IAB

IAB·Discussion Paper 12/2008

The effects of active labor market programs in Germany

An investigation using different definitions of non-treatment

Gesine Stephan

The effects of active labor market programs in Germany

An investigation using different definitions of non-treatment

Gesine Stephan (IAB)

Contents

Abstract

This paper estimates the effects of several German labor market programs – starting in March 2003 – on the employment outcomes of participants using propensity score matching. The main objective is to compare estimated average treatment effects for treatment and comparison groups, which vary in the choice of the classification window that defines treatment and non-treatment. The first approach does not put any restrictions on the future of the treated as well as of the comparison group. This approach has become more and more common in the evaluation of European labor market policies. In contrast, the second approach considers only potential comparison group members, who have not entered any labor market program during the entire observation period of 3 ½ years. The third approach additionally restricts itself to participants, who have not participated in further labor market programs during the observation period. The results differ considerably; program effectiveness is estimated to be much lower using the second approach. The paper highlights the fact that program careers are a non-trivial issue that deserves more attention in future research.

1 Introduction

During 2006, the German Public Employment Service ("Bundesagentur für Arbeit") and the German government spent 18.5 billion Euros on active and 46.8 billions on passive labor market policies (Bundesagentur für Arbeit 2006). The question, whether the considerable funds devoted to active labor market policies have been used effectively gains more and more attention in Germany. An increasing number of careful evaluation studies investigate the effectiveness of labor market programs in Germany. Starting with large evaluation projects on public job creation schemes (see for instance Caliendo et al. 2005a, 2005b, 2006) and further vocational training measures (see for instance Fitzenberger et al. 2006, Fitzenberger/Völter 2007, Lechner et al. 2005, 2007a), a further step was reached with the comprehensive evaluation of the so called "Hartz" reforms of German labor market policies. This huge evaluation project involved the majority of German socio-economic research institutes and has been supervised by the Federal Ministry for Labor and Social Affairs (BMAS). Furthermore, the German Public Employment Service started the TrEffeR-project (Treatment Effects and Prediction) to develop its own evaluation and monitoring tool (Stephan et al. 2006).

This paper adds to this literature by analyzing empirically a particular aspect – the definition of non-participation – arising in evaluation studies that utilize non-experimental data. For participants in five active labor market programs, starting during March 2003, I apply statistical matching techniques to estimate program effects on cumulated days spent in regular employment during the 3 ½ years after program entry. I compare the results for a very narrow and a very wide "classification window" defining non-treatment.

In the former case all persons are categorized as non-treated who did not enter a program during March 2003, while they may be "waiting" to participate at a later date. In the latter case the non-treatment group consists of persons who have never been in any program during the entire observation period of 3 ½ years. Additional estimates also include only participants, who did not join any further labor market program during the observation period. This is no minor issue, since typically around 40 to 50 percent of comparison group members and around 50 to 70 percent of those taking-up a program in March 2003 also participated in at least one (further) program starting after March 2003.

A similar analysis has been conducted by Steiger (2004) for Switzerland, who did, however, not distinguish between different labor market programs. Her main result was that participation is much more effective compared to "waiting" than compared to "never in any program".

A recent example for the possible importance of the topic is provided by three studies of German further vocational training programs, covering program entries from 2000 to 2002: Wunsch/Lechner (2008) estimated that further vocational training – and other programs – had generally negative or insignificant effects on employment probabilities of participants and their cumulated days spent in regular employment 30 months after program start. Small positive effects were found for few subgroups of participants only. This was taken up by a German

newspaper claiming that active labor market policy in Germany wastes billions of Euros (Frankfurter Allgemeine Zeitung, August 28, 2007). Biewen et al. (2007) point out that – also 30 months after program start – they found positive effects of further vocational training with short or medium duration on the employment prospects of certain subgroups in West Germany. Rinne et al (2007), who investigated entries in further vocational training programs of medium duration during 2002, obtained – two years after program start – even positive effects on employment probabilities for all subgroups investigated.

One of the underlying reasons might be that Wunsch/Lechner (2008) "define participants as unemployed who participate at least once in a program in the 18 months following the inflow date into our sample. Accordingly, non-participants are all persons who do not enter a program during this period." (p. 143). Thus "non-treatment" requires that no treatment has occurred up to a period of 18 months after entry into unemployment. Biewen et al. (2007) perform separate estimates by duration of unemployment at the beginning of a treatment (up to three months, four to six months, seven to twelve months of unemployment) and require only that non-participants have not entered a measure during an accordingly chosen classification window of three or six months. Rinne et al. (2007) stratify estimates by duration of unemployment in months and require that non-participants had not taken-up the same type of training program before and during the quarter of the participant's program entry (p. 10). Sensitivity analysis conducted by these authors (Section 4.4 of their paper) showed also that estimated program effects decreased considerably if they use a similar definition of non-treatment than Wunsch/Lechner (2008) did.

The following Section 2 briefly describes German active labor market policies and recent evaluation results. Section 3 discusses the evaluation problem investigated. Section 4 introduces data, variables and the particular method and evaluation design used. The empirical results are presented in Section 5, while Section 6 lists the conclusions.

2 Active labor market policies in Germany and recent evaluation results

The "Hartz" reforms, enacted in 2002, caused a fundamental revision of active and passive labor market policies in Germany: Since 2003 the design of several active measures was modified and a number of new instruments were introduced. Regarding passive labor market policies, since 2005 former unemployment assistance for long-term unemployed persons and former social assistance has been merged into a new variant of basic social care for needy employable persons.

This section gives a short overview of the main instruments of active labor market policies in Germany. Furthermore, results from selected recent evaluation studies on the effectiveness of active labor market policies in Germany will be sketched. Most of these studies are based on statistical matching techniques, while a few apply duration analysis. Statistical matching techniques compare a group of treated individuals with a selected group of similar, but un-

treated comparison persons; the estimated program effect is then simply given by the difference in outcome variables – mostly employment rates – between the two groups (see Section 3 for more details). Duration analysis parametrically estimates the effect of treatment on the "risk" of entering employment or leaving unemployment. The increasingly popular timing-of-events model (Abbring/van den Berg 2003, 2004) additionally takes into account the risk of entering a program during an unemployment spell.

Table 1 shows entries and populations in selected programs of active labor market policies administered by the German Public Employment Service for the period 2000 to 2006. The empirical results presented in section 5 in this paper restrict themselves to short programs up to six months duration and include variants of further vocational training, short-term training measures, wage subsidies and public job creation schemes.

Table 1
Entries and average stock of participants in selected labor market programs (in 1000)

	2000	2001	2002	2003	2004	2005	2006
Entries							
Further vocational training ("Förderung beruflicher Weiterbildung")	523	442	455	255	185	132	247
Short-term training ("Trainingsmaßnahmen")	485	551	865	1064	1188	894	978
Wage subsidy ("Eingliederungszuschüsse")	152	127	188	183	157	134	217
Start-up subsidy I ("Überbrückungsgeld")	93	96	125	159	183	157	108
Start-up subsidy II ("Existenzgründungszuschuss")	-	-	-	95	168	91	43
Public job creation I ("Arbeitsbeschaffungsmaßnahmen, SAM")	318	246	215	179	161	80	80
Public job creation II ("Arbeitsgelegenheiten")	-	-	-	-	-	630	742
Contracting-out to private placement agencies ("Beauftragung Dritter")*	-	-	-	-	635	426	301
Temporary help-firms ("Personal-Service-Agenturen")	-	-	-	45	56	27	16
Average stock of participants							
Further vocational training ("Förderung beruflicher Weiterbildung")	343	352	340	260	184	114	119
Short-term training ("Trainingsmaßnahmen")	52	60	74	93	95	69	70
Start-up subsidy I ("Überbrückungsgeld")	105	118	136	153	110	60	82
Start-up subsidy II ("Existenzgründungszuschuss")	43	46	56	73	84	83	63
Start-up subsidy ("Existenzgründungszuschuss")	-	-	-	40	151	234	210
Public job creation I ("Arbeitsbeschaffungsmaßnahmen, SAM")	266	237	193	144	117	61	50
Public job creation II ("Arbeitsgelegenheiten")	-	-	-	-	-	201	293
Contracting-out to private placement agencies ("Beauftragung Dritter")*	-	-	-	-	95	103	100
Temporary help-firms ("Personal-Service-Agenturen")	-	-	-	10	25	13	6

*) Figures are available since 2004, while different variants started already in 1998 (contracting-out of subtasks) respectively 2002 (contracting-out of entire placement).

Source: Statistics of the German Public Employment Service (Date-Warehouse).

For a long time so called "further" vocational training ("Förderung beruflicher Weiterbildung") belonged to the most important programs in Germany. It encompasses a range of different treatments, which can be broadly classified in qualification programs, training within "practice firms" (that offer practical occupational training without trainees actually working in a real company) and long retraining programs. However, during the first half of this decade entries as well as the duration of these measures were shrinking, whereas the number of entries increased again in 2006. The effectiveness of these programs has been analyzed in a considerable number of studies applying statistical matching techniques. Lechner et al. (2005, 2007a)

investigated program entries during the years 1993 and 1994, their observation period covers about seven years after program start. Fitzenberger et al. (2006) analyze program entries inflows in unemployment during the years 1986/87 and 1993/94 in West Germany, while Fitzenberger/Völter (2007) focus on unemployment entries during 1993/94 in East Germany. Generally, these studies obtained the result that – in the longer run – further vocational training programs had mostly significant positive effects on the employment prospects of participants. However, since program effects are rather weak, it may take some years until the estimated program effect turns positive. More recent program entries have been investigated by Biewen et al. (2007), Rinne et al. (2007) and Wunsch/Lechner (2008), who analyzed programs starting in the years 2000 to 2002. As has already been mentioned in the introduction, the estimates of Wunsch/Lechner indicated no positive effects of further vocational training on employment prospects of participants in West Germany. In contrast, Biewen et al. found positive effects for programs of short and medium duration in West Germany (but not in East Germany) and particular groups of unemployed. Also Rinne et al. (2007) estimated positive effects of participation in medium length programs on the employment probabilities of participants in all subgroups investigated. Hujer et al. (2006b) applied duration analysis to East German data from the years 1999 to 2002 – their result was that participation in further vocational training prolonged unemployment duration.

Short-term training programs ("Trainingsmaßnahmen"), whose duration does in most cases not exceed two months, have been increasing in number in particular during 2003 and 2004. These short-term measures are utilized to train qualifications and abilities (firm-internal and firm-external), to test the availability of unemployed persons, to check whether unemployed are suited for further longer-term measures and to provide job search assistance through application training. These short-term measures have been analyzed also in the already cited studies of Biewen et al. (2007) and Wunsch/Lechner (2008). Similar to further vocational training, Biewen et al. find mostly positive and Wunsch/Lechner found insignificant effects of short training-programs. Hujer et al. (2006a) showed – using duration analysis – that the risk of entering employment is significantly higher for individuals participating in a short-training program. These studies, however, do not take into account the substantial heterogeneity of this class of programs. Wolff/Jozwiak (2007) distinguished between short classroom training and short training within firms for individuals; they investigated the effect on the employment prospects of unemployed receiving means-tested unemployment benefits ("Arbeitslosengeld II"). They obtained the result that both variants have positive effects, which are much larger for short training within firms. These are, however, also much more prone to deadweight losses. Büttner (2007) used data from a social experiment on short-training programs to test the availability of the unemployed. He showed that it is the notification of treatment rather than participation that had an effect on leaving unemployment.

A number of programs foster a direct integration of the unemployed in the first labor market. Of particular importance are a variant of targeted wage subsidies ("Eingliederungszuschüsse"), paid to employers for a fixed period of time. They gained importance first in East Germany

following the reunification and thereafter again in the late nineties, but lost importance until 2005, like further vocational training schemes. From 1998 to 2003 three variants were in place: One was characterized by a rather low level of targeting, while one was aimed at hard-to-place unemployed with severe problems of reintegration and one at workers of age 50 and older. The "Hartz" reforms collapsed these into a single wage subsidy for hard-to-place workers, with a looser definition of target groups and less generous financial support. Jaenichen/Stephan (2007) estimated average treatment effects of a subsidy on previously unemployed hard-to-place individuals. In line with the international literature on wage subsidies, they showed that subsidies have a favorable effect on the employment prospects of participants. Boockmann et al. (2007) utilized a "natural experiment": They computed the effect of changes in the legislation on wage subsidies for older workers on the employment prospects of this group, using a difference-in-differences estimator to compare changes in transition probabilities between the affected group and a comparison group comprised of slightly younger workers. The authors found nearly no significant effects and concluded that deadweight effects - those subsidized would have been hired anyway - are a major problem of wage subsidies.

Two programs offering financial support for unemployed persons founding their own businesses grew in numbers until 2006. A first variant ("Überbrückungsgeld") encouraged unemployed persons to start-up a new business by proceeding to pay unemployment benefits as well as a subsidy to social security contributions for six months. The "Hartz" reforms in 2003 additionally introduced a second variant of a start-up subsidy ("Existenzgründungszuschuss"), which provided a fixed, but time-decreasing amount for up to three years and was attractive for unemployed persons who received comparatively few unemployment benefits. In August 2006 both programs were collapsed in a new variant of a start-up subsidy ("Gründungszuschuss"). Baumgartner/Caliendo (2007) analyzed the effect of both previous variants on unemployed persons, who founded a subsidized business in 2003. Their results showed that unemployment rates were lower 28 months after program start and rates of regular employment or self-employment were significantly higher across those subsidized than in an unsubsidized comparison group.

Previously very important programs for job creation in the public sector ("Arbeitsbeschaffungsmaßnahmen und Strukturanpassungsmaßnahmen") nearly disappeared until 2004. However, since 2005 a new variant of public job creation for long-term unemployed ("Arbeitsgelegenheiten") is the most important program for unemployed receiving social basic care. The latter provide mostly only a modest additional reimbursement for work ("Ein-Euro-Jobs"). Caliendo et al. (2005a, 2005b, 2006) and Hujer/Thomsen (2006c) investigated entries into public job creation schemes in 2000, using the matching method. They estimated heterogeneous effects on participants, which are in the longer-run mostly negative or insignificant. Exceptions are long-term unemployed, highly qualified men and older women in West Germany. Hujer/Zeiss (2006d) evaluated these programs in East Germany schemes also with the timing-of-events method. Their main conclusion is that participation in these schemes increased individual unemployment duration of participants. Recent results on the new "One-Euro-Jobs"

(Hohmeyer/Wolff 2007), introduced in 2005, highlighted the effect heterogeneity of this program. The authors found slightly positive effects in particular for participants from West Germany and individuals out of regular employment for a longer time period.

Contracting-out to private placement services gained importance during the last years in Germany. On the one hand unemployed persons may be assigned to private agencies that compete on a quasi-market ("Beauftragung privater Dritter"). On the other hand unemployed persons may ask for a voucher that entitles them to use the services of a private placement agency. Also temporary help firms may employ previous unemployed persons, while receiving financial reimbursement by the Public Employment Service ("Personal-Service-Agenturen"). Using statistical matching methods, Winterhager (2008) showed that most subgroups of unemployed persons contracted out to private placement agencies in 2004 did not experience an increase in employment prospects during a period of nine months after treatment start. For unemployed persons receiving basic social care and contracted-out at the beginning of 2005, Bernhard/Wolff (2008) obtained evidence of slightly positive effects of contracting-out on the employment prospects two years later. Winterhager et al. (2006) found positive treatment effects on the employment probability of individuals who received a job voucher during May and June 2003; Hess et al. (2006) identified no positive effects of employment in a temporary help firm on the latter likelihood of being in a regular job.

A number of studies (for instance Lechner et al. 2005, 2007a, Jaenichen/Stephan 2007) showed that estimated treatment effects of labor market programs are more positive for the outcome variable "regular employment" than for the outcome variable "not unemployed". Both outcome variables may differ since not every person avoiding unemployment has to be regularly employed – several persons might withdraw temporarily or permanently from the labor force (another option might for instance be to take up unsubsidized self-employment). The observed difference between both outcome variables results from the fact that a higher share of the comparison group than of the treatment group withdraws from the labor force. Thus one effect of the participation in labor market programs might be to activate individuals who would otherwise have withdrawn from the labor market.

Since the main topic of the paper is the classification window underlying definitions of non-treatment using statistical matching techniques, Overview 1 summarizes selected information on several studies cited above. The "maximum restriction" in the last column shows how long after program start of a treated person its matched comparison person may not be "allowed" to enter a program. It depends first on the time window over which treatments are sampled: If program entry occurs at the beginning of this time window the maximum restriction applies, while the restriction is not binding if a program is taken up at the end of this time window. It depends second on the decision to stratify the sample by duration of unemployment: If the sample is stratified, the restriction occurs only within each strata. Authors applying no stratification typically condition their choice of comparison groups on duration of unemployment.

Overview 1
Definitions of treatment and non-treatment in selected recent studies for Germany

Study	Sampling condition	Definition of treatment	Definition of non-treatment	Maximal restriction
Baumgartner/Caliendo (2007)	Unemployed during the 3rd quarter of 2003	Program entry during 3rd quarter of 2003	No program entry during 3rd quarter of 2003	3 months
Bernhard/Wolff (2008)	Unemployed recipients of basic social care for jobseekers at January 31, 2005	Program entry from 2/2005 to 4/2005	No entry into the program investigated from 2/2005 to 4/2005	3 months*
Biewen et al. (2007)	Inflow into unemployment from 2/2000 to 1/2002	First program entry within 12 months after unemployment entry, stratified by unemployment duration (0-3 months, 4-6 months, 7-12 months)	No program entry while in particular unemployment strata	3 or 6 months
Caliendo et al. (2005a)	Unemployed during 1/2000	Program entry during 2/2000	No program entry during 2/2000	1 month
Caliendo et al. (2005b)	Unemployed during 1/2000	Program entry during 2/2000	No program entry during 2/2000	1 month
Caliendo et al. (2006)	Unemployed during 1/2000	Program entry during 2/2000	No program entry during 2/2000	1 month
Fitzenberger et al. (2006)	Inflow into unemployment during 1986/87 and 1993/94 (West Germany)	First program entry within two years after unemployment entry, stratified by unemployment duration (1-2 quarters, 3-4 quarters, 5-8 quarters)	No program entry while in particular unemployment strata	6 or 12 months
Fitzenberger/Völter (2007)	Inflow into unemployment during 1993/94 (East Germany)	First program entry within two years after unemployment entry, stratified by unemployment duration (1-2 quarters, 3-4 quarters, 5-8 quarters)	No program entry while in particular unemployment strata	6 or 12 months
Hohmeyer/Wolff (2007)	Unemployed recipients of basic social care for jobseekers at Januar 31, 2005	Program entry from 2/2005 to 4/2005	No entry into the program investigated from 2/2005 to 4/2005	3 months*
Hujer/Thomsen (2006c)	Unemployed during 6/2000, 8/2000, 10/2000, 12/2000, 2/2001 and 4/2001	Program entry during 7/2000, 9/2000, 11/2000, 1/2001, 3/2001 and 5/2001, stratified by unemployment duration (1, 2,…12 quarters)	No program entry while in particular unemployment strata	3 months
Jaenichen/Stephan (2007)	Unemployed during the 2rd quarter of 2002	First program entry during 2nd quarter of 2002	No program entry during 2nd quarter of 2002	3 months
Lechner et al. (2005)	Inflow into unemployment during 1993/94 (West Germany)	First program entry during 1993/94	No program entry until 1995	24 months
Lechner et al. (2007)	Inflow into unemployment during 1993/94 (East Germany)	First program entry during 1993/94	No program entry until 1995	24 months
Rinne et al. (2007)	Unemployed during 2002	Program entry during 2002, stratified by unemployment duration (in months) and quarter of program entry	No entry into the program investigated before and during the quarter of program entry while in particular unemployment strata	3 months*
Winterhager et al. (2006)	Unemployed during 5/2003 or 6/2003	Receipt of a voucher during 5/2003 or 6/2003	No receipt of a voucher until 6/2004	12 months
Wolff/Jozwiak (2007)	Unemployed recipients of basic social care for jobseekers at Januar 31, 2005	Program entry from 2/2005 to 4/2005	No entry into the program investigated from 2/2005 to 4/2005	3 months*
Wunsch/Lechner (2008)	Inflow into unemployment during 1/2000 to 12/2002 (West Germany)	First program entry during the 18 months after inflow into unemployment and before 2003	No program entry during the 18 months after inflow into unemployment	18 months

Note: See Section 2 for information on the instruments investigated and the main results.

*) Restriction applies only to the program investigated.

Note finally that for instance Lechner et al. (2005, 2007a) use a rather wide participation window, but assigned for each non-participant a hypothetical starting date, drawn from the distribution of observed starting dates of treatment. To obtain comparable samples of participants and non-participants, all non-participants who were already employed at this date were then excluded from further analysis. However, Fitzenberger/Speckesser (2007, Footnote 11) criticize this approach for adding additional noise to the data.

3   The evaluation problem

The majority of micro studies using non-experimental data estimate average treatment effect on the treated (Heckman et al. 1999). In our case this implies estimating the average effect of participation in active labor market programs on subsequent days spent in regular employment. I will first sketch the framework and then introduce a classification window for the definition of non-treatment.

Let $D = 0$ indicate that no measure has started during a certain time interval in calendar time (in the empirical analysis this will be March 2003), while $D = 1$ indicates that an entry in a particular labor market program took place. The effectiveness of the treatment is measured by the outcome variable Y, which takes the value Y_1 under treatment and Y_0 under non-treatment. Independent of the exact definition of non-treatment, we observe either Y_0 or Y_1 for each individual. If the program does not have effects on the labor market outcomes of non-participants – this is the „Stable Unit Treatment Value Assumption" (SUTVA) – the average treatment effect on the treated (ATT) is very generally given by

(1) $\Delta_{ATT} = E(Y_1 - Y_0 \mid D = 1) = E(Y_1 \mid D = 1) - E(Y_0 \mid (D = 1)$

While we observe $E(Y_1 \mid D = 1)$, the average outcome of the treated with treatment, we cannot observe the average outcome of the treated without treatment $E(Y_0 \mid D = 1)$. Using non-experimental data one thus has to find a comparison group of non-treated individuals to impute the counterfactual outcome of the treated without treatment (Rubin 1974).

One standard approach to this problem applies statistical matching techniques to balance the distribution of individual characteristics between the groups of treated and non-treated individuals. This method – that underlies also the empirical analysis conducted in this paper – requires that all variables X, which determine the decision to join a program and the expected success of a program, are known and available. Conditioning on those variables, the expected outcome under non-treatment should not depend on the decision to join $Y_0 \perp D \mid X$. If this „Conditional Independence Assumption" (CIA) holds, the ATT is given by

(2) $\Delta_{ATT} = E(Y_1 - Y_0 \mid D = 1) = E(Y_1 \mid X, D = 1) - E(Y_0 \mid X, D = 0)$.

Furthermore, the "Common Support Condition" requires that each treated individual has a positive probability not to be in a program, which guarantees that all of them have a counterpart in the group of non-participants.

The most common approach used in the evaluation of European labor market policies has been suggested by Sianesi (2002, 2004) and uses a narrow classification window. The underlying idea is that European labor market programs are ongoing and any unemployed is a potential participant. Individuals will join sooner or later provided they are still eligible. The treatment group consists of all individuals "joining" a program during a chosen time interval. Non-participants are defined as "waiting" in the sense that they do not take up treatment until the beginning of the evaluation period, but eventually at a later date. Frederiksson/Johansson (2004) characterize this as a time-varying treatment indicator. To formalize this, let the ATT be given more specifically as

$$(3) \qquad \Delta_{ATT}^{t+h, JW} = E(Y_1^{t+h} \mid X, D^t = 1) - E(Y_0^{t+h} \mid X, D^t = 0)$$

where t is the timing of treatment and t+h the point of time when the outcome is observed (in this paper t will be March 2003 and h will be 3 ½ years). This "joining versus waiting" approach has been adopted for instance in a comprehensive evaluation of recent German labor market reforms (Deutscher Bundestag 2006). The estimated effects display the advantage of joining at a given time compared to waiting longer and are useful for testing for the existence of a treatment effect (see Frederiksson/Johansson 2004, Proposition 4). But they do not isolate the effect of a single intervention and are thus not suited for a cost-benefit analysis.

Note that Sianesi (2004) in implementing this approach stratifies her sample by duration of unemployment in months. An alternative is to condition the choice of the comparison group on the duration of unemployment up to program entry. This requires, however, computing a "hypothetical" program entry date for potential comparison persons.

An entirely different approach uses an extremely wide classification window for the choice of potential control persons and to define non-treatment as no treatment during the entire observation period. The corresponding average treatment effect on the treated has been characterized by Steiger (2004) as "joining versus never in any program" and might be written as

$$(4) \qquad \Delta_{ATT}^{t+h, JN} = E(Y_1^{t+h} \mid X, D^t = 1)$$
$$- E(Y_0^{t+h} \mid X, D^t = 0, D^{t+1} = 0, ..., D^{t+h} = 0).$$

The idea of this approach is to imitate a social experiment, in which randomly chosen participants are assigned to treatment while members of a randomly chosen control group will not receive treatment. It has been used for instance by US-studies where a program is administered at fixed point in time and individuals are either treated or not treated. With non-experimental data, however, the timing of program participation in an unemployment spell cannot be assumed to be random. Suppose that people do not enter a program because they expect to find a job soon. Then selecting a comparison group of individuals, who never participated in any program, bases selection on expected (successful) future outcomes. Thus matching conditional on observable individual characteristics might not suffice to remove selectivity (Sianesi 2002, 2004, Fredriksson/Johansson 2004, Proposition 1).

As has already been mentioned in the introduction, a comparison between the "waiting" and the "never in any program" concepts has been conducted by Steiger (2004), who analyzed entries in Swiss active labor market programs. She obtained positive effects of most programs compared to non-treatment in the "joining versus waiting" context, but negative effects of almost all programs compared to non-treatment in the "joining versus never in any program" context. There are at least two possible explanations. On the one hand, those who were never in any program may be in fact a positive selection from all individuals. This is the aspect highlighted in the theoretical literature. However, on the other hand they may have the advantage to experience no lock-in effects at all from later participation. In contrast, the treatment group is not only locked-in during participation, but partly also during additional later treatment, if "program careers" evolve. This aspect has rather been neglected in the literature, but might be in particular important if programs are rather long and the observation period is rather short.

Thus, I will also apply a third concept in the following, suggested by the empirical fact that program careers are a non-trivial issue. The underlying idea is that the "joining versus never in any program" applies a restriction only to future outcomes of comparison group members, but not to those of participants. Both problems mentioned above might be partly resolved if we introduce a weaker, but similar restriction also to the group of participants, thus applying the classification window for the definition of non –treatment also to the future of the treatment group. In other words we will estimate the effect of "joining once versus never in any program". Let s be the duration of the first program, with $s < h$. Then the effect is given by

$$(5) \qquad \Delta_{ATT}^{t+h, ON} = E(Y_1^{t+h} \mid X, D^t = 1, D^{t+s+1} = 0, ..., D^{t+h} = 0)$$
$$- E(Y_0^{t+h} \mid X, D^t = 0, D^{t+1} = 0, ..., D^{t+h} = 0).$$

While the second approach bases the selection of comparison persons on their (successful) future outcomes, the third concept additionally bases the selection of participants on the fact that they have not participated in a further program during the observation period. The drawbacks are, however, that the analysis is restricted on selected subgroups of individuals and that the approach neglects that program participation itself might induce participation in further programs.

Finally, it should be noted that the identifying assumption of statistical matching techniques is that no unobserved heterogeneity correlated with the selection into programs and with outcome variables remains after accounting for observable variables. In contrast, duration analysis, in particular the timing-of-events approach (Abbring/van den Berg 2003, 2004), allows also for selection on unobserved characteristics. The drawback is that – in contrast to the non-parametrical statistical matching approach – these models impose the identifying assumption that transition processes into labor market programs as well as across labor market states can be modeled as a multivariate mixed proportional hazard model. An applied comparison between a dynamic matching estimator and the timing-of-events approach has been conducted by Lalive et al. (2008) for Swiss labor market programs.

4   Data, variables, and applied method

The empirical analysis utilizes the TrEffeR data set (Stephan et al. 2006), which has been constructed for monitoring purposes of the German Public Employment Service. The current version merges data flows from the distinct computer based operative systems of the Public Employment Service on periods of registered job search, registered unemployment, participation in labor market programs and employment for the period 2000 to 2007. Note that the data of the Public Employment Service are partly incomplete since 2005 as a consequence of the already mentioned last "Hartz" labor market reform, which re-allocated responsibilities for long-term unemployed persons between the Public Employment Service and local municipalities. Data flows from several local municipalities ("Optierende Kommunen") opting out of the co-operation with the Public Employment Service have not been entirely integrated yet. For individuals from these municipalities we cannot distinguish between times of unemployment and programs and other times out of the labor force. However, information on times in regular employment – the outcome variable utilized in this paper – is available also for these persons.

The sample analyzed here covers all individuals who were unemployed for up to one year in March 2003 and of age 25 to 59. This selection excludes individuals eligible for specific programs for youth unemployed, while older workers may be eligible for early retirement schemes. The distance to the previous unemployment spell has to amount to at least one month. All estimates are performed separately for West and East Germany.

As has already been mentioned in the introduction, the treatment groups consist of individuals who took up one of the following labor market programs of short or medium length during March 2003: a) Participation in the most common variant of further vocational training, aimed at the provision of specific professional skills (in the Tables and Figures abbreviated as "provision of skills"). b) Treatment within a practice firm, another variant of further vocational training that offers practical training without trainees working in a real company. c) Short firm-external classroom training imparting qualifications and abilities, with an average of about 1 ½ months. d) Short firm-internal training, also imparting qualifications and abilities, with an average duration of about one month. e) A wage subsidy for hard-to-place workers, covering up to 50 percent of the wage and paid to the employers. f) Participation in a public job creation scheme ("Arbeitsbeschaffungsmaßnahme"). Note that case managers in local employment agencies have latitude in the decision to grant participation in one of these programs. For all programs except the short training programs only treatments with duration of 4 to 6 months are taken into account. The underlying assumption is that a program is not characterized by its type but also by its length. Following the majority of the literature, only the effect of the first treatment during an unemployment spell is analyzed.

Overview 2 gives a graphical impression of the evaluation designs used. In the first and second design the treatment group includes all individuals who took-up one of the programs investigated in March 2003 ("joining"), while in the third design considers only those, who started their only program during the observation period of 3 ½ years in March 2003 ("joining once"). Non-treatment is defined as not entering treatment during March 2003 – but maybe later

during the observation period – in the first design ("waiting") and as not entering any program during the 3 ½ years after March 2003 in the second and third design ("never in any program"). I do not apply stratification by unemployment duration, but condition on unemployment duration until program entry in month (see below).

Overview 2
Evaluation design

Joining vs. waiting				*Outcome: Cumulated days in regular employment*
Treated	2-years-history	up to 1 year of unemployment	joining a labor market program	no restrictions
Controls	2-years-history	up to 1 year of unemployment	not joining a labor market program	no restrictions
			March 03	*Sep 06*

Joining vs. never in any program				*Outcome: Cumulated days in regular employment*
Treated	2-years-history	up to 1 year of unemployment	joining a labor market program	no restrictions
Controls	2-years-history	up to 1 year of unemployment	not joining a labor market program	never in any program
			March 03	*Sep 06*

Joining once vs. never in any program				*Outcome: Cumulated days in regular employment*
Treated	2-years-history	up to 1 year of unemployment	joining a labor market program	no further program
Controls	2-years-history	up to 1 year of unemployment	not joining a labor market program	never in any program
			March 03	*Sep 06*

Note that program participation will be slightly under-recorded in the sample investigated since we have no information on program participation for long-term unemployed individuals from local municipalities opting out of the co-operation with the Public Employment Service since the beginning of 2005. However, results are very similar, if the sample investigated is restricted to local labor markets without such local municipalities. Furthermore, while different variants of contracting-out to private placement agencies started already in the year 1998 (contracting-out of subtasks) respectively 2002 (contracting-out of entire placement tasks), no information on participation has been included in the data prior to 2004. Information on receipt of a placement voucher is not available in the data at all.

The outcome variable utilized here are cumulated days spent in regular, unsubsidized employment during the 3 ½ years after program entry in March 2003 for participants. Since the classification window encompasses only one month, cumulated days for non-treatment groups are measured since March 15, 2003, and all individuals who had already left unemployment at this date were excluded from the sample.

The choice of comparison groups is based on the following variables, which are all categorized as dummy variables: a) Socio-demographic characteristics, measured at the start of an unemployment spell: Age, marital status, nationality, education and degree of disablement. b) Unemployment duration in the current spell, measured in months, until program entry in March 2003 for participants respectively until March 15, 2003 for matched comparison persons. c) Variables describing the employment-history in the both years preceding the analyzed un-

employment spell, measured at the start of the unemployment spell: Former unemployment, participation in labor market programs, sanctions and periods of illness. d) Information on the regional labor market situation: Performance cluster of the regional labor market (Blien et al. 2004). Mean values of selected explaining variables can be found in Table A.1 in the Appendix.

A non-testable assumption underlying the statistical matching technique is the Conditional Independence Assumption (CIA), which states that – conditioning on the explaining variables – the expected outcome under treatment and under non-treatment should not differ across treated and controls. The variables available allow controlling for a wide range of individual characteristics, including past experiences with the Public Employment Service. Thus unobserved individual heterogeneity should not be a serious problem for the analysis, since the data include also information on previous employment histories, which should capture most of the effects of unobserved individual factors (Heckman et. al 1999). However, for wage subsidies as well as short firm-internal training measures also labor demand side factors will influence access to subsidized jobs as well as the probability of further employment. The matching of workers to heterogeneous firms, together with the lack of information on employer characteristics may therefore produce a bias in the estimates because of firm heterogeneity (Jaenichen/Stephan 2007). Thus results for both programs should be taken with caution.

The applied method is a so called nearest neighbor matching on the propensity score with ties and replacement. A useful simplification of matching on a high-dimensional vector of X-variables is given by propensity-score matching. Rosenbaum/Rubin (1983) have shown that it is sufficient to match on the propensity score $Pr(X) = Pr(D = 1|X)$ to obtain the same probability distribution for treated and non-treated individuals. Thus, if $(Y_0, Y_1 \perp D) \mid X$ holds, $(Y_0, Y_1 \perp D) \mid Pr(X)$ will also be satisfied. For treated persons, all non-treated individuals, whose propensity scores have the smallest distance to that of the treated person, are chosen as comparisons. Replacement implies that potential comparison group members can be used as comparisons for more than one treated person. The program impact is estimated as the mean difference in the weighted outcomes of both groups, where a weighting is applied to comparison persons if one treated person has more than one comparison person.

Estimates are performed using the stata-module *psmatch2* (Leuven/Sianesi 2003). To test for the quality of matching the mean standardized bias (MSB) (Rosenbaum/Rubin 1983) between each treated group and its matched comparison group is computed across all variables from X. The standardized bias of a covariate is defined as the difference of means in the treated and matched control sample, divided by the square root of the average sample variance. Thus a lower value of the MSB indicates more similarity between the two groups.

5   Empirical results

Figure 1 summarizes the main results of the analysis – the average treatment effect on cumulated days spent in regular employment 3 ½ years after program entry in March 2003. The change of effects over time is displayed in detail in Figure A.1 in the Appendix, which shows clearly in particular the lock-in effects arising during program participation.

Figure 1
Estimated effects on cumulated days in regular employment 3 ½ years after program entry in March 2003

ATT: Cumulated days in regular employment

West | East — Provision of skills
West | East — Practice firm
West | East — Short classroom training
West | East — Short training within firm
West | East — Wage subsidy hard-to-place
West | East — Public job creation

■ Joining vs. waiting ■ Joining vs. never in any program ■ Joining once vs. never in any program
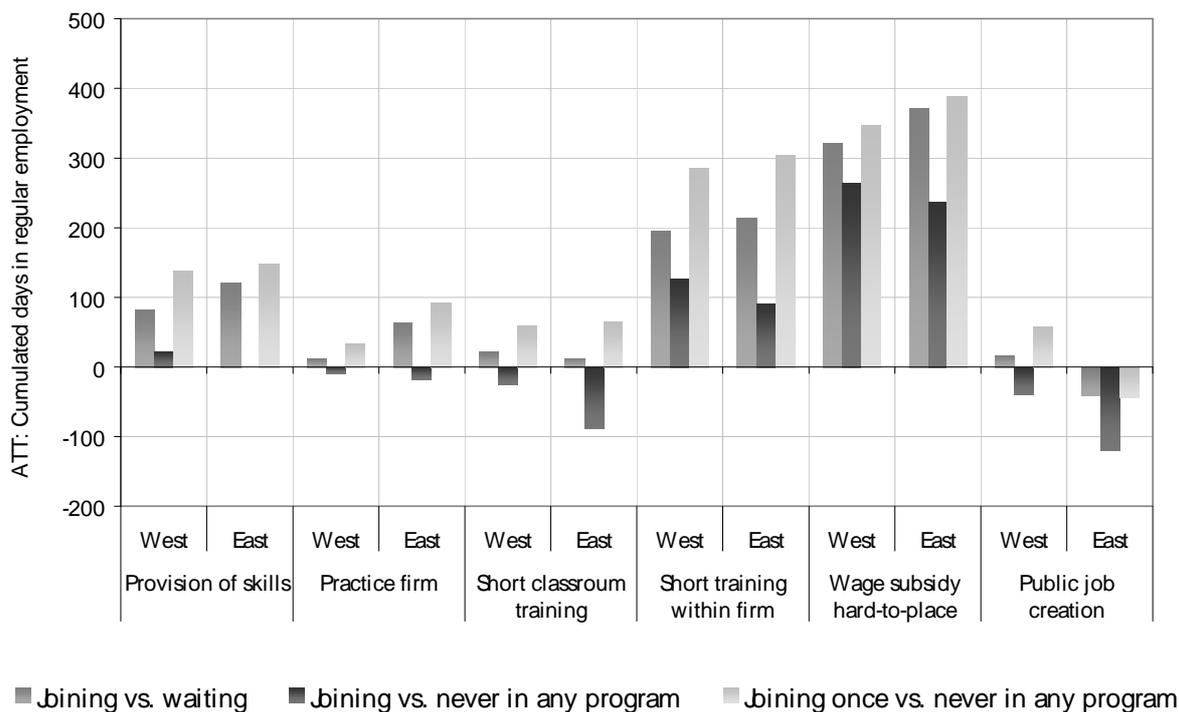
Source: Own calculations, based on the TrEffeR data.

Additional information can be found in Table 2, which also shows cumulated days spent in regular employment during the observation period for treated persons (T) and matched comparison persons (C). For instance, if we apply the "joining versus waiting" approach, individuals participating in further vocational training providing skills in West Germany were 483 days of 1260 days in regular employment, while comparable persons not entering a program during March 2003 were employed for 400 days. The estimated treatment effect is significant positive and amounts to 83 days – nearly three months. Furthermore, Table 2 contains the mean standardized bias (MSB) before and after matching as well as the bias reduction through matching. The latter results will not be discussed further, but indicate a rather good quality of the matching.

As can be seen most clearly from Figure 1, estimated effects differ systematically across programs, slightly across both regions of Germany – and they depend strongly on the classification window chosen. In what follows I will first discuss results by program, starting with the most "positive" programs. Second, results for West and East Germany are compared briefly. Third, the importance of the classification window chosen will be explored in detail.

Table 2
Cumulated days in regular employment during the 3 ½ years after program entry in March 2003 for treated persons (T) and matched comparison persons (C), average estimated treatment effects on the treated (ATT), mean standardized bias (MSB) before and after matching, bias reduction through matching, average duration of treatment and number of observations

			Days in regular employment			Mean standardized bias			Average duration	Obs.
			T	C	ATT	Before	After	Red.		
Provision of skills (4-6 months)	West	Joining vs. waiting	483	400	83 **	10.0	2.7	0.73	149	1414
		Joining vs. never	483	461	22	10.8	2.8	0.75	149	1414
		Joining once vs. never	585	447	138 **	11.8	2.4	0.80	151	713
	East	Joining vs. waiting	571	450	121 **	12.1	3.4	0.72	147	432
		Joining vs. never	572	571	0	12.5	3.7	0.71	147	432
		Joining once vs. never	720	572	148 **	14.0	5.3	0.62	146	171
Practice firm (4-6 months)	West	Joining vs. waiting	435	423	12	13.8	3.7	0.73	155	401
		Joining vs. never	435	444	-9	14.3	3.1	0.78	155	401
		Joining once vs. never	499	464	34	14.0	6.5	0.54	158	188
	East	Joining vs. waiting	460	397	63	10.4	4.5	0.57	150	163
		Joining vs. never	460	479	-19	11.7	5.5	0.53	150	163
		Joining once vs. never	599	507	92	14.0	8.1	0.42	145	60
Short classroom training	West	Joining vs. waiting	417	395	22 **	9.8	1.7	0.83	46	5716
		Joining vs. never	417	443	-26 **	10.4	1.8	0.83	46	5716
		Joining once vs. never	502	443	60 **	11.4	1.9	0.83	48	2698
	East	Joining vs. waiting	411	399	13	6.1	2.3	0.63	39	2403
		Joining vs. never	411	500	-89 **	7.3	2.1	0.72	39	2403
		Joining once vs. never	555	490	65 **	6.1	3.0	0.51	40	808
Short training within firm	West	Joining vs. waiting	636	441	195 **	9.2	1.9	0.79	33	3135
		Joining vs. never	636	510	126 **	10.5	1.7	0.84	33	3135
		Joining once vs. never	806	521	285 **	10.8	3.0	0.72	34	1502
	East	Joining vs. waiting	638	424	214 **	9.5	2.7	0.72	23	1057
		Joining vs. never	638	547	92 **	10.8	2.8	0.74	23	1057
		Joining once vs. never	853	549	304 **	10.5	3.7	0.65	23	383
Wage subsidy for hard-to-place worker (4-6 months)	West	Joining vs. waiting	673	352	321 **	13.4	2.8	0.79	167	734
		Joining vs. never	673	408	265 **	15.4	3.0	0.81	167	734
		Joining once vs. never	787	440	347 **	14.9	3.2	0.79	169	506
	East	Joining vs. waiting	670	299	372 **	13.8	3.9	0.72	172	345
		Joining vs. never	670	433	237 **	15.5	3.4	0.78	172	345
		Joining once vs. never	823	434	389 **	14.9	4.6	0.69	177	230
Public job creation (4-6 months)	West	Joining vs. waiting	296	279	17	18.7	3.4	0.82	172	449
		Joining vs. never	296	336	-40	19.1	3.7	0.81	172	449
		Joining once vs. never	409	351	58	18.1	6.1	0.66	169	186
	East	Joining vs. waiting	237	278	-42 *	14.6	2.7	0.82	175	792
		Joining vs. never	237	356	-119 **	14.0	3.4	0.76	175	792
		Joining once vs. never	293	338	-44	15.7	3.4	0.78	177	351

*) α = 0.05   **) α = 0.01.
Source: Own calculations, based on the TrEffeR data.

These results regarding the particular programs are so far in line with the literature cited in Section 2. Persistent positive treatment effects are found for hard-to-place individuals who received a wage subsidy. Short firm-internal training programs also exert a significant positive impact on cumulated days of regular employment. However, as has already been mentioned in Section 4, the estimates for both programs may partly suffer from selectivity at the firms' side. For further vocational training providing skills, the estimated treatment effects are – depending on the classification window chosen – either significantly positive or insignificant. For those who receive a treatment within a practice firm, insignificant effects are found. Short classroom training programs has partly positive, partly insignificant and partly negative effects on cumulated days of regular employment – again dependent on the underlying classification window. Finally, estimated treatment effects are insignificant or negative for those working in public job creation schemes.

Most programs investigated perform slightly better for persons treated in East Germany than for participants in West Germany. This is to a huge extent the result of the poor labor market situation in East Germany (especially for women), with generally worse opportunities to leave unemployment for comparison group members compared to West Germany. One noticeable exception is treatment in public job creation schemes, whose effectiveness for the treated is lower in East Germany.

Let us now turn to the main topic of the paper, the relationship between the size of the classification window and estimated program effectiveness. To give an example, comparing participants in further vocational training providing skills in West Germany with a matched group of "waiting" persons, we estimate that treated persons have spent a significant length of time, specifically 83 days or almost 3 months, more in regular employment during the 3 ½ years after program start. However, the program effect shrinks to an insignificant 22 days if we compare participants with a group of comparison persons who have never entered any program during the observation period. If, however, we restrict the analysis additionally to treated persons, who took-up no further program during the observation period, the estimated treatment effect increases again – to significantly positive 138 days spent more in regular employment.

The overall picture is similar: Table 2 shows that estimated effects are always lowest and in most cases negative if the classification window encompasses the entire observation period, while no restriction is put on the future of the treatment group ("joining versus never in any program"). In contrast, estimated effects are more positive, when the classification window encompasses only one month and no further restrictions on the future of the treated and the comparison persons are applied ("joining versus waiting").

Why does the "never in any program" comparison group perform better in the market than the "waiting" comparison group? First, as has been already discussed in Section 3, Sianesi (2004) and Fredriksson/Johansson (2004) argue that the former are a selected group of unemployed who did not enter a program because they did expect to find a job soon anyway. Also case-workers might find that these persons are not in need of participation. Hence the decision to participate may be the result of an adverse selection process, and future unemployment prospects partly explain the decision to participate during the observation period. This would imply that the Conditional Independence Assumption might not hold.

Second, participation in active labor market programs is in most cases associated with a significant lock-in-effect due to reduced search activity. In the "waiting" context the treatment group would then be compared with a comparison group that partly experiences lock-in effects from treatment also – only later. The later during the observation period the entry into a program occurs, the lower is the probability that positive program effects will outweigh these lock-in effects. The "never in any program" group has ex post avoided such lock-in effects. This argument will hold even if ex ante no unobserved heterogeneity is present.

Both effects cannot be disentangled empirically. Both raise, however, the question, how many of those "waiting" do in fact enter a program later during the observation period. Also individuals entering a program in March 2003 may enter another program during this period. That in particular short programs are often followed by further program participation has already been shown by Wunsch/Lechner (2008), who described further program participation of participants and non-participants in their sample.

Table 3 shows for treated and comparison persons, which share of each group enters further treatment during the 3 ½ years after program start (as has been depicted in Section 4, these shares might be even slightly underestimated). For instance, 49 percent of those taking-up further vocational training providing skills in West Germany in March 2003 received at least one further treatment during the 3 ½ years following. The first further program starting after March 2003 was in most cases – 22 percentage points – a short training program (while even 32 percentage points of the treatment group participated in such a measure during the entire observation period sooner or later). In contrast, "only" 37 percent of the matched comparison group took part in a labor market program after March 2003. Most of them – 18 percentage points – also participated first in a short training measure, while altogether 25 percentage points joined such a program at all.

The main contents of Table 3 are also summarized in Figure 2. It gives the impression that "program careers" occur rather often: Around 50 to 70 percent of those taking-up a program in March 2003 also had at least one further program during the observation period. In contrast the share of matched comparison persons participating in a program later on is generally 10 to 20 percentage points lower. The only exception is found for those who took up subsidized employment – the share found in further measures is only around 30 percent and 10 to 20 percentage points lower than across their matched control persons.

The results show clearly that both treatment effects discussed so far for those "joining" a program do not isolate the effect of a single intervention. If we restrict the sample to treated individuals who joined a program only once (in March 2003) and to potential comparison persons who never entered any program during the observation period, we obtain the third treatment effect displayed in Table 1 ("joining once versus never in any program"). Estimated treatment effects are typically highest if this wide classification window for the definition of non-treatment as well as for the identification of an "isolated" treatment is used. This effect is not distorted by further interventions for treated as well as for comparison persons and makes both groups investigated more similar. However, now also the treatment group will consist of a selected group of individuals, who did expect or were expected by caseworkers to find a job without entering a further program.

Table 3
Entry into further programs for treated and matched comparison persons during the 3 ½ years after program entry in March 2003

Treatment	(Further) treatment			At least one program after 3/2003	Variant of further vocational training		Variant of short training		Variant of wage subsidy		Variant of start-up subsidy		Variant public job creation		Other program	
					First program after 3/2003	Any program until 9/2006	First program after 3/2003	Any program until 9/2006	First program after 3/2003	Any program until 9/2006	First program after 3/2003	Any program until 9/2006	First program after 3/2003	Any program until 9/2006	First program after 3/2003	Any program until 9/2006
Provision of skills (4-6 months)	West	Treated		0.49	0.03	0.05	0.22	0.32	0.05	0.08	0.08	0.10	0.03	0.07	0.09	0.15
		Comparison		0.37	0.03	0.04	0.18	0.25	0.02	0.03	0.04	0.05	0.03	0.08	0.08	0.12
	East	Treated		0.60	0.08	0.10	0.25	0.40	0.09	0.17	0.03	0.05	0.08	0.15	0.08	0.16
		Comparison		0.49	0.03	0.05	0.22	0.35	0.02	0.06	0.04	0.05	0.11	0.19	0.07	0.11
Practice firm (4-6 months)	West	Treated		0.54	0.04	0.06	0.25	0.38	0.06	0.13	0.05	0.08	0.04	0.10	0.09	0.14
		Comparison		0.36	0.03	0.06	0.19	0.25	0.01	0.02	0.03	0.05	0.03	0.06	0.07	0.10
	East	Treated		0.63	0.07	0.12	0.33	0.52	0.07	0.18	0.03	0.07	0.08	0.15	0.06	0.11
		Comparison		0.48	0.03	0.04	0.16	0.22	0.03	0.05	0.08	0.09	0.13	0.21	0.05	0.13
Short classroom training	West	Treated		0.53	0.07	0.10	0.23	0.35	0.03	0.07	0.08	0.12	0.03	0.08	0.07	0.12
		Comparison		0.35	0.03	0.04	0.17	0.24	0.01	0.03	0.03	0.04	0.03	0.06	0.07	0.11
	East	Treated		0.66	0.10	0.14	0.27	0.46	0.08	0.15	0.06	0.10	0.11	0.21	0.05	0.13
		Comparison		0.48	0.04	0.05	0.21	0.31	0.03	0.07	0.03	0.04	0.12	0.22	0.06	0.11
Short training within firm	West	Treated		0.52	0.04	0.07	0.23	0.37	0.11	0.16	0.07	0.10	0.02	0.07	0.05	0.10
		Comparison		0.37	0.03	0.05	0.19	0.26	0.02	0.03	0.03	0.05	0.04	0.08	0.06	0.10
	East	Treated		0.64	0.04	0.07	0.24	0.44	0.22	0.30	0.04	0.08	0.06	0.11	0.04	0.09
		Comparison		0.49	0.04	0.06	0.22	0.33	0.03	0.07	0.03	0.05	0.11	0.19	0.07	0.12
Wage subsidy for hard-to-place worker (4-6 months)	West	Treated		0.31	0.03	0.04	0.13	0.19	0.02	0.05	0.05	0.07	0.01	0.03	0.07	0.08
		Comparison		0.44	0.03	0.05	0.20	0.31	0.02	0.03	0.03	0.04	0.06	0.15	0.09	0.13
	East	Treated		0.33	0.02	0.02	0.13	0.19	0.03	0.07	0.03	0.04	0.04	0.10	0.07	0.11
		Control		0.56	0.02	0.05	0.27	0.37	0.04	0.07	0.02	0.04	0.14	0.28	0.07	0.13
Public job creation (4-6 months)	West	Treated		0.59	0.05	0.06	0.19	0.30	0.03	0.07	0.04	0.07	0.16	0.31	0.10	0.15
		Comparison		0.38	0.02	0.03	0.15	0.22	0.02	0.03	0.03	0.04	0.11	0.20	0.07	0.11
	East	Treated		0.57	0.01	0.02	0.20	0.28	0.05	0.10	0.03	0.05	0.20	0.34	0.08	0.13
		Control		0.48	0.01	0.04	0.18	0.28	0.03	0.07	0.02	0.02	0.18	0.29	0.06	0.11

Source: Own calculations, based on the TrEffeR data.

Figure 2
Share of treated persons entering at least one further program and of comparison persons entering at least one program during the 3 ½ years after March 2003
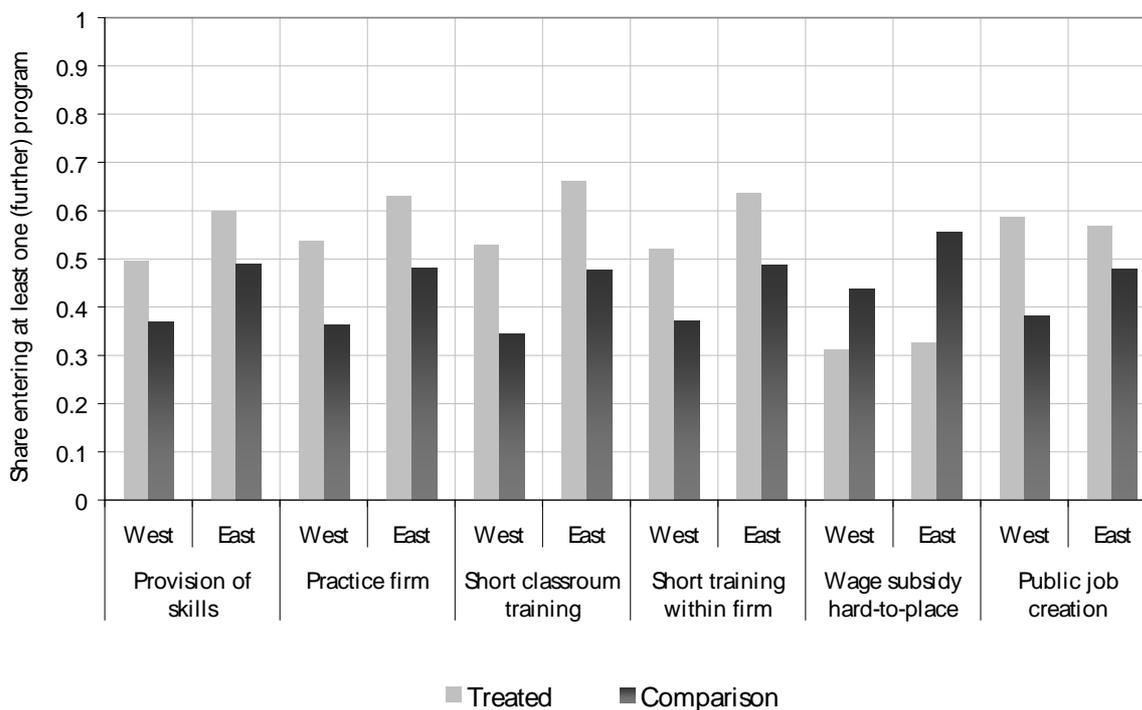
Source: Own calculations, based on the TrEffeR data.

All estimates have been replicated also for the outcome variable "cumulated days not in un-employment and not in a labor market program", restricting the analysis on the group of indi-viduals from local labor markets without a municipality opting out of the co-operation with the Public Employment Service (see section 4). Even though estimated treatment effects are smaller than those found for the outcome variable "cumulated days in regular employment", as is often the case (see section 2), the results show exactly the same pattern: Estimated treat-ment effects depend strongly on the classification window chosen for the definition of non-treatment, with more positive results for a small than for a wide classification window.

6   Conclusions

The decision for a classification window that defines treatment and non-treatment is a crucial issue in evaluation studies of active labor market programs. Whereas a social experiment al-lows for a random assignment of individuals to treatment and control group, a major chal-lenge of observation studies is to control for selection mechanisms. A broader classification window might easily introduce selection on future expected outcomes of participants. This paper estimated the program effects for participants in different German labor market pro-grams empirically, applying one very small (one month) and one very wide (3 ½ years) classifi-cation window.

Since results are sensitive to the choice of the classification window, a main conclusion of the paper is that researchers should be careful in interpreting the results of evaluation studies of active labor market programs using statistical matching techniques: The meaning of estimated

treatment effects as well as the size of estimated effects hinge strongly on the choice of treatment and comparison group. This is in particularly obvious for training programs. These programs have been shown to have typically rather weak effects. Thus the design of the study could make the difference between obtaining significant positive, insignificant or negative effects. Thus authors of evaluation studies should spend some effort to describe their classification window chosen very clearly and to discuss the implications of this aspect of their evaluation design.

The paper demonstrates also that none of the approaches applied here is appropriate to isolate the effect of a certain program and at the same time to avoid selectivity: The "joining versus waiting" design has become more and more common in the evaluation of active labor market policies in Europe. Its drawback is that it cannot be used to isolate the effects of a single intervention, which also prevents a clear analysis of efficiency of the program – members of the treatment as well of the comparison group may enter other programs later, which might induce further costs. In contrast, the "joining versus never in any program" design, which compares treated persons with those entering no program at all during the observation period, might be prone to unobserved heterogeneity on the side of the comparison group. Finally, the "joining once versus never in any program" design introduces a similar restraint for program participants. Thus unobserved selectivity will be partly removed across treatment and comparison groups investigated – but estimates are not valid for the entire number of program participants.

A further result of the paper is that program careers are an important and still under-researched issue. For most programs investigated here, the majority of participants enter at least one further program during the 3 ½ years after taking-up a program in March 2003. Several recent papers deal with dynamic treatment assignment (Adda et al. 2006, Lechner 2004, 2006, Lechner/Wiehler 2007b), the literature in this field is, however, still sparse.

References

Abbring, J.H, Van den Berg, G. (2003), The non-parametric identification of treatment effects in duration models, Econometrica 71, 1491-1518.

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

Adda, J., Costa Dias, M., Meghir, C., Sianesi, B. (2006), Evaluating Dynamic Treatment: The Swedish Active Labor Market Programmes, mimeo

Baumgartner, H., Caliendo, M. (2007), Turning Unemployment into Self-Employment, Effectiveness and Efficiency of Two Start-Up Programmes, DIW Discussion Paper 671.

Bernhard, S., Wolff, J. (2008), Contracting out placement services in Germany. Is assignment to private providers effective for needy job-seekers?, IAB-Discussion-Paper 5/2008.

Biewen, M., Fitzenberger, B., Osikominu, A., Waller, M. (2007), Which Program for Whom? Evidence on the Comparative Effectiveness of Public Sponsored Training Programs in Germany, IZA Discussion Paper 2885, Bonn.

Blien, U., Hirschenauer, F., Arendt, M., Braun, H.-J., Gunst, D.-M., Kilcioglu, S., Kleinschmidt, H., Musati, M., Roß, H., Vollkommer, D., Wein, J. (2004), Typisierung von Bezirken der Agenturen für Arbeit, Zeitschrift für Arbeitsmarktforschung 37, 146-175.

Boockmann, B., Zwick, T., Ammermüller, A., Maier, M. (2007), Do Hiring Subsidies Reduce Unemployment Among the Elderly? Evidence from Two Natural Experiments, ZEW-Discussion Paper 07-001.

Bundesagentur für Arbeit (2006), Arbeitmarkt 2006, Amtliche Nachrichten der Bundesagentur für Arbeit 55, Sondernummer 1.

Büttner, T. (2007), Ankündigungseffekt oder Maßnahmewirkung? Eine Evaluation von Trainingsmaßnahmen zur Überprüfung der Verfügbarkeit, IAB Discussion Paper 25/2007.

Deutscher Bundestag (2006), Bericht der Bundesregierung zur Wirksamkeit moderner Dienstleistungen am Arbeitsmarkt, Unterrichtung durch die Bundesregierung, Bundestagsdrucksache 16/3982, Berlin.

Caliendo, M., Hujer, R., Thomsen, S. (2005a), The Employment Effect of Job Creation Schemes in Germany, A Microeconometric Evaluation, IZA Discussion Paper 1512, forthcoming in: Advances in Econometrics.

Caliendo, M., Hujer, R., Thomsen, S. (2005b), Identifying Effect Heterogeneity to Improve the Efficiency of Job Creation Schemes in Germany, IAB Discussion Paper No. 8/05, forthcoming in: Applied Economics.

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

Fitzenberger, B., Osikominu, A., Völter, R. (2006), Get training or wait? Long-run employment effects of training programs for the unemployed in West Germany, IAB Discussion Paper 17/2006.

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

Fitzenberger, B., Völter, R. (2007), Long-run effects of training programs for the unemployed in East Germany, Labor Economics 14, 370-755.

Frankfurter Allgemeine Zeitung (August 28, 2007), Milliardenloch am Arbeitsmarkt, 13.

Frederiksson, P., Johansson, P. (2004), Dynamic Treatment Assignment – The Consequences for Evaluations Using Observation Data, IZA Discussion Paper 1062.

Heckman, J. J., Lalonde, R. L., Smith, J. A. (1999), The Economics and Econometrics of Active Labor Market Programs. In: Ashenfelter, O., Card, D. (Hrsg.) Handbook of Labor Economics 3, Amsterdam, S. 1865-2097.

Hess, D., Kaps, P., Mosley, H. (2006), Implementations- und Wirkungsanalyse der Personal-Service-Agentur, Vierteljahreshefte zur Wirtschaftsforschung 75, 9-31.

Hohmeyer, K., Wolff, J. (2007), A fistful of Euros, Does One-Euro-Job participation lead means-tested benefit recipients into regular jobs and out of unemployment benefit II receipt?, IAB-Discussion Paper 32/2007.

Hujer, R., Thomsen, S., Zeiss, C. (2006a), The Effects of Short-Term Training Measures on the Individual Unemployment Duration in West Germany, ZEW-Discussion Paper 06-065, Mannheim.

Hujer, R., Thomsen, S., Zeiss, C. (2006b), The Effects of Vocational Training Programmes on the Duration of Unemployment in Eastern Germany, Allgemeines Statistisches Archiv 90, 299-322.

Hujer, R., Thomsen, S. (2006c), How Do Employment Effects of Job Creation Schemes Differ with Respect to the Foregoing Unemployment Duration, ZEW Discussion Paper 06-047.

Hujer, R., Zeiss, C. (2006d), The Effects of Job Creation Schemes on the Unemployment Duration in East Germany, IAB Discussion Paper 16/2006.

Jaenichen, U., Stephan, G. (2007), The Effectiveness of Targeted Wage Subsidies for Hard-to-Place Workers, IAB Discussion Paper 16/2007.

Lalive, R., van Ours, J., Zweimüller, J. (2008), The impact of active labor market programmes on the duration of unemployment in Switzerland, The Economic Journal 118, 235-257.

Lechner, M. (2004), Sequential Matching Estimation of Dynamic Causal Models, IZA Discussion Paper 1042

Lechner, M., Miquel, R., Wunsch, C. (2005), Long-Run Effects of Public Sector Sponsored Training in West Germany, IAB Discussion Paper 03/2005.

Lechner, M. (2006), Matching Estimating of Dynamic Treatment Models: Some Practical Issues, University of St. Gallen, mimeo.

Lechner, M., Miquel, R., Wunsch, C. (2007a), The Curse and Blessing of Training the Unemployed in a Changing Economy: The Case of East Germany after Unification, German Economic Review 8, 468-509.

Lechner, M., Wiehler, S. (2007b), Does the Order and Timing of Active Labor Market Programs matter?, University of St. Gallen, mimeo.

Leuven, E., Sianesi, B. (2003), psmatch2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing, http://ideas.repec.org/c/boc/bocode/s432001.html

Rinne, U., Schneider, M., Uhlendorff, A. (2007), Too Bad to Benefit? Effect Heterogeneity of Public Training Programs. IZA Discussion Paper 3240.

Rosenbaum, P. R., Rubin, D. B. (1983), The Central Role of the Propensity Score in Observational Studies for Causal Effects, Biometrika 70, 41-55.

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

Rubin, D.B. (1978), Bayesian Inference for Causal Effects: The Role of Randomization, Annals of Statistics 6, 34-58.

Sianesi, B. (2002), Swedish Active Labor Market Programmes in the 1990s: Overall Effectiveness and Differential Performance, Institute for Fiscal Studies Working Paper 02/03

Sianesi, B. (2004), An Evaluation of the Swedish System of Active Labor Market Programs in the 1990s, The Review of Economics and Statistics 86, 133-155.

Steiger, H. (2004) Is less more? A look at nonparticipation in Swiss active labor market programmes, University of St. Gallen, mimeo.

Stephan, G., Rässler, S., Schewe, T. (2006), Das TrEffeR-Projekt der Bundesagentur für Arbeit. Die Wirkung von Maßnahmen aktiver Arbeitsmarktpolitik, Zeitschrift für Arbeitsmarktforschung 39, 447-465.

Winterhager, H., Heinze, A., Spermann, A. (2006), Deregulatin job placement in Europe: A microeconometric evaluation of an innovative voucher scheme in Germany, Labor Economics 13, 505-517.

Winterhager, H., (2008), Öffentlich geförderte Beschäftigung von Arbeits-losen, ZEW Wirtschaftsanalysen 84, Baden Baden.

Wolff, J., Jozwiak, E. (2007), Does short-term training activate means-tested unemployment benefit recipients in Germany? IAB Discussion Paper 29/2007 Nürnberg.

Wunsch, C. und Lechner, M. (2008), What Did All the Money Do? On the General Ineffectiveness of Recent West German Labor Market Programmes, Kyklos 61, 134–174.

Appendix: Additional Figures and Tables

Figure A.1
Estimated effects on cumulated days in regular emplyoment during the 3 ½ years after program entry in March 2003
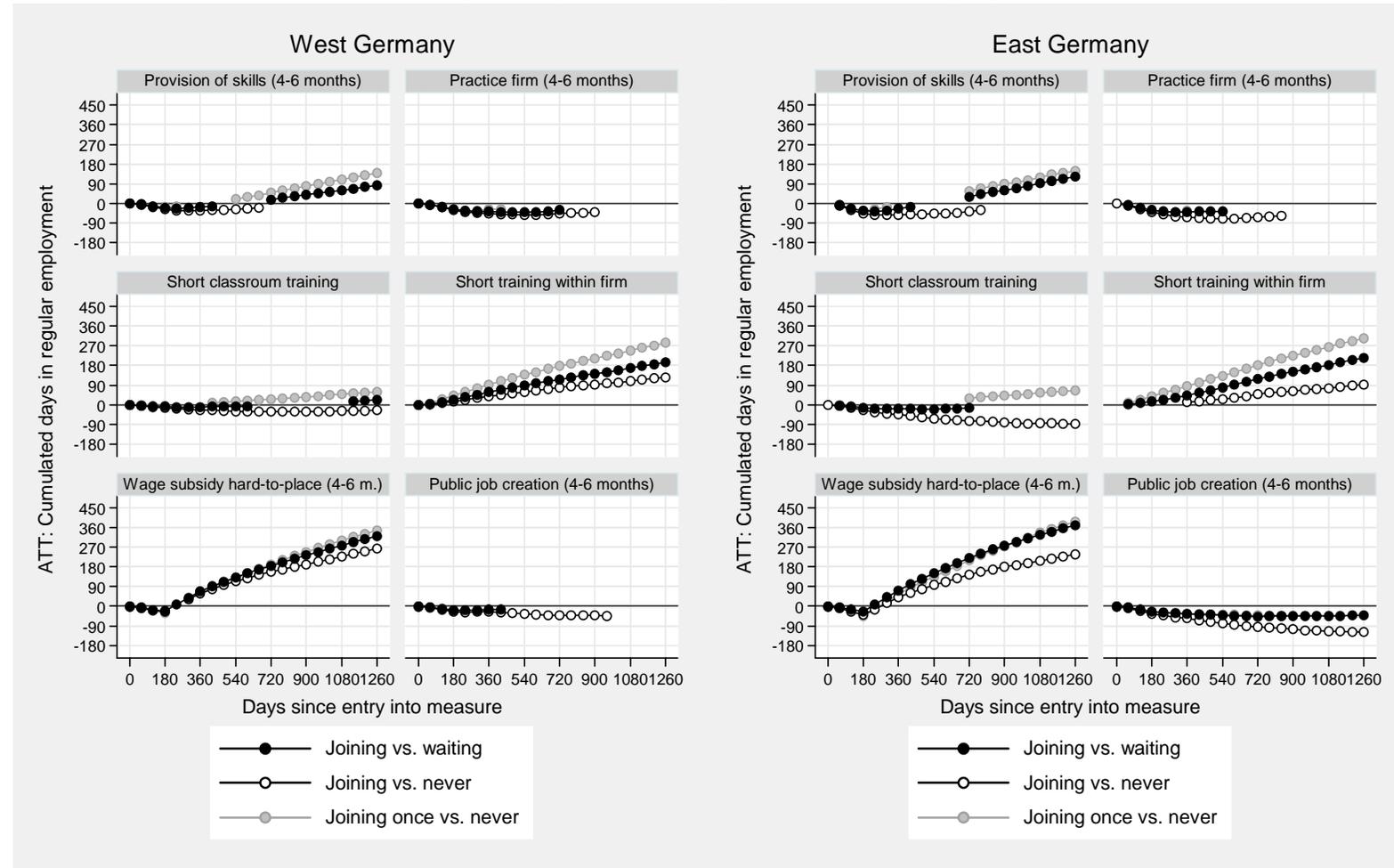
Source: Own calculations, based on the TrEffeR data. Only effects significant at α = 0.05 are displayed.

Table A.1
Variable means of selected attributes (0 = no, 1 = yes) for treated and potential comparison persons with (I) and without (II) participation in labor market measures during the 3 ½ years after March 2003

		Provision of skills (4 to 6 months)		Practice firm (4 to 6 months)		Short classroom training		Short training within firm		Wage subsidy for hard-to-place (4 to 6 months)		Public job creation scheme (4 to 6 months)		Potential controls I		Potential controls II	
		West	East	West	East	West	East	West	East	West	East	West	East	West	East	West	East
a) Individual characteristics	Female	0.47	0.32	0.56	0.40	0.55	0.44	0.33	0.33	0.33	0.44	0.47	0.42	0.41	0.43	0.44	0.45
	Age 26-30	0.16	0.16	0.15	0.15	0.16	0.15	0.24	0.19	0.17	0.13	0.12	0.05	0.15	0.12	0.14	0.11
	Age 31-35	0.21	0.17	0.19	0.17	0.19	0.17	0.20	0.18	0.21	0.20	0.11	0.07	0.17	0.14	0.16	0.13
	Age 36-8	0.24	0.22	0.18	0.13	0.22	0.19	0.20	0.22	0.22	0.25	0.15	0.10	0.18	0.16	0.17	0.15
	Age 41-45	0.19	0.19	0.21	0.21	0.19	0.19	0.17	0.18	0.20	0.23	0.15	0.14	0.16	0.17	0.15	0.16
	Age 46-50	0.13	0.15	0.18	0.17	0.13	0.16	0.11	0.12	0.16	0.18	0.16	0.20	0.13	0.17	0.13	0.16
	Age 51-54	0.06	0.10	0.07	0.13	0.09	0.11	0.06	0.09	0.02	0.01	0.21	0.25	0.12	0.15	0.13	0.14
	Age 55-59	0.01	0.00	0.01	0.04	0.02	0.04	0.02	0.02	0.01	0.00	0.11	0.19	0.09	0.10	0.12	0.14
	Health problems	0.06	0.04	0.05	0.03	0.06	0.06	0.08	0.04	0.10	0.06	0.15	0.09	0.13	0.10	0.13	0.11
	Slightly disabled	0.01	0.00	0.01	0.01	0.02	0.02	0.02	0.01	0.02	0.01	0.02	0.02	0.03	0.02	0.03	0.02
	Severly disabeled	0.02	0.01	0.03	0.02	0.02	0.02	0.04	0.02	0.01	0.00	0.10	0.07	0.04	0.02	0.05	0.03
	Married	0.54	0.58	0.55	0.64	0.51	0.54	0.46	0.54	0.45	0.51	0.48	0.71	0.53	0.55	0.57	0.59
	Married and female	0.26	0.20	0.33	0.25	0.29	0.26	0.13	0.21	0.14	0.27	0.24	0.31	0.24	0.27	0.28	0.30
	Foreigner	0.13	0.03	0.13	0.02	0.15	0.02	0.11	0.02	0.13	0.01	0.11	0.00	0.18	0.02	0.18	0.02
	Secondary degree (Hauptschule)	0.40	0.16	0.46	0.13	0.40	0.21	0.47	0.18	0.51	0.21	0.34	0.31	0.52	0.30	0.53	0.30
	Secondary degree (Realschule)	0.29	0.67	0.34	0.68	0.27	0.65	0.28	0.67	0.23	0.69	0.34	0.56	0.19	0.57	0.19	0.56
	Secondary degree (Gymnasium)	0.23	0.13	0.13	0.17	0.24	0.12	0.19	0.11	0.15	0.07	0.20	0.08	0.14	0.07	0.14	0.08
	Vocational training	0.62	0.85	0.71	0.83	0.61	0.84	0.69	0.87	0.60	0.89	0.59	0.89	0.54	0.82	0.54	0.83
	University degree	0.11	0.08	0.03	0.13	0.12	0.07	0.08	0.06	0.05	0.03	0.12	0.05	0.07	0.04	0.07	0.04
b) Timing of program entry	1st month of unemployment	0.09	0.13	0.09	0.10	0.11	0.08	0.17	0.14	0.15	0.21	0.10	0.04	0.13	0.12	0.13	0.12
	2nd month of unemployment	0.11	0.12	0.09	0.10	0.14	0.11	0.15	0.13	0.09	0.06	0.09	0.07	0.13	0.11	0.13	0.12
	3rd month of unemployment	0.13	0.14	0.16	0.15	0.18	0.17	0.15	0.17	0.11	0.10	0.08	0.08	0.16	0.17	0.17	0.18
	4th month of unemployment	0.11	0.11	0.11	0.12	0.10	0.12	0.11	0.11	0.09	0.10	0.10	0.10	0.11	0.12	0.11	0.12
	5th month of unemployment	0.10	0.09	0.08	0.10	0.10	0.11	0.09	0.09	0.07	0.05	0.10	0.08	0.09	0.09	0.09	0.09
	6th month of unemployment	0.10	0.09	0.12	0.06	0.09	0.08	0.07	0.08	0.08	0.09	0.09	0.10	0.08	0.08	0.08	0.07
	7th month of unemployment	0.08	0.07	0.07	0.03	0.07	0.07	0.06	0.07	0.07	0.08	0.09	0.09	0.07	0.06	0.06	0.06
	8th month of unemployment	0.06	0.08	0.07	0.06	0.05	0.05	0.04	0.04	0.07	0.05	0.07	0.10	0.05	0.05	0.05	0.05
	9th month of unemployment	0.06	0.06	0.08	0.09	0.05	0.07	0.06	0.07	0.08	0.08	0.09	0.10	0.06	0.06	0.05	0.05
	10th month of unemployment	0.06	0.04	0.04	0.07	0.04	0.05	0.03	0.04	0.05	0.06	0.06	0.09	0.04	0.05	0.04	0.05
	11th month of unemployment	0.05	0.03	0.02	0.06	0.03	0.04	0.03	0.04	0.05	0.05	0.07	0.06	0.04	0.05	0.04	0.04
	12th month of unemployment	0.05	0.04	0.06	0.06	0.03	0.05	0.03	0.04	0.07	0.08	0.06	0.10	0.04	0.05	0.04	0.04
c) 2-years-history	Unemployed up to 1 month	0.57	0.53	0.57	0.55	0.63	0.49	0.53	0.42	0.37	0.30	0.37	0.35	0.45	0.28	0.47	0.30
	Unemployed 1-6 months	0.19	0.20	0.23	0.19	0.19	0.19	0.24	0.22	0.16	0.12	0.22	0.16	0.19	0.18	0.19	0.19
	Unemployed 7-12 months	0.14	0.19	0.14	0.15	0.12	0.18	0.16	0.24	0.24	0.27	0.20	0.22	0.17	0.23	0.17	0.22
	Unemployed 13-18 months	0.06	0.08	0.06	0.12	0.06	0.14	0.08	0.12	0.22	0.31	0.12	0.16	0.10	0.18	0.09	0.17
	Unemployed 19-24 months	0.04	0.06	0.03	0.08	0.04	0.10	0.03	0.06	0.12	0.20	0.10	0.11	0.08	0.13	0.08	0.13
	Labor market measure	0.21	0.35	0.19	0.39	0.21	0.42	0.30	0.47	0.46	0.71	0.35	0.39	0.21	0.42	0.17	0.36
	Period of sickness	0.08	0.15	0.11	0.13	0.08	0.17	0.09	0.14	0.18	0.25	0.16	0.19	0.13	0.20	0.13	0.20
	Sanctions	0.02	0.01	0.02	0.01	0.03	0.01	0.02	0.01	0.03	0.02	0.02	0.00	0.04	0.02	0.05	0.02
	Observations	1414	432	401	163	5716	2403	3135	1057	734	345	449	792	144406	57435	96153	30921

Source: Own calculations, based on the TrEffeR data.

Recently published

No.	Author(s)	Title	Date
24/2007	Bruckmeier, K. Schnitzlein, D.	Was wurde aus den Arbeitslosenhilfeempfängern? : Eine empirische Analyse des Übergangs und Verbleibs von Arbeitslosenhilfeempfängern nach der Hartz-IV-Reform	9/07
25/2007	Büttner, T.	Ankündigungseffekt oder Maßnahmewirkung? : Eine Evaluation von Trainingsmaßnahmen zur Überprüfung der Verfügbarkeit	10/07
26/2007	Brücker, H. Defoort, C.	Inequality and the (self-)selection of international migrants: Theory and novel evidence	10/07
27/2007	Brücker, H. Schröder, P.J.H.	International migration with heterogeneous agents: Theory and evidence	10/07
28/2007	Krug, G.	In-work benefits for low wage jobs: Can additional income hinder labor market integration?	11/07
29/2007	Wolff, J. Jozwiak, E.	Does short-term training activate means-tested unemployment benefit recipients in Germany?	11/07
30/2007	König, M. Möller, J.	Mindestlohneffekte des Entsendegesetzes? : Eine Mikrodatenanalyse für die deutsche Bauwirtschaft	11/07
32/2007	Hohmeyer, K. Wolff, J.	A fistful of Euros: Does One-Euro-Job participation lead means-tested benefit recipients into regular jobs and out of unemployment benefit II receipt?	12/07
1/2008	Schnitzlein, D.	Verbunden über Generationen: Struktur und Ausmaß der intergenerationalen Einkommensmobilität in Deutschland	1/08
2/2008	Bauer, T. K. Schmucker, A. Vorell, M.	KMU und Arbeitsplatzdynamik: Eine Analyse auf Basis der Beschäftigten-Historik-Datei	1/08
3/2008	Baas, T. Brücker, H.	Macroeconomic consequences of migration diversion: A CGE simulation for Germany and the UK	1/08
4/2008	Amend, E. Herbst, P.	Labor market pooling and human capital investment decisions	1/08
5/2008	Bernhard, S. Wolff, J.	Contracting out placement services in Germany: Is assignment to private providers effective for needy jobseekers?	2/08
6/2008	Lehmer, F. Ludsteck, J.	The returns to job mobility and inter-regional migration	2/08
7/2008	Klinger, S. Wolf, K.	What explains changes in full-time and part-time employment in Western Germany? : A new method on an old question	2/08
8/2008	Hohmeyer, K. Jozwiak, E.	Who is targeted by One-Euro-Jobs? : A selectivity analysis	2/08
9/2008	Schnabel, C. Kohaut, S. Brixy, U.	Employment stability of entrants in newly founded firms: A matching approach using linked employer-employee data from Germany	2/08
10/2008	Konle-Seidl, R.	Changes in the governance of employment services in Germany since 2003	2/08
11/2008	Ludsteck, J.	Adjustment from the bottom: Zoom evidence on the cyclicality of wages and the wage curve in West Germany	2/08

As per: 27.02.2008

For a full list, consult the IAB website
http://www.iab.de/de/publikationen/discussionpaper.aspx

Imprint

IAB-Discussion Paper 12/2008

Editorial address
Institute for Employment Research
of the Federal Employment Agency
Regensburger Str. 104
D-90478 Nuremberg

Editorial staff
Regina Stoll, Jutta Palm-Nowak

Technical completion
Jutta Sebald

Website
http://www.iab.de

Download of this Discussion Paper
http://doku.iab.de/discussionpapers/2008/dp1208.pdf

For further inquiries contact the author:

Gesine Stephan
Phone +49.911.179 5850
E-mail gesine.stephan@iab.de