Employment Effects of Short and Medium Term Further Training Programs in Germany in the Early 2000s[1]

Martin Biewen,[♭] Bernd Fitzenberger,[♮] Aderonke Osikominu,[♯] and Marie Waller[§]

[♭]Goethe University Frankfurt, IZA, DIW, biewen@wiwi.uni-frankfurt.de

[♮]Goethe University Frankfurt, ZEW, IZA, IFS, fitzenberger@wiwi.uni-frankfurt.de

[♯]Goethe University Frankfurt, osikominu@wiwi.uni-frankfurt.de

[§]Goethe University Frankfurt, CDSEM University of Mannheim, ZEW, f-waller@zew.de

Correspondence: Department of Economics, Goethe University, PO Box 11 19 32 (PF 247), 60054 Frankfurt am Main, Germany.

This version: March 21, 2006

Preliminary! Please do not quote without permission of the authors!

Abstract: We use a new and exceptionally rich administrative data set for Germany to evaluate the employment effects of several short and medium term further training programs in the early 2000s. Building on the work of Sianesi (2003, 2004), we apply propensity score matching methods in a dynamic, multiple treatment framework in order to address program heterogeneity and dynamic selection into programs. Our results suggest that in West Germany both short-term and medium-term programs may have considerable employment effects for certain population subgroups, and that short-term programs are surprisingly effective when compared to the traditional and more expensive longer-term programs. With a few exceptions, we find little evidence for significant treatment effects in East Germany.

Keywords: evaluation, multiple treatments, dynamic treatment effects, local linear matching, active labor market programs, administrative data

JEL: C 14, J 68, H 43

[1]This study is part of the project "Employment effects of further training programs 2000-2002 – An evaluation based on register data provided by the Institute of Employment Research, IAB (Die Beschäftigungswirkungen der FbW-Maßnahmen 2000-2002 auf individueller Ebene – Eine Evaluation auf Basis der prozessproduzierten Daten des IAB)" (IAB project number 6-531.1A). The project is joint with the Swiss Institute for Internationale Economics and Applied Economic Research at the University of St. Gallen (SIAW) and the Institut für Arbeitsmarkt- und Berufsforschung (IAB). We gratefully acknowledge financial support by the IAB. All errors are our sole responsibility.

Contents

1 Introduction

Each year, the German government spends about 20 billion Euros[2] on active labor market policies. A considerable part of these resources flows into different types of training programs aimed at providing general and specific skills to unemployed individuals. In contrast to public sector sponsored training in other countries, German training schemes are traditionally long-term programs requiring full-time participation. In fact, the average length of such programs ranges from 6 to 12 months. Following criticism that such programs may not be effective as they "lock-in" the participants for a long time, there has been a shift towards short-term programs recently.[3] Lasting typically some weeks only, these schemes are less expensive and do not release participants from continuing their job search. However, due to lacking empirical evidence, it remains an open question how such short-term programs compare to the traditional medium to long-term programs in terms of effectiveness.

In this paper, we employ a dynamic multiple treatment framework to compare the employment effects of short-term training to those of traditional medium to long-term programs (called further training in the following). Within the group of further training programs we distinguish between pure classroom training and programs that include "hands-on experience" in the form of internships or working in practice firms. Building on the work of Sianesi (2003, 2004), we apply propensity score matching methods in a dynamic, multiple treatment framework. In order to take account of the dynamic sorting process among the unemployed, treatment effects are stratified by elapsed duration of unemployment and estimated using local linear matching based on the propensity score as well as on the calendar month of the beginning of the unemployment spell. Treatment status is defined subject to the time window of elapsed unemployment duration. The treatment parameters we estimate thus mirror the decision problem of the case worker and the unemployed who recurrently during the unemployment spell decide whether to start any of the programs now or to postpone participation to the future.

For a long time, it was difficult to obtain informative statistical results on the effectiveness of active labor market policies in Germany. Existing data sets were either too small or lacked detailed information on program participation. Evaluation results therefore often did not find statistically significant effects and were not able to address aspects such as program heterogeneity or dynamic selection into pro-

[2]19.52 billion Euros in 2004, see Bundesagentur für Arbeit (2005a).

[3]See Bundesagentur für Arbeit (2005b), p. 83.

grams.[4] It has not been until very recently, that sufficiently large and informative data sets have become available. Using such a data set, Lechner, Miquel and Wunsch (2005a,b), Fitzenberger and Speckesser (2005), Fitzenberger, Osikominu and Völter (2005) have found positive employment effects for some programs and some subgroups of participants in further training programs in the 1980s and 1990s. In this evaluation study we use another new and exceptionally rich data set, the Integrated Employment Biographies Sample (IEBS). The IEBS is based on administrative records and comprises detailed daily information on employment and benefit reception histories from 1990 onwards as well as detailed information on unemployment and participation in different programs of active labor market policy from 2000 onwards. Moreover, our data set includes a rich set of covariates that allow to control for the selection into the different programs, in particular information on health, education and family characteristics, as well as detailed regional and sectoral information. As for instance Lechner, Miquel and Wunsch (2005b) have found that treatment effects may differ for different population subgroups, we carry out our analysis for West and East Germany, and for male and female participants separately. In particular, we analyze treatment effects for an inflow sample of individuals who became unemployed between February 2000 and January 2002, and some of those participated in one of the training programs described above at some point of their unemployment spell.

The remainder of the paper is structured as follows: Section 2 provides a short description of the institutional regulations for active labor market policies in Germany. Section 3 focuses on the data and the different training programs analyzed in this study. Section 4 describes our methodological approach to estimating treatment effects. The empirical results are discussed in section 5. Section 6 concludes.

2 Training Schemes within German Active Labor Market Policy

2.1 Basic Regulation

The central goal of active labor market policy in Germany is to permanently reintegrate unemployed individuals or individuals who are at risk of becoming unemployed

[4] A recent survey can be found in Speckesser (2004, chapter 1).

back into the labor market.[5] Special attention is paid to problem groups such as long-term unemployed, the elderly, disabled persons, and women reentering the labor market after parental leave. German active labor market policy comprises a variety of measures ranging from subsidized employment to training programs aimed at improving the qualifications of the participants. For an overview over different kinds of policies and their quantitative importance, see tables 1 to 3 in the appendix.

As to training measures, German legislation distinguishes three main types of training: further training *(Berufliche Weiterbildung)*, retraining *(Umschulung)*, and short-term training *(Trainingsmaßnahmen und Maßnahmen der Eignungsfeststellung)*.[6] In general, all three types of training measures require full-time participation. The different kinds of training measures differ considerably in length and contents. The label 'further training' subsumes different medium term training measures that last several months. On the one hand, there are advanced training or refresher courses imparting professional skills and techniques through classroom or on-the-job training. On the other hand, the heading 'further training' also includes practical training programs, the so-called practice firms or workshops, that provide rather general skills. Participants in these programs have the opportunity to practice everyday working activities in a simulated work environment. The most comprehensive training scheme is retraining. This program type lasts two to three years and typically leads to a new vocational education degree within the German apprenticeship system. In general, it comprises periods of classroom training as well as internships. By contrast, short-term training courses last only two to twelve weeks. The aim of this type of training is twofold. On the one hand, it provides skills that facilitate job search, e.g. through job application training or basic computer courses. On the other hand, it is used to assess and to monitor the abilities and the willingness to work of the unemployed.

To become eligible for participation in an active labor market program, job seekers have to register personally at the local labor office. This involves a counseling interview with the caseworker. Besides being registered as unemployed or as a job seeker at risk of becoming unemployed, candidates for short-term training measures

[5] This study focuses on active labor market policies in Germany in the years 2000 to 2002, the period just before the so-called Hartz-reforms became effective. The following paragraphs give a brief overview over the institutional setting in this period. In 2003, there was a major change in the regulation of the assignment into training measures (the Hartz I-reform, see Biewen and Fitzenberger (2004)).

[6] Furthermore, there are specific training schemes for adolescents and disabled persons, as well as German language courses for asylum seekers or ethnic Germans returning from former German settlements in Eastern Europe. These training measures are not analyzed here.

do not have to fulfil any additional eligibility criteria. As regards to medium- and long-term training measures, individuals are in principal eligible only if they also fulfil a minimum work requirement of one year and are entitled to unemployment compensation. However, there are several exceptions to these requirements. The really binding criterium is that the training scheme has to be considered necessary in order for the job seeker to find a new job. This is for instance the case if the employment chances in the target occupation of a job seeker are good but require an additional adjustment of skills. Training measures are usually assigned by the caseworker. The registered job seekers may also take the initiative, but their proposition has to be approved by the caseworker. Suitable training programs are chosen from a pool of certified public or private institutions or firms.

Active labor market policy is complemented by passive measures. The unemployment compensation system distinguishes three kinds of transfers: unemployment benefits (*Arbeitslosengeld*), unemployment assistance (*Arbeitslosenhilfe*) and subsistence allowance (*Unterhaltsgeld*). Unemployment compensation, in contrast to social assistance, is granted to individuals who contributed to the unemployment insurance in the past and who are able and available to work. Registered unemployed who fulfil a minimum work requirement of twelve months within the last three years are entitled to unemployment benefits. The amount and the entitlement period of unemployment benefits depend on age, previous earnings, previous employment experience, and family status. After expiration of their unemployment benefits, unemployed individuals may receive the lower, means tested unemployment assistance. Subsistence payments are paid to participants in further and retraining programs who fulfil the eligibility criteria stated above. It is usually of the same amount as the transfer payment the unemployed received before. Once granted, subsistence allowance is paid at least as long as the unemployed participates in the program. Overall, there are no significant financial incentives for unemployed to participate in a training program.[7]

[7]The regulation of unemployment benefits, unemployment assistance and subsistence allowance changed in 2005. Subsistence payments during training measures have been abolished. If entitled, participants in training programs now continue drawing unemployment benefits. In addition, unemployment assistance and social assistance are combined into a second kind of unemployment benefit (*Arbeitslosengeld II*). As we only evaluate programs starting in the period February 2000 to January 2003, our results are only marginally affected by these changes.

2.2 Participation

Traditionally, training measures are the most important part of active labor market policy in Germany. Since the introduction of the Social Code III *(Sozialgesetzbuch III)* in 1998, several reforms were introduced, leading to a focus on measures considered particularly effective in activating the unemployed in the short run and in preventing long-term unemployment. In recent years, allocation of resources has shifted from the comprehensive and expensive medium- and long-term training schemes to less expensive short-term measures.

In fact, tables 1 and 2 in the appendix show a clear decline in average stocks as well as in entries into longer-term training programs, as opposed to an increasing trend for short-term programs. As evident from table 3, the average monthly training costs per participant are much lower for short-term training courses than for the longer-term measures. In addition to these higher direct costs, participants in longer-term training schemes usually receive subsistence allowance. However, the subsistence payments simply replace the ordinary unemployment compensation the participants would have otherwise received. Most striking is the considerable difference in average duration of the courses (see column (2) of table 3). While short-term training courses last on average one month, the duration of longer-term programs, where the average is taken over both further and retraining schemes, lies between eight and ten months.

3 Data

3.1 Structure of the Integrated Employment Biographies Sample (IEBS)

In this paper, we use a new and particularly rich administrative data set, the Integrated Employment Biographies Sample. The IEBS is a 2.2% random sample of individual data drawn from the universe of data records collected in four different administrative processes.[8] The individuals in the IEBS are thus representative for the population made up by those who have data records in any of the four administrative processes. The IEBS contains detailed daily information on employment subject to social security contributions, receipt of transfer payments during unemployment,

[8]For more information on the IEBS, see Osikominu (2005, section 3) and Hummel et al. (2005). We use a version of the IEBS that has been supplemented with additional information not publicly available.

job search, and participation in different programs of active labor market policy. In addition, the IEBS comprises a large variety of covariates including socio-economic characteristics (information on family, health and educational qualifications), occupational and job characteristics, extensive firm and sectoral information, as well as details on individual job search histories such as assessments of case workers. The advantage of this rich set of covariates is that it can be used to reconstruct the circumstances that did or did not lead to the participation in a particular program thus making it possible to control for the factors that drive the selection of individuals into participants and non-participants of a given program.

The IEBS collects information from four different administrative sources: the Employment History (*Beschäftigten-Historik*), the Benefit Recipient History (*Leistungsempfänger-Historik*), the Supply of Applicants (*Bewerberangebot*), and the Data Base of Program Participants (*Massnahme-Teilnehmer-Gesamtdatenbank*).

The first data source, the Employment History, consists of social insurance register data for employees subject to contributions to the public social security system. It covers the time period 1990 to 2003. The main feature of this data is detailed daily information on the employment status of each recorded individual. We use this information to account for the labor market history of individuals as well as to measure employment outcomes. For each employment spell, in addition to start and end dates, data from the Employment History contains information on personal as well as job and firm characteristics such as wage, industry or occupation.

The second data source, the Benefit Recipient History, includes daily spells of all unemployment benefit, unemployment assistance and subsistence allowance payments individuals in our sample received between January 1990 and June 2004. It also contains information on personal characteristics. The Benefit Recipient History is important as it provides information on the periods in which individuals were out of employment and therefore not covered by the Employment History. In particular, the Benefit Recipient History includes information about the exact start and end dates of periods of transfer receipt. We expect this information to be very reliable since it is, at the administrative level, directly linked to flows of benefit payments. Information on benefit payments allow us to construct individual benefit histories reaching back several years. Moreover, we use additional information contained in the Benefit Recipients History describing penalties and periods of disqualification from benefit receipt that may reveal that unemployed individuals showed signs of lacking motivation.

The third data source included in the IEBS is the so-called Supply of Applicants, which contains diverse data on individuals searching for jobs. The Supply of Applicants data covers the period January 1997 to June 2004. In our study it is used in two ways. First, it provides additional information about the labor market status of a person, in particular if the person in question is searching for a job but is not (yet) registered as unemployed or whether he or she is sick while registered unemployed. Second, the spells of job search episodes contained in the Supply of Applicants file include detailed information about personal characteristics, in particular about educational qualifications, nationality, marital status. They also provide information about the labor market prospects of the applicants as assessed by the case worker, about whether the applicant wishes to change occupations, and about health problems that might influence employment chances. Finally, the data on applicants include regional information, which we supplement with unemployment rates at the district level.

The fourth and final data source of the IEBS is the Data Base of Program Participants. This data base contains diverse information on participation in public sector sponsored labor market programs covering the period January 2000 to July 2004. Similar to the other sources, information comes in the form of spells indicating the start and end dates at the daily level, the type of the program as well as additional information on the program such as the planned end date, whether the participant entered the program with a delay, and whether the program was successfully completed. The Data Base of Program Participants not only contains information on the set of training measures evaluated in this paper, but also on other programs such as employment subsidies. This is important, as it enables us to distinguish between different types of employment when measuring evaluation outcomes.

3.2 Reliability of the Data

Being among the first to use the IEBS, we checked the reliability of the data very carefully. We ran extensive consistency checks of the records coming from the different sources, making use of additional information on the data generating process provided to us by the Institute for Employment Research.[9] In addition, we consulted experts in local labor agencies. Our conclusion is that the employment and benefit data are highly reliable concerning employment status, wage and transfer payments, and the start and end dates of spells. The reason for this seems to be that contribution rates and benefit entitlements are directly based on this information. On

[9] This work is documented in Bender et al. (2004, 2005).

the other hand, information not needed for these administrative purposes seems less reliable. For example, in the employment data base the educational variable appears to be affected by measurement error as it is not directly relevant for social security entitlements.[10] Personal characteristics exhibit a higher degree of reliability in the program participation and job seeker data, because they are relevant for the purpose of assigning job offers or programs to the unemployed. For our evaluation, we exploit the available information as efficiently as possible by choosing the data source that is most reliable for a given purpose.

Although the data generally seem very reliable, we saw some need for minor corrections and imputations. As mentioned before, start and end dates are highly reliable in the employment and benefit payments data. Unfortunately, this does not seem to be the case for the end dates in the data on job seekers and program participants. With regard to the job search data, we circumvent this problem by using the transfer payments and employment spells in order to correctly assess the labor market status of an individual. However, the limited reliability of the end dates in the program participation data – mostly due to non-attendance or early drop-outs not always correctly registered in the data – is a problem for the evaluation. We therefore devised a correction procedure for the end dates of program spells. Whenever possible, we aligned the program spells with the corresponding subsistence allowance spells. The latter are highly reliable because they are directly linked to benefit payments. In addition, we corrected the end dates of program spells if there was an implausible overlap with subsequent regular employment or if variables regarding the status or the success of participation indicated that a given participant dropped out of a particular program.

3.3 Evaluation Sample and Training Programs

In the following, we focus on an inflow sample into unemployment consisting of individuals who became unemployed between the beginning of February 2000 and the end of January 2002 after having been continuously employed for at least three months. Entering unemployment is defined as quitting regular (not marginal), non-subsidized employment and subsequently being in contact with the labor agency (not necessarily immediately), either through benefit receipt, program participation or a

[10] Fitzenberger, Osikominu and Völter (2006) analyze the quality of the education variable in German employment register data and provide imputation methods for improving it.

job search spell.[11] In order to exclude individuals eligible for specific labor market programs for young people and individuals eligible for early retirement schemes, we only consider persons aged between 25 and 53 years at the beginning of their unemployment spell. Concentrating on three different types of training programs, we focus on the first program that is taken up during an unemployment spell.

We classify training programs into three different types:[12] short-term training (STT), classroom further training (CFT) and practical further training (PFT).[13] These three programs differ in length as well as in contents. Short-term training lasts on average several weeks and aims at providing general skills that facilitate job search. At the same time these courses can be employed to assess and monitor the abilities and the willingness to work of the unemployed. The two other types of programs considered in this paper are longer (typically several months). Classroom further training is aimed at refreshing existing as well as training new professional skills. In contrast to classroom further training, practical further training includes "hands-on" experience in training workshops or firms. Their aim is to improve work habits and to provide practical working experience. Descriptive statistics for the evaluation sample and the programs can be found in the appendix.

4 The Multiple Treatment Framework

Our empirical analysis is based upon the potential-outcome-approach to causality, see Roy (1951), Rubin (1974), and the survey of Heckman, LaLonde, and Smith (1999). Lechner (2001) and Imbens (2000) extend this framework to the case of multiple, exclusive treatments, while Lechner (2001) and Gerfin and Lechner (2002)

[11] Note that this implies that the same individual may appear more than once in our evaluation sample. Approximately ten percent of the individuals in our sample are represented by more than one unemployment spell according to the above definition.

[12] A fourth category of training program is retraining, which lasts two to three years on average. It typically leads to a new occupational training degree within the German apprenticeship system. Currently, it is not possible to estimate the effects of these long programs with IEBS data, because the program participation data are only available from January 2000 onwards, whereas the employment data end in December 2003. We plan to include retraining programs in future research.

[13] Our classification does not correspond one-to-one to the categories distinguished by legislation. Instead, we are led by economic criteria and classify the qualification programs according to their similarity. Two programs that are similar in duration and contents may belong to the same category in our classification although they represent different program types under the law. In particular, depending on the length and contents of the program under consideration, we also group measures of "discretionary support" *(Freie Förderung)* and measures financed through the European Social Fund *(Europäischer Sozialfond)* into one of the three program types defined in the text.

show how to extend standard propensity score matching estimators for this purpose. For the following, let $\{Y^0, Y^1, ..., Y^K\}$ be $K+1$ potential outcomes, where $Y^k, k = 1, ..., K$, denotes the outcome associated with treatment k and Y^0 the outcome when receiving none of the K treatments. To simplify the discussion, we will from now on refer to the nontreatment outcome Y^0 as one of the $K+1$ treatment outcomes. For each individual, only one of the $K+1$ potential outcomes is observed and the remaining K outcomes are counterfactual.

Given these counterfactual outcomes, one can define pairwise average treatment effects on the treated (ATT)

$$(1) \qquad \theta(k,l) = E(Y^k - Y^l | T = k) \text{ with } k, l = 0, 1, ..., K \text{ and } k \neq l,$$

where $T = 0, 1, ..., K$ represents the treatment actually received. The individual treatment effect is the difference between the outcome Y^k and the outcome Y^l, where the latter is not observed for individuals who received treatment $T = k$. In the following, we call the individuals who undergo treatment $T = k$ the k–group and individuals who undergo treatment $T = l$ the l–group. Note that in general $\theta(k,l) \neq \theta(l,k)$ because the characteristics of participants in treatment k differ from those of participants in treatment l.

4.1 Extension to Dynamic Setting

We use the static multiple treatment framework in a dynamic context. Our basic samples consist of individuals who start an unemployment spell as defined above between February 2000 and January 2002. These individuals can participate in any of the three training programs at different points of time in their unemployment spell. Both the type of treatment and the selectivity of the treated may depend upon the exact starting date of the program. Abbring and van den Berg (2003) and Fredriksson and Johansson (2003, 2004) interpret the start of the program as an independent random variable in the "timing of events". In a similar vein, Sianesi (2003, 2004) argues for Sweden that all unemployed individuals are potential future participants in active labor market programs, a view which is particularly plausible for countries with comprehensive systems of active labor market policies like Sweden or Germany. Unemployed individuals are not observed to participate in a program either because their participation takes place after the end of the observation period or because they leave the state of unemployment either by finding a job or by moving out of the labor force.

Fredriksson and Johansson (2003, 2004) argue that it is incorrect to undertake a static evaluation analysis by assigning unemployed individuals to a treatment and a nontreatment group based on the treatment information observed in the data up to certain point in time. The reason is that, if one defines a fixed classification window during which participation is recorded, one effectively conditions on future outcomes. For example, consider the case of analyzing treatment irrespective of its actual starting date in the unemployment spell. On the one hand, individuals who find a job later during the fixed observation period are assigned to the control group. This may lead to a downward bias in the estimated treatment effect. On the other hand, future participants whose participation starts after the end of the observation period are also assigned to the control group, which may cause an upward bias.

The above discussion implies that a purely static evaluation of the different training programs is not appropriate.[14] We therefore extend the static framework presented above in the following way. We analyze the employment effects of a training program conditional on the starting date of the treatment.[15] We distinguish between treatment starting during months 0 to 3 of the unemployment spell (stratum 1), treatment starting during months 4 to 6 (stratum 2), and treatment starting during months 7 to 12 (stratum 3).[16] In each of these time windows, we define individuals to be undergoing treatment if they start one of the three programs during the time period defined by the window. Individuals not starting any program during the time window in question, are in a "waiting" state because they are not treated at this point but may be treated later (in one of the following windows). We then carry out the evaluation for each window separately, circumventing the problems

[14] In a static setting one has to deal, in addition, with the problem that the potential starting dates of the nonparticipants are unobserved. In this situation, drawing random starting times of the program is a way to proceed, see e.g. Lechner (1999) and Lechner et al. (2005a,b). However, this strategy does not overcome the problems discussed above and we prefer to consider the timing of events explicitly. We do not introduce a random timing of the program starts among the nonparticipants for the following three reasons. First, random starting dates add noise to the data. Second, the starting time drawn might be infeasible in the actual situation of the nontreated individual. Third, drawing random starting dates does not take the timing of events seriously.

[15] We only consider the first treatment during the unemployment spell and thus do not analyze multiple sequential treatments as in Bergemann et al. (2004), Lechner and Miquel (2005), or Lechner (2004). In fact, the treatment effect we analyze can be expressed as a special case of Lechner and Miquel (2005) and Lechner (2004). What we call a stratum is comparable to a period in the Lechner-Miquel framework. For instance, the effect of training versus waiting in the second stratum corresponds to the following effect in the Lechner-Miquel framework: nonparticipation in period one and training in period two versus nonparticipation in period one and nonparticipation in period two for the population of those who do not participate in period one and enroll in training in period two.

[16] We do not analyze treatments starting later than month 12 because, at present, our employment data ends in 2003 thus making it impossible to evaluate programs that start too late after January 2002.

discussed above.

4.2 Propensity Score Matching

In order to evaluate the differential effects of multiple treatments we assume that the Conditional Independence Assumption (CIA) holds, i.e. that, conditional on individual characteristics X, the potential outcomes $\{Y^0, Y^1, ..., Y^K\}$ are independent of treatment status T. Building on Rosenbaum and Rubin's (1983) result on the balancing property of the propensity score in the case of a binary treatment, Lechner (2001) shows that the conditional probability of treatment k, given that the individual receives treatment k or l, exhibits an analogous balancing property for the pairwise estimation of the ATT's $\theta(k,l)$ and $\theta(l,k)$. Formally, we have

$$E(Y^l|T=k, P^{k|kl}(X)) = E(Y^l|T=l, P^{k|kl}(X)) \tag{2}$$

and analogously,

$$E(Y^k|T=l, P^{k|kl}(X)) = E(Y^k|T=k, P^{k|kl}(X)) \ .$$

$P^{k|kl}(X)$ is the conditional probability of treatment k, given that the individual receives treatment k or l, i.e.

$$P^{k|kl}(X) = \frac{P(T=k|X)}{P(T=k|X)+P(T=l|X)} \equiv \frac{P^k(X)}{P^k(X)+P^l(X)} \ .$$

The balancing property in equation (2) allows one to apply standard binary propensity score matching based on the sample of individuals participating in either program k or l (compare Lechner (2001), Gerfin and Lechner (2002), Sianesi (2003)). In this subsample of the data, one simply estimates the probability of treatment k versus l, yielding an estimate of the conditional probability $P^{k|kl}(X)$, and then applies standard matching techniques known from the binary case.

In order to account for the dynamic nature of the treatment assignment process, we estimate the probability of treatment k versus l given that unemployment lasts long enough to make an individual 'eligible'. For the treatment during months 0 to 3, we take the total sample of unemployed, who participate in k or l during months 0 to 3, and estimate a Probit model for participation in k. If the comparison involves nonparticipation in any treatment, then this group includes those unemployed who either never participate in any treatment or who start treatment after month 3. For the treatment during months 4 to 6 or months 7 to 12, the sample consists of those still unemployed at the beginning of the time window considered. Using a Probit model, we then estimate the propensity of beginning a program within the time interval of elapsed unemployment duration defined by the respective window, using all individuals still unemployed at the beginning of the window and participating in either k or l during the time interval. This is in contrast to Sianesi (2004) who estimates separate Probit models for each of the different program starting dates. In our case, the number of observations would be too small for such an approach. However, even if we had enough observations, we think that it would not be advisable to estimate Probit regressions by month. The reason is that the starting date of the treatment is somewhat random (relative to the elapsed duration of the unemployment spell) due to available programs starting only at certain calendar dates. We therefore pool the treatment Probit for all eligible persons in unemployment assuming that the exact starting date is random within the time interval considered. However, when matching treated and nontreated individuals, we align them in elapsed unemployment duration by month at the start of the program.

As already mentioned, we aggregate relative starting dates into three strata, while employment status is measured at a monthly frequency. To account for this difference of scale we impose as a matching requirement that the comparison group with treatment l for an individual receiving treatment k is still unemployed in the month before treatment k starts. In the following, we refer to this subset of the l-group as the eligible l-group. In this way, we only match participants in l who may have started a treatment k in the same month as the respective participant in treatment k. Second, within this group of eligible l-matches, we match individuals based on the similarity of calendar month in which unemployment starts, and based on the similarity of the estimated propensity score. As a matching procedure we chose local linear matching. The prediction for the counterfactual outcome in treatment l for an individual undergoing treatment k is thus given by the prediction of a local linear regression of the treatment outcome in l on the estimated propensity score and the starting month of unemployment, evaluated at the estimated propensity score and the starting month of unemployment of the k-individual. The local linear regression is estimated in the subset of eligible l-individuals matched to the individual receiving treatment k. In this way, we obtain a close alignment in calendar time as well as elapsed unemployment duration thus avoiding drawing random starting times for the programs. For weighting, we use a bivariate product kernel. Technical details of our matching procedure are given below.

4.3 Interpretation of Estimated Treatment Effect

Our estimated ATT parameter has to be interpreted in a dynamic context. We analyze treatment conditional upon unemployment lasting at least until the start of treatment k and this being the first treatment during the unemployment spell. The treatment parameter we estimate is therefore given by

$$(3) \qquad \theta(k,l;u,\tau) = E(Y^k(u,\tau)|T_u = k, U \geq u-1, T_1 = ... = T_{u-1} = 0)$$

$$-E(Y^l(\tilde{u},\tau-(\tilde{u}-u))|T_u = k, u \leq \tilde{u} \leq \bar{u}, U \geq u-1, T_1 = ... = T_{u-1} = 0) \ ,$$

where T_u is the treatment variable for treatment starting in month u of unemployment, $Y^k(u,\tau)$, $Y^l(u,\tau)$ are the treatment outcomes for treatments k and l, respectively, in periods $u+\tau$, $\tau = 0,1,2,...$, counts the months since the beginning of treatment, U is the duration of unemployment, and $\bar{u} = 3,6,12$ is the last month of the stratum of elapsed unemployment considered. Then, $Y^l(\tilde{u},\tau-(\tilde{u}-u))$ is the outcome of individuals who start treatment l in month $\tilde{u} \in [u;\bar{u}]$. For starts of l later than u, we have $\tilde{u}-u > 0$, and therefore, before l starts, $\tau-(\tilde{u}-u)) < 0$. This implies that these individuals are still unemployed, i.e. $Y^l(\tilde{u},\tau-(\tilde{u}-u)) = 0$ when the second argument of $Y^l(.,.)$ is negative. Taken together, this accounts for the fact that the alternative treatment l, for which the individual receiving treatment k in period u is eligible, might not start in the same month u. In this way, we make sure that each member of the eligible subset of the l-group is used in the pairwise comparisons for treatment k.[17]

Conditioning on past treatment decisions and outcomes, the treatment parameter for a later treatment period (months 4 to 6 or months 7 to 12) is not invariant to changes in the determinants of the unemployment exit rate and the treatment propensity in the earlier phase of the unemployment spell. This is a direct consequence of modeling heterogeneity with respect to the starting time of treatment relative to the length of elapsed unemployment. Both the k-group and the l-group at the start of treatment are affected by the dynamic sorting effects taking place before, see Abbring and van den Berg (2004) for discussion of this problem in the context of duration models.[18] Taking the timing of events seriously, estimated treatment

[17]Based on monthly data, Sianesi (2003) restricts the comparison to treatment l starting in the same month as treatment k. In our setup, where starting times are aggregated into three strata, that would leave a large number of eligible individuals for comparison with treatment k starting in period u not being used in any pairwise treatment combination, because, if u lies before the end of the time window of elapsed unemployment considered, then some individuals in the eligible subset of the l-group receive treatment after u.

[18]Heckman and Navarro (2006) consider the identification of dynamic treatment effects in a

parameters thus depend dynamically on treatment decisions and outcomes in the past (Abbring and van den Berg (2003), Fredriksson and Johansson (2003), Sianesi (2003, 2004)). To avoid this problem, one often assumes that treatment effects are constant over the duration of elapsed unemployment at program start. Alternatively, other suitable uniformity or homogeneity assumptions for the treatment effect can be made. Such assumptions are not attractive in our context.[19] Because of the dynamic sorting effects taking place before treatment, there is no simple relationship between our estimated treatment parameter in equation (3) and the static ATT in equation (1), the literature typically attempts to estimate.[20]

Using propensity score matching in a stratified manner, we estimate the treatment parameter in (3) allowing for heterogeneity in the individual treatment effects and for an interaction of individual treatment effects with dynamic sorting processes. In order for our approach to be valid we need to make the following dynamic version of the conditional independence assumption (DCIA)

$$(4)\quad E(Y^l(\tilde{u}, \tau - (\tilde{u} - u))|T_u = k, u \leq \tilde{u} \leq \bar{u}, U \geq u - 1, T_1 = ... = T_{u-1} = 0, X)$$

$$= E(Y^l(\tilde{u}, \tau - (\tilde{u} - u))|T_{\tilde{u}} = l, u \leq \tilde{u} \leq \bar{u}, U \geq u - 1, T_1 = ... = T_{u-1} = 0, X) \ ,$$

where X are time constant as well as time-varying (during the unemployment spell) characteristics, $T_{\tilde{u}} = l$ indicates treatment l between u and $\bar{u}$, and $\tau \geq 0$, see equation (3) above and the analogous discussion in Sianesi (2004, p. 137). We effectively assume that conditional on X, and conditional on being unemployed until period $u-1$, individuals are comparable in their outcome for treatment l occurring between u and $\bar{u}$.

For $l = 0$, i.e. the comparison to the nontreatment alternative, the treatment parameter in (3) is interesting if one is in the situation to decide each time period whether to start treatment in the next month or to postpone possible treatment to the future (treatment now versus "waiting", see Sianesi (2004)). By contrast, for $l \neq 0$ and $k \neq 0$, treatment parameter (3) is interesting in the situation in which one decides whether to start treatment k in the next month against the alternative to receive treatment l at some point in the near future, i.e. before the end of

discrete time dynamic discrete choice framework.

[19]Sianesi (2003) reports 'synthetic' averages over the relative starting dates u, $\sum_u \frac{N_{k,u}}{N_k}\theta(k, l; u, \tau)$, to provide a summary statistic of the u specific treatments. These estimated averages have by themselves no causal interpretation.

[20]Fitzenberger and Speckesser (2005) provide a more detailed discussion of the relationship between the static and dynamic treatment parameter in the binary treatment case $K = 1$.

the current time window (treatment k versus l in a dynamic context, see Sianesi (2003)). In addition, exits from unemployment must not be known until the period in which they take place, i.e. job arrivals or the start of some treatment must not be anticipated for sure. The former would introduce a downward bias in the estimated treatment effect, while the latter would induce an upward bias. This is a problem in any analysis based on the timing-of-events approach. Note however, that anticipation effects are no problem if they only refer to the probability that one of these events occur, and if this happens in the same way for all individuals with characteristics X and elapsed unemployment duration $u - 1$.

As a test of conditional independence assumption, we implement a pre-program test. By construction, treated individuals in the k-group and their matched counterparts in the l-group exhibit the same unemployment duration until the beginning of the treatment k. However, one can test whether they differ in time-invariant unobserved characteristics by analyzing employment differences during 13 months before the start of the unemployment spell. Significant differences between matched employment outcomes before treatment would indicate a violation of the conditional independence assumption.

4.4 Details of the Matching Approach

Estimating the ATT for treatment k versus l requires constructing the counterfactual outcome of individuals in treatment k, had they instead received treatment l during the period defined by the respective time window. As indicated above, this counterfactual outcome can be constructed using a matching approach (Rosenbaum and Rubin (1983), Heckman, Ichimura, Todd (1998), Heckman, LaLonde, Smith, (1999), Lechner (1999)) based on the estimated dynamic propensity score. We apply local linear matching to estimate the average counterfactual outcome.

4.4.1 Local Linear Regression

Effectively, we run a nonparametric local linear kernel regression (Heckman, Ichimura, Smith, Todd (1998), Pagan, Ullah (1999), Bergemann et al. (2004)). The idea of this regression is to predict the counterfactual outcome of an individual i undergoing treatment k as the weighted average of outcomes in the subset of eligible l-individuals. In this weighted average, the weight $w_{N_l}(i, j)$ of an individual j receiving treatment l is the higher the closer this individual is in terms of estimated propensity score and starting month of unemployment to individual i receiving treatment k whose counterfactual outcome is to be determined. The ATT is then estimated as the average difference between the actual outcomes of individuals i receiving treatment k and their predicted counterfactual outcomes under treatment l

$$\frac{1}{N_k} \sum_{i \in \{T_u = k\}} \left\{ Y^k_{i,u,\tau} - \sum_{j \in \{T_{\tilde{u}} = l, u \leq \tilde{u} \leq \bar{u}\}} w_{N_l}(i,j)\, Y^l_{j,\tilde{u},\tilde{\tau}} \right\}, \tag{5}$$

where N_k is the number of participants i in treatment k (this group of individuals is denoted as $\{T_u = k\}$), N_l the number of eligible participants in starting treatment l in month $\tilde{u}$ (this group is denoted as $\{T_{\tilde{u}} = l, u \leq \tilde{u} \leq \bar{u}\}$). The variables $Y^k_{i,u,\tau}$ and $Y^l_{j,\tilde{u},\tilde{\tau}} = Y^l_j(\tilde{u}, \tau - (\tilde{u} - u))$ are the outcomes in post treatment period $u + \tau$, where $\tilde{\tau} = \tau - (\tilde{u} - u)$.

Kernel matching has a number of advantages compared to nearest neighbor matching, which is widely used in the literature (Lechner (1999), Lechner et al. (2005a,b), Sianesi (2003, 2004)). The asymptotic properties of kernel based methods are relatively easy to analyze and it has been shown that bootstrapping provides a consistent estimator of the sampling variability of the estimator in (5) even if matching is based on closeness in generated variables such as the estimated propensity score (see Heckman, Ichimura, Smith, and Todd (1998) or Ichimura and Linton (2001) for an asymptotic analysis of kernel based treatment estimators.) Abadie and Imbens (2006) have shown that matching methods based on a fixed number of matches are not root-N consistent and that the bootstrap is in general not valid due to their extreme nonsmoothness.

4.4.2 Kernel Function and Bandwidth Choice

As a kernel function in our local linear regression, we use a product kernel (see Racine and Li (2004)) in the estimated propensity score and the calendar month of entry into unemployment

$$KK(p,c) = K\left(\frac{p - p_j}{h_p}\right) \cdot h_c^{|c - c_j|}, \tag{6}$$

where $K(z) = exp(-z^2/2)/\sqrt{2\pi}$ is the Gaussian kernel function, p and c are the propensity score and the calendar month of entry into unemployment of a particular individual $i \in \{T_u = k\}$ whose counterfactual outcome is to be predicted, p_j and c_j are the estimated propensity score and the calendar month of entry into unemployment of an individual j belonging to the comparison group of individuals treated with l, and h_p and h_c are the bandwidths which are determined by the cross-validation procedure described in the next paragraph.[21]

For the local linear kernel regression using the product kernel in equation (6), standard bandwidth choices for pointwise estimation are not applicable because we are ultimately interested in predicting as good as possible the average expected outcome in treatment l for individuals treated with k. In order to choose the bandwidths h_p and h_c, we employ the leave-one-out cross-validation procedure suggested in Bergemann et al. (2004) and Fitzenberger, Osikominu, and Völter (2005). This procedure mimics the estimation of the average expected outcome in the alternative treatment l for each period. First, for each participant i in the k-group, we identify the nearest neighbor $nn(i)$ in the eligible subset of the l-group, i.e. the individual in that group whose propensity score is closest to that of i. Second, we choose the bandwidths to minimize the sum of the period-wise squared prediction errors

$$(7) \qquad \sum_{\tau=0}^{\tau_{\max}} \left[\frac{1}{N_k} \sum_{i=1}^{N_k} \left(Y^l_{nn(i),u,\tau} - \sum_{j \in \{T_{\tilde{u}(i)}=l, u \leq \tilde{u} \leq \bar{u}\} \setminus nn(i)} w_{(N_l(i)-1)}(i,j) Y^l_{j,\tilde{u},\tilde{\tau}} \right) \right]^2$$

where $\tau_{\max} = 33 - u$, $u = 0, 4, 7$ is the first month of the time window 0–3, 4–6, and 7–12 months during which treatment starts, $u(i)$ is the month in which treatment for i starts, τ counts the number of months since month u, and $N_l(i)$ represents the size of the eligible l-group for i, $\{T_{\tilde{u}(i)} = l\}$. In the estimation of the employment status for $nn(i)$, observation $nn(i)$ itself is not used. However, individual $nn(i)$ is used for the local linear regression for other treated individuals in the k-group, provided it is in the eligible l-group, and provided it does not happen to be also the nearest neighbor in this case. Therefore, the local linear regression in (7) always depends on $N_l(i) - 1$ observations. The optimal bandwidths h_p and h_c determining the weights $w_{(N_l(i)-1)}(i,j)$ through the local linear regression are identified in a two-dimensional search procedure.[22] The resulting bandwidths are sometimes larger and sometimes smaller than a rule-of-thumb value for pointwise estimation, see Ichimura, Linton (2001) for similar evidence in small samples based on simulated data.

[21] Note that $h_c \in [0,1]$, where $h_c = 0$ amounts to only considering matches whose unemployment spell starts in the same calendar month.

[22] When the control group consists of the large group of nonparticipants in the respective stratum it often turns out that there is no need to smooth in addition over the calendar month of entry into unemployment, i.e. $h_c = 0$.

4.4.3 Bootstrapping

We take account of the sampling variability in estimated propensity scores by computing bootstrap standard errors of the estimated treatment effects. Our bootstrap procedure is partly parametric as we resample the coefficients of the probit estimates for the propensity scores based on their estimated asymptotic distribution. To account for clustering on the individual level and autocorrelation over time, we use the entire time path for each individual as a block resampling unit. All the bootstrap results reported in this paper are based on 200 resamples. As the cross-validation in (7) is computationally expensive, the sample bandwidths are used in all resamples.

4.4.4 Balancing Test

In order to test whether covariates are balanced sufficiently by matching on the estimated propensity score $\hat{P}(X)$, we carry out the balancing test suggested by Smith and Todd (2005). The test involves regressing each covariate X_g on a flexible polynomial in $\hat{P}(X)$ of order δ and interactions with the treatment dummy variable

$$X_g = \sum_{d=0}^{\delta} \beta_d \, \hat{P}(X)^d + \sum_{d=0}^{\delta} \gamma_d \, D_k \, \hat{P}(X)^d + \eta_{kl} \, , \tag{8}$$

where X_g is one component of the covariate vector X, and $D_k = I(T = k)$ is a dummy variable for treatment k. The regression in (8) is estimated separately based on the sample of those individuals receiving either treatment k or l in the respective interval for unemployment duration (0–3, 4–6, 7–12 months). If the estimated propensity score balances the covariate X_g in the treatment and the control sample, then the coefficients on all terms involving the treatment dummy γ_d should be zero. We test this joint hypothesis both for cubic ($\delta = 3$) and quartic ($\delta = 4$) polynomials in order to see whether our test results are sensitive to the choice of δ, a problem mentioned by Smith and Todd (2005, p. 373). If the test does not reject, then the treatment dummy D_k does not provide any significant information about the covariate X_g conditional on the estimated propensity score. For each specification of the propensity score, we report the number of covariates for which the balancing test passes, i.e. the zero hypothesis is not rejected.

5 Empirical Results

5.1 Estimation of Propensity Scores

For matching to be a valid exercise we need variables that jointly influence participation and outcomes such that the DCIA condition (equation 4) holds. In a dynamic context, the participation decision involves the following two components. First, one needs to consider determinants that are relevant for the timing of the participation during the unemployment spell. Second, one needs to model the factors that drive the selection into the different programs. As we show below, our data base allows us to account for all of these components. In particular, we are able to construct both time constant and time-varying variables that reflect the motivation, plans and labor market prospects of the unemployed and the way they are perceived by the caseworker in the labor office. Thus, we have at our disposal the relevant information to model the decision process governing the selection into treatment over time. As a consequence, we are confident that – conditional on the propensity score and the calender month of the beginning of the unemployment spell – there is no unobserved heterogeneity left causing a correlation between treatment indicators and outcomes.

Based on the information contained in all four data sources of the IEBS, we construct a large set of time constant as well as time-varying (within the unemployment spell) variables to model the selection into treatments. Time-varying covariates are updated at the beginning of each stratum. For this purpose, we use information of a spell with a start date as near as possible to the first day of the stratum in question. For time-varying variables, information from spells starting more than a few days later than the beginning of the respective time window is not used in order to avoid endogeneity problems. To model the propensity of participating in a particular program as opposed to not participating, we use the following variables and their interactions.[23]

In order to account for differences in individual labor market histories, we consider the following variables: occupation and industry of the last job before unemployment, whether this last job was less than full-time, whether it was a white-collar or blue-collar position, the reason why this last job was ended, the quarter of the beginning of the unemployment period, whether there were any periods of incapacity in the last three years, the number of days in employment during the last

[23] See the appendix for summary statistics and a more detailed description of the variables used.

three years, the number of days when transfer payments were received during the last three years (i.e. unemployment benefit, unemployment assistance, subsistence allowance), the number of days without any information in the data set, the number of days in contact with the labor agency during the last three years before unemployment, whether the person was employed 6, 12, 24 months before the beginning of the unemployment period, log daily wage in the last job before unemployment, an indicator whether this wage was censored, the log average wage in the year before unemployment and censoring dummies related to this variable.

As to personal characteristics driving the selection into the different programs, we considered: age, disability status, schooling and professional qualification, family status, whether there are children, whether there are children under 10 years, nationality other than German, as well as whether someone is an ethnic German who has migrated back into Germany (usually from Eastern European countries).

As to the assessment of the case workers with regards to the motivation, plans and labor market prospects of the unemployed, the following information is considered: current health status, past health problems, information on whether a program has been canceled within the last three years, penalties and disqualification from benefits within the last three years, participation in a program with a social work component, indication of lack of motivation within the last three years, the number of proposals the unemployed has received, as well as the characteristics of the desired job.

In addition, we consider regional information in form of the following variables: different kinds of unemployment rates in the home district of an individual, region type according to classification of the labor market characteristics of the district, the federal state, the region of Germany.

For each propensity score, we ran an extensive specification search. In each case, the final specification was chosen based on economic considerations and significance tests. In order to test the balancing property of the covariates, we carried out different variants of the balancing test by Smith and Todd (2005). The final specification usually includes 20 to 35 covariates covering all important aspects that drive the selection into treatments.[24]

[24] The estimation results for the different propensity scores are available from the authors upon request.

5.2 Treatment Effects

The results from our econometric evaluation are shown in figures 3 to 7. Each graph shows the average treatment effect on the treated (ATT), i.e. the difference between the actual and the counterfactual employment outcome averaged over those individuals who participate in the program under consideration. Here, we compare the actual employment outcome of the treated to the employment outcome these individuals would have had, *had they not taken part in any other program in the respective time window of their unemployment spell*. In general, we distinguish between programs starting in three different time windows (strata) of elapsed unemployment: 0 to 3 months (stratum 1), 4 to 6 months (stratum 2), and 7 to 12 months (stratum 3). Due to the smaller number of treated individuals, we only consider one time window ranging from month 0 to 12 for participants in practical further training (PFT).

We evaluate treatment effects at different points in time. On the time axis in our graphs, positive values denote months since the program start, while negative values represent pre-unemployment months. We omit the period between the start of unemployment and the start of the program where both control and treatment group are unemployed. The dashed lines around the estimated ATT are bootstrapped 95 percent confidence bands. Treatment effects for a particular month are statistically significant if zero is not contained in the confidence band.

Figure 3 shows estimated treatment effects for short-term training programs (STT) in West Germany. The results for men are given in the left column, while those for women are shown in the right column. The figures suggest short and not very pronounced lock-in effects of short-term training measures of -2 to -5 percentage points (i.e. during the program, participants had a 2 to 5 percentage points lower monthly employment probability than they would have had if they had not participated in the program). These lock-in effects do not last more than 2 or 3 months, which is not surprising given the average length of such programs. After the short lock-in period, the difference between actual and counterfactual employment outcomes of participants becomes more and more positive, suggesting positive treatment effects. However, results seem to depend strongly on elapsed unemployment duration. While there is no evidence for statistically significant treatment effects for individuals participating in the first three months of their unemployment spell (stratum 1), treatment effects for individuals starting a short-term training program in months 4 to 6 (stratum 2) or months 7 to 12 (stratum 3) of their unemployment spell are positive and statistically significant. According to these estimates, the monthly employment probability of West German men participating in short-term training is increased by 5 percentage points. At some 10 percentage points, this effect is even larger for female participants. Interestingly, these employment effects are not short-lived but persist over time. For women, they even seem to grow within our observational window (see last graph of figure 3).

Figure 4 presents the corresponding results for East Germany. They suggest that short-term training measures in East Germany generally do not have any positive effects on the employment probability of the participants. Measured average treatment effects are mostly small and statistically insignificant. The only exception are men who receive treatment in months 7 to 12 of their unemployment spell. For these individuals, participating in short-term training increases their long-term employment probability by about 5 percentage points. In the latter case, it is remarkable that the effects take some time to kick in. They are not statistically significant until about 12 months after the end of the program, suggesting a dynamic mechanism that leads to positive employment effects in future periods.

Results for the more substantive classroom further training measures (CFT) are given in figures 5 and 6. The first column of figure 5 shows average treatment effects for West German men participating in CFT, results for West German women are given in the right column. The most conspicuous difference between these results and those for short-term training programs is the long and pronounced lock-in effect. During the first months of their participation in the program, participants have an employment rate that is up to 20 percentage points lower than it would have been if they had not taken part in the program. The lock-in period lasts up to 12 months for individuals who take up their treatment during the first 6 months of their unemployment spell. Interestingly, lock-in effects are less deep and shorter for individuals that have been unemployed for more than 6 months (stratum 3).

There are several possible reasons for this finding. First, it might be that individuals with a longer elapsed unemployment duration are assigned shorter measures within the group of CFT programs. Second, it is possible that such individuals drop out of the program more often or earlier. Third, a reason for pronounced lock-in effects in the first stratum may be that a considerable number of those just having become unemployed find a new job quickly. If these individuals are assigned to CFT measures anyway, they will be "locked-in" in the program, while many of their counterparts in the control group may already have found employment again. This would mean that some of the short-term unemployed receive training even though they do not need it to overcome unemployment. In addition, there would be a tendency towards finding less pronounced lock-in effects for the late program starts if many of the long-term unemployed in the control group abandon their job search and move out of labor force. Hence, an additional channel through which training programs work may consist in keeping the long-term unemployed in the labor force.

While there is little evidence for statistically significant treatment effects for West German men receiving their treatment in months 0 to 6 of their unemployment spell (strata 1 and 2) or West German women starting CFT in the first 3 months of unemployment (stratum 1), treatment effects for longer-term unemployed men (stratum 3) and medium to longer-term unemployed women (strata 2 and 3) are large and statistically significant. After the initial lock-in phase, they amount to about 7 percentage points for men and to some 10 percentage points for women. For men these employment effects are persistent within the time window permitted by our data. For women they are even rising.

The corresponding results for classroom further training measures in East Germany are shown in figure 6. As in West Germany, there are long and deep lock-in effects of up to 20 percentage points in the first 12 months after treatment start. With the exception of men starting their program relatively early in their unemployment spell (stratum 1), there is no evidence for positive treatment effects after this initial lock-in phase. In some cases employment effects remain negative throughout our observation window.

In contrast to pure classroom further training, practical further training (PFT) also includes practical elements such as internships or working in a practice firm. Evaluation results for these measures are given in figure 7. The results for West Germany shown in the first row of figure 7 suggest considerable positive employment effects of about 10 percentage points for women after a lock-in period of up to 10 months. There are no such effects for men. A reason for this finding could be that particularly in practice-related jobs, men and women select into different occupations. If women disproportionately often take part in training measures for occupations in the service sector, and job chances are better in this sector than in the industrial or the construction sector then this may be an explanation for gender differences in the employment effects of PFT.[25]

Similar as for other types of training, there are no employment effects for participants in PFT in East Germany (second row of figure 7). The negative picture

[25] Lechner, Miquel and Wunsch (2005b) consider gender specific target professions for public sector sponsored training in East Germany in the nineties and study the associated differences in treatment effects.

the employment effects reveal for East Germany probably reflects the difficult labor market situation in large parts of East Germany. In districts where open jobs are extremely rare in all sectors, the potential employment effects of training programs are very limited. In addition to this, it is possible that the group of participants in East Germany differs to some extent from that in West Germany. In regions with very high unemployment rates, training programs are to a certain extent used to reduce the frustration of those who want to work, but have very few job chances. In fact, differences in selection into treatments may induce differences in treatment effects.

In a few cases participants seem to have higher employment probabilities than the control group of non-participants even before the program (e.g. first graph of figure 4). In these cases our matching procedure seems unable to fully account for differences in employment probabilities between participants and the control group. This might lead to findings of spurious treatment effects if participants, even after propensity-score matching, represent a positive selection of unemployment risks when compared to non-participants. However, a closer look reveals that in our results, this only concerns a very small number of cases in which there are no significant treatment effects anyway. Our interpretation of positive treatment effects therefore remains unaffected.

Summing up, we find that the effectiveness of the different programs in terms of monthly employment rates depends on many factors. For West Germany, we find statistically significant and sizeable positive employment effects for male and female participants in short-term training programs if these programs are not begun too early in the unemployment spell. Moreover, there is evidence for positive employment effects for men and women starting classroom further training programs after having been unemployed for more than 6 months, while there is little evidence that starting such a program earlier has similar effects. It is a general finding for West Germany that the effectiveness of training tends to be the larger the later it starts in the unemployment spell. However, as the composition of participants changes over time these findings do not imply that programs should start later in the unemployment spell. We also find significant employment effects for women taking part in practical further training measures but no such effects for men. In general, we find that training programs in West Germany are considerably more effective for women than for men. The results for East Germany reveal a much bleaker picture, suggesting no positive treatment effects in the majority of cases. The only exception are moderate positive effects of short-term training for men who start training after having been unemployed for more than 6 months and positive effects for men who take part in classroom further training directly after they become unemployed. We find no positive effects in other cases, in particular there seems little to be gained from participating in a training program for East German women.

6 Conclusion

The aim of this paper was to analyze and compare the employment effects of three types of public sector sponsored training in Germany in the early 2000s. The three types of training programs considered here were short-term training (STT), class-room further training (CFT), and practical further training (PFT). Building on the work of Sianesi (2003, 2004), we applied propensity score matching methods in a dynamic, multiple treatment framework. We were particularly interested in the question of whether short-term training programs can be compared in terms of effectiveness to traditional medium-term further training schemes.

Our results suggest that the effectiveness of the different programs strongly depends on the personal characteristics of the participants and the circumstances of program participation. For West Germany, we find statistically significant positive treatment effects for both men and women taking part in short-term training as well as for male and female participants in longer-term further training schemes. A closer look reveals that, within the time window permitted by our data set, employment effects of short-term training programs are of a similar magnitude as those of traditional medium-term measures, but, due to their shorter length, take effect much earlier. According to our results, West German men taking part in short-term or medium-term training may increase their medium-term employment rate by some 5 to 10 percentage points. The effect for women is even larger, leading to increases in employment probabilities of 10 percentage points or more. We also find that in West Germany, practical further training is more effective than class-room further training for men but that it is completely ineffective for women. A limitation of our results is that, at present, we cannot answer the question whether employment effects of medium-term programs do not fully unfold within our observation window.

Another interesting finding is that for both short-term and medium-term programs in West Germany, employment effects are larger for individuals who start their program at a later point during their unemployment spell. In fact, in many cases we do not find any significant effects for individuals who start their treatment very early in their unemployment spell. It would be wrong to conclude from this that treatment is the more effective the later it is provided to the participants as individuals who are long-term unemployed will differ in observed and unobserved characteristics from those who are short-term unemployed. However, the result is remarkable because it suggests that, especially in the case of long-term unemployment, training measures can help to provide the human capital to improve the employment chances.

In contrast to the encouraging results for West Germany, we find only little evidence for positive treatment effects in East Germany. Apart from positive effects for East German men taking part in short-term training after having been unemployed for more than six months, and positive effects for men beginning classroom further training in the first three months of their unemployment spell, we see little benefits from short- or medium-term training measures in East Germany. In particular, we do not find any positive effects for women. Our results for East Germany reflect the generally difficult labor market situation in the East, especially for women. High unemployment rates seem to render both short and medium-term training programs ineffective to a large extent.

References

Abadie, and G. Imbens (2006). "Large Sample Properties of Matching Estimators for Average Treatment Effects." *Econometrica* 74, 235-267.

Abbring, J. and G.J. van den Berg (2003). "The Nonparametric Identification of Treatment Effects in Duration Models." *Econometrica* 71, 1491-1517.

Abbring, J. and G.J. van den Berg (2004). "Social Experiments and Instrumental Variables with Duration Outcomes." Unpublished Manuscript, Free University Amsterdam and Tinbergen Institute.

Bender, S., M. Biewen, B. Fitzenberger, M. Lechner, S. Lischke, R. Miquel, A. Osikominu, T. Wenzel, C. Wunsch (2004). "Die Beschäftigungswirkung der FbW-Maßnahmen 2000 - 2002 auf individueller Ebene: Eine Evaluation auf Basis der prozessproduzierten Daten des IAB – Vorläufiger Zwischenbericht Oktober 2004", Goethe University Frankfurt and SIAW St. Gallen.

Bender, S., M. Biewen, B. Fitzenberger, M. Lechner, R. Miquel, A. Osikominu, M. Waller, C. Wunsch (2005). "Die Beschäftigungswirkung der FbW-Maßnahmen

2000 - 2002 auf individueller Ebene: Eine Evaluation auf Basis der prozessproduzierten Daten des IAB – Zwischenbericht Oktober 2005", Goethe University Frankfurt and SIAW St. Gallen.

Biewen, M. and B. Fitzenberger (2004). "Neuausrichtung der Förderung beruflicher Weiterbildung." In: Hagen, T. and A. Spermann (Ed.): *Hartz-Gesetze - Methodische Ansätze zu einer Evaluierung*, ZEW Wirtschaftsanalysen, Bd. 74, Baden-Baden, 190-202.

Bundesagentur für Arbeit (2001). *Daten zu den Eingliederungsbilanzen 2000*, Nürnberg.

Bundesagentur für Arbeit (2002a). *Arbeitsmarkt 2001*, Nürnberg.

Bundesagentur für Arbeit (2002b). *Daten zu den Eingliederungsbilanzen 2001*, Nürnberg.

Bundesagentur für Arbeit (2003a). *Arbeitsmarkt 2002*, Nürnberg.

Bundesagentur für Arbeit (2003b). *Daten zu den Eingliederungsbilanzen 2002*, Nürnberg.

Bundesagentur für Arbeit (2004a). *Arbeitsmarkt 2003*, Nürnberg.

Bundesagentur für Arbeit (2004b). *Daten zu den Eingliederungsbilanzen 2003*, Nürnberg.

Bundesagentur für Arbeit (2005a). *Arbeitsmarkt 2004*, Nürnberg.

Bundesagentur für Arbeit (2005b). *Geschäftsbericht 2004*, Nürnberg.

Fitzenberger, B., A. Osikominu and R. Völter (2005). "Get Training or Wait? Long Run Employment Effects of Training Programs for the Unemployed in Germany." Working Paper, Goethe University Frankfurt.

Fitzenberger, B., A. Osikominu, and R. Völter (2006). "Imputation Rules to Improve the Education Variable in the IAB Employment Subsample." Forthcoming *Journal of Applied Social Science Studies*.

Fitzenberger, B. and S. Speckesser (2005) "Employment Effects of the Provision of Specific Professional Skills and Techniques in Germany." ZEW Discussion Paper.

Fredriksson, P. and P. Johansson (2003) "Program Evaluation and Random Program Starts." Institute for Labour Market Policy Evaluation (IFAU), Uppsala, Working Paper, 2003:1.

Fredriksson, P. and P. Johansson (2004) "Dynamic Treatment Assignment – The Consequences for Evaluations Using Observational Data." IZA Discussion Paper No. 1062.

Gerfin, M. and M. Lechner (2002) "Microeconometric Evaluation of the Active Labor Market Policy in Switzerland." *Economic Journal*, 112(482), 854-893.

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

Heckman, J. H. Ichimura, J.A. Smith and P. Todd (1998). "Characterizing Selection Bias using Experimental Data." *Econometrica* 65, 1017-1098.

Heckman, J. R.J. LaLonde, and J.A. Smith (1999). "The Economics and Econometrics of Active Labor Market Programs." In: O. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics*, Vol. 3 A, Amsterdam: Elsevier Science, 1865-2097.

Heckman, J.J. and S. Navarro (2006). "Dynamic Discrete Choice and Dynamic Treatment Effects." Forthcoming *Journal of Econometrics.*

Hummel, E., P. Jacobebbinghaus, A. Kohlmann, M. Oertel, C. Wübbeke, M. Ziegerer (2005). "Stichprobe der Integrierten Erwerbsbiographien, IEBS 1.0", FDZ Datenreport 6/2005, Institut für Arbeitsmarkt- und Berufsmarktforschung (IAB), Nürnberg.

Ichimura, H. and O. Linton (2001). "Asymptotic Expansions for some Semiparametric Program Evaluation Estimators." Discussion paper, London School of Economics and University College London.

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

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

Lechner, M. (2001). "Identification and Estimation of Causal Effects of Multiple Treatments under the Conditional Independence Assumption." In: M. Lechner and F. Pfeifer (eds.) (2000), *Econometric Evaluation of Active Labor Market Politics in Europe*, Heidelberg: Physica-Verlag.

Lechner, M. (2004). “Sequential Matching Estimation of Dynamic Causal Models.” Discussion Paper 2004–06, University of St. Gallen.

Lechner, M. and R. Miquel (2005). “Identification of the Effects of Dynamic Treatments by Sequential Conditional Independence Assumptions” Discussion Paper, SIAW, University of St. Gallen.

Lechner, M., R. Miquel, and C. Wunsch (2005a). “Long-Run Effects of Public Sector Sponsored Training in West Germany.” IZA Discussion Paper No. 1443.

Lechner, M., R. Miquel, and C. Wunsch (2005b). “The Curse and Blessing of Training the Unemployed in a Changing Economy: The Case of East Germany after Unification.” Discussion Paper, University of St. Gallen.

Osikominu, A. (2005). “Eine Analyse der Teilnehmerselektion in die berufliche Weiterbildung auf Basis der Integrierten Erwerbsbiografien (IEB)”. IAB Forschungsbericht Nr. 23/2005, Institut für Arbeitsmarkt- und Berufsmarktforschung (IAB), Nürnberg.

Pagan, A. and A. Ullah (1999). *Nonparametric Econometrics.* Cambridge: Cambridge University Press.

Racine, J. and Q. Li (2004). “Nonparametric estimation of regression functions with both categorical and continuous data.” *Journal of Econometrics* 119, 99-130.

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

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

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

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

Sianesi, B. (2003) "Differential Effects of Swedish Active Labour Market Programmes for Unemployed Adults in the 1990s." Discussion Paper, Institute for Fiscal Studies, London.

Speckesser, S. (2004). "Essays on Evaluation of Active Labour Market Policy." Ph.D. Dissertation, Department of Economics, University of Mannheim.

Smith, J.A. and P. Todd (2005). "Rejoinder." *Journal of Econometrics* 125, 365-375.

Appendix

Participation in Active Labor Market Programs in Germany

Table 1: Average Stocks of Participants in Active Labor Market Programs in Germany from 2000 - 2004

	2000	2001	2002	2003	2004
Qualification schemes	469,463	420,943	456,269	516,782	434,901
– further/retraining	351,960	344,816	331,586	259,922	184,418
– short-term training	47,492	51,266	61,950	92,681	94,748
Employment subsidies on the first labor market	227,813	211,111	222,349	310,533	398,472
Specific measures for young adults	382,433	369,149	443,949	510,602	488,352
Employment on the second labor market	260,766	219,859	179,525	144,734	119,029
Other	57,920	66,471	62,899	34,984	73,005
Total	1,398,395	1,287,533	1,364,991	1,521,800	1,618,879

Source: Bundesagentur für Arbeit, Arbeitsmarkt 2001-2004, own calculations.

Table 2: Entries into Active Labor Market Programs in Germany from 2000 - 2004

	2000	2001	2002	2003	2004
Qualification schemes	1,153,720	1,069,409	1,457,047	1,502,166	1,548,439
– further/retraining	551,534	449,622	456,301	254,718	185,041
– short-term training	476,672	565,132	877,038	106,4293	1,188,369
Employment subsidies on the first labor market	458,557	464,904	538,312	807,682	950,109
Placement and advisory services	601,281	742,065	947,098	1,460,170	2,566,780
Specific measures for young adults	445,823	457,724	447,265	388,810	408,168
Employment on the second labor market	314,291	246,084	219,626	193,999	170,107
Other	391,122	515,670	453,224	21,283	309,446
Total	3,364,794	3,495,856	4,062,572	4,565,010	5,953,049

Source: Bundesagentur für Arbeit, Arbeitsmarkt 2001-2004, own calculations.

Table 3: Average Expenditures per Participant in Short-term, Further and Retraining in Germany from 2000-2003

	2000		2001		2002		2003	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Short-term training	580	1,2	570	1,1	658	0,9	538	1
Further/retraining	1627	8,2	1668	9,3	1686	9,1	1555	10,5
– subsistence allowance	1152		1178		1188		1156	
– training costs	640		664		681		631	

Note: Columns labeled with a (1) contain the average monthly expenditures (in Euro) per participant, columns labeled with a (2) display the average duration of the program in months. Source: Bundesagentur für Arbeit, Daten zu den Eingliederungsbilanzen 2000-2003.

Variable Definitions

Table 4: Variable Definitions

Name	Definition
east	1 if place of residence is in East Germany (Berlin included), 0 otherwise
female	1 if female, 0 otherwise
agegroup	age in 6 groups
foreigner	1 if citizenship is not German, 0 otherwise
ethnicgerman	1 if ethnic German, i.e. returned settler from former German settlements, 0 otherwise
qualification	1 no degree, 2 vocational training degree, 3 university or technical college degree
schooling	1 no schooling degree, 2 Hauptschulabschluss or Mittlere Reife /Fachoberschule (degrees reached after completion of the 9th or 10th grade), 3 Fachhochschulreife or Abitur/Hochschulreife (degrees reached after completion of the 12th or 13th grade)
health	1 no health problems mentioned, 2 health problems, but considered without impact on placement, 3 health problems considered to have an impact on placement
pasthealth	same categories as health, but referring to the past two years before the beginning of the unemployment spell
disabled	1 if disabled, 0 otherwise
land	16 categories for the German Bundesländer
area	German Bundesländer aggregated into 6 categories. 1 SH, NI, HB, HH; 2 NW, 3 HE, RP, SL; 4 BY, BW; 5 MV, BB, BE; 6 SN, ST, TH
region	classification of the districts of residence according to local labor market conditions in 5 groups
family	1 missing, 2 living alone, 3 not married, but living together with at least one person, 4 single parent, 5 married
married	1 missing, 2 married, 3 not married
child	1 if at least one child, 0 otherwise

<continued on next page>

Table 4: Variable Definitions <continued>

Name	Definition
youngchild	1 if at least one child younger than 10 years, 0 otherwise
occupation	occupation of last employment in 7 categories
industry	industry of last employment in 6 categories
occhange	1 missing, 2 if the person wishes to work in the same occupation as in the last employment, 3 otherwise
parttime	1 if the person worked less than full-time in the last employment, 0 otherwise
whitecollar	2 if the previous employment was a white-collar job, 3 if it was a blue-collar job, 1 missing
problemgroup	1 if participation in a program with a social work component within the last three years, 0 otherwise
onlyparttime	1 if information available that only part-time job is desired, 0 otherwise
endlastjob	2 termination of last occupation by employer, 3 by employee, 4 limited in time, 5 other and missing
quarter	quarter of the end of the last employment (from 1 to 9)
penalty	1 if the unemployed had a period of disqualification from benefits within the last three years, 0 otherwise
motivationlack	1 if within the last three years there is information, that the person did not appear regularly at the labor office, on lack of cooperation, availability or similar
pasttreatcancel	1 if abandonment of a program in the past according to the benefit data, 0 otherwise
pastincapacity	1 if incapapacity of work due to illness, parental leave, cure or therapy within the last three years
proposals	number of placement proposals divided by the days since the beginning of the unemployment spell and the start date of the spell from which the information is taken
dapp	1 if employed as apprentice within the last three years before the beginning of the unemployment spell, 0 otherwise

<continued on next page>

Table 4: Variable Definitions <continued>

Name	Definition
countemp, countub, countua, countsp, countoos, countcontact	number of days within the last three years before the beginning of unemployment spent in regular employment, receiving unemployment benefits, unemployment assistance, subsistance payment, out of sample, in contact with the labor agency, respectively
demp6, demp12, demp24, demp6_12, demp12_24	1 if in regular employment 6, 12, 24, 6 and 12 and 12 and 24 months, respectively, before the beginning of the unemployment spell
waged	daily wage in the last job(s) before the beginning of the unemployment spell
ddssec, ddcens, ddmarg	dummies if waged is censored: ddsec is 1 if earnings are within the social security thresholds, ddcens is 1 if earnings are above the social security threshold, ddmarg is 1 if earnings are below the social security threshold
lnwage, lnwagedsq	log(waged) and log(waged) squared interacted with ddssec
wage	total wage in the last year before the beginning of the unemployment spell
dssec, dcens, dmarg	censoring dummies referring to wage (see above)
lnwage, lnwagesq	log(wage) and log(wage) squared interacted with dssec
ur_yb, ur_qb, ur_qb3, ur_qb6, ur_qb12, ur_qb24	unemployment rate in the individual's home district in the calendar year before the beginning of unemployment, in the last month of the quarter before the beginning of unemployment, and in the last month of the quarter before the beginning of the stratum, respectively

Note: If not mentioned otherwise, variables are defined relative to the beginning of the time window of elapsed unemployment duration.

Sample Sizes and Descriptive Statistics for Selected Variables

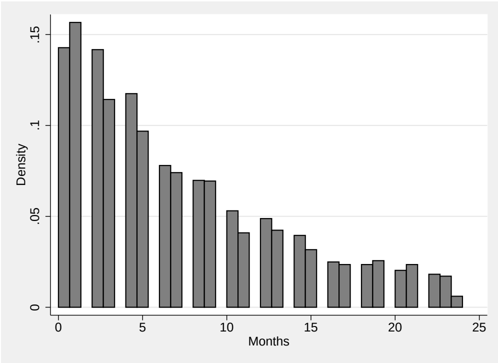
Table 5: Sample Sizes

	East=0, Fem.=0	East=0, Fem.=1	East=1,Fem.=0	East=1, Fem.=1
Stratum 1 (0-3 Months)				
Waiting	29351	18409	15505	8538
STT	912	693	621	368
CFT	389	344	265	136
Stratum 2 (4-6 Months)				
Waiting	18529	12572	10270	6450
STT	547	409	339	286
CFT	251	194	218	143
Stratum 3 (7-12 Months)				
Waiting	10996	8421	5810	4277
STT	662	497	471	353
CFT	270	201	264	218
Aggregated Stratum 1 for PFT (0-12 Months)				
Waiting	25854	16060	12636	6614
STT	2120	1593	1432	1013
CFT	915	741	742	495
PFT	263	234	145	98

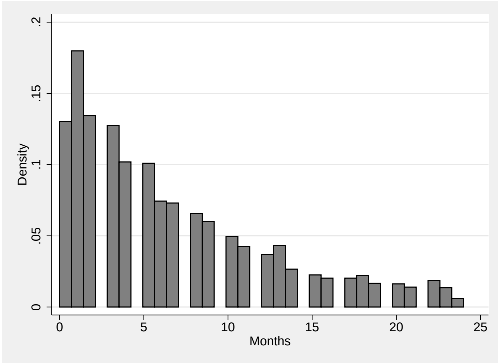
Table 6: Descriptive Statistics for Selected Variables

Variable	Mean	SD	Min	Median	Max
		East=0, Female=0			
age	36.90053	7.718731	25	36	53
foreigner	.1736603	.378823	0	0	1
schooling1	.1237411	.3292911	0	0	1
schooling2	.7414522	.4378433	0	1	1
schooling3	.1348067	.3415223	0	0	1
qualification1	.3572672	.4792019	0	0	1
qualification2	.5908243	.4916894	0	1	1
qualification3	.0519085	.2218457	0	0	1
countemp	801.0523	287.0826	86	873	1096
		East=0, Female=1			
age	37.94112	7.889209	25	37	53
foreigner	.1079787	.3103611	0	0	1
schooling1	.0723604	.25909	0	0	1
schooling2	.7323496	.4427451	0	1	1
schooling3	.19529	.3964335	0	0	1
qualification1	.3214931	.467061	0	0	1
qualification2	.6059022	.488668	0	1	1
qualification3	.0726047	.2594928	0	0	1
countemp	776.9238	312.8626	86	845	1096
		East=1, Female=0			
age	38.3997	7.844472	25	38	53
foreigner	.0340613	.1813919	0	0	1
schooling1	.0546255	.2272544	0	0	1
schooling2	.8529804	.3541357	0	1	1
schooling3	.0923941	.2895899	0	0	1
qualification1	.1134218	.3171169	0	0	1
qualification2	.8383827	.3681101	0	1	1
qualification3	.0481956	.2141855	0	0	1
countemp	829.9092	272.4151	86	914	1096
		East=1, Female=1			
age	39.21149	7.836369	25	39	53
foreigner	.0256919	.1582228	0	0	1
schooling1	.038329	.1919993	0	0	1
schooling2	.8185901	.3853776	0	1	1
schooling3	.1430809	.3501737	0	0	1
qualification1	.1095561	.312352	0	0	1
qualification2	.8129504	.3899717	0	1	1
qualification3	.0774935	.2673868	0	0	1
countemp	748.3454	309.1351	86	762	1096

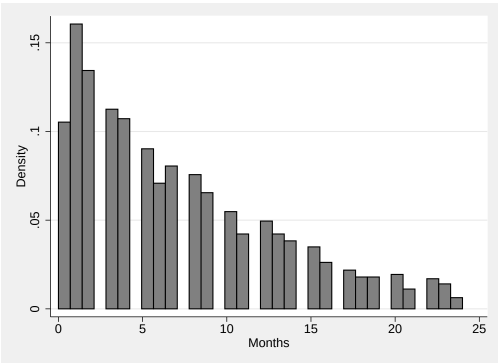
Figure 1: Density of Duration of Unemployment until Program Start



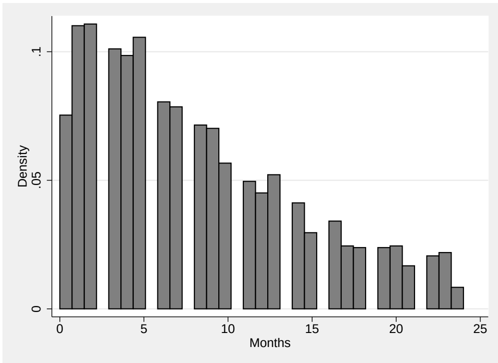
East=0, Female=0



East=0, Female=1



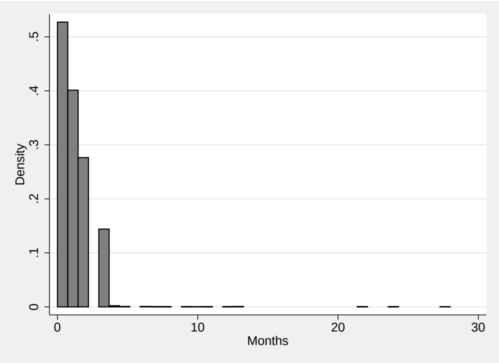
East=1, Female=0



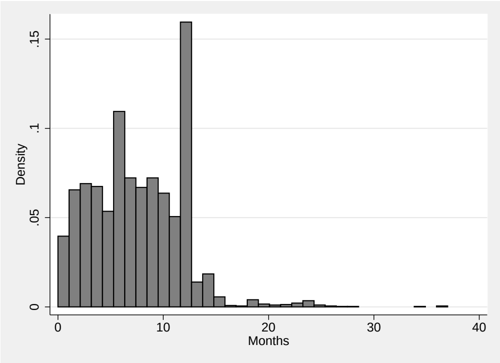
East=1, Female=1

Note: Calculations based on the sample of entrants into STT, CFT, and PFT.

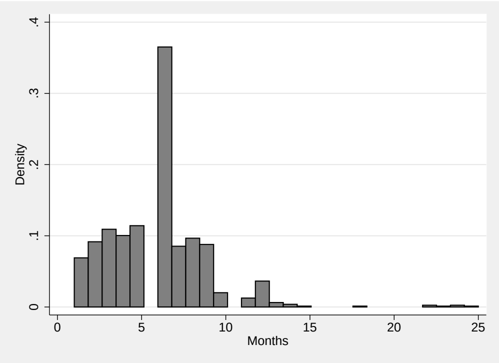
Figure 2: Density of Program Duration



Short Term Training



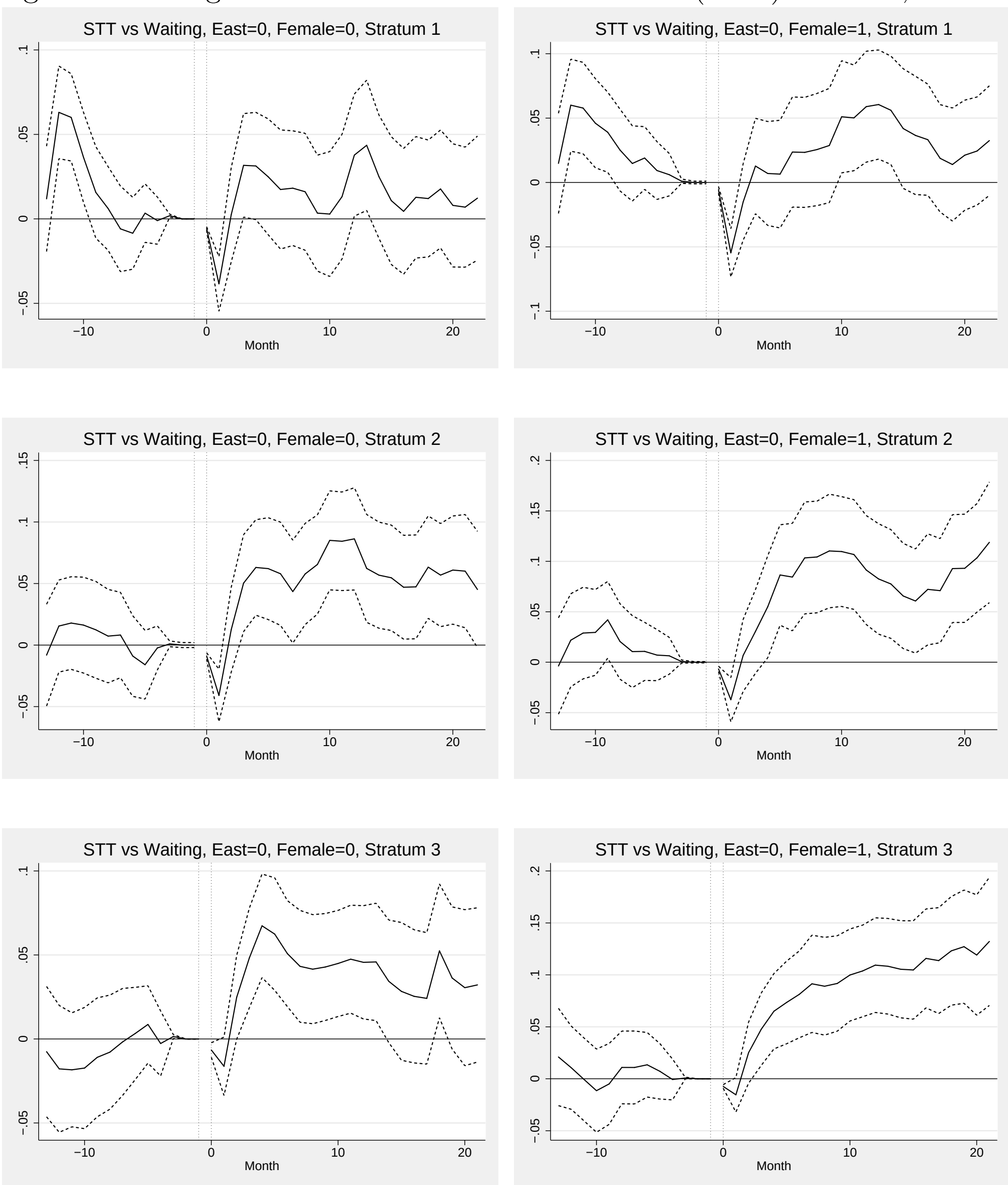
Classroom Further Training



Practical Further Training

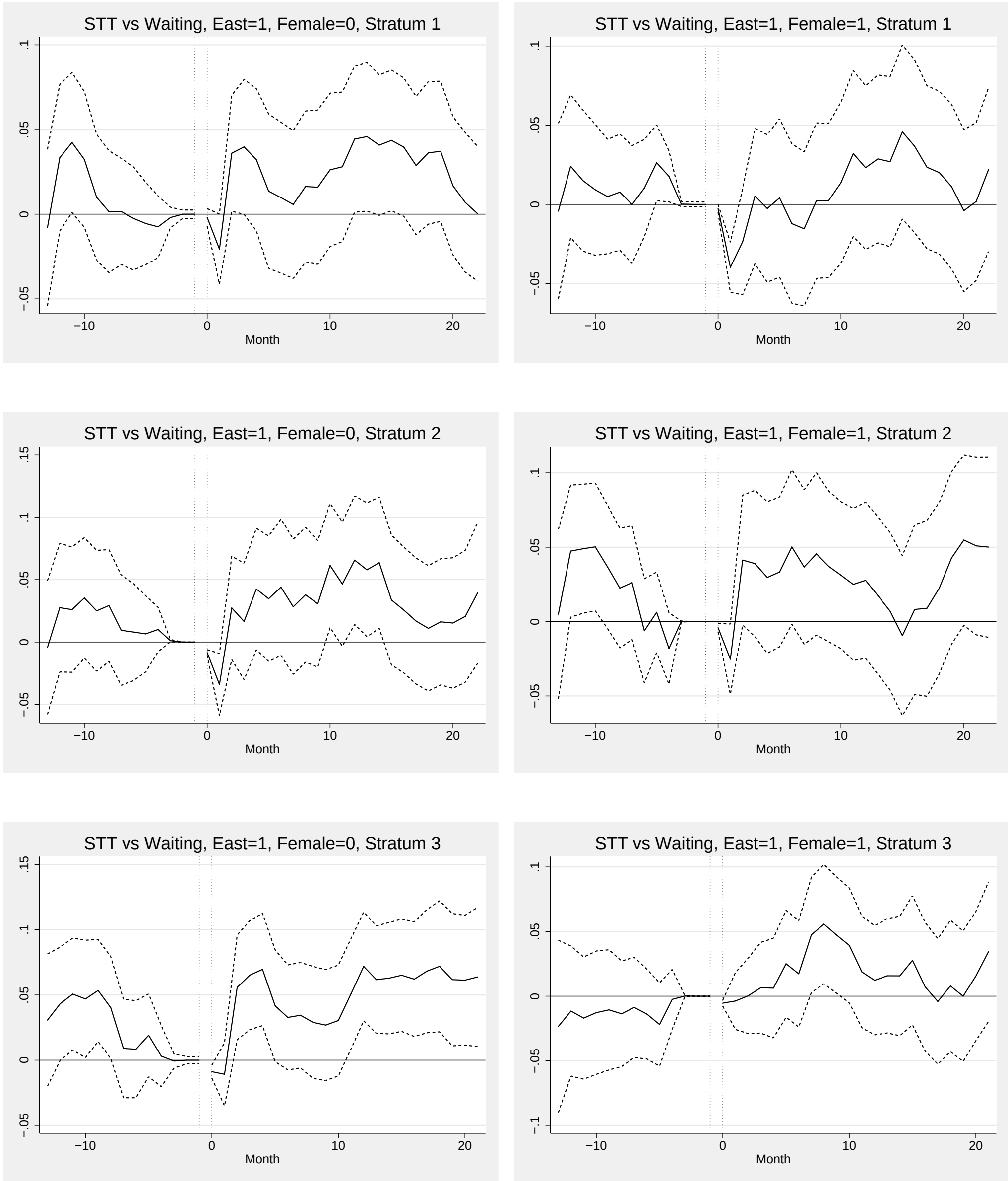
Estimated Employment Effects

Figure 3: Average Treatment Effect on the Treated (ATT) for STT, West Germany



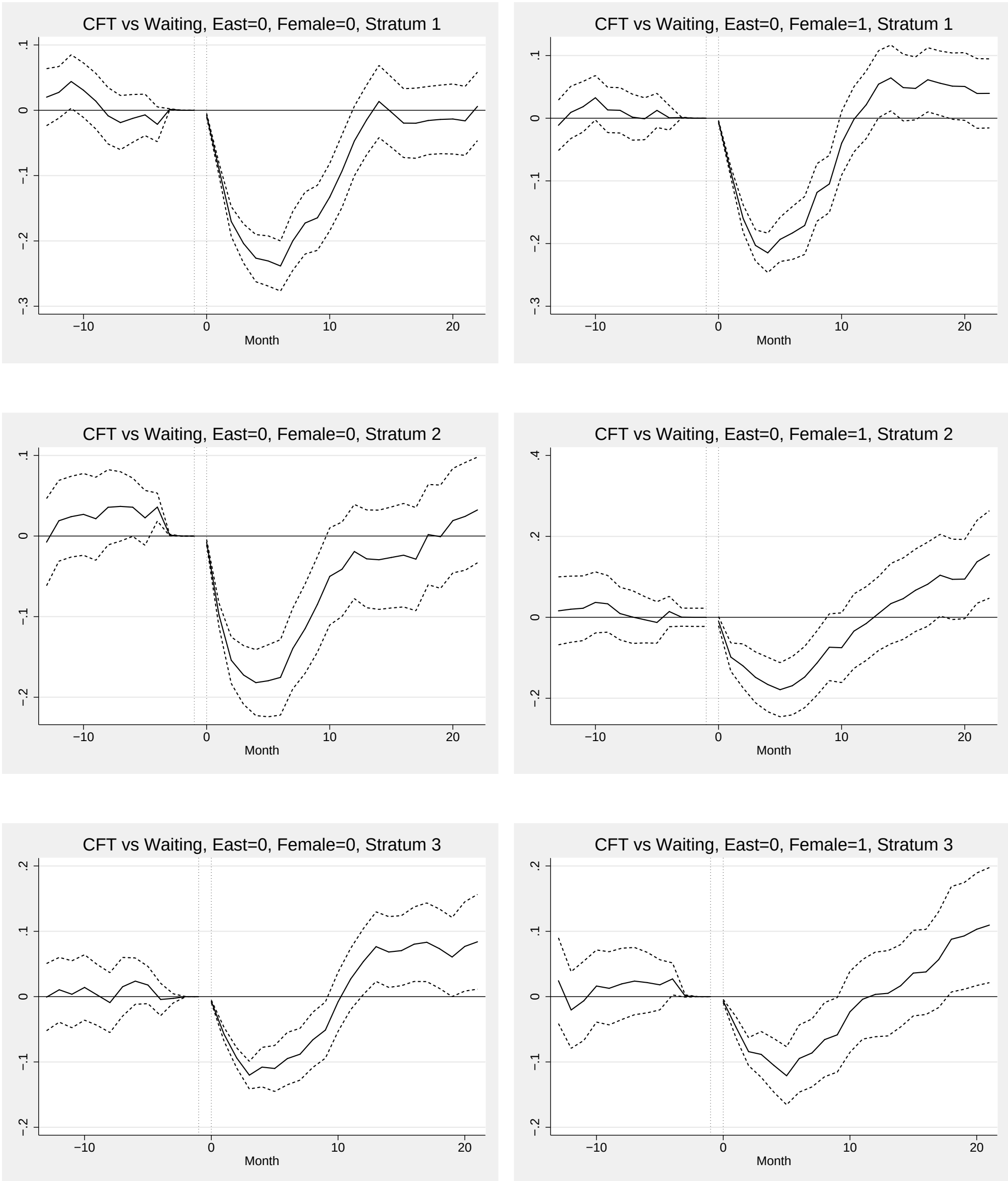
Difference in employment rates is measured on the ordinate, pre-unemployment (< 0) and post-treatment (≥ 0) months on the abscissa.

Figure 4: Average Treatment Effect on the Treated (ATT) for STT, East Germany



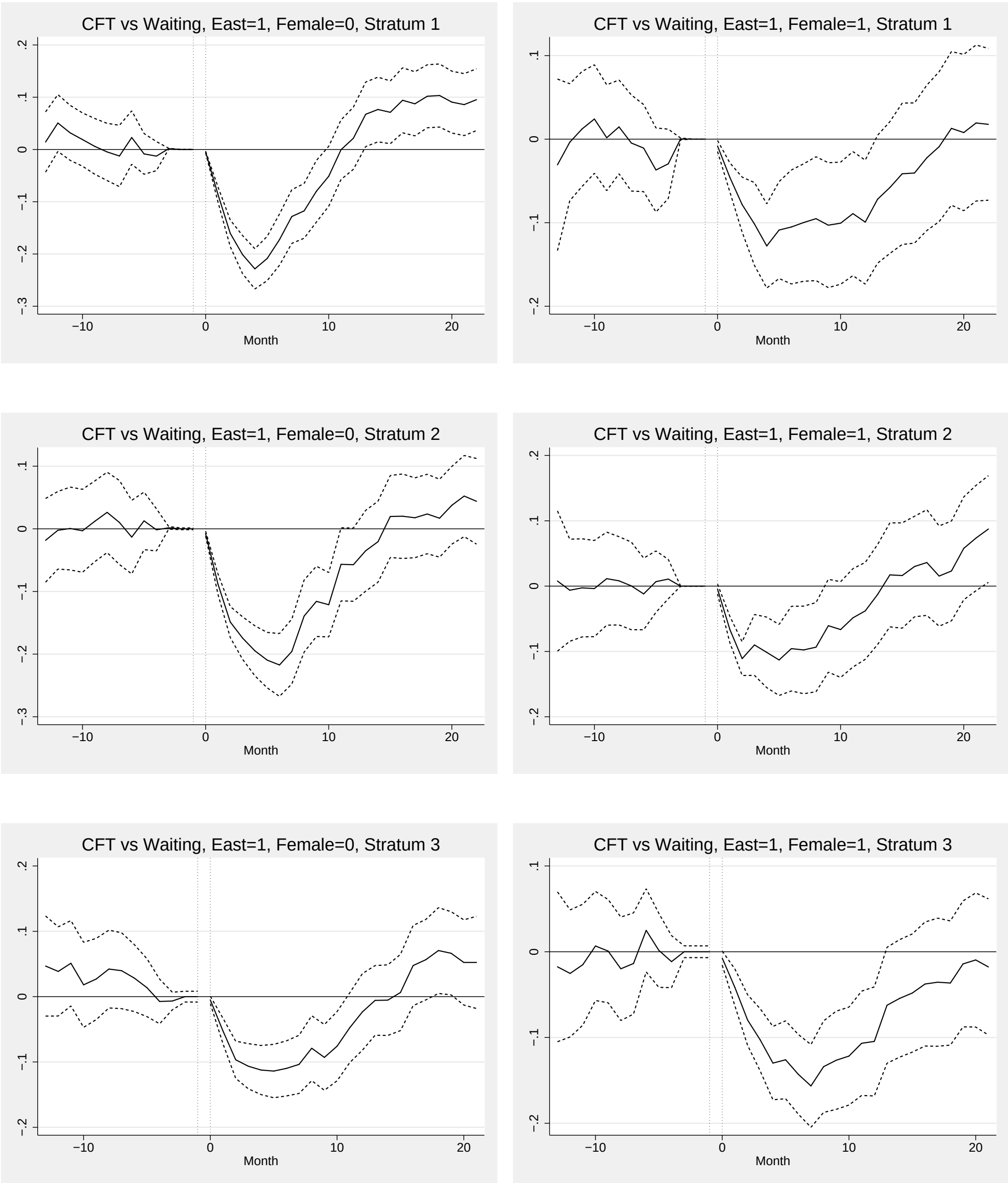
Difference in employment rates is measured on the ordinate, pre-unemployment (< 0) and post-treatment (≥ 0) months on the abscissa.

Figure 5: Average Treatment Effect on the Treated (ATT) for CFT, West Germany



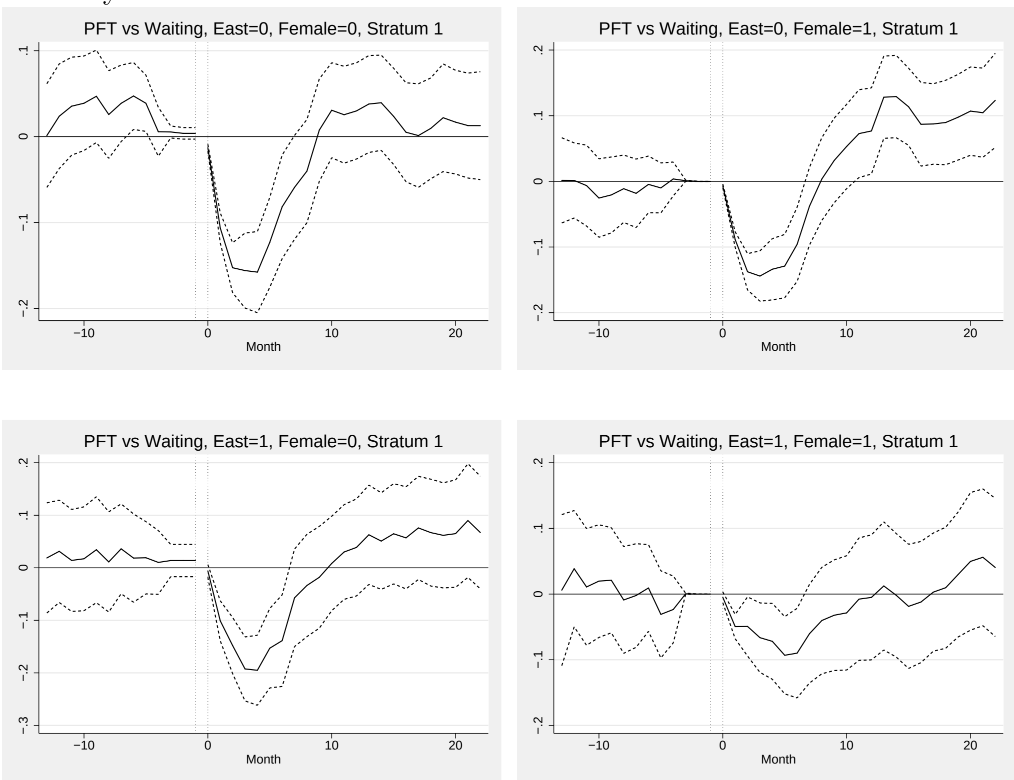
Difference in employment rates is measured on the ordinate, pre-unemployment (< 0) and post-treatment (≥ 0) months on the abscissa.

Figure 6: Average Treatment Effect on the Treated (ATT) for CFT, East Germany



Difference in employment rates is measured on the ordinate, pre-unemployment (< 0) and post-treatment (≥ 0) months on the abscissa.

Figure 7: Average Treatment Effect on the Treated (ATT) for PFT, West and East Germany



Difference in employment rates is measured on the ordinate, pre-unemployment (< 0) and post-treatment (≥ 0) months on the abscissa.