat effect of participation on employment rate, treatment effect of participation on employment rate	West Germany: jsa + internship: + positive threat effect on employment rate, 0 effect of participation, + positive effect of participation for young unemployed	no cba

continued on next page

Table 3: Evaluation Studies on Job Search Assistance Programs in Europe (continued)

Country	Name/Paraphrase	Authors	Program Period	Design	Method	Outcome(s)	Results	Cost-Benefit-Analysis
Hungary	Job search monitoring in six counties	Micklewright and Nagy (2005)	2003	experimental	experimental methods, nonparametric duration model	transition to employment	0 insignificant effects, + positive effects for women aged over 30 years	no cba
Netherlands	Counseling & Monitoring at 7 sites	Gorter and Kalb (1996)	1989 to 1990	experimental	experimental methods, ordered regression, flexible baseline hazard-rate model	job finding rate, application intensity, matching probability	- negative effect on finding a permanent job when person had a temporary job before, 0 no (+ weak positive) effect when person had a permanent job before + positive effect on application intensity - small negative effect on matching probability	no cba
	Counseling & Monitoring at 2 sites	van den Berg and van der Klaauw (2006)	1998 to 1999	experimental	nonparametric duration model, mixed proportional hazards model, binary probit, bivariate ordered probit	exit rate to work with explicit consideration of the effects of formal and informal job search channels	0 insignificant effects on exit rate to employment, (+) at best small positive effects	+ program (seems to be) cost-effective
Portugal	REAGE program for long-term unemployed and INSER-JOVEM for young unemployed persons	Centeno, Centeno, and Novo (2006)	1998	experimental	difference-in-differences matching estimator	duration of unemployment	+ positive effect of REAGE (long-term unemployed), 0 insignificant effect of INSER-JOVEM	no cba

continued on next page

Table 3: Evaluation Studies on Job Search Assistance Programs in Europe (continued)

Country	Name/Paraphrase	Authors	Program Period	Design	Method	Outcome(s)	Results	Cost-Benefit-Analysis
United Kingdom	RESTART	Dolton and O'Neill (2002)	1989 to 1994	experimental, but controls are allowed to participate after six months	experimental, cross-section estimator, bivariate duration model	unemployment rate six months to five years after program	+ positive effects for men in terms of reduced unemployment rates, 0 no effects for women; threat component associated with the program responsible for short-run effects, services provided determine the long-run effects	+ cost-efficient for males, tends to be cost-efficient for women
	New Deal for Young People	Blundell, Costa Dias, Meghir, and van Reenen (2004)	1998 to 1999 with pilot period in 1997	non-experimental	combined dif-in-dif and matching methodologies	proportion leaving unemployment within four months, outflows into employment	+ positive effects for men, stronger treatment effect in the first quarter after introduction "program introduction effect", + positive effects for women, but results more sensitive to evaluation method due to small no. of obs	no cba