Biewen, Martin; Fitzenberger, Bernd; Osikominu, Aderonke; Waller, Marie

Working Paper

Which Program for Whom? Evidence on the Comparative Effectiveness of Public Sponsored Training Programs in Germany

ZEW Discussion Papers, No. 07-042

Provided in cooperation with:
Zentrum für Europäische Wirtschaftsforschung (ZEW)

Suggested citation: Biewen, Martin; Fitzenberger, Bernd; Osikominu, Aderonke; Waller, Marie (2007) : Which Program for Whom? Evidence on the Comparative Effectiveness of Public Sponsored Training Programs in Germany, ZEW Discussion Papers, No. 07-042, http://hdl.handle.net/10419/24609

Discussion Paper No. 07-042

Which Program for Whom? Evidence on the Comparative Effectiveness of Public Sponsored Training Programs in Germany

Martin Biewen, Bernd Fitzenberger,
Aderonke Osikominu, and Marie Waller

ZEW
Zentrum für Europäische Wirtschaftsforschung GmbH
Centre for European Economic Research

Discussion Paper No. 07-042

Which Program for Whom? Evidence on the Comparative Effectiveness of Public Sponsored Training Programs in Germany

Martin Biewen, Bernd Fitzenberger,
Aderonke Osikominu, and Marie Waller

Non–technical Summary

Based on a new and exceptionally rich administrative data set for Germany, we evaluate the employment effects of four different further training programs in the early 2000s. We consider dynamic aspects of selection into different programs as well as effect heterogeneity with respect to different population subgroups. From an economic policy point of view, we address the following two important questions: (1) the question whether relatively short training measures can compete in effectiveness with more involved medium- to longterm measures, and (2) the question whether practically oriented training programs have advantages over theoretical classroom training.

Our analysis distinguishes the following four types of training programs: short term training, classroom further training, practical further training, and retraining. Short-term training programs last on average several weeks and aim at providing skills that facilitate job search. At the same time, these programs are used to assess and to monitor the abilities and the willingness to work of the unemployed. Classroom and practical further training programs last typically six to twelve months. They provide specific professional skills, either at a theoretical or at a practical level. In contrast to classroom further training, practical further training is usually conducted in a training firm or in the context of an internship. Retraining is with an average duration of two to three years the most comprehensive type of training. It provides a new professional degree according to the German apprenticeship system.

In order to take account of the dynamic sorting process among the unemployed, the treatment status is defined subject to the time window of elapsed unemployment duration. The treatment parameter we estimate thus mirrors 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. We evaluate the employment effects of the four training programs against each of the possible alternatives, i.e. waiting further in open unemployment or participating in one of the other three programs.

For the comparison of participating against waiting, we find statistically significant positive employment effects for male and female participants in short-term training and classroom further training in West Germany who started their training not too early during their unemployment spell. Moreover, West German women but not West German men benefited from practical further training programs. A closer look

reveals that, within the time window permitted by our data set, employment effects of short-term training are of a similar magnitude as those of traditional medium-term programs. However, due to the shorter length, the positive effects of the former materialize much earlier. Whereas for the medium-term programs the initial lock-in period, characterized by strongly negative employment effects, lasts eight to 14 months, it takes only up to two months for short-term training. According to our results, West German men taking part in short-term training or medium-term training increase their medium-term employment rate by some 5 to 10 percentage points. The effect for women is even larger, lying somewhat above 10 percentage points.

The surprising effectiveness of short-term training when compared to the different forms of medium-term training is also confirmed in the pairwise comparisons of the training programs. In particular, this holds for the comparison of short-term training against classroom further training. However, the results are less clear for the comparison with practically oriented further training, where it appears that participants of practical further training courses would have reduced their employment chances if they had taken part in short-term training instead. As to the comparison of classroom and practical further training, our results suggest that practical training may have advantages over pure classroom training. Furthermore, the effects of training programs may be very different across different subgroups. Employment effects are generally larger for individuals who start their program later in the unemployment spell and for women than for men. In some cases older individuals and individuals with low qualifications benefit less. Finally, in contrast to the somewhat positive picture for West Germany, we find only little evidence for positive treatment effects in East Germany.

Which program for whom? Evidence on the comparative effectiveness of public sponsored training programs in Germany[1]

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

[♭]Johannes Gutenberg-University Mainz, IZA, DIW, biewen@uni-mainz.de

[♮]Albert Ludwigs-University Freiburg, ZEW, IZA, IFS, bernd.fitzenberger@vwl.uni-freiburg.de

[♯]Albert Ludwigs-University Freiburg, aderonke.osikominu@vwl.uni-freiburg.de

[§]Albert Ludwigs-University Freiburg, CDSE University of Mannheim, marie.waller@vwl.uni-freiburg.de

Correspondence:

Dept. of Economics, Albert Ludwigs-University Freiburg, 79085 Freiburg, Germany

This version: July 17, 2007

Abstract: We use a new and exceptionally rich administrative data set for Germany to evaluate the employment effects of a variety of public sponsored training programs in the early 2000s. Building on the work of Sianesi (2003, 2004), we employ 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 show considerable employment effects for certain population subgroups but in some cases the effects are zero in the medium run. 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 positive treatment effects in East Germany. There is some evidence that the employment effects decline for older workers and for low–skilled workers.

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 International Economics and Applied Economic Research at the University of St. Gallen (SIAW) and the Institut für Arbeitsmarkt- und Berufsforschung (IAB). We gratefully acknowledge financial and material support by the IAB. We especially thank Stefan Bender for his efforts to make available the data on which this study is based. The usual caveat applies.

Contents

1 Introduction

Recent years have witnessed an enormously increased interest in the evaluation of active labor market policies, both in the US and Europe (for comprehensive overviews see Heckman, LaLonde and Smith (1999), Martin (2000), Martin and Grubb (2001), Kluve and Schmidt (2002) and Kluve (2006)). While, due to methodological and data limitations, earlier studies typically focussed on the evaluation of a single program, recent developments in evaluation methodology and data access have made it possible to gain deeper insights into the possibly very heterogenous effects of different types of programs and their comparative effectiveness. Prominent examples of recent evaluations involving multiple comparisons of different programs are Lechner (2002), Gerfin and Lechner (2002), Sianesi (2003), Hardoy (2005) and Dyke et al. (2006). This progress has been made possible by both methodological developments, in particular the extension of propensity score matching methods to the case of multiple treatments (Imbens (2000), Lechner (2001)), and the increasing availability of large, administrative data sets that provide the necessary sample sizes and program information to carry out in-depth evaluations of narrowly defined sub-programs. Given these new data sources, it is possible not only to evaluate the differential effects of the classical instruments of active labor market policy such as public employment services, job creation in the public sector, or public training programs. It has also become possible to evaluate different sub-programs within these categories, for example to study the comparative effectiveness of different forms of employment subsidies or different forms of public training programs.

This paper contributes to the growing evidence on the comparative effects of public sponsored training programs. We focus on the differential effects of public training programs in Germany. The case of Germany provides ideal conditions to study differential effects of public sponsored training for several reasons. First, the country has a long tradition of extensive active labor market programs covering all kinds of approaches.[2] As to public training programs, the Federal Employment Office of Germany has been offering a wide range of different programs ranging from very short measures aimed at minor skill adjustments and job search assistance to medium- and long-term programs with the explicit goal of increasing the human capital of the participants. In fact, the range of programs offered is much wider than in most other countries and the durations of typical programs vary between one or two weeks to

[2]The total expenditure on active labor market policies was over 20 billion in 2004 (see Bundesagentur für Arbeit (2005a)). Programs include, among others, job search assistance, employment subsidies, job creation in the public sector, youth measures, measures to promote self-employment, and public training programs.

several months or even several years. Another reason for using Germany is that the country has recently developed a growing awareness for the need to evaluate active labor market policies, which helped to open up existing administrative data bases to rigorous scientific research.[3] This has led to large, informative data sets merging different administrative sources. These data sets not only contain precise information on individual employment and transfer receipt histories but also comprehensive and detailed information on participation in all public sponsored measures of active labor market policy.[4] Large sample sizes make it possible to address aspects that have hitherto been difficult or impossible to address such as the heterogeneity of programs, the heterogeneity of effects across different groups of participants and the dynamic selection into different programs.

This paper provides a comprehensive and detailed econometric evaluation of public training programs conducted in Germany during the period February 2000 to January 2002. We distinguish different types of programs and consider effect heterogeneity with respect to population subgroups. Building on the work of Sianesi (2003, 2004) on dynamic treatments and on the work of Lechner (2001) on pairwise comparison of multiple treatments, we employ a stratified matching approach based on the propensity score, the elapsed duration of unemployment, and the calendar time. In order to take account of dynamic sorting processes, we stratify treatment effects by elapsed duration of unemployment. Our results show that average effects for too broad populations may hide statistically and economically significant treatment effects for individual subgroups and therefore help to understand why previous evaluation studies often yielded inconclusive results. While in many cases there is no discernible effect heterogeneity between subgroups, there is some evidence that the effects decline for older workers and for low–skilled workers. In these cases, the differences in treatment effects are very pronounced.

From an economic policy point of view, we address two important questions that have recently attracted considerable attention: the question whether relatively short training measures can compete in effectiveness with more involved medium- to long-term measures, and the question whether practically oriented training programs have advantages over theoretically oriented class-room training. Our main motivation

[3] As a part of major labor market reforms, the so-called *Hartz-Reforms*, the need for rigorous scientific evaluation of program effectiveness was explicitly encoded into the law, see e.g. Jacobi and Kluve (2006).

[4] In fact, part of the project leading to this paper was the design and the validation of a merged administrative data base in cooperation with the Institut für Arbeitsmarkt- und Berufsforschung of the Federal Employment Office. This data base has subsequently been used for most of the policy evaluations in the context of the *Hartz-Reforms*.

for the first question is that, traditionally, the focus of German public training programs was on medium- to long-term measures lasting several months to several years. Following criticism that such programs may not be effective as they 'lock-in' the participants for a long time, there has been a drastic shift towards short-term programs recently.[5] In terms of the number of participants, short-term training measures are by now the largest program of German active labor market policy. One of our aims is to evaluate whether or not this policy change can be justified *ex-post*. The specific form of short-term measures in Germany is also interesting from another point of view, as these measures often comprise elements of job search assistance, profiling or monitoring of the unemployed, apart from the provision of specific skills. By evaluating these kinds of programs we therefore also contribute to the literature that has focused on these specific forms of active labor market policy (see e.g. Martin (2000), Dolton and O'Neill (2002) and OECD (2005)).

The second question we address is also of considerable policy interest. It concerns the contents of training programs and focuses on the aspect of whether practically oriented training measures are better suited to provide unemployed workers with the skills and qualifications needed to improve labor market chances. Our results support hypotheses put forward in the literature (see e.g. Martin and Grubb (2001) and OECD (2005)) that practically oriented training may have advantages over pure classroom training. In this regard, our findings are in contrast to earlier findings for Germany during the 1990's, see Lechner et al. (2005a), Fitzenberger et al. (2006a), and Fitzenberger and Völter (2007).

A key advantage of our methodology is that it allows us to directly *compare* training programs, i.e. to ask the question of what would have happened if participants in short-term programs had participated in longer-term programs, or if participants in classroom training had taken part in more practically oriented training. This leads to more informative results than if one compares the effectiveness of different types of training when compared to not taking part in training at all. These results can directly be used for policy purposes, as they provide information on which programs are most advantageous for whom.

The remainder of the paper is structured as follows. Section 2 reviews the related literature. In section 3 we describe the main institutional features of the German system of public sponsored training. Section 4 presents details on the data used in this study. In section 5, we describe our econometric evaluation strategy. Section 6 discusses our empirical results, and section 7 concludes.

[5] See e.g. Bundesagentur für Arbeit (2005b), and figure 1 below.

2 Literature Review

Although there exists a vast literature evaluating different aspects of active labor market policies in different countries (see the overview studies by Heckman, LaLonde and Smith (1999), Martin (2000), Martin and Grubb (2001), Kluve and Schmidt (2002) and Kluve (2006)) there are relatively few studies that focus on the comparative effects of different forms of training programs.

One of the first studies to consider differences in the outcomes of training programs was Gerfin and Lechner (2002). Using data for Switzerland, Gerfin and Lechner distinguished between five forms of public sponsored training programs with durations ranging between 5 and 13 weeks. Their results were negative in the sense that, one year after program start, the employment rate of participants was lower than that of comparable non-participants. However, longer, more involved training courses seemed to produce less negative results than shorter ones.

Most recent studies that focus on the differential effects of training programs use data for Germany. For example, Lechner et al. (2005a,b) evaluate the effects of a variety of training programs employed in East and West Germany in the 1990s. They distinguish between medium-term programs (mean duration 4 months), longer programs (mean duration 9 to 12 months) and long programs with specific contents such as retraining or training in a practice firm. Lechner et al. conclude that most of the programs had positive effects in the long run, even in East Germany. An important finding is that medium-term programs seem to outperform longer programs as they exhibit a much shorter lock-in period with otherwise similar employment effects after the end of the program. These findings are shared by Fitzenberger and Speckesser (2007), Fitzenberger et al. (2006a), and Fitzenberger and Völter (2007) who use the same data source but different econometric methods. Contrary to common hypotheses about the effectiveness of more practically oriented training programs (see e.g. Martin and Grubb (2001) or OECD (2005)), Lechner et al. (2005a), Fitzenberger et al. (2006a), and Fitzenberger and Völter (2007) do not find that practical training as implemented in the 1990s dominates other kinds of training.

Using more recent and more informative data, Hujer et al. (2004) study the effectiveness of training programs in the early 2000s depending upon the duration of the programs. The study distinguishes programs of short (1-3 months), medium (6-12 months), and long (over 12 months) duration and estimates a multivariate mixed proportional hazard model. The results imply strong lock-in effects for the time the

programs are attended but no significant effects on the exit rate from unemployment after completion of the program. Schneider et al. (2006) present policy evaluation results commissioned by the federal government in the context of the *Hartz-Reforms*. Although their focus is on the changes caused by these reforms, they also provide some results on the comparative effectiveness of a number of medium-term and long-term training programs. Their results also confirm the finding that shorter programs may be more effective than longer ones.

A drawback of all of these studies is that they omit the by now most important type of public sponsored training in Germany, so-called short-term training ('Trainingsmaßnahmen') – this program is not to be confused with short further training programs as analyzed by Hujer et al. (2004) or Lechner et al. (2005a,b). Short-term training courses typically last only 2 to 12 weeks and often combine elements of job search assistance with the provision of specific skills (see more detailed description below). In light of the policy debate (Martin and Grubb (2001) or OECD (2005)), short-term training seems attractive since it may serve the purpose of activating the unemployed without locking them in lengthy training programs. Furthermore, a number of recent contributions from the evaluation literature suggest that increased job search assistance may be an inexpensive way to help unemployed individuals back into employment (see e.g. Blundell et al. (2004), Weber and Hofer (2004), Fougère et al. (2005), Hujer et al. (2005), Crépon et al. (2005), and Van den Berg and Van der Klaauw (2006)).

The only other two studies we are aware of that consider short-term training in Germany are Hujer et al. (2006), and Lechner and Wunsch (2006). Hujer et al. examine whether participation in short-term training measures reduces the unemployment duration of West German job-seekers. They do not compare short-term training to other measures of active labor market policy. Lechner and Wunsch (2006) evaluate a large number of different training and non-training measures in East Germany, among them short-term training. Their results suggest no or even negative effects for all programs considered. Lechner and Wunsch explain their finding by the difficult situation in the East German labor market.

3 Training as Part of Active Labor Market Policy

The main goal of German active labor market policy is to permanently reintegrate unemployed individuals (and individuals who are at risk of becoming unemployed)

back into employment. The policy instruments cover a wide range of different measures such as employment subsidies, job creation in the public sector, measures directed at youth unemployment, measures to promote self-employment, and public training programs. For an overview over the different kinds of policies and their quantitative importance, see figure 1 and table 3 in the appendix.[6]

— Figure 1 about here —

As shown in figure 1, public training programs have traditionally been the most important part of German active labor market policy. There are three main categories of training programs: *short-term training* ('Trainingsmaßnahmen'), *further training* ('Berufliche Weiterbildung'), and *retraining* ('Umschulung').[7] Apart from the fact that all three types of training require full-time participation, they differ considerably in length and contents. Recently, *short-term training* has become the largest training program regarding the number of participants – for the following, see Kurtz (2003). Short-term training measures last only two to twelve weeks (the mean duration is slightly over four weeks, see table 1) and typically pursue one or several of the following three aims. A first potential aim is aptitude and qualification testing, i.e. the program is used to assess job seekers' labor market opportunities and their suitability for different jobs. This may also entail profiling activities on the side of the Federal Employment Office and preparation of more detailed work plans to reintegrate the job seeker into the labor market. A second aim is to test the job seeker's willingness to work and to improve job search skills. This may be achieved through activities such as job-application training, simulation of job interviews or general counseling on job search methods. The third and final aim of short-term training measures is the provision of specific skills that are necessary to improve the job seeker's labor market prospects. Typical examples for this type of measures are computer courses or courses providing commercial training. In 2001, 22 percent of short-term training measures belonged to the first type, 19 percent to the second type, and some 28 percent to the third type. About 31 percent were combinations of the different types. In most cases, these were combinations of job search assistance

[6]This paper focuses on public training programs attended in the period 2000 to 2002. The following paragraphs describe the relevant institutional settings up to the end of 2002, before the *Hartz-Reforms* were enacted. The reforms also changed some of the rules on public training programs. These changes are not relevant to our study but they will be important for future evaluations (see e.g. Biewen and Fitzenberger (2004) or Schneider et al. (2006)).

[7]In addition, there are specific training schemes for youth unemployed and disabled persons, as well as German language courses for asylum seekers and ethnic Germans returning from former German settlements in Eastern Europe. These training measures are not considered here.

and the provision of specific skills, or aptitude testing and the provision of specific skills (Kurtz, 2003, tables A3 and A6).

In comparison to short-term training, the more substantial *further training* programs typically take much longer and are more involved. With durations ranging between several months and one year, further training measures can be classified as medium-term programs. Their aim is to maintain, update, adjust, and extend professional skills and qualifications. Further training programs cover a wide range of courses in a variety of fields and may also comprise practical elements such as on-the-job training, internships or working in practice firms. In our evaluation we will distinguish between practically-oriented further training programs (which are typically of shorter duration) and pure class-room training. Apart from short-term and further training, employment offices also offer *retraining*. Retraining programs last two to three years and typically lead to a new vocational education degree within the German apprenticeship system. Retraining may involve vocational training in a profession that was not the original profession of the job seeker. In addition, retraining may be granted to job seekers who face difficult labor market prospects because they lack a vocational degree in the first place. In general, retraining programs are similar to regular apprenticeships and typically combine class-room training with on-the-job training.

To become eligible for participation in one of the training programs, 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 do not have to fulfil any additional eligibility criteria. In the case of medium- and long-term training, individuals are typically eligible only if they also fulfil a minimum work requirement of one year and if they 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 example, 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. Depending on regional and local circumstances, caseworkers may exercise a great deal of discretion when allocating the different programs. Suitable programs are chosen from a pool of certified public or private institutions or firms.

If a person is admitted to one of the training measures, the employment office pays all direct training costs. In addition, the participants of short-term training may con-

tinue to receive unemployment benefits or means-tested unemployment assistance, if they are eligible for such transfer payments. Participants of short-term training are still registered unemployed during the program. In contrast, participants of further training or retraining do not remain registered unemployed during the program. Participants of further training and retraining usually also receive a subsistence allowance provided they fulfill a minimum work requirement of twelve months within the last three years. This subsistence allowance is usually of the same amount as unemployment benefits or unemployment assistance. Overall, there are no significant financial incentives for unemployed individuals to participate in a training program, in contrast to the situation in Germany before 1998, see Fitzenberger et al. (2006a).

— Table 1 about here —

Table 1 shows that the average monthly training costs per participant are lower for short-term training courses (about 570 Euros in 2001) than for the longer-term measures (664 Euros). Given the that average length of short-term measures is only 1.1 months while that of longer-term measures is some 9.3 months, this results in training costs for short-term measures (627 Euros) that amount to only about one tenth of those for medium- and long-term measures (6175 Euros).[8] Since 2002, in light of huge differences in costs, the Federal Employment Office has been drastically increasing the share of short-term training measures at the expense of longer-term measures (see figure 1). Of course, the higher training costs may be justified if the medium- to long-term measures lead to correspondingly higher gains in employment probabilities. This one of the main questions motivating our evaluation.

4 Data

4.1 Integrated Biographies Sample

Our study uses a new and exceptionally rich administrative data base, the so-called Integrated Biographies Sample (IEBS). This data base has only recently been made available by the Federal Employment Office of Germany.[9] The IEBS is a merged

[8] In addition to the direct costs, participants in longer-term training schemes usually receive the subsistence allowance. However, the subsistence payments simply replace the ordinary unemployment compensation the participants would have otherwise received.

[9] For more information on the IEBS, see Osikominu (2005, section 3) and Hummel et al. (2005).

2.2% random sample of individual data drawn from the universe of data records collected through four different administrative processes. Our version of the IEBS has been supplemented with additional information which is not publicly available (especially information on health). The IEBS contains detailed daily information on employment subject to social security contributions, receipt of transfer payments during unemployment, 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 and assessments of case workers. For evaluation purposes, a rich set of covariates is essential as 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 selection of individuals into programs.

We give a brief description of the IEBS in order to underscore its value for evaluation purposes. The IEBS is based on four different administrative sources the so-called 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 *Employment History* involves register data comprising employment information for all employees subject to contributions to the public social security system. It covers the time period 1990 to 2004. 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 between January 1990 and June 2005. 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. The Information on benefit payments allow us to construct individual benefit histories dating back several years. Moreover, we use additional information contained in the Benefit Recipients History involving sanctions and periods of disqualification from benefit receipt that may serve as indicators for a lack of motivation.

The third administrative data source of the IEBS is the so-called *Supply of Applicants*, which contains data on individuals searching for jobs. The Supply of Applicants data cover the period January 1997 to June 2005. In our study they are used in two ways. First, they provide additional information about the labor market status of a person, in particular whether the person in question searches for a job but is not (yet) registered as unemployed or whether he or she is sick while registered unemployed. Second, the job search episodes include additional information about personal characteristics, in particular about educational qualifications, nationality, and marital status. They also provide information about whether the applicant wishes to change occupations, about health problems that might influence employment chances, and about the labor market prospects of the applicants as assessed by the case worker. Finally, the data on applicants include regional and local identifiers, which we use to link regional and local information, for example unemployment rates at the district level.

The fourth data source in the IEBS is the *Data Base of Program Participants*, which is particularly important for evaluation purposes. This data base contains detailed information on participation in public sector sponsored labor market programs covering the period January 2000 to July 2005. 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 regular and subsidized employment when evaluating employment outcomes.[10]

[10]A disadvantage of the data covering labor market training in German in the 1990s used in studies such as Fitzenberger et al. (2006a), Fitzenberger and Speckesser (2007), Fitzenberger and Völter (2007), and Lechner et al. (2005a,b) is that it is not possible to distinguish whether participants found employment in the regular labor market or whether they took part in job creation measures. Note that for the time period from the year 2000 onwards, Lechner et al. (2005a,b) use the information based on the IEBS whether an individual is employed in a subsidized job.

Being among the first to use the IEBS, we were involved in comprehensive data checks.[11] 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.[12] Our conclusion is that on the one hand the employment and benefit data are highly reliable concerning employment status, wage and transfer payments, and the start and end dates of spells. The likely reason for this is that contribution rates and benefit entitlements are directly based on this information. On the other hand, information not needed for these administrative purposes can be less reliable. For example, in the employment data base the educational variable appears to be affected by non-negligible measurement error as it is not directly relevant for social security entitlements (see Fitzenberger et al. (2006b) for imputation methods to correct the education variable). 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. In our evaluation, we exploited the available information as efficiently as possible by choosing the data source that is most reliable for a given purpose.

Although the data in the IEBS generally seem very reliable, there is some need for data corrections. In particular, we corrected in some cases the end dates of program spells if there was evidence that the end dates recorded in the data base of program participation was wrong. For details on measurement error in program end dates in the IEBS and correction procedures, see Waller (2007).

4.2 Evaluation Sample and Training Programs

We follow an evaluation strategy (see below) that is based on comparisons with (multiple) control groups. A common feature of control group approaches is that they partition the group of potential participants into a group of participants and a group of non-participants. As a consequence, the first question that has to be answered when selecting the evaluation sample is that of who is a potential program participant.

For several reasons, we decide to focus on individuals who become unemployed after having been continuously employed for at least three months, instead of individuals

[11] Given the non-trivial task of merging four large scale administrative data sources of very different designs such checks were indispensable.

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

who are *observed* unemployed at a given point of time. This is to avoid the case of individuals registering as unemployed from being out of labor force because they want to participate in a training program. In interviews, case workers told us that especially women returning from maternity leave, divorcees, or university graduates who have difficulty finding a job may contact the local employment office inquiring about the possibility of participating in public training programs. However, these individuals often only register as unemployed if the chances of actual participation are high enough. An evaluation sample based on observed unemployment status (instead of an inflow sample into unemployment) would therefore suffer from the problem of an incompletely observed control group, because it would be difficult to find comparable non-participants for those individuals who endogenously register as unemployed (due to their non-registering as unemployed, non-participating counterparts would not appear in the sample). Analyzing an inflow sample into unemployment, we focus on individuals who have been attached to the labor market, which helps to construct the control group based on the labor market relevant information in the data. Furthermore, the beginning of unemployment defines a natural time scale to align treated and nontreated individuals.

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 employment office (not necessarily immediately), either through benefit receipt, program participation or a job search spell.[13] In order to exclude individuals eligible for specific labor market programs for the youth and individuals eligible for early retirement schemes, we only consider persons aged between 25 and 53 years at the beginning of their unemployment spell. Our evaluation focusses on the first training program that is attended in the course of an unemployment spell.

Based on the description of program types in section 3, we analyze four different types of training, which closely follows the legal grouping of program types:

- short-term training (STT),
- classroom further training (CFT),

[13]Note that this implies that the same individual may appear more than once in our evaluation sample. About ten percent of the individuals in our sample are represented by more than one unemployment spell according to the above definition. We take account of multiple inclusion of the same individual in the sample when calculating standard errors, see section 5.

- practical further training (PFT), and
- retraining (RT).

In some cases, we grouped programs whose planned duration and contents did not really fit into the category defined by the law into the category that was most appropriate from an economic point of view. According to the same criteria, we also grouped measures of 'discretionary support' (Freie Förderung) and measures financed through the European Social Fund (Europäischer Sozialfond, ESF) into one of the four program categories. We carry out our evaluations for men and women, for East and West Germany, and (for reasons explained in the next section) for different durations of elapsed unemployment separately. This results in a total number of twelve evaluation samples, the sample sizes of which are shown in table 4 in the appendix. Table 7 and figure 16 provide descriptive information on the duration of different program types. STT is the shortest program and RT the longest program. Durations for CFT are fairly uniformly distributed between 1 and 12 months with a strong spike at 12 months. PFT is shorter than CFT and shows a strong spike at 7 months.

5 Econometric Implementation

Our goal is to analyze the effect of the $K = 4$ different training programs (STT, CFT, PFT, RT) on monthly employment at the individual level. In a situation where individuals have multiple treatment options, we estimate the average treatment effect on the treated (ATT) of one training program against nonparticipation in any of the three programs and of pairwise comparisons of two programs. Extending the static multiple treatment approach to a dynamic setting, we follow Sianesi (2003, 2004) and apply the standard static treatment approach recursively depending on the elapsed unemployment duration. The implementation builds upon the approach for binary treatment in Fitzenberger and Speckesser (2007) and for multiple treatments in Fitzenberger et al. (2006a). In contrast to these earlier papers, we also analyze the heterogeneity of the estimated ATT by various socio-economic characteristics of the treated individuals.

5.1 Multiple Treatments in a Dynamic Context

Our empirical analysis is based upon the potential-outcome-approach to causality, see Roy (1951), Rubin (1974), and the survey of Heckman et al. (1999). Lechner (2001) and Imbens (2000) extend this framework to allow for multiple, exclusive treatments. Let the potential outcome $Y^k, k = 1, ..., 4$, represent the outcome associated with training program k and Y^0 is the outcome when participating in none of the 4 training programs. For each individual, only one of the $K + 1$ potential outcomes is observed and the remaining K outcomes are counterfactual. We estimate the average treatment effect on the treated (ATT) of participating in treatment $k = 1, 2, 3, 4$ against nonparticipation $k = 0$ (treatment versus waiting) and the differential effects of the programs (program k versus program l where $k, l \neq 0$), see Lechner (2001).

Fredriksson and Johansson (2003, 2004) argue that a static evaluation analysis, which assigns unemployed individuals to a treatment group and a nontreatment group based on the treatment information observed in the data, yields biased treatment effects. This is because the definition of the control group conditions on future outcomes or future treatment. For Sweden, Sianesi (2004) argues 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 Germany). This discussion implies that a purely static evaluation of the different training programs is not warranted. Following Sianesi (2003, 2004), we analyze the effects of the first participation in a training program during the unemployment spell considered *conditional on the starting date of the treatment*. 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).

We analyze treatment conditional upon the unemployment spell lasting at least until the start of the treatment k and this being the first treatment during the unemployment spell considered. Therefore, the ATT parameter (comparing treatments k and l) of interest is

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

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

where T_u is the treatment variable for treatment starting in month u of unemploy-

ment. $Y^k(u,\tau)$, $Y^l(u,\tau)$ are the potential treatment outcomes for treatments k and l, respectively, in periods $u+\tau$, where treatment starts in period u and $\tau = 0, 1, 2, ...$, counts the months since the beginning of treatment. When $l = 0$, we compare treatment k versus waiting (nonparticipation in the stratum) and when $l \geq 1$, we do a pairwise comparison between treatment k and l. U is the duration of unemployment, $\tilde{u}$ is the random month when alternative treatment l starts, and $\bar{u} = 2, 4, 8$ is the last month in the stratum of elapsed unemployment considered. Then, $\tau - (\tilde{u} - u)$ counts the months since start of treatment l yielding alignment of unemployment experience, because $u + \tau = \tilde{u} + (\tau - (\tilde{u} - u))$, and $Y^l(\tilde{u}, \tau - (\tilde{u} - u))$ is the outcome of individuals who receive treatment l between period u and $\bar{u}$. For starts of l later than u, we have $\tilde{u} - u > 0$ and therefore, before l starts, $\tau - (\tilde{u} - u) < 0$. Then, 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. This way, we account for the fact that alternative treatments, for which the individual receiving treatment k in period u is eligible, might not start in the same month u. The treatment parameter we actually estimate is the average within a stratum

$$\theta(k, l; \tau) = \sum_u g_u \theta(k, l; u, \tau) \ ,$$

with respect to the distribution g_u of starting dates u within the stratum.

Our estimated treatment parameter (1) mirrors 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.

We evaluate the differential effects of multiple treatments assuming the following dynamic version of the conditional mean independence assumption (DCIA)

$$(2) \quad E(Y^l(\tilde{u}, \tau - (\tilde{u} - u))|T_u = k, u \leq \tilde{u} \leq \bar{u}, U \geq u-1, T_1 = ... = T_{u-1} = 0, X)$$

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

where X are time-varying as well as time-invariant (during the unemployment spell) characteristics, $T_{\tilde{u}} = l$ indicates treatment l between u and $\bar{u}$ ($\bar{u}$ is the end of the stratum of elapsed unemployment considered), and $\tau \geq 0$, see equation (1) above and the analogous discussion in Sianesi (2004, p. 137). We effectively assume that conditional on X, conditional on being unemployed at least until period $u-1$, and conditional on not receiving any treatment before u (both referring to treatment in period u) individuals are comparable in their outcome for treatment l occurring

between u and $\bar{u}$.

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 treatment l, $P^{k|kl}(X)$, exhibits an analogous balancing property for the pairwise estimation of the ATT's of program k versus l. This allows to apply standard binary propensity score matching based on the sample of individuals participating in either program k or in program l. For this subsample, we simply estimate the probability of treatment k and then apply a bivariate extension of standard propensity matching techniques. Implicitly, we assume that the actual beginning of treatment within a stratum is random conditional on X.

To account for the dynamic treatment assignment, we estimate the probability of treatment k given that unemployment lasts long enough to make an individual 'eligible'. For treatment during months 0 to 3, we take the total sample of unemployed, who participate in k or l during months 0 to 3 (stratum 1), and estimate a Probit model for participation in k. For $l = 0$, the group of nonparticipants in k includes those unemployed who either never participate in any program or who start some treatment after month 4. For treatment during strata 2 and 3, the basic sample consists of those unemployed who are still unemployed in the first month of the stratum.

We implement a stratified local linear matching approach by imposing that the matching partners for an individual receiving treatment k are still unemployed in the month before treatment k starts, i.e. we exactly align treated and nontreated individuals by elapsed unemployment duration in months. For the comparison of training against waiting, we align treated and controls in addition by the elapsed duration of unemployment benefit receipt in months. The expected counterfactual employment outcome for nonparticipation is obtained by means of a bivariate local linear regression on the propensity score and the starting month of the unemployment spell. We use a product kernel in the estimated propensity score and the calendar month of entry into unemployment

$$KK(p,c) = K\left(\frac{p-p_j}{h_p}\right) \cdot h_c^{|c-c_j|}, \tag{3}$$

where $K(z)$ is the Gaussian kernel function, p and c are the propensity score and the calendar month of entry into unemployment of a particular treated individual, 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. h_p and h_c are the bandwidths. Taken together, we impose three matching requirements: i) similarity of the pairwise propensity score, ii) *exact* match of the elapsed unemployment (and benefit receipt) duration, and iii) similarity of beginning of unemployment.

We use a bivariate crossvalidation procedure to obtain the bandwidths h_p and h_c by minimizing the squared prediction error for the average of the l-outcome for the nearest neighbors of the participants in program k.[14] An estimate for the variance of the estimated treatment effects is obtained through bootstrapping based on 250 resamples. This way, we take account of the sampling variability in the estimated propensity score.

As a balancing test, we use the regression test suggested in Smith and Todd (2005) to investigate whether the covariates are balanced sufficiently by matching on the estimated propensity score using a flexible polynomial approximation. Furthermore, we investigate whether treated and matched nontreated individuals differ significantly in their outcomes before the beginning of unemployment, in addition to those variables already used as arguments of the propensity score. We estimate these differences in the same way as the treatment effects after the beginning of the program. By construction, treated individuals and their matched counterparts exhibit the same unemployment duration until the beginning of treatment.

5.2 Specification of the Propensity Scores

First, we need to discuss the plausibility of the DCIA (2) for our application. For propensity score matching to be a valid procedure one needs to control for the variables that jointly influence participation and outcomes such that, when conditioning on these variables, potential outcomes are mean independent of treatment status. It is therefore essential to base the estimation of the propensity scores on all relevant information. Given our data base, we are in the lucky position to construct a large set of time-constant as well as time-varying (within the unemployment spell) variables to model the selection into the different training programs.

As Sianesi (2004), we argue that the participation probability depends upon the variables determining re-employment prospects once unemployment began. Conse-

[14] This method is also used in Fitzenberger et al. (2006a) and it is an extension of the crossvalidation procedure suggested in Bergemann et al. (2004).

quently, all individuals are considered who have left employment in the same two years (matching controls for beginning of unemployment) and who have experienced the same unemployment duration before program participation. Furthermore, observable individual characteristics and information from the previous employment and benefit history have been included in the propensity score estimation. E.g., we consider skill information, regional information, occupational status, and industry which should be crucial for re-employment chances. In addition, we use subjective assessments of the unemployed by case workers, which should proxy for further relevant unobserved characteristics. In addition to matching on the begínning of unemployment and the elapsed duration of unemployment, we argue that the variables used in the estimation of the propensity score are rich enough to control for the selection into treatment. This is particularly plausible because participation occurred at a fairly large scale, assignment was not very targeted and driven by the supply of programs, and case workers had little guidance on 'what works for whom'. Supporting our point of view, Schneider et al. (2006) argue that until 2002 assignment to training was strongly driven by the supply of available courses.

Concretely, we use the following variables and their interactions for the specification of the propensity score.[15]

Personal characteristics

As personal characteristics, we consider age, disability status, schooling and professional qualification, family status, whether there are children, whether there are children under 10 years, nationality other than German, and whether the person in question is an ethnic German who has migrated back into Germany (usually from Eastern European countries).

Labor market and benefit histories

We use information on 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 total length of employment (all durations are measured in days) during the last three years, the duration of transfer payments during the last three years (i.e. unemployment benefits, unemployment assistance, subsistence

[15]See the appendix for summary statistics and a more detailed description of the variables used. Time-varying covariates are updated at the beginning of each stratum. 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.

allowance), times without any information in the data set, times of contact with the employment office 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.

Case worker reported assessments

As to the assessment of the case workers with regards to the motivation, plans and labor market prospects of the unemployed, we consider current health status, past health problems, information on whether a program was 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 job proposals made the unemployed received from the employment office, and information on the desired job.

Regional information

We use different unemployment rates in the home district of an individual, the districts which share the labor market classification of the region, the federal state, and all of Germany.

Using these variables as possible regressors, we fit the propensity scores separately for each of the twelve evaluation samples (men/women, East/West, stratum 1/2/3), and each treatment comparison pair. In each case, we run an extensive specification search. The final specification is chosen based on economic considerations, statistical significance of the variables included, and the balancing tests described above.[16] The final specification typically includes 20 to 35 covariates.

5.3 Estimating Effect Heterogeneity

The estimation of the ATT provides a semiparametric, aggregate impact measure for a possibly heterogeneous treatment group. However, it is conceivable that a zero average hides positive and negative treatment effects for different subgroups of

[16]Estimation results are available from the authors upon request.

the treated. Although, one could estimate the ATT for each subgroup of interest[17] using the dynamic matching estimator as described in section 5.1, such a strategy is limited by the curse–of–dimensionality. The sample sizes of the subgroups need to be sufficiently large to do so with reasonable precision.

As a simple alternative, we propose to run linear regressions of the estimated individual cumulated treatment effects in the matched samples on a covariates which could cause effect heterogeneity. Such regressions after matching are used in the literature to adjust for possible remaining mismatch between treated individuals and matched controls (see e.g. Lechner (1999)). However, we are not aware of any recent study in the evaluation of active labor market policy that uses regressions after matching to study effect heterogeneity. We focus on the cumulated treatment effect after the end of the lock–in period.

To be specific, we run the following regressions in the matched sample for the cumulated treatment effect over the months $\tau = T_1, ..., T_2$

$$\sum_{\tau=T_1}^{T_2} (Y_i^k(\tau) - \hat{Y}_i^l(\tau)) = \alpha + x_i\beta + (x_i - \hat{x_i})\gamma + u_i \,, \tag{4}$$

where individual i with covariates x_i receives treatment k with observed outcome Y_i^k, $\hat{Y}_i^l$ and $\hat{x_i}$ are the predicted, counterfactual l-outcome and the predicted covariates based on the local linear regression in the matching procedure, and $x_i - \hat{x_i}$ represents the mismatch between the average covariates of the matched comparison individuals and the covariates of indvidual i. We estimate the cumulated effect over the time interval $[T_1, T_2]$ after the beginning of treatment. T_1 and T_2 are chosen specifically for the program and the stratum under consideration. T_1 is a proxy for the end of the lock–in period and T_2 is the last post treatment month observed. As a benchmark, if $\beta = \gamma = 0$, then the estimated α corresponds to the estimated ATT for the entire treatment sample and there is no systematic effect heterogeneity by the level of covariates.

We test systematically for significant effect heterogeneity by covariates. The standard errors of the estimated regression coefficients are obtained through the bootstrap procedure for the matching estimator by rerunning the regression (4) for all resamples. In some cases, the regression (4) in the matched sample suffers from multicollinearity problems due to the mismatch terms being highly correlated with the covariates. In cases of strong multicollinearity, we exclude the mismatch from

[17] E.g. Lechner et al. (2005a) do so for a small number of subsets of the treatment sample to investigate effect heterogeneity.

the regression. As final results we only report those specification of the effect heterogeneity regressions with significant covariates.

6 Empirical Results

6.1 Training vs. 'Waiting'

The evaluation results for training vs. not participating in any measure of active labor market policy ('waiting') are shown in figures 2 to 6. Each graph displays 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. More precisely, 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.* As already mentioned, 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) and one ranging from month 0 to 3 for participants in retraining (RT).

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 2 about here —

Figure 2 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 programs of minus five percentage

points (i.e. during the program, participants had a five 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 two or three 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 turn positive. 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 men starting a short-term training program in months 7 to 12 (stratum 3) of their unemployment spell, and women starting one in month 4 or later are positive and statistically significant (except for men in stratum 2). According to these estimates, the monthly employment probability of West German men participating in short-term training is increased by about 5 percentage points. At some 10 percentage points, this effect is larger for women.

— Figure 3 about here —

Figure 3 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 their participants. Measured average treatment effects are mostly small and statistically insignificant. The only exception are men who receive treatment in months 7 to 12 (stratum 3) of their unemployment spell. For these individuals, participating in short-term training increases their long-term employment probability by about 5 percentage points. However, this effect is only marginally statistically significant and does not seem to last in the long run.

— Figure 4 about here —

Results for the more substantive classroom further training measures (CFT) are given in figures 4 and 5. 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 25 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 to shorter measures within the group of CFT programs. Second, it is possible that such individuals drop out of the program more often or earlier. A third reason may be that a large number of those just having become unemployed easily find new jobs if they do not take part in a training program. If these individuals are assigned to CFT measures anyway, they will be 'locked-in', while many of their counterparts in the control group have already found employment. This would imply that some of the short-term unemployed receive training even though they do not need it to overcome unemployment. In addition, there may 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.

— Figure 5 about here —

While there is little evidence for statistically significant employment effects for West German men starting classroom further training in months 0 to 6 of their unemployment spell (strata 1 and 2) or West German women starting it 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 some 8 percentage points for men and to some 10 percentage points for women. The corresponding results for classroom further training measures in East Germany are given in figure 5. As in West Germany, there are long and deep lock-in effects of up to twenty five percentage points in the first 12 to 15 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 initial lock-in.

— Figure 6 about here —

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 6. The results for West Germany shown in the first row of figure 6 suggest considerable positive employment

effects of about 10 percentage points for women after a lock-in period of up to 8 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 themselves into different occupations. If women more often participate in training for occupations in the service sector, where employment chances are generally better than in manufacturing or construction jobs, this will lead to more positive effects of practical training measures for women.[18]

Similar as for other types of training, there are no employment effects for participants in PFT in East Germany (second row of figure 6). The negative picture our results draw 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 may be very limited. In addition to this, it is likely 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 may to a certain extent be used to reduce the frustration of those who want to work, but have no employment prospects. In fact, differences in selection into treatments may induce differences in treatment effects, if treatment effects are heterogeneous.

— Figure 7 about here —

Finally, figure 7 shows estimated treatment effects for the very long retraining measures. Although the large majority of these programs do not last longer than two years (see figure 16), no statistically positive employment effects can be observed up to thirty months after program start. On the contrary, retraining measures cause a grave lock-in effect of minus forty percentage points during most of the program's duration. Just in order to break even, employment gains after completion of the program have to be very large and long-enduring given the large loss in employment probability caused by the participation in the program (see discussion of cumulated effects below).

6.2 Pairwise Evaluation of Training Programs

Given that in many cases, especially in West Germany, training programs may have considerable employment effects when compared to attending no program, the ques-

[18] See Lechner et al. (2005b) or Fitzenberger and Völter (2007) who consider gender specific target professions for public sector sponsored training in East Germany in the 1990s.

tion arises which of the different training programs is the most effective for a given subpopulation. The results presented in the previous section already suggest that short-term training may have similar positive effects as classroom further training or practical further training when each of the programs is compared to attending no program. However, this does not necessarily mean that participants in short-term training could not have improved their employment chances by attending classroom or practical further training instead, or that participants in classroom or practical further training would not have lost from taking part in short-term training instead. This is the question we address next.

— Figures 8 and 9 about here —

Figures 8 and 9 show that participants in short-term training generally would *not* have improved their employment chances by attending class room further training, neither in East nor in West Germany.[19] On the contrary, the much shorter and less pronounced lock-in effect of short-term measures makes this form of training seem more effective than the longer-term classroom training. This applies especially to East German participants in the later strata who would have significantly lowered their employment probability even in the long-run if they had attended classroom further training instead of short-term training. Furthermore, figures 10 and 11 show that participants in classroom further training would *not* have lost from attending short-term training instead. This is remarkable since it means that these individuals could have been assigned to the much less expensive short-term measures without reducing their employment chances. Again, taking the shorter lock-in effect of short-term measures at program start into account, short-term training would have even been preferable for these individuals. Taken together, classroom further training is on balance not more effective than short-term training.

— Figures 10 and 11 about here —

How does practical further training compare to short-term training? Figure 12 shows that individuals who took part in short-term training would not have gained from attending practical training instead. However, figure 13 indicates that practical further training was significantly *more effective* for West Germans taking part in this kind of training than short-term training would have been. This means that

[19]The evidence is not so clear for West German women who started the program after having been unemployed for more than 4 months.

it would not necessarily have been advisable for these individuals to substitute the longer practical training courses by the shorter programs. It also means that, to a certain extent, participants were well allocated to courses, as neither individuals in short-term training nor individuals in practical further training would have gained by reallocating them to the other program. As the last row of figure 12 shows, this does not necessarily apply to East German participants in practical training whose employment chances would not have been significantly reduced if they had been reallocated to short-term training.

— Figures 12 and 13 about here —

How do the practical training courses compare to the more theoretical classroom further training courses? Evidence on this comparison is given in figures 14 and 15. Figure 14 suggests that practical training was better for West German participants of practical training than classroom training would have been. This holds especially for female participants in West Germany whose employment probability was significantly higher than it would have been in classroom training even long after the programs ended. Note that in West Germany practical training programs also exhibited significantly smaller lock-in effects which is not surprising given their shorter length. The lower row of figure 15 shows that East German participants in practical training would neither have gained nor would they have lost from taking part in classroom further training instead.

— Figures 14 and 15 about here —

We omit comparisons with the long retraining programs as these comparisons are entirely determined by the extensive lock-in effect of retraining (see figure 7). The general conclusion is that shorter programs with positive employment effects outperform longer programs due to the difference in the length of the lock-in period.

6.3 Cumulated Effects

The higher effectiveness of shorter programs is confirmed in table 2 which shows gains and losses in months employed for all pairwise comparisons cumulated over two years after program start. For example, for West German men participating in short-term training after having been unemployed for at least 7 months (stratum 3) the net gain

of the program versus waiting is 0.903 employment months during the first 24 months after the program start. The effects are even stronger for West German women who gain on average 1.767 (stratum 2) and 2.122 (stratum 3) employment months. In East Germany, only long-term unemployed men gain from taking part in short-term training (plus 0.947 employment months, stratum 3). The results suggest that short-term training is the only form of training that has positive and statistically significant cumulated employment effects during the first two years, and this only for the individuals who are not treated too early after entering unemployment.

— Table 2 about here —

Table 2 also nicely summarizes the comparative effectiveness of the different programs. Rows 2, 3, 16 and 17 show that short-term training was better in terms of cumulated employment months for those participating in it than classroom further training or practical further training would have been (STT vs. CFT). On the other hand, rows 6 and 20 (CFT vs. STT) suggest that for participants in classroom further training, short-term training would have been more effective. However, row 10 (PFT vs. STT) indicates that short-term training is not uniformly better than the longer further training programs as West German participants in *practical* further training would have lost if they had been assigned to short-term training measures instead.[20] On the one hand, the practical training programs stand out as quite effective, as e.g. participants of these courses fared significantly better than if they had taken part in classroom further training courses (rows 11 and 25). On the other hand, participants of classroom further training would not necessarily have gained from switching to more practical courses (rows 7 and 21).

6.4 Effect Heterogeneity

The results presented so far reveal considerable heterogeneity in the effects of the programs in different subpopulations. A consistent finding for West Germany seems to be that programs are only effective for those individuals who start a program after having been unemployed for some time. In East Germany this seems to be reversed (see first two graphs in figure 5), suggesting that selection of subpopulations into treatment may differ between East and West. Another finding is that training effects are generally larger for women than for men (see e.g. figures 2 and 4).

[20] Of course, practical training courses are much more costly so that short-term training may still have been the better alternative.

Women also seem to benefit from practical training, while men do not (figure 6). These results show that treatment effects averaged over too broad subpopulations or heterogenous programs may hide statistically and economically significant effects for particular subprograms or subpopulations.

In order to investigate effect heterogeneity also *within* the subgroups defined above, we regressed cumulated individual treatment effects on a number of observed personal characteristics, see equation (4). Concretely, we considered individual treatment effects (i.e. the difference between the actual and the nonparametrically predicted counterfactual outcome) cumulated over the period after lock-in (STT: after month 8, PFT: after month 12, and CFT: after month 18) and divided by the number of months after lock-in. We are interested in how the treatment effects vary with personal characteristics. We also account for the mismatch of the treated individual and the control group with respect to particular individual characteristics. This allows us to compute 'mismatch corrected' treatment effects (by omitting the mismatch when calculating average treatment effects based on the regression of individual treatment effects on personal characteristics). If our matching approach works well, the mismatch-corrected average treatment effects have to coincide with uncorrected average treatment effects.

The results are given in tables 8 to 10 in the appendix. Generally, we only report specifications with significant covariates (we omit some covariates because of likely multicollinearity problems). Therefore, a case with no coefficients reflects a situation where no significant covariates was be found. We find some evidence for effect heterogeneity as in many cases older participants benefited less or not at all from training programs. This holds specifically for East and West German participants in short-term training (see table 8), West German women and East German men and women participating in classroom further training (table 9) and West German women participating in practical further training (table 10). In some cases, treatment effects also vary with educational qualifications, especially in the case of West German men taking part in short-term training (top panel of table 8) and West German women taking part in classroom further training (middle panel of table 9). In both cases low educational qualifications harm the effect of training on employment outcomes. We also note that in all cases, the mismatch-corrected estimates of average treatment effects coincide with the uncorrected estimates which supports the validity of our matching procedure.

We also investigated effect heterogeneity in pairwise program comparisons, but the

results were less clear-cut.[21] The general finding of both the comparisons of program vs. non-participation and the cross-program comparisons seems to be that – in addition to the heterogeneity found between the subgroups defined by gender, region, and unemployment duration – there is generally little heterogeneity along observed characteristics. However, when there is such heterogeneity, it can be very strong. This suggests that a better targeting of the programs could be achieved along these lines.

7 Conclusions

This paper analyzes and compares the employment effects of the four important types of public sector sponsored training in Germany in the early 2000s. These are short-term training (STT), classroom further training (CFT), practical further training (PFT), and retraining (RT). In light of recent policy reforms fostering shorter training programs, we are particularly interested in the question of how short-term training programs compare in terms of effectiveness to traditional medium-term further training schemes. Our econometric approach uses nonparametric kernel matching methods in a dynamic, multiple treatment framework.

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 employment effects for male and female participants in short-term training and classroom further training who started their training not too early during their unemployment spell. Moreover, West German women but not West German men benefited from practical further training measures. A closer look reveals that, within the time window permitted by our data set, employment effects of short-term training were of a similar magnitude as those of traditional medium-term measures, but, due to the shorter length, these positive effects materialized much earlier. According to our results, West German men taking part in short-term training 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.

The surprising effectiveness of short-term training when compared to the different forms of medium-term training is also confirmed in pairwise comparisons. According

[21] These results are available upon request.

to our results, participants in short-term training would in general not have gained if they had taken part in medium-term further training, and participants in the latter programs would generally not have lost if they had been assigned to short-term courses instead. In particular, this holds for the comparison of short-term training and classroom further training. However, this is less clear in comparison to practically oriented further training, where it appeared that participants of practical further training courses may have reduced their employment chances if they had taken part in short-term training instead. As to the comparison of classroom vs. practical further training, our results suggest that practical training may have advantages over pure classroom training, a finding that is consistent with the international evidence (see Martin and Grubb (2001) and OECD (2005)), but not with evidence for Germany in the 1990's (Lechner et al. (2005a), Fitzenberger et al. (2006a), Fitzenberger and Völter (2007)).

We do not find any positive employment effects for the long retraining measures during the time window permitted by our data. Even more than six months after the completion of such a program, the employment rate of participants is below than or equal to the employment rate of a comparable control group of non-participants. Given the strong lock-in effect of these programs – during participation participants have employment rates that may be 40 percentage points lower than those of non-participants – it seems extremely unlikely that these measures justify their large costs. Long-run evidence on retraining during the 1990's in Lechner et al. (2005a,b), Fitzenberger et al. (2006a), and Fitzenberger and Völter (2007) suggests that, although there may be positive long-run effects, cumulated employment effects are always lower than those of shorter programs.

The general ineffectiveness of long-term retraining and the only moderate effectiveness of medium-term training when compared to short-term training suggests that the policy shift that took place in 2002 and that massively substituted long-term and medium-term training courses by inexpensive short-term courses was justified. In fact, given our results, it may be well the case that further substitution of long- and medium-term programs by short-term training may be warranted. Although we lack detailed information on training costs, the results in table 2 combined with the information on average costs in table 1 suggest that an average short-term training course costing only 627 Euros may lead to an employment gain of one month within twenty four months after program start, while an average further training or retraining course costing 6175 Euros leads to *no gain* or even a loss in months employed within the first twenty four months after program start.

While it is a common finding shared by studies such as Lechner et al. (2005a,b), Fitzenberger et al. (2006a) and Hujer et al. (2004) that, especially if lock-in effects are taken into account, shorter training programs may outperform longer ones,[22] it seems surprising that short-term programs lasting only two to twelve weeks may have employment effects at all. Given that it is hard to believe that such short programs lead to substantial increases in human capital, other aspects may be more relevant. Looking at the particular contents of short-term training analyzed here, it seems more plausible that these programs help to activate their participants who may otherwise not look as intensively for new jobs as they do when they are assigned short-term training programs that often comprise elements of profiling, job search assistance, or monitoring. Our results are therefore in line with a number of recent studies that focus on the positive effects of increased job search assistance and activation, see e.g. Blundell et al. (2004), Weber and Hofer (2004), Fougère et al. (2005), Hujer et al. (2005), Crépon et al. (2005), and Van den Berg and Van der Klaauw (2006).

Furthermore, our results show that the effects of training programs may be very different across different subgroups. One result is that employment effects are usually larger for individuals who start their program at a later point during their unemployment spell. In fact, in many cases we do not find significant employment 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 may differ in observed and unobserved characteristics from those who are short-term unemployed. However, the result is remarkable because it suggests that in cases of long-term unemployment, training programs may help to restore the employment chances of their participants. We also find that training effects may be heterogenous with respect to gender, age, and qualification. In line with other results in the literature (see Bergemann and van den Berg (2006) for an overview) we find that employment effects of training are generally larger for women than for men. Moreover, it seems that older individuals benefit less from the training programs analyzed here. In some cases, this also applies to individuals with low educational qualifications. Based on these results, it seems advisable that the targetting in program assignment should be improved.

Finally, in contrast to the result of positive treatment effects in a number of cases for West Germany, we find only little evidence for positive treatment effects in East

[22]However, note that neither of these studies consider the very short training programs analyzed here.

Germany. Apart from positive effects for East German men taking part in short- and medium-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-, medium-, or long-term training 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, showing that the effect of training may strongly depend on the specific circumstances of the labor market under consideration. The ineffectiveness of training in East Germany in the early 2000's is in line with results by Lechner and Wunsch (2006), although the latter paper takes a different methodological approach. However, the results are in contrast to the somewhat more positive findings for the 1990's (Lechner et al. (2005b), Fitzenberger and Speckesser (2007), Fitzenberger and Völter (2007)).

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.

Bergemann, A. and G.J. van den Berg (2006). "Active Labor Market Policy Effects for Women in Europe: A Survey". IZA Discussion Paper No. 2365.

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.

Blundell, R., M. Costas Dias, C. Meghir, J. Van Reenen (2004). "Evaluating the Employment Impact of a Mandatory Job Search Program". *Journal of the European Economic Association* 2, 569-606.

Bundesagentur für Arbeit (2001, 2002a, 2003a, 2004a). *Daten zu den Eingliederungsbilanzen 2000, 2001, 2002, 2003*, Nürnberg (various issues).

Bundesagentur für Arbeit (2002a, 2003a, 2004a, 2005a). *Arbeitsmarkt 2001, 2002, 2003, 2004*, Nürnberg (various issues).

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

Crépon, B., M. Dejemeppe and M. Gurgand (2005). "Conseling the unemployed: does it lower unemployment duration and recurrence?". IZA Discussion Paper 1796.

Dolton, P. and D. O'Neill (2002). "The Long-Run Effects of Unemployment Monitoring and Work-Search Programs: Experimental Evidence from the United Kingdom". *Journal of Labor Economics* 20, 381-403.

Dyke, A., C.J. Heinrich, P.R. Mueser, K.R. Troske, and K.-S. Jeon (2006): "The Effects of Welfare-to-Work Program Activities on Labor Market Outcomes", *Journal of Labor Economics* 24, 567-607.

Fitzenberger, B., A. Osikominu and R. Völter (2006a). "Get Training or Wait? Long Run Employment Effects of Training Programs for the Unemployed in Germany." ZEW Discussionpaper No. 06-039.

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

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

Fitzenberger, B., R. Völter (2007): "Long Run Effects of Training Programs for the Unemployed in East Germany." Forthcoming *Labour Economics*.

Fougère, D., J. Pradel and M. Roger (2005). "Does Job Search Assistance Affect Search Effort and Outcomes?". IZA Discussion Paper 1825.

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.

Hardoy, I. (2005): "Impact of Multiple Labour Market Programmes on Multiple Outcomes: The Case of Norwegian Youth Programmes". *Labour* 19, 425-467.

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*.

Hujer, R., S.L. Thomsen, C. Zeiss (2004). "The Effects of Vocational Training Programmes on the Duration of Unemployment in Germany". IZA Discussion Paper No. 1117.

Hujer, R., M. Caliendo and S. Thomsen (2005). "Mikroökonometrische Evaluation des Integrationserfolges", *Beiträge zur Arbeitsmarkt- und Berufsforschung (BeitrAB)* 291, Nürnberg.

Hujer, R., S.L. Thomsen and C. Zeiss (2006).“The Effects of Short-Term Training Measures on the Individual Unemployment Duration in West Germany”. ZEW Discussion Paper No. 06-065.

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.

Jacobi, L. and J. Kluve (2006). “Before and After the Hartz Reforms: The Performance of Active Labour Market Policy in Germany”. IZA Discussion Paper No. 2100.

Kluve, J. (2006): “The Effectiveness of European Active Labor Market Policy”. IZA Discussion Paper No. 2018.

Kluve, J. and C.M. Schmidt (2002): “Can training and employment subsidies combat European unemployment?”. *Economic Policy* 35, 409-448.

Kurtz, B. (2003). “Trainingsmaßnahmen - Was verbirgt sich dahinter?”. IAB Werkstattbericht Nr. 8, Nürnberg.

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. Pfeiffer (eds.) (2000), *Econometric Evaluation of Active Labor Market Politics in Europe*, Heidelberg: Physica-Verlag.

Lechner, M. (2002). “Program heterogeneity and propensity score matching: An Application to the evaluation of active labor market policies”. *Review of Economics and Statistics* 84, 205-220.

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.

Lechner, M. and C. Wunsch (2006). "Active Labour Market Policy in East Germany: Waiting for the Economy to Take Off". IZA Discussion Paper No. 2363.

Martin, J.P. (2000). "What works among active labour market policies: Evidence from OECD countries' experiences". *OECD Economic Studies* No. 30, 79-113.

Martin, J.P. and D. Grubb (2001). "What works for whom: A review of OECD countries' experiences with active labour market policies". *Swedish Economic Policy Review* 8, 96.

OECD (2005). "Labour Market Programmes and Activation Strategies: Evaluating the Impacts", *OECD Employment Outlook*, Ch. 4.

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.

Schneider, H., K. Brenke, D. Hess, L. Kaiser, J. Steinwedel, A. Uhlendorff (2006). "Evaluation der Maßnahmen zur Umsetzung der Vorschläge der Hartz-Kommission - Modul 1b: Förderung beruflicher Weiterbildung und Transferleistungen". IZA Research Report No. 7, Bonn.

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.

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

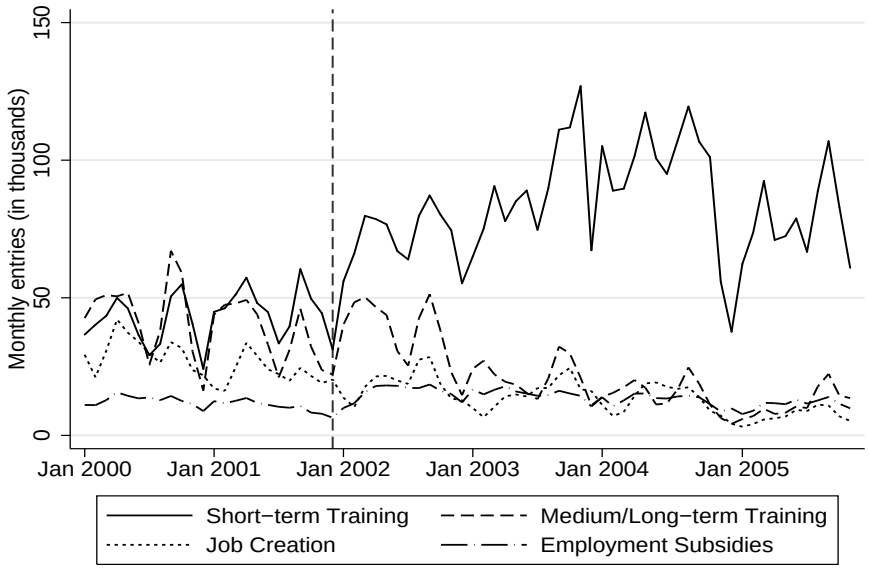
Van den Berg, G.J. and B. van der Klaauw (2006). "Counseling and monitoring of unemployed workers: theory and evidence from a controlled social experiment". Forthcoming in: *International Economic Review*

Waller, M. (2007). "Do Reported End Dates of Treatments Matter for Evaluation Results? – An Investigation Based on the German Integrated Employment Biographies Sample", FDZ Methodenreport No. 01-07, Institut für Arbeits- und Berufsforschung (IAB), Nürnberg.

Weber, A. and H. Hofer (2004). "Employment effects of early interventions on job search programs". IZA Discussion Paper No. 1076.

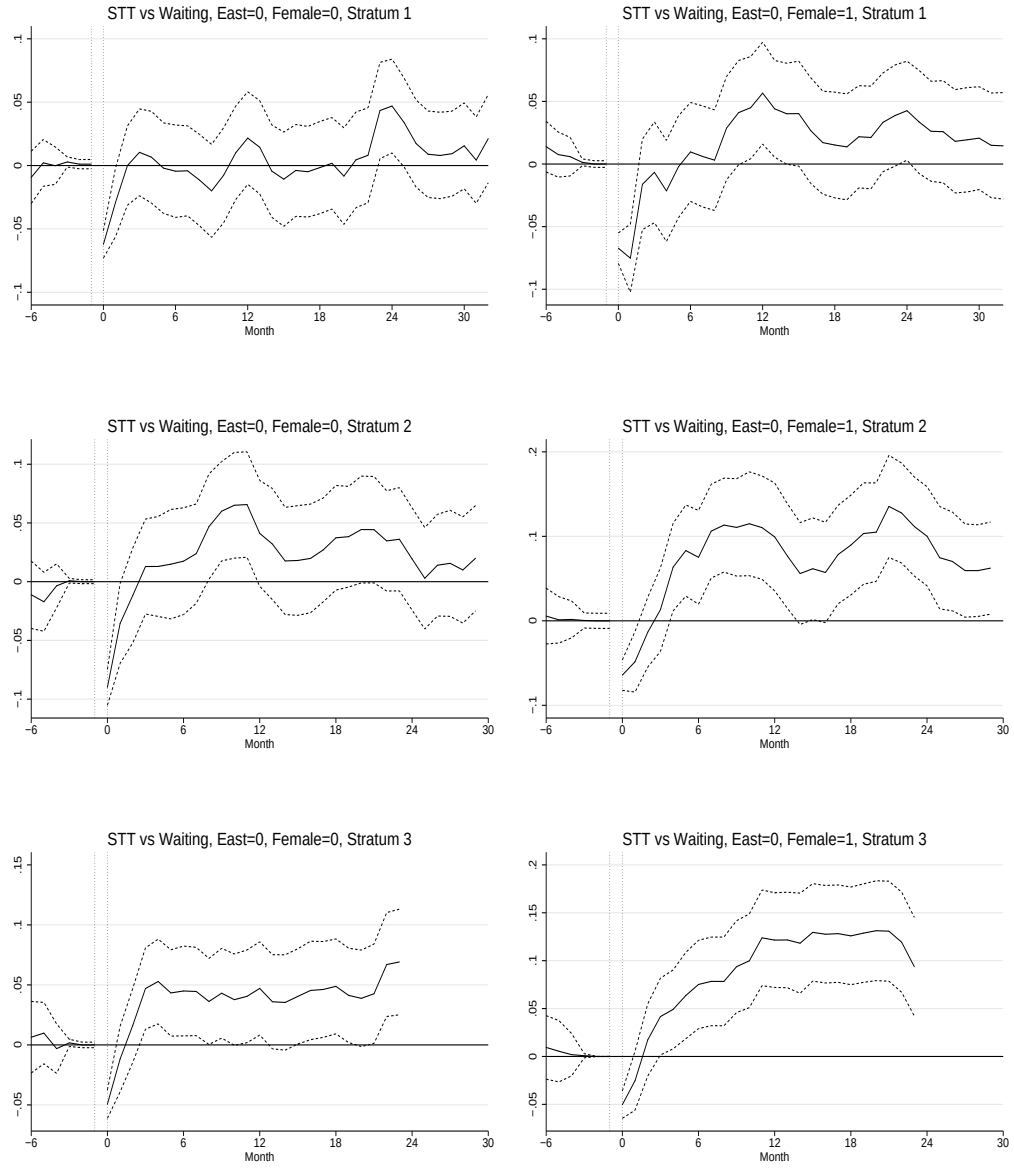
Figures

Figure 1: Active Labor Market Policies in Germany



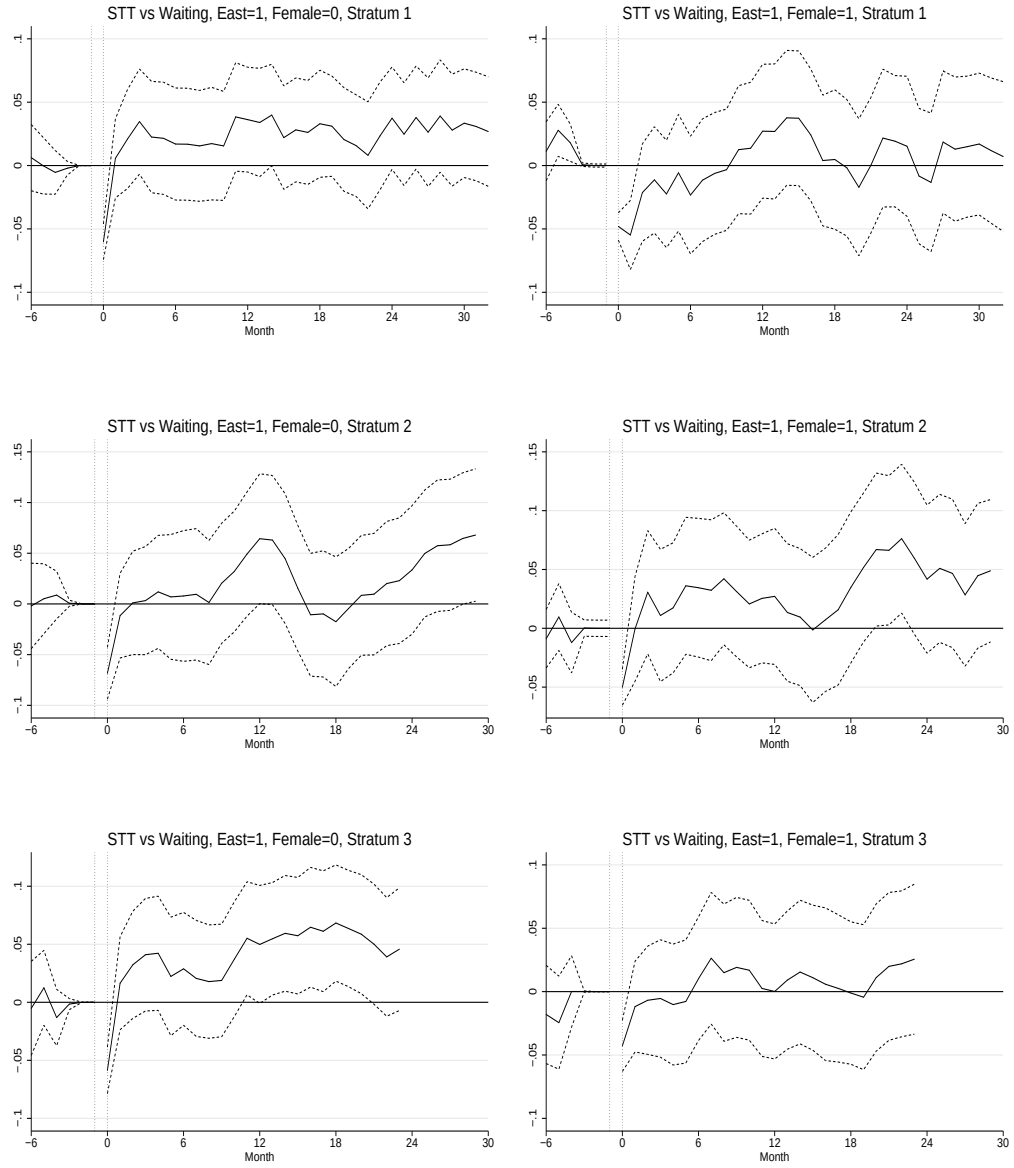
Source: Statistics of the Federal Employment Office of Germany.

Figure 2: Treatment Effect STT vs. Waiting, West Germany



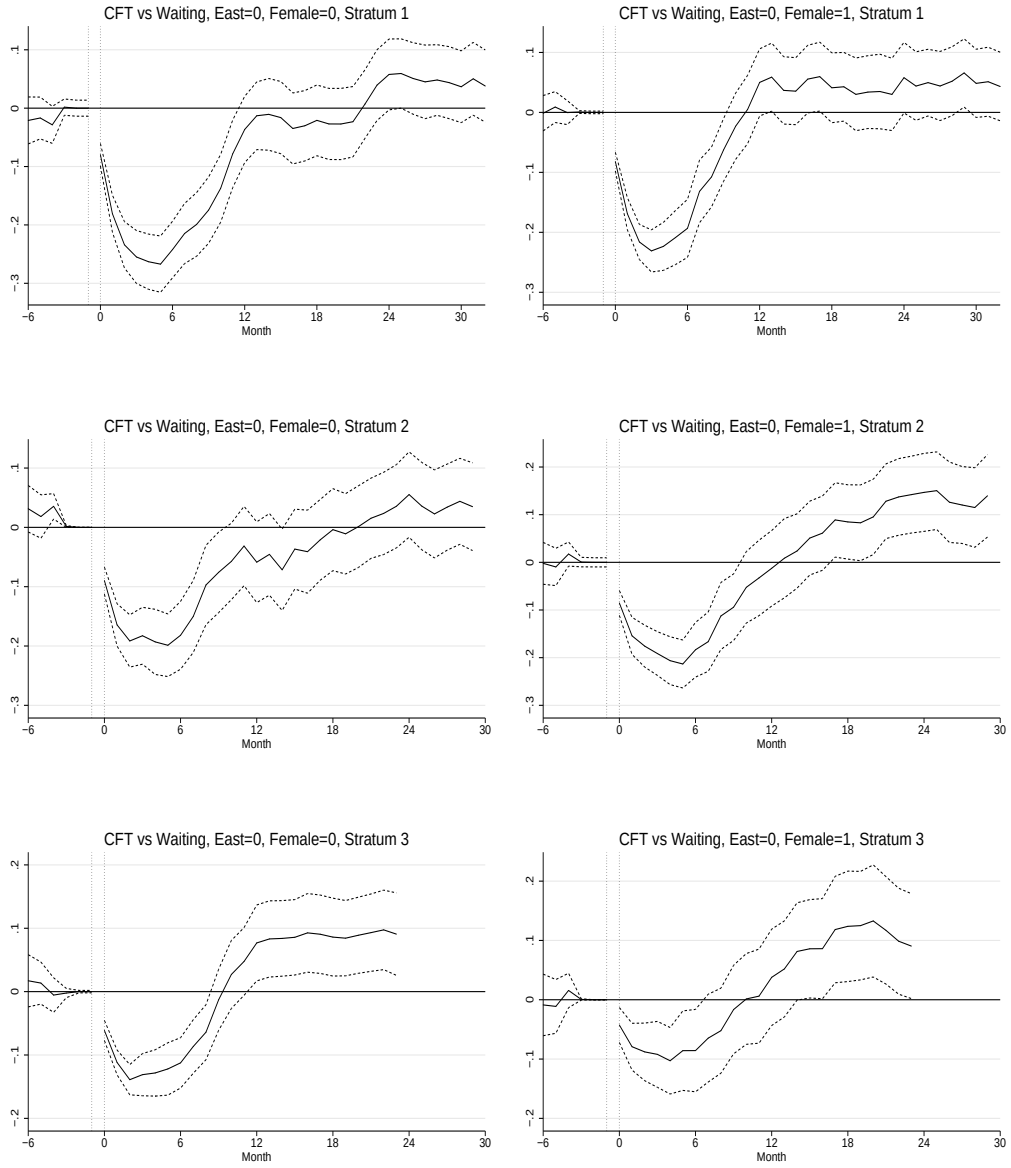
Difference in employment rates is measured on the ordinate, pre-unemployment (< 0) and post-treatment (≥ 0) months on the abscissa. Stratum 1 denotes entry into program during months 0 to 3 of unemployment, stratum 2 during months 4 to 6, and stratum 3 during months 7 to 12.

Figure 3: Treatment Effect STT vs. Waiting, East Germany



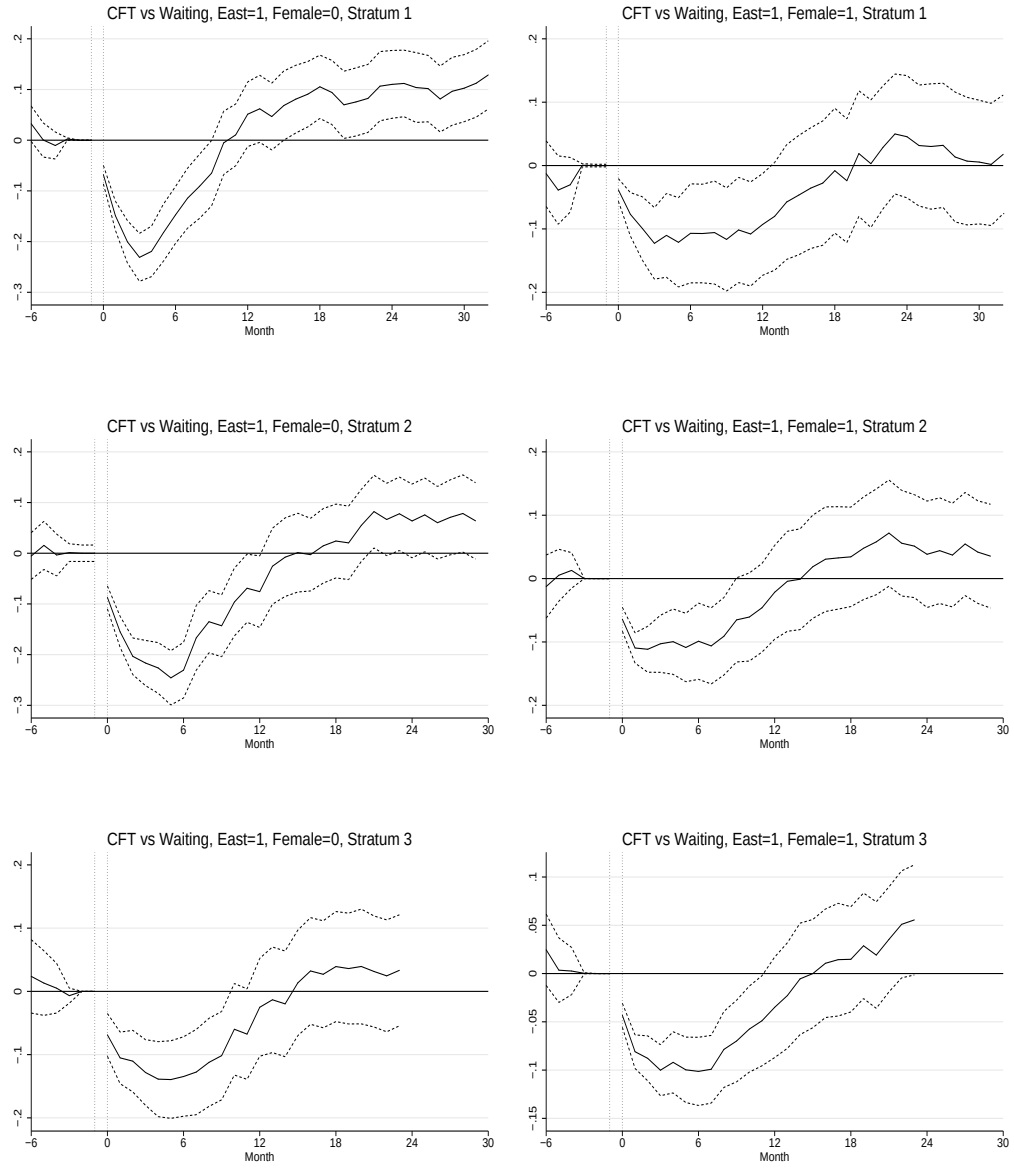
Difference in employment rates is measured on the ordinate, pre-unemployment (< 0) and post-treatment (≥ 0) months on the abscissa. Stratum 1 denotes entry into program during months 0 to 3 of unemployment, stratum 2 during months 4 to 6, and stratum 3 during months 7 to 12.

Figure 4: Treatment Effect CFT vs. Waiting, West Germany



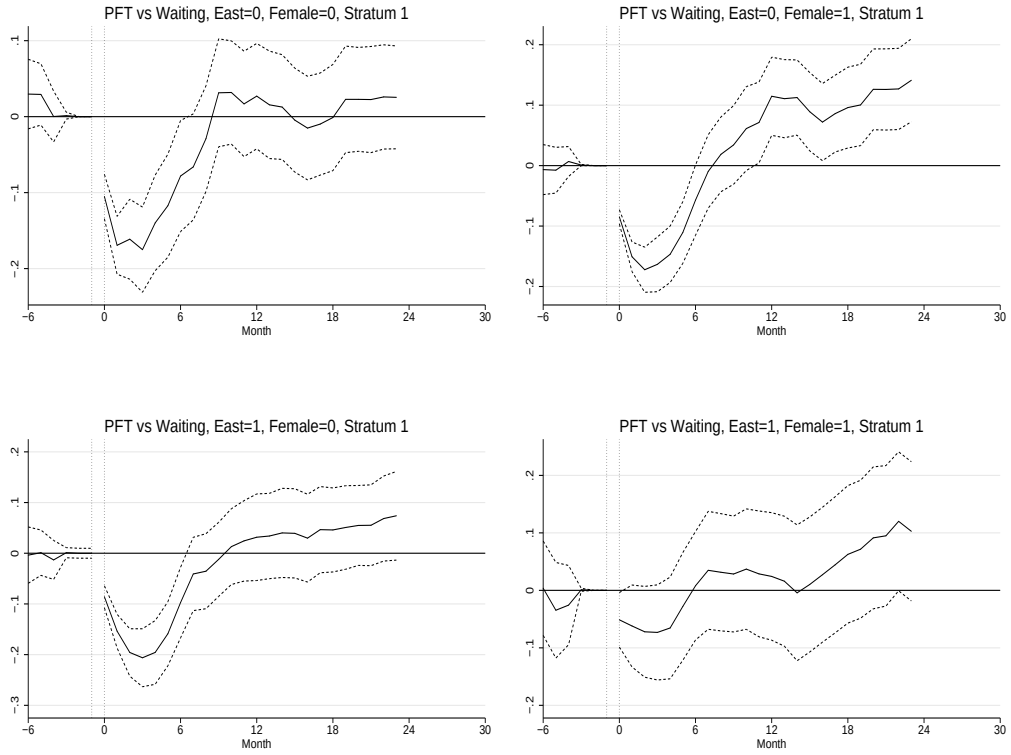
Difference in employment rates is measured on the ordinate, pre-unemployment (< 0) and post-treatment (≥ 0) months on the abscissa. Stratum 1 denotes entry into program during months 0 to 3 of unemployment, stratum 2 during months 4 to 6, and stratum 3 during months 7 to 12.

Figure 5: Treatment Effect CFT vs. Waiting, East Germany



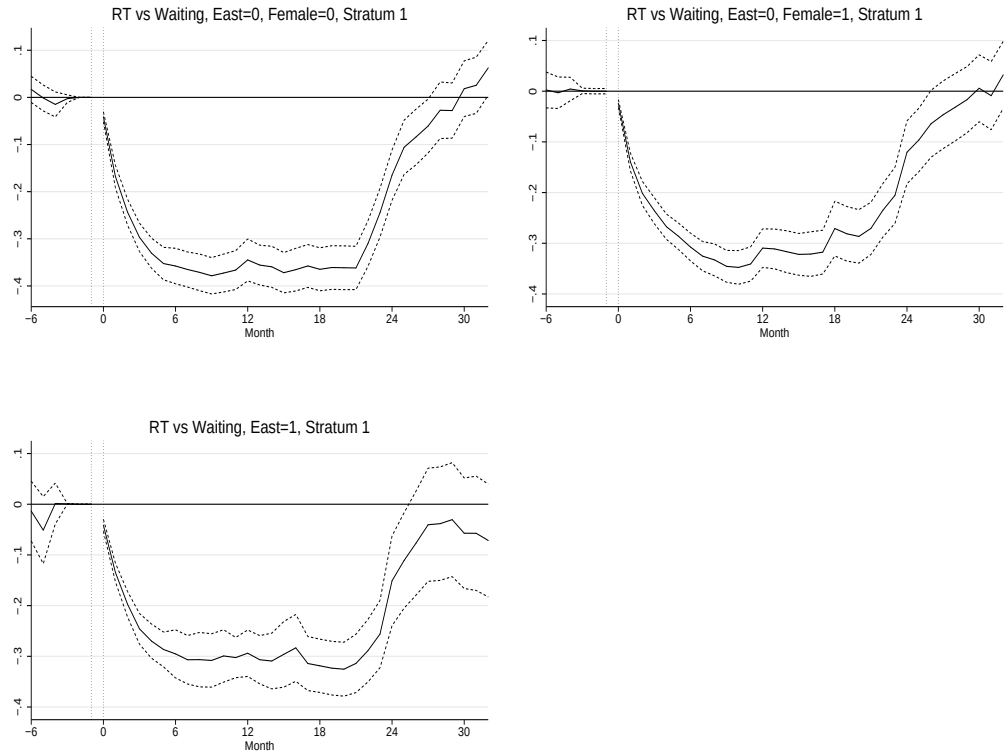
Difference in employment rates is measured on the ordinate, pre-unemployment (< 0) and post-treatment (≥ 0) months on the abscissa. Stratum 1 denotes entry into program during months 0 to 3 of unemployment, stratum 2 during months 4 to 6, and stratum 3 during months 7 to 12.

Figure 6: Treatment Effect PFT vs. Waiting, West and East Germany



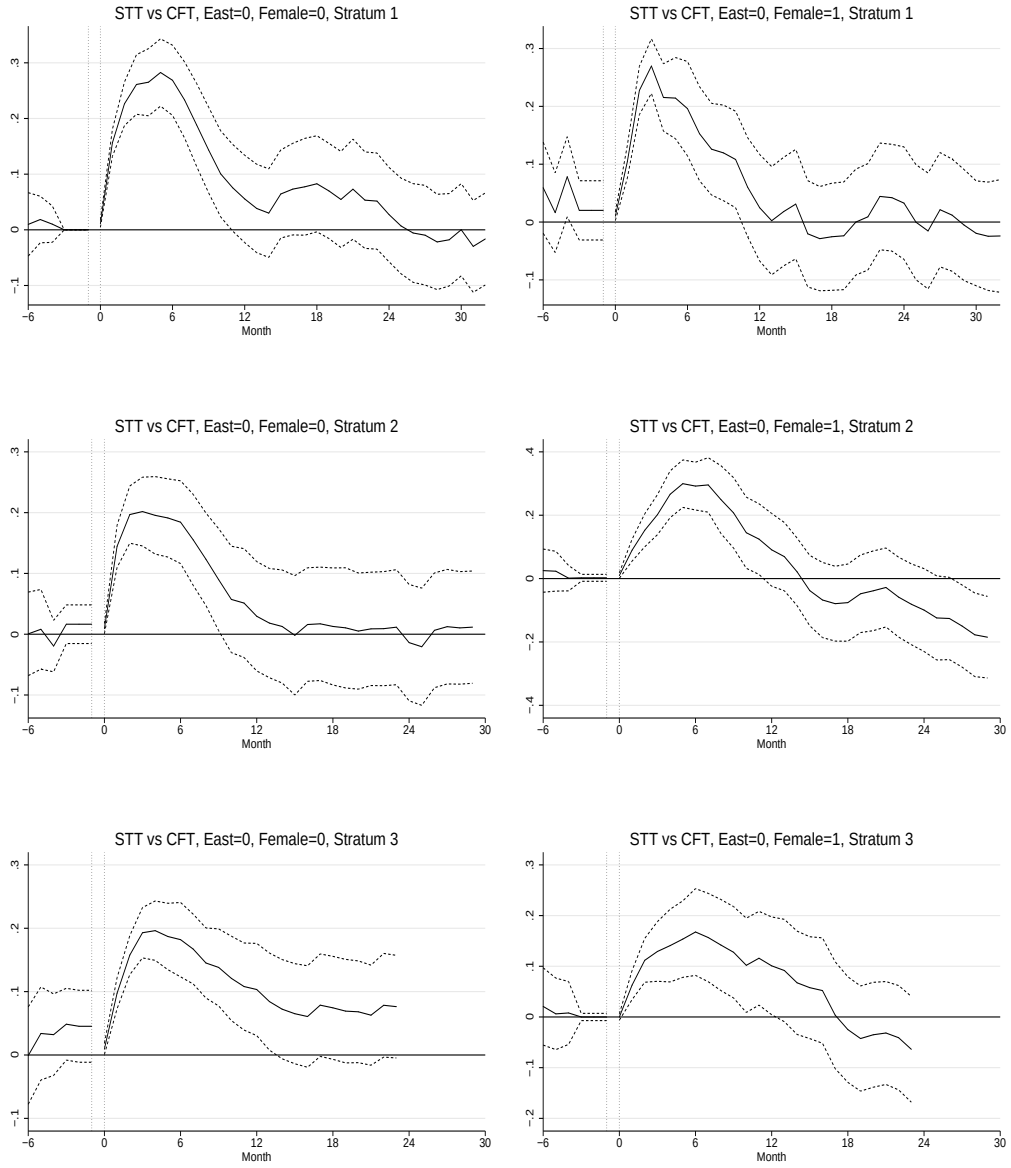
Difference in employment rates is measured on the ordinate, pre-unemployment (< 0) and post-treatment (≥ 0) months on the abscissa. Entry into program during months 0 to 12 of unemployment spell.

Figure 7: Treatment Effect RT vs. Waiting, West and East Germany



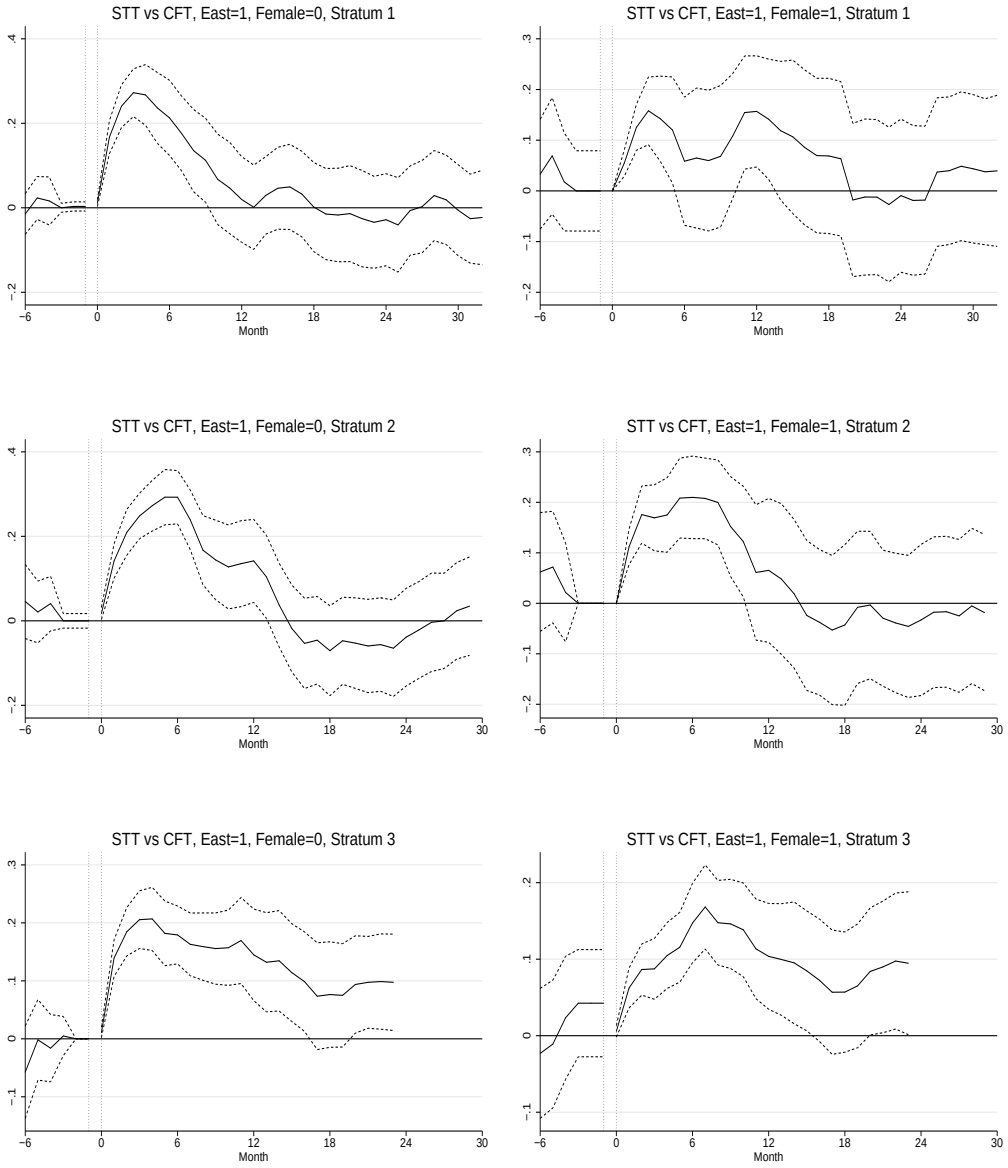
Difference in employment rates is measured on the ordinate, pre-unemployment (< 0) and post-treatment (≥ 0) months on the abscissa. Entry into program during months 0 to 12 of unemployment spell. East Germany: pooled sample containing male and female participants.

Figure 8: Treatment Effect STT vs. CFT, West Germany



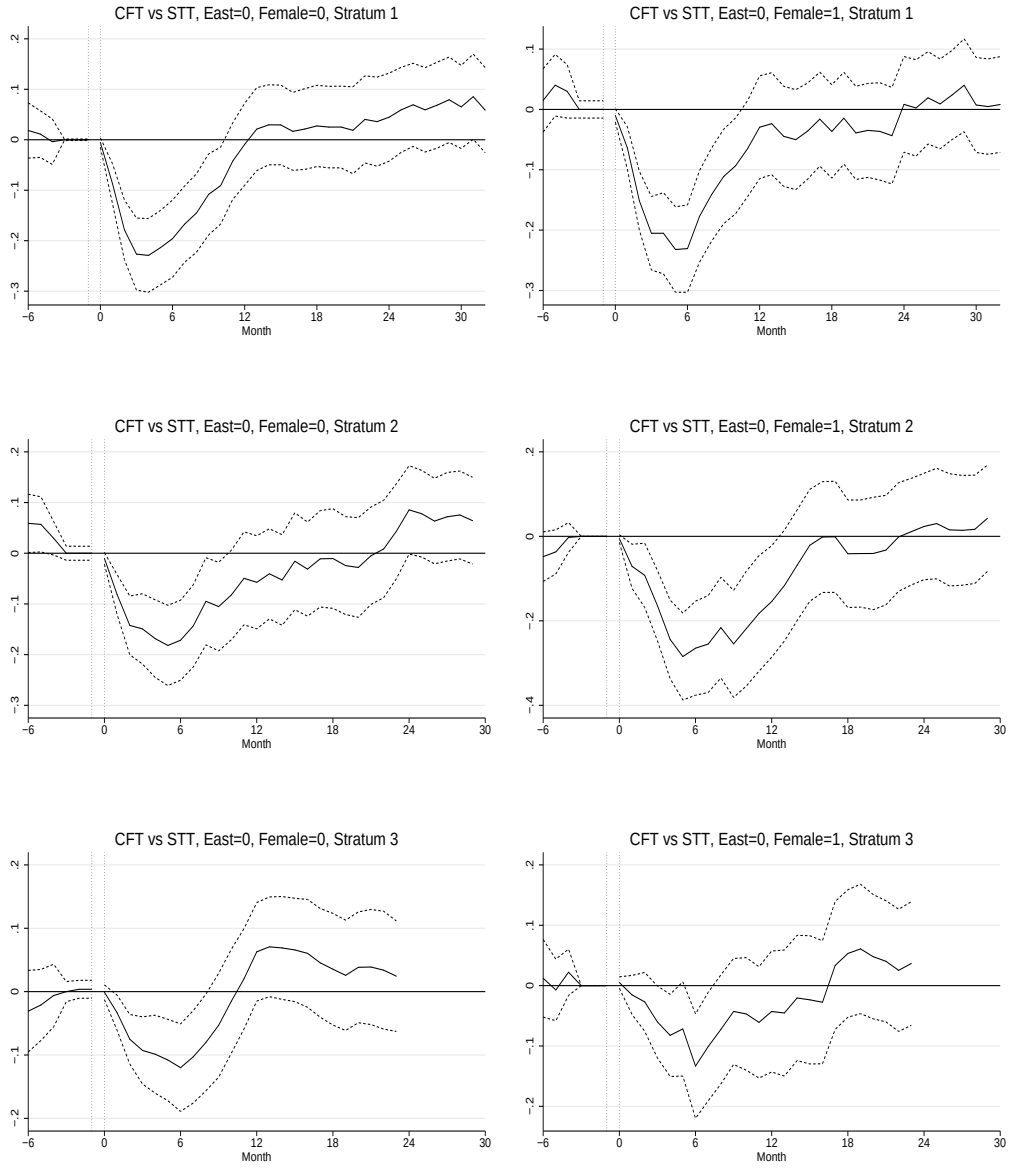
Difference in employment rates is measured on the ordinate, pre-unemployment (< 0) and post-treatment (≥ 0) months on the abscissa. Stratum 1 denotes entry into program during months 0 to 3 of unemployment, stratum 2 during months 4 to 6, and stratum 3 during months 7 to 12.

Figure 9: Treatment Effect STT vs. CFT, East Germany



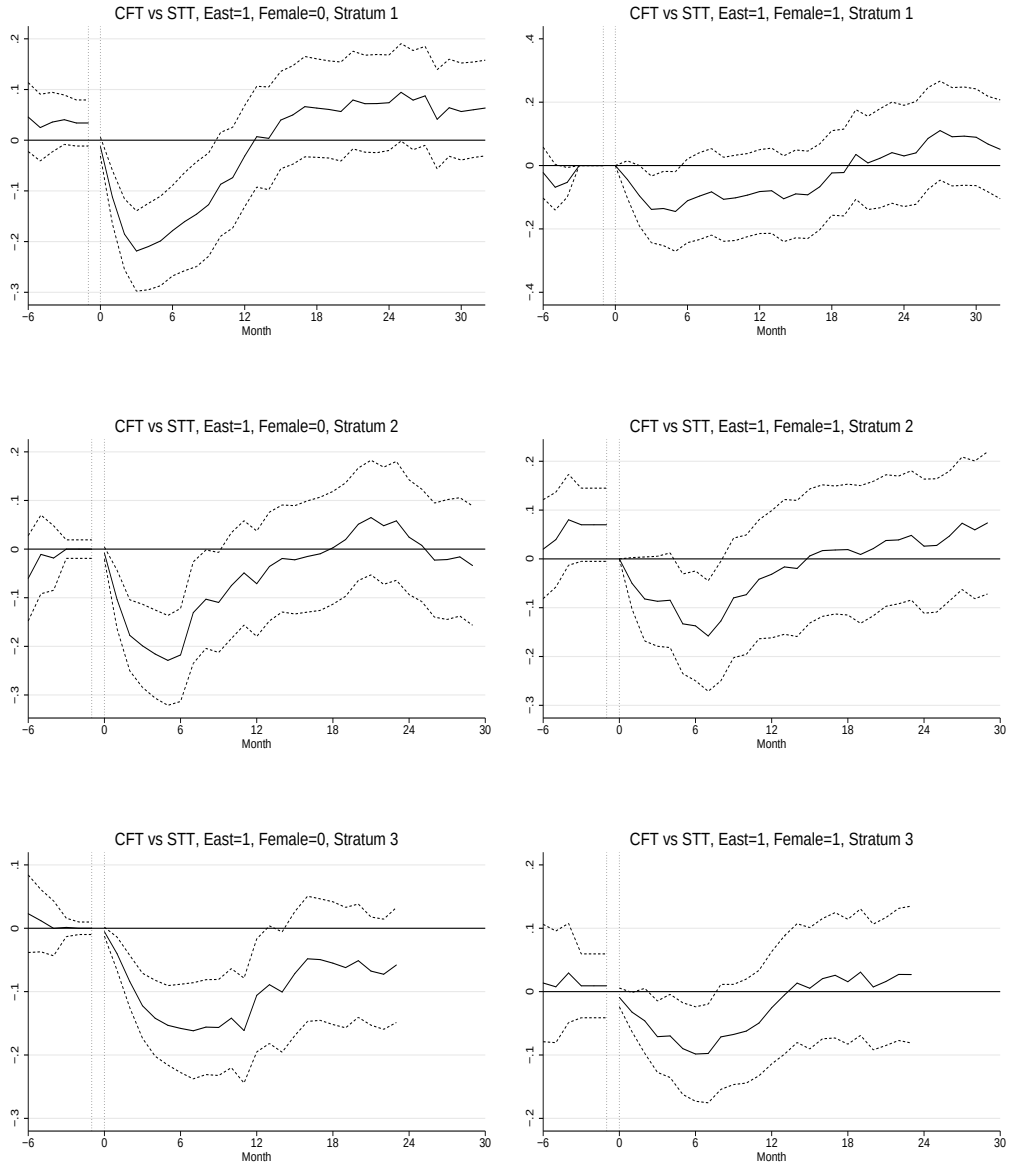
Difference in employment rates is measured on the ordinate, pre-unemployment (< 0) and post-treatment (≥ 0) months on the abscissa. Stratum 1 denotes entry into program during months 0 to 3 of unemployment, stratum 2 during months 4 to 6, and stratum 3 during months 7 to 12.

Figure 10: Treatment Effect CFT vs. STT, West Germany



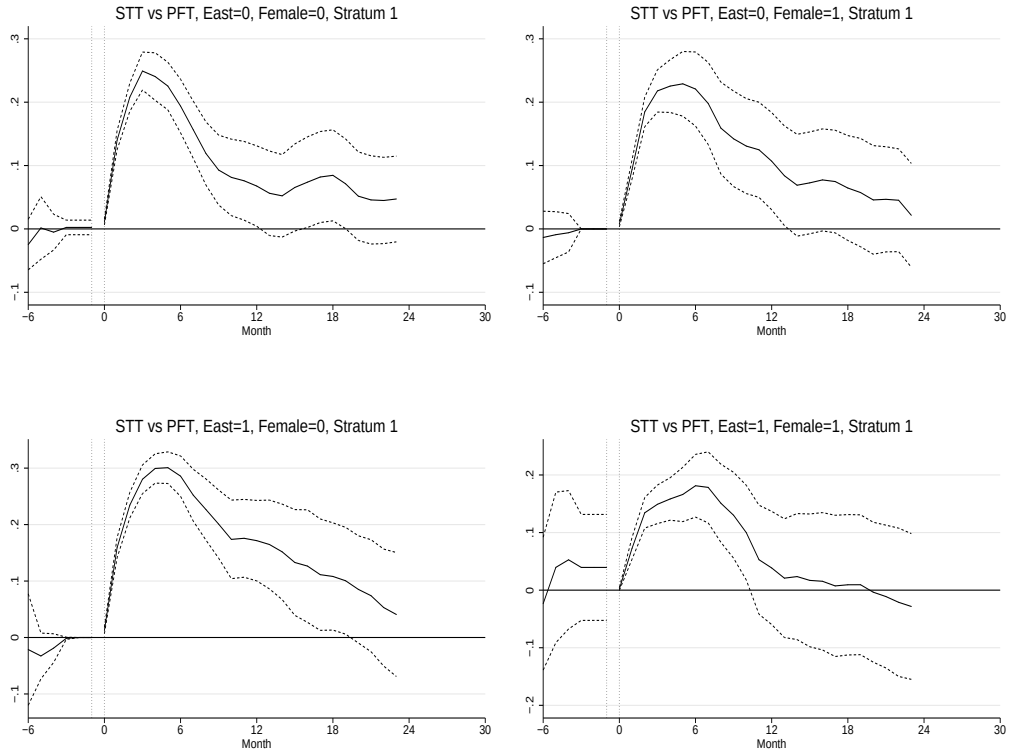
Difference in employment rates is measured on the ordinate, pre-unemployment (< 0) and post-treatment (≥ 0) months on the abscissa. Stratum 1 denotes entry into program during months 0 to 3 of unemployment, stratum 2 during months 4 to 6, and stratum 3 during months 7 to 12.

Figure 11: Treatment Effect CFT vs. STT, East Germany



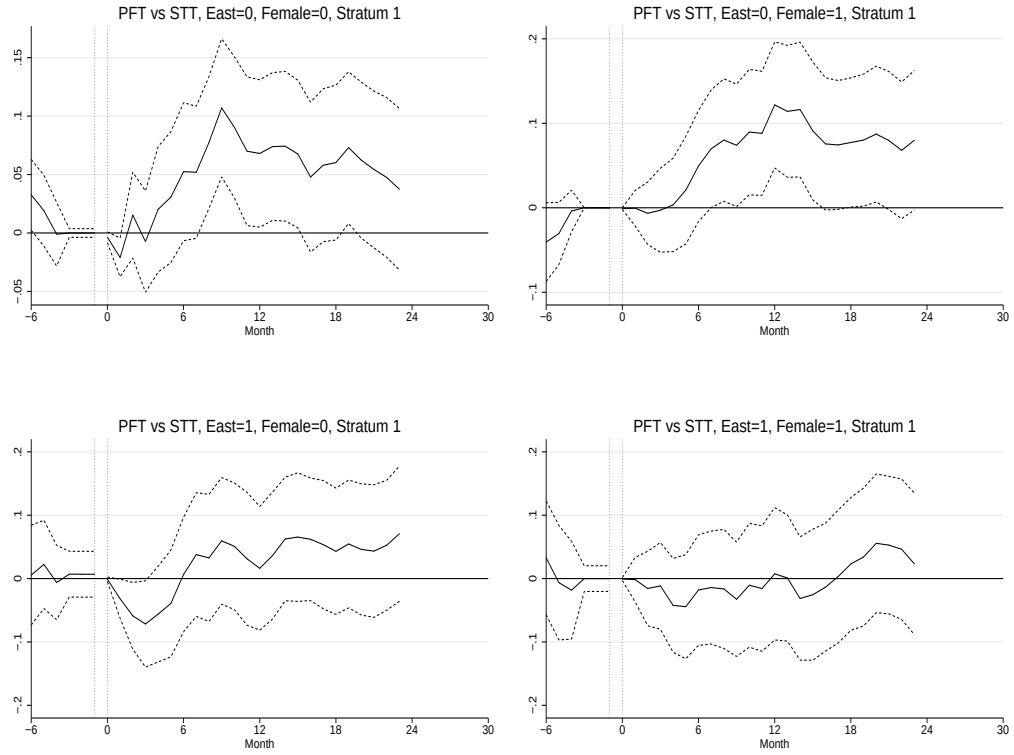
Difference in employment rates is measured on the ordinate, pre-unemployment (< 0) and post-treatment (≥ 0) months on the abscissa. Stratum 1 denotes entry into program during months 0 to 3 of unemployment, stratum 2 during months 4 to 6, and stratum 3 during months 7 to 12.

Figure 12: Treatment Effect STT vs. PFT, West and East Germany



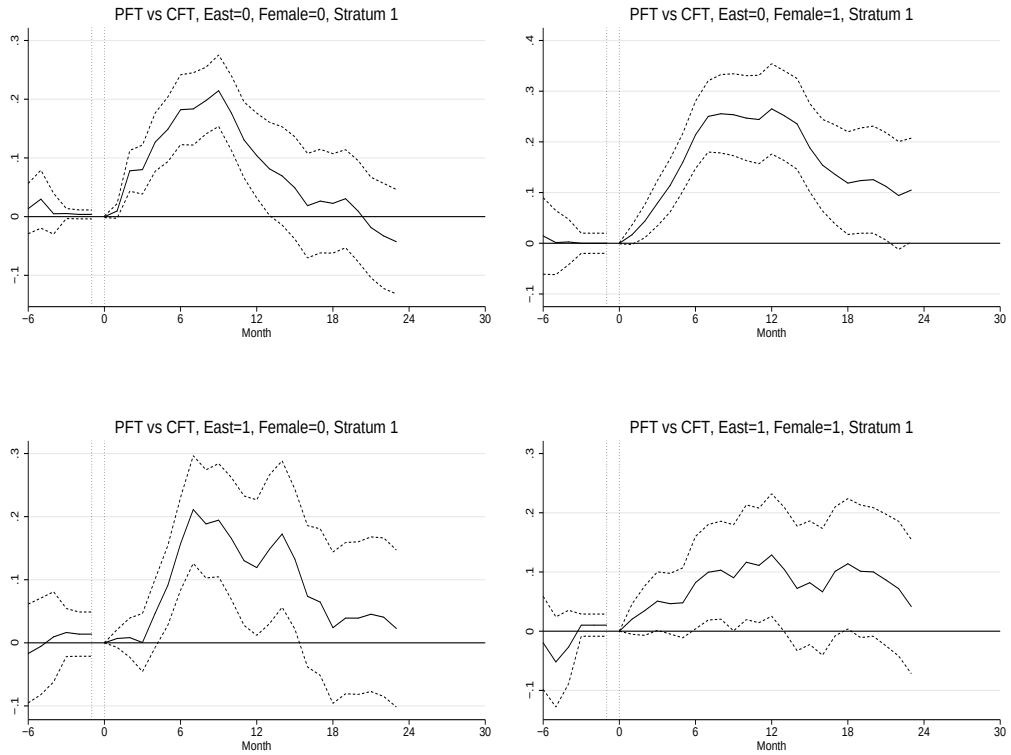
Difference in employment rates is measured on the ordinate, pre-unemployment (< 0) and post-treatment (≥ 0) months on the abscissa. Entry into program during months 0 to 12 of unemployment spell.

Figure 13: Treatment Effect PFT vs. STT, West and East Germany



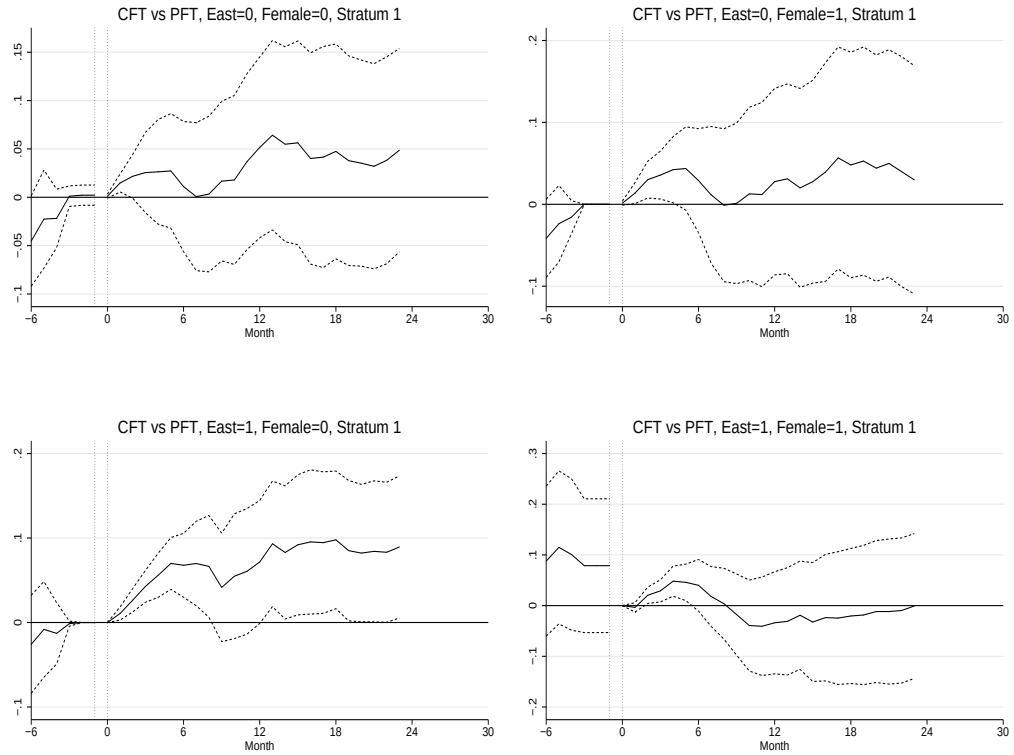
Difference in employment rates is measured on the ordinate, pre-unemployment (< 0) and post-treatment (≥ 0) months on the abscissa. Entry into program during months 0 to 12 of unemployment spell.

Figure 14: Treatment Effect PFT vs. CFT, West and East Germany



Difference in employment rates is measured on the ordinate, pre-unemployment (< 0) and post-treatment (≥ 0) months on the abscissa. Entry into program during months 0 to 12 of unemployment spell.

Figure 15: Treatment Effect CFT vs. 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. Entry into program during months 0 to 12 of unemployment spell.

Tables

Table 1: 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 2001,2002a-2004a.

Table 2: Cumulated Treatment Effects after 24 Months

	Men			Women		
	Stratum 1	Stratum 2	Stratum 3	Stratum 1	Stratum 2	Stratum 3
	West Germany					
STT vs Waiting	-0.054 (0.323)	0.573 (0.399)	0.903 (0.357)**	0.314 (0.368)	1.767 (0.532)***	2.122 (0.429)***
STT vs CFT	2.937 (0.638)***	1.750 (0.716)**	2.597 (0.573)***	1.878 (0.683)***	2.000 (0.877)**	1.545 (0.849)*
STT vs PFT	2.536 (0.518)***	.	.	2.700 (0.643)***	.	.
STT vs RT	9.173 (0.391)***	.	.	8.983 (0.429)***	.	.
CFT vs Waiting	-2.518 (0.488)***	-1.827 (0.535)***	0.159 (0.446)	-1.134 (0.446)**	-0.776 (0.608)	0.448 (0.687)
CFT vs STT	-1.407 (0.693)**	-1.604 (0.780)**	-0.189 (0.684)	-2.090 (0.665)***	-2.761 (1.069)***	-0.571 (0.821)
CFT vs PFT	0.752 (0.822)	.	.	0.699 (1.003)	.	.
CFT vs RT	7.205 (0.471)***	.	.	7.519 (0.487)***	.	.
PFT vs Waiting	-0.817 (0.588)	.	.	0.593 (0.553)	.	.
PFT vs STT	1.208 (0.520)**	.	.	1.530 (0.657)**	.	.
PFT vs CFT	1.845 (0.596)***	.	.	3.788 (0.727)***	.	.
RT vs Waiting	-7.800 (0.378)***	.	.	-6.601 (0.351)***	.	.
RT vs STT	-6.781 (0.912)***	.	.	-9.229 (1.406)***	.	.
RT vs CFT	-5.780 (0.819)***	.	.	-9.198 (1.174)***	.	.
	East Germany					
STT vs Waiting	0.484 (0.365)	0.270 (0.517)	0.947 (0.448)**	0.003 (0.453)	0.658 (0.547)	0.123 (0.508)
STT vs CFT	2.020 (0.795)**	2.104 (0.670)***	3.150 (0.613)***	1.857 (1.111)*	1.643 (1.053)	2.324 (0.594)***
STT vs PFT	3.922 (0.627)***	.	.	1.554 (0.873)*	.	.
STT vs RT	7.607 (0.705)***	.	.	.	.	.
CFT vs Waiting	-0.528 (0.497)	-1.744 (0.540)***	-1.075 (0.665)	-1.487 (0.784)*	-0.690 (0.609)	-0.793 (0.390)**
CFT vs STT	-1.170 (0.843)	-1.546 (1.027)	-2.314 (0.720)***	-1.603 (1.257)	-0.906 (1.150)	-0.607 (0.776)
CFT vs PFT	1.619 (0.579)***	.	.	-0.133 (0.875)	.	.
CFT vs RT	6.140 (0.607)***	.	.	.	.	.
PFT vs Waiting	-0.573 (0.668)	.	.	0.478 (0.996)	.	.
PFT vs STT	0.567 (0.843)	.	.	-0.047 (0.871)	.	.
PFT vs CFT	2.123 (0.822)***	.	.	1.872 (0.792)**	.	.
RT vs Waiting	-6.623 (0.446)***	.	.	-6.623 (0.446)***	.	.
RT vs STT	-7.559 (1.077)***	.	.	-7.559 (1.077)***	.	.
RT vs CFT	-4.901 (0.800)***	.	.	-4.901 (0.800)***	.	.

*** = statistically significant at 1 %, ** = at 5 %, * = at 10 %, bootstrapped standard errors

Appendix

Table 3: 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	1,064,293	1,188,369
Employment subsidies	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
Public Job Creation	314,291	246,084	219,626	193,999	170,107
Other	391,122	515,670	453,224	212,183	309,446
Total	3,364,794	3,495,856	4,062,572	4,565,010	5,953,049

Source: Bundesagentur für Arbeit, Arbeitsmarkt 2002b-2005b, own calculations.

Table 4: 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
RT	263	262	86	
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 5: 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
youngchild	1 if at least one child younger than 10 years, 0 otherwise

<continued on next page>

Table 5: Variable Definitions <continued>

Name	Definition
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 5: 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 office, 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.

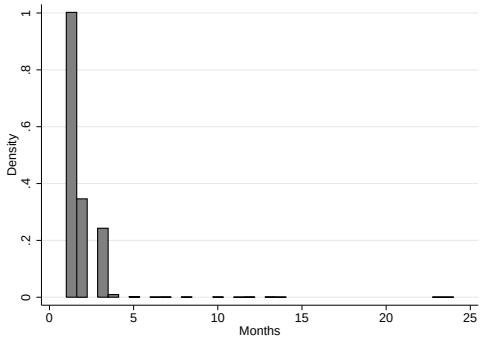
Table 6: Descriptive Statistics for Selected Variables

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

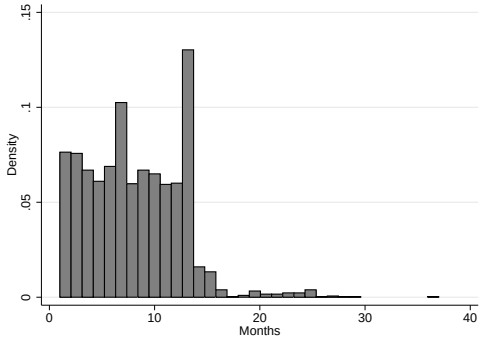
Table 7: Program Duration of Analyzed Participations

	N	Min	Mean	Median	75th Perc.	95th Perc.	Max
STT	6158	1	1.6	1	2	3	24
CFT	2893	1	8.2	8	12	14	37
PFT	740	1	6.5	7	8	12	26
RT	1066	1	23.5	24	26	37	53
Total	10857	1	5.8	2	8	24	53

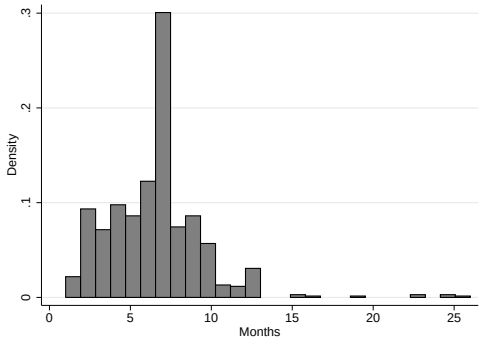
Figure 16: Densities of Program Duration



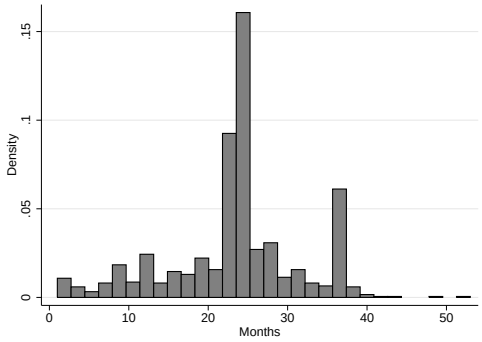
Short Term Training



Classroom Further Training



Practical Further Training



Retraining

Table 8: Effect Heterogeneity, STT vs. Waiting

East=0, Sex=0			
	Stratum 1	Stratum 2	Stratum 3
agegroup6	-0.107 (0.051)**		
qualification1			-0.083 (0.031)***
_cons	0.015 (0.016)		0.082 (0.022)***
N	908	547	662
comparable uncorr. ATT	0.008	0.033	0.047

East=0, Sex=1			
	Stratum 1	Stratum 2	Stratum 3
agegroup5			-0.139 (0.058)**
agegroup6			-0.124 (0.063)*
qualification1	-0.059 (0.159)		
mismqualification1	-0.079 (0.152)		
_cons	0.045 (0.053)		0.151 (0.026)***
N	693	409	492
corrected ATT with CI	0.026 [-0.006, 0.059]		
comparable uncorr. ATT	0.028	0.090	0.113

East=1, Sex=0			
	Stratum 1	Stratum 2	Stratum 3
agegroup6	-0.208 (0.057)***	-0.282 (0.140)**	-0.133 (0.132)
family5		0.127 (0.075)*	0.160 (0.053)***
mismagegroup6		0.077 (0.124)	-0.041 (0.113)
mismfamily5		-0.014 (0.061)	-0.055 (0.038)
mismqualification3			0.117 (0.068)*
_cons	0.045 (0.017)***	-0.001 (0.048)	-0.005 (0.035)
N	619	332	470
corrected ATT with CI		0.026 [-0.024, 0.075]	0.055 [0.013, 0.097]
comparable uncorr. ATT	0.027	0.027	0.053

East=1, Sex=1			
	Stratum 1	Stratum 2	Stratum 3
agegroup6			-0.200 (0.105)*
mismagegroup1			-0.091 (0.043)**
mismagegroup6			0.020 (0.087)
_cons			0.045 (0.030)
N	368	286	350
corrected ATT with CI			0.018 [-0.031, 0.067]
comparable uncorr. ATT	0.011	0.038	0.014

Empty entries indicate that the regressor is not included in the final specification due to lack of significance or due to a multicollinearity problem. Specifications without significant covariates are not reported (columns in which no coefficients are reported).
*** = statistically significant at 1 %, ** = at 5 %, * = at 10 %, bootstrapped standard errors

Table 9: Effect Heterogeneity, CFT vs. Waiting

East=0, Sex=0			
	Stratum 1	Stratum 2	Stratum 3
occupation1		0.178 (0.094)*	
_cons		-0.003 (0.033)	
N	385	249	265
comparable uncorr. ATT	0.023	0.021	0.095

East=0, Sex=1			
	Stratum 1	Stratum 2	Stratum 3
agegroup4			-0.209 (0.110)*
agegroup5			-0.166 (0.245)
agegroup6			-0.210 (0.243)
qualification1		-0.168 (0.095)*	
mismagegroup5			0.025 (0.201)
mismagegroup6			0.025 (0.216)
_cons		0.154 (0.044)**	0.187 (0.073)**
N	344	192	199
corrected ATT with CI			0.099 [0.016, 0.181]
comparable uncorr. ATT	0.045	0.121	0.098

East=1, Sex=0			
	Stratum 1	Stratum 2	Stratum 3
agegroup6		-0.300 (0.093)***	-0.219 (0.121)*
mismagegroup6			-0.008 (0.069)
_cons		0.085 (0.033)**	0.078 (0.045)*
N	265	218	252
corrected ATT with CI			0.052 [-0.028, 0.131]
comparable uncorr. ATT	0.096	0.060	0.046

East=1, Sex=1			
	Stratum 1	Stratum 2	Stratum 3
agegroup6			-0.187 (0.101)*
_cons			0.035 (0.032)
N	136	143	217
comparable uncorr. ATT	0.014	0.045	0.020

Empty entries indicate that the regressor is not included in the final specification due to lack of significance or due to a multicollinearity problem. Specifications without significant covariates are not reported (columns in which no coefficients are reported).
*** = statistically significant at 1 %, ** = at 5 %, * = at 10 %, bootstrapped standard errors

Table 10: Effect Heterogeneity, PFT vs. Waiting

East=0, Sex=0	
N	258
comparable uncorr. ATT	0.023
East=0, Sex=1	
agegroup6	-0.279 (0.111)**
_cons	0.121 (0.031)***
N	233
comparable uncorr. ATT	0.094
East=1, Sex=0	
agegroup6	-0.196 (0.244)
mismagegroup6	-0.009 (0.237)
_cons	0.072 (0.057)
N	145
corrected ATT with CI	0.045 [-0.030, 0.119]
comparable uncorr. ATT	0.045
East=1, Sex=1	
N	97
comparable uncorr. ATT	0.068

Empty entries indicate that the regressor is not included in the final specification due to lack of significance or due to a multicollinearity problem. Specifications without significant covariates are not reported (columns in which no coefficients are reported).
*** = statistically significant at 1 %, ** = at 5 %, * = at 10 %, bootstrapped standard errors