

Discussion Paper No. 07-009

**Long-Run Effects of Training Programs  
for the Unemployed in East Germany**

Bernd Fitzenberger and Robert Völter

**ZEW**

Zentrum für Europäische  
Wirtschaftsforschung GmbH

Centre for European  
Economic Research

Discussion Paper No. 07-009

## **Long-Run Effects of Training Programs for the Unemployed in East Germany**

Bernd Fitzenberger and Robert Völter

Download this ZEW Discussion Paper from our ftp server:

**<ftp://ftp.zew.de/pub/zew-docs/dp/dp07009.pdf>**

Die Discussion Papers dienen einer möglichst schnellen Verbreitung von neueren Forschungsarbeiten des ZEW. Die Beiträge liegen in alleiniger Verantwortung der Autoren und stellen nicht notwendigerweise die Meinung des ZEW dar.

---

Discussion Papers are intended to make results of ZEW research promptly available to other economists in order to encourage discussion and suggestions for revisions. The authors are solely responsible for the contents which do not necessarily represent the opinion of the ZEW.

## Non-technical Summary

Active labor market policy (ALMP) has been used at an unprecedentedly high scale during the transition process in East Germany in the 1990s. Public sector sponsored training has been a major part of ALMP with the goal to adjust the skills of the East German workforce to the needs of a Western market economy. Annual entries into training programs were around 250 thousand during the years 1993 to 1996. In comparison to public sector sponsored training in other countries, the East German experience shows the following five specific aspects. First, participants had fairly high levels of formal education. Second, access to treatment was easy since targeting was very low. Third, the market for training provision had to be established and in the early 1990s case workers had no practical experience on what works. Fourth, predictions about the catching up process of East Germany and about future labor market trends proved to be wrong. Fifth, the duration of training programs is fairly long.

It is often argued, that long-term public sector sponsored training programs show little or negative short-run employment effects and often it is not possible to assess whether positive long-run effects exist. For Germany, appropriate data for an evaluation of the long-term effects of public sector sponsored training were not available for a long time. Based on unique administrative data, which have only recently become available, this paper estimates the long-run differential effects on employment and benefit recipiency of three different types of training programs in East Germany. Using data on employment, periods of transfer payments, and participation in training programs, we carefully identify three types of public sector sponsored training programs for the unemployed. These programs are not associated with a regular job. The largest program among the three is the Provision of Specific Professional Skills and Techniques (SPST). SPST programs provide additional skills and specific professional knowledge in medium-term courses. The two other training programs are working in a Practice Firm (PF) and Retraining (RT). Typically, RT involves a two-year program providing complete vocational training in a new occupation and lasts longer than an SPST program. PF involves training in a work environment simulating a real job. PF tends to be a slightly shorter treatment than SPST.

We use inflows into unemployment for the years 1993/94 and apply local linear matching based on the estimated propensity score to estimate the effects of training programs starting during 1 to 2, 3 to 4, and 5 to 8 quarters of unemployment.

Specifically, we estimate the average treatment effect on the treated (ATT) against the alternative of nonparticipation in any program. We evaluate medium- and long-run treatment effects both for employment and benefit recipiency up to 24-30 quarters after the beginning of the treatment depending on the starting date of the treatment. The analysis is performed separately for males and females.

Our results imply positive medium- and long-run employment effects for the largest program, Provision of Specific Professional Skills and Techniques (SPST), a program which involves sizeable off-the-job class room training. In contrast, practice firms show no positive employment effects and this holds also for retraining (the longest program) in four out of six cases. Furthermore, we do not find any of the three programs to reduce significantly the benefit recipiency rate in the medium and long run and in the short run all programs show the lock-in effect with an increase in the benefit recipiency rate, thus providing evidence for 'benefit churning'. The fact that we see increased employment rates and constant benefit recipiency rates in the long run for SPST means that nonparticipation in the labor market went down. This suggests that such programs prevent its participants from leaving the labor force. Overall, the treatment effects are quite similar for females and males. Our evidence confirms the necessity to analyze long-term effects of sizeable training programs because all programs show strong negative lock-in effects in the short run. The positive assessment of SPST compared to practice firms is in contrast to the conventional wisdom in most of the literature.

# Long–Run Effects of Training Programs for the Unemployed in East Germany\*

Bernd Fitzenberger<sup>†</sup> and Robert Völter<sup>‡</sup>

January 2007

**Abstract:** Public sector sponsored training was implemented at a large scale during the transition process in East Germany. Based on new administrative data, we estimate the differential effects of three different programs for East Germany during the transition process. We apply a dynamic multiple treatment approach using matching based on inflows into unemployment. We find positive medium– and long–run employment effects for the largest program, Provision of Specific Professional Skills and Techniques. In contrast, the programs practice firms and retraining show no consistent positive employment effects. Furthermore, no program results in a reduction of benefit recipiency and the effects are quite similar for females and males.

**Keywords:** multiple treatments, training programs, East Germany

**JEL:** C 14, J 68, H 43

---

\*We are grateful for very helpful comments by two anonymous referees. We thank Aderonke Osikominu and Stefan Speckesser for helpful discussions as well as Stefan Bender for support in obtaining the data used here. This study is part of the project “On the effectiveness of further training programs. An evaluation based on register data provided by the Institute of Employment Research, IAB” (“Über die Wirksamkeit von Fortbildungs- und Umschulungsmaßnahmen. Ein Evaluationsversuch mit prozessproduzierten Daten aus dem IAB”); IAB project number 6–531A. The data were compiled as part of the project jointly with the SIAW St. Gallen (Michael Lechner, Ruth Miquel, Conny Wunsch) and the IAB Nürnberg (Stefan Bender). We gratefully acknowledge financial support by the IAB. All errors are our sole responsibility.

<sup>†</sup>Goethe–University Frankfurt, ZEW, IZA, IFS. Address: Department of Economics, Goethe–University, PO Box 11 19 32 (PF 247), 60054 Frankfurt am Main, Germany. E–mail: fitzenberger@wiwi.uni-frankfurt.de

<sup>‡</sup>Goethe University Frankfurt, CDSEM University of Mannheim. E–mail: voelter@wiwi.uni-frankfurt.de.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Basic Regulation and Programs</b>	<b>4</b>
2.1	Basic Regulation . . . . .	4
2.2	Evaluated Programs . . . . .	4
2.3	Financial Incentives for Participation . . . . .	6
<b>3</b>	<b>Data</b>	<b>6</b>
<b>4</b>	<b>Evaluation Approach</b>	<b>9</b>
<b>5</b>	<b>Empirical Results</b>	<b>13</b>
5.1	Estimation of Propensity Scores . . . . .	13
5.2	Estimated Treatment Effects . . . . .	14
<b>6</b>	<b>Conclusions</b>	<b>18</b>
	<b>References</b>	<b>19</b>
	<b>Appendix</b>	<b>22</b>
	Descriptive Statistics and Description of Data . . . . .	22
	Estimated Effects of Further Training Measures . . . . .	26
	<b>Additional Appendix</b>	<b>35</b>
<b>A</b>	<b>Estimation Results for the Propensity Score</b>	<b>35</b>
A.1	Sample Sizes . . . . .	35
A.1.1	Sample Sizes by Stratum . . . . .	35
A.1.2	Sample Sizes by Quarter . . . . .	36
A.2	Variable Definitions . . . . .	37
A.3	Summary Statistics . . . . .	38
A.4	Results of Propensity Score Estimations and Balancing Tests . . . . .	46
A.5	Common Support . . . . .	54
<b>B</b>	<b>Information about the data</b>	<b>57</b>

B.1	Other types of further training . . . . .	57
B.2	Construction of the monthly panel . . . . .	58
B.3	Identifying program participation . . . . .	58
B.4	Coding plan for the treatment information . . . . .	61
B.5	Sample construction in comparison to LMW . . . . .	66
<b>C</b>	<b>Heterogeneous Treatment Effects by Target Profession</b>	<b>67</b>
C.1	Retraining for Men . . . . .	67

# 1 Introduction

Active labor market policy (ALMP) has been used at an unprecedentedly high scale during the transition process in East Germany in the 1990s. Public sector sponsored training has been a major part of ALMP with the goal to adjust the skills of the East German workforce to the needs of a Western market economy. Annual entries into training programs were around 250 thousand during the years 1993 to 1996 (BA 1993, 1997, 2001). In comparison to public sector sponsored training in other countries, the East German experience shows the following five specific aspects. First, participants had fairly high levels of formal education. Second, access to treatment was easy since targeting was very low. Third, the market for training provision had to be established and in the early 1990s case workers had no practical experience on what works. Fourth, predictions about the catching up process of East Germany and about future labor market trends proved to be wrong. Fifth, the duration of training programs is fairly long.

During the last decade, there were a lot of pessimistic assessments regarding the usefulness of public sector sponsored training programs in raising employment chances of the unemployed (see the surveys in Fay, 1996; Heckman et al., 1999; Martin and Grubb, 2001; Kluve and Schmidt, 2002). These studies doubt that large scale training programs, which are not well targeted, are successful in raising employment. However, evidence for Eastern European transition economies (other than East Germany) has often shown positive effects (Kluve et al., 2004; Lubyova and van Ours, 1999; Puhani, 1999). Recently, OECD (2005) has argued that long-term labor market programs, such as training, often have little or negative short-run effects on outcomes, which can be attributed to lock-in effects. However, in some cases, positive long-term effects exist for long training programs, for which lock-in effects are worse than for short programs (see also Fay, 1996). Therefore, it is crucial to assess program impacts in a longer term perspective in order to investigate whether the sizeable lock-in effects in the short run are compensated by positive long run effects.

For East Germany, appropriate data for a long term evaluation of public sector sponsored training were not available for a long time and, until recently, the available evidence has been quite mixed.<sup>1</sup> Detailed administrative data have been used in

---

<sup>1</sup>See Bergemann et al. (2004), Fitzenberger and Prey (2000), Kraus et al. (1999), or Lechner (2000) for exemplary studies based on survey data. Speckesser (2004, chapter 1) and Wunsch (2006, section 6.5) provide comprehensive surveys of this literature, which is not reviewed here for the sake of brevity, and discuss critically the data used.



the recent studies of Lechner et al. (2005b) ( $\equiv$ LMW), Fitzenberger and Speckesser (2007) ( $\equiv$ FS), and Hujer et al. (2006) ( $\equiv$ HTZ), where the first two studies are based on the same data as this study, while the third uses administrative data since 2000. HTZ find negative short-run effects which are probably driven by lock-in effects, while their data do not allow to investigate long-run effects. LMW and FS find positive medium- and long-run employment effects for some treatments considered in this paper. LMW evaluate effects of three training programs (long training, short training, retraining) on employment and benefit reciprocity. They find strong evidence that, on average, the training programs under investigation increase long-term employment prospects and do not change benefit reciprocity. As important exceptions, long training and retraining show no positive employment effects for males. FS estimate the employment effects of one major training program (Provision of Specific Professional Skills and Techniques, SPST) against nonparticipation in SPST for 36 months after the beginning of the treatment. The analysis is performed only for the 1993 inflow sample into unemployment. The analysis finds positive medium-run employment effects, but it does not distinguish between genders.

The vast majority of the existing evaluation studies for East Germany uses a static evaluation approach, which contrasts receiving treatment during a certain period of time against the alternative of not receiving treatment during this period of time (FS and HTZ are recent exceptions). In a dynamic setting, the timing of events becomes important, see Abbring and van den Berg (2003), Fredriksson and Johansson (2003, 2004), and Sianesi (2003, 2004). Static treatment evaluations run the risk of conditioning on future outcomes, leading to possibly biased treatment effects. This paper follows Sianesi (2003, 2004) and estimates the effects of treatment starting after some unemployment experience against the alternative of not starting treatment at this point of time and waiting longer. The actual implementation of the estimator builds and extends upon FS and Fitzenberger et al. (2006). The estimated dynamic treatment effects mirror the decision problem of the case worker and the unemployed who decide recurrently during the unemployment spell, whether to begin any program now or to postpone participation to the future.

Using a dynamic multiple treatment framework, this study analyzes the effects of three exclusive training programs (practice firms, SPST, retraining) for inflow samples into unemployment for the two years 1993/94. We evaluate medium- and long-run treatment effects both for employment and benefit reciprocity up to 24–30 quarters after the beginning of the treatment depending on the starting date of the treatment. The analysis is performed separately for males and females to reexamine the evidence in LMW and the two studies differ substantially regarding the exact

treatment definition, the choice of valid observations, and the econometric methods used. Our results confirm the positive employment effects for SPST reported in FS after the initial negative lock-in effect to hold for a much longer time period and to apply for both males and females. Our study finds no positive employment effects for practice firms and in four out of six cases for retraining. We do not find systematic gender differences and, similar to the study Fitzenberger et al. (2006) for West Germany, our assessment of retraining is considerably worse than of SPST. Furthermore, we do not find any of the three programs to reduce significantly the benefit reciprocity rate in the medium and long run. In the short run, all programs show the lock-in effect with an increase in the benefit reciprocity rate, thus providing evidence for 'benefit churning' as in Kluve et al. (2004).

Our analysis differs considerably from the recent work of LMW, FS, and HTZ. LMW use a static multiple treatment evaluation approach. They find gender differences for long training and retraining, which we can not replicate using our dynamic evaluation approach. We explore potential reasons for the different results. LMW analyze the effects of treatments starting during the years 93/94 for unemployed whose unemployment spells start during the years 93/94. We also analyze the inflows into unemployment for the years 93/94 but we analyze the ATT effects of all treatments taking place during the first two years of unemployment. We investigate whether the estimated treatment effects differ for treatments during three different time windows of elapsed unemployment durations. Furthermore, there are a number of important differences in the definition of the treatments, the selection of samples, and the implemented methods. FS use a dynamic treatment evaluation approach for SPST only. We estimate the effects of three training programs for a much longer time period after the beginning of the program, and our analysis distinguishes between genders. In addition to the employment effects, we also analyze the effects on benefit reciprocity. Furthermore, we use a larger inflow sample than FS. In contrast to HTZ, who estimate a duration model and focus on exits from unemployment, we estimate medium- and long-run effects on both employment and benefit reciprocity, which we distinguish from lock-in effects. Estimating a duration model, it would be very difficult to take account of the large number of exits into and out of employment observed after the first exit from unemployment.

The remainder of this paper is structured as follows: Section 2 gives a short description of the institutional regulation and participation figures for Active Labor Market Policy. Section 3 focuses on the different options of further training, their target groups, and course contents. Section 4 describes the methodological approach to estimate the treatment effects. The empirical results are discussed in section 5.

Section 6 concludes. The final appendix provides further information on the data and detailed empirical results. An additional appendix, which is available on our webpage, includes further details on the data and the empirical results.

## 2 Basic Regulation and Programs

### 2.1 Basic Regulation

For the time period considered here, public sector sponsored training in Germany is regulated by the Labor Promotion Act (*Arbeitsförderungsgesetz*, AFG) and is offered and coordinated by the German Federal Employment Office (formerly *Bundesanstalt für Arbeit*, BA). We consider the two main training programs: **Further training** (*Weiterbildung*) includes the assessment, maintenance and extension of skills, including technical development and career advancement. The duration of the courses depends on individual predispositions and adequate courses provided by the training suppliers. **Retraining** (*Umschulung*) enables vocational re-orientation if a completed vocational training does not lead to adequate employment. Retraining is supported for a period up to 2 years and aims at providing a new certified vocational education degree.

### 2.2 Evaluated Programs

Further training is a very broad legal category and consists of quite heterogeneous programs. Hence we utilize a classification developed in FS and evaluate two specific further training programs: Practice Firms (PF) and provision of specific professional skills and techniques (SPST).

**Practice Firms (PF)** are simulated firms in which participants practice everyday working activities. The areas of practice are whole fields of profession, not specific professions. Hence, practice firms mainly train general skills while provision of new professional skills is of less importance. Some of the practice firms are technically oriented, the practice studios, whereas others are commercially oriented, the practice enterprises. One of the practice firm's goals is to evaluate the participant's aptitude for a field of profession. The programs usually last for six months and do not provide official certificates.

Provision of **Specific Professional Skills and Techniques (SPST)** intends to improve the starting position for finding a new job by providing additional skills and specific professional knowledge in medium-term courses. It involves refreshing specific skills, e.g. computer skills, or training on new operational practices. SPST mainly consists of classroom training but an acquisition of professional knowledge through practical work experience may also be provided. After successfully completing the course, participants usually obtain a certificate indicating the contents of the course, i.e. the refreshed or newly acquired skills and the amount of theory and practical work experience. Such a certificate is supposed to serve as an additional signal to potential employers and to increase the matching probability since the provision of up to date skills and techniques is considered to be a strong signal in the search process. The provision of specific professional skills and techniques aims at sustained reintegration into the labor market by improving skills as well as providing signals.

Compared to retraining, which is a far more formal and thorough training on a range of professional skills and which provides a complete vocational training degree, the role of SPST for a participant's occupational knowledge is weaker. However, the amount of occupation specific knowledge imparted in SPST certainly exceeds the level provided in short-term programs (not evaluated here) that usually aim at improving job search techniques or general social skills. Thus, SPST ranges in the middle between very formal (and very expensive) courses and very informal and short courses (improving general human capital).

**Retraining (RT)** consists of the provision of a new and comprehensive vocational training according to the regulation of the German apprenticeship system. It is targeted to individuals who already completed a first vocational training and face severe difficulties in finding a new employment within their profession. It might however also be offered to individuals without a first formal training degree if they fulfil additional eligibility criteria.

Retraining provides widely accepted formal certificates. It comprises both, theoretical training and practical work experience. The theoretical part of the formation takes place in the public education system. The practical part is often carried out in firms that provide work experience in a specific field to the participants, but sometimes also in interplant training establishments. This type of treatment leads to a certified job qualification in order to improve the job match. Ideally, the training occupation in retraining corresponds to qualifications which are in high demand in the labor market.

## 2.3 Financial Incentives for Participation

Participants in the training programs considered are granted an income maintenance (IM, *Unterhaltsgeld*). To qualify, they must have been employed for at least one year or they must be entitled to unemployment benefits or subsequent unemployment assistance.<sup>2</sup>

Since 1994, IM is equal to the standard unemployment benefits (UB, *Arbeitslosengeld*). It amounts to 67% of previous net earnings for participants with at least one dependent child and 60% otherwise (note that in 1993 replacement ratios for IM were higher at 73% and 65%, respectively). In contrast, unemployed, whose UB expired, can receive the lower, means tested unemployment assistance (UA, *Arbeitslosenhilfe*) which amounts to 57% (with children) and 53% (without children). This means that for these unemployed IM during the program is higher than UA. Additionally, participants could defer the transition from UB to the lower UA and, in some cases, even requalify for the higher UB.

Concluding, there are positive financial incentives for the unemployed to join a program. In addition, the BA bears all costs directly incurred through participation in a further training scheme, especially course fees.

## 3 Data

We use a database which integrates administrative individual data from three different sources (see Bender et al. (2005) for a detailed description). The data contain spells on

- employment subject to social insurance contributions,
- transfer payments by the BA,
- and participation in training programs.

Further details on the compilation of the data can be found in the additional appendix.

---

<sup>2</sup>For a more detailed description of the institutions, see Bender et al. (2005), Fitzenberger, Osikominu and Völter (2006), or Wunsch (2006).

The basic data source is the **IAB Employment Subsample** (*IAB Beschäftigtenstichprobe*, IABS) for the time period 1975–97, see Bender et al. (2000) and Bender et al. (2005, chapter 2.1). The IABS is a 1% random sample drawn from employment register data for all employees subject to social insurance contributions. Therefore, we restrict the analysis to inflows from employment to unemployment. For this study, we merge additional information for 1998–2002 to the basic data.

The second important source is the **Benefit Payment Register** (*Leistungsempfängerdatei*, LED) of the Federal Employment Office (BA), see Bender et al. (2005, chapter 2.2). These data consist of spells on periods of transfer payments granted by the BA to unemployed and program participants. Besides unemployment benefit or assistance, these data also record very detailed information about income maintenance payments related to the participation in training programs.

The third data source records training participation (**FuU-data**). The BA collects these data for all participants in further training, retraining, and other training programs for internal monitoring and statistical purposes, see Bender et al. (2005, chapter 2.3). For every participant the FuU-data contains detailed information about the program and about the participant.

The FuU-data were merged with the combined IABS–LED data by social insurance number and additional covariates. Numerous corrections have been implemented in order to improve the quality of the data, see Bender et al. (2005, chapters 3–4), FS, and the additional appendix for more information. The IABS provides information on personal characteristics and employment histories. The combination of the transfer payment information and the participation information is used to identify the likely participation status regarding the different types of training programs.

When an individual is not observed in any of the three spell types (employment, transfers, training participation), we interpret this as being out of the labor force. The spell information on the employment state of an individual is first transformed into monthly dummy variables (based on the dominating state). We construct separately monthly dummy variables for training status. Then, for our analysis, the data is aggregated to a quarterly frequency.

**Inflow Sample into Unemployment:** To analyze the effect of training programs on employment and benefit reciprocity of unemployed individuals, we base our empirical analysis on the sample of inflows into unemployment during the years 1993/94 in East Germany, omitting Berlin. We consider individuals who experience a transition from employment to nonemployment and for whom a spell with transfer payments

from the Federal Employment Office starts during the first 12 months of nonemployment or for whom the training data indicate program participation before a new job is found.<sup>3</sup> The start of the nonemployment spell is denoted as the beginning of the unemployment spell. We condition on receipt of unemployment compensation or program participation to exclude individuals who move out of the labor force.<sup>4</sup> This rule concerns almost exclusively individuals who do not participate in any training program during their nonemployment spell. A treatment is only considered if the unemployed does not start employment before the second month of treatment (to omit training while holding a job). Furthermore, we restrict our samples to the 25 to 55 years old in order to rule out periods of formal education or vocational training as well as early retirement. For RT, we restrict the sample to the 25 to 50 years old.

We choose the years 1993/94 because data for East Germany start in 1992 and we want to control for one year of labor market experience before the beginning of unemployment. Our merged data allow to follow individuals until the end of 2002. Table 1 gives information about the size of the inflow samples and the incidence of training.

**Participation by Type of Training:** We focus on the three types of training programs PF, SPST, and RT, as described in section 2.2 above. These programs are targeted to the unemployed and do not involve on-the-job training (training while working in a regular job). The total inflow sample comprises 6,135 spells for women and 5,911 spells for men. There are 1,550 training spells for females and 835 for men. Thus, about 25% of the females and 14% of the men participate in one of the three training programs considered, which reflects the large scale of training programs during the East German transition process. Among these programs, SPST represents the largest with 78% and 63% of the training spells, respectively for females and males. For females 13% and for men 28% of all training spells are RT, and PF represents the smallest group in both samples. In absolute numbers, there are 145 (73) PF spells in the female (male) inflow sample, 1,210 (528) SPST spells and 195 (234) RT spells. Table 2 shows the frequency of training by elapsed duration of unemployment.

Table 3 provides descriptive statistics on the elapsed duration of unemployment at the beginning of treatment. Our discussion focuses on quantiles because averages may be misleading. The median entrant in PF has been unemployed for 10 months

---

<sup>3</sup>This design allows the same individual to be in the sample more than once if it has more than one transition from employment to unemployment in 1993/94.

<sup>4</sup>Only 1% of training participants do not receive transfer payments during the first 12 months.

for females and only for 5 months for males. Late starts (75%-quantile) of PF occur after 14 months for females and after 11 months for men. SPST is the program which starts latest with a median of 11 months for females and 7.5 months for men. RT is the program which starts the earliest for females. The median is 8 months. The median for males of 6 months is higher than the value for PF. In general, females start later than men.

Table 4 provides descriptive information on the duration of training spells. The average durations are quite different between the programs but comparable across genders. Participation in PF is shortest. On average woman stay 6.5 months in PF and men 6.1 months. Participation in SPST has an average duration of 9.1 months for females and 8.8 months for males. Participation in RT lasts almost twice as long as in SPST with an average of 18.7 months for women and 17.3 months for men.

## 4 Evaluation Approach

Our goal is to analyze the effect of  $K = 3$  different training programs on two outcome variables, namely the individual quarterly employment rate (ER) and the individual quarterly benefit reciprocity rate (BR), both measured as quarterly averages of monthly dummy variables.<sup>5</sup> 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. 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. This dynamic evaluation approach is implemented for our problem as in FS and Fitzenberger et al. (2006). The estimated dynamic ATT parameters mirror the decision problem of the case worker and the unemployed who decide recurrently during the unemployment spell, whether to begin any program now or to postpone participation to the future.

Our empirical analysis is based upon the potential–outcome–approach to causality, see Roy (1951), Rubin (1974), and the survey of Heckman, LaLonde, Smith (1999). Lechner (2001) and Imbens (2000) extend this framework to allow for multiple, exclusive treatments. Let the 4 potential outcomes be  $\{Y^0, Y^1, Y^2, Y^3\}$ , where  $Y^k$ ,  $k = 1, \dots, 3$ , represents the outcome associated with training program  $k$  and  $Y^0$  is the outcome when participating in none of the 3 training programs. For each

---

<sup>5</sup>These quarterly rates can take the four values 0, 1/3, 2/3, and 1.



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  $k = 1, 2, 3$  against nonparticipation  $k = 0$ .<sup>6</sup>

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).<sup>7</sup> 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 quarters 1 to 2 of the unemployment spell (stratum 1), treatment starting during quarters 3 to 4 (stratum 2), and treatment starting during quarters 5 to 8 (stratum 3).

Our estimated ATT parameter has to be interpreted in a dynamic context. 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 estimated treatment parameter is

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

where  $T_u$  is the treatment variable for treatment starting in quarter  $u$  of unemployment and  $U$  is the completed duration of the unemployment spell.  $Y^k(u, \tau)$  and  $Y^0(u, \tau)$  are the potential treatment outcomes for treatments  $k$  and 0, respectively, in periods  $u + \tau$ , where treatment starts in period  $u$  and  $\tau = 0, 1, 2, \dots$ , counts the quarters since the beginning of treatment. The nontreatment outcome  $Y^0(u, \tau)$  refers to the case where the individual does not receive any treatment until

---

<sup>6</sup>Using the same approach, a pairwise comparison of the differential effects of the programs would be feasible, see Lechner (2001) or Fitzenberger et al. (2006). Such a pairwise comparison is not pursued in this paper for the sake of space.

<sup>7</sup>In East Germany, active labor market programs were implemented after unification at an unprecedented scale.

the end of the stratum considered. Actually, we estimate the treatment parameter  $\theta(k; \tau) = \sum_u g_u \theta(k; u, \tau)$ , which is averaged within a stratum with respect to the distribution  $g_u$  of starting dates  $u$ .

We evaluate the differential effects of multiple treatments assuming the following dynamic version of the conditional mean independence assumption (DCIA)<sup>8</sup>

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

where  $X$  are time-invariant (during the unemployment spell) characteristics,  $ben(u)$  is the number of months the unemployed were receiving benefits during the unemployment spell before the start of the treatment  $u$ , and  $\bar{u}$  denotes the last quarter of the stratum considered. We effectively assume that conditional on  $X$ , conditional on being unemployed until period  $u-1$ , conditional on having received benefits the same number of months before  $u$ , and conditional on not having received a treatment before  $u$ , individuals treated in  $u$  are comparable in their nontreatment outcome to individuals who do not start any treatment until  $\bar{u}$  (recall from above, that  $Y^0(u, \tau)$  involves no treatment until  $\bar{u}$ ).

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 no treatment 0,  $P^{k|k0}(X)$ , exhibits an analogous balancing property for the pairwise estimation of the ATT's of program  $k$  versus no participation 0. This allows to apply standard binary propensity score matching based on the sample of individuals participating in either program  $k$  or in no program 0 (Lechner, 2001; Gerfin and Lechner, 2002; Sianesi, 2003). 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 quarters 1 to 2, we take the total sample of unemployed,

---

<sup>8</sup>In addition to DCIA, we also assume that the probability of treatment is less than one conditional on the conditioning variables in equation 2 and that the Stable Unit Treatment Value assumption holds. These are further assumptions needed to estimate an ATT parameter, see Heckman, LaLonde, Smith (1999).

who participate in  $k$  or in no program during quarters 1 to 2 (stratum 1), and estimate a Probit model for participation in  $k$ . This group includes those unemployed who either never participate in any program or who start some treatment after quarter 2. For treatment during strata 2 and 3, the basic sample consists of those unemployed who are still unemployed at the beginning 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 quarter (of elapsed unemployment duration) before treatment  $k$  starts and have received benefits the same number of months until the quarter before treatment starts. The expected counterfactual employment outcome for nonparticipation is obtained by means of a local linear regression on the propensity score and the starting month of the unemployment spell to match on calendar time. We use a bivariate crossvalidation procedure to obtain the bandwidths in both dimensions (propensity score and beginning of unemployment spell). An estimate for the variance of the estimated treatment effects is obtained through bootstrapping based on 200 resamples.<sup>9</sup> 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 time-invariant (during the unemployment spell) covariates are balanced sufficiently by matching on the estimated propensity score  $P^{k|k0}(X)$  using a flexible polynomial approximation. Furthermore, we investigate whether treated and matched nontreated individuals differ significantly in their outcomes before the beginning of treatment, in addition to those 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.

---

<sup>9</sup>Abadie and Imbens (2006) show that the bootstrap fails for nearest neighbor matching because of a lack of smoothness resulting in local convergence not being uniform (see also Heckman et al., 1998, p. 276). In contrast, local linear matching with appropriate trimming to guarantee common support and under a weak convergence condition for the bandwidth parameters, is shown by Heckman et al. (1998, p. 278) to exhibit sufficiently smooth convergence for standard asymptotic distribution theory to hold. In particular, the estimated ATT parameter has a standard asymptotically linear representation and it is asymptotically normally distributed with  $\sqrt{N}$  convergence rate. Although we are not aware of a formal proof, the bootstrap is therefore likely to be valid for local linear matching. Horowitz (2001, section 2) discusses the consistency of the bootstrap for  $\sqrt{N}$  asymptotically normal estimators with an asymptotically linear representation. Although local linear matching involves an intermediate nonparametric estimation step, a similar result is likely to hold.

Finally, we need to discuss why we think that the DCIA (2) is plausible for our application. As Sianesi (2004), we argue that the participation probability depends upon the variables determining re-employment prospects once unemployment began. Consequently, 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 and the same number of months receiving benefits before program participation. Furthermore, observable individual characteristics and information from the previous employment spell 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. Unfortunately, our data lack subjective assessments of labor market chances of the unemployed (e.g. by case workers). We argue that these are proxied sufficiently by the observed covariates in so far as they affect selection into the program. This is particularly plausible, since participation occurred at a large scale, assignment was not very targeted, and case workers lacked practical experience on 'what works' in a quickly changing economic environment. 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.

## 5 Empirical Results

### 5.1 Estimation of Propensity Scores

Our empirical analysis is performed separately for females and males. To estimate the propensity scores, we run Probit regressions for each of the three programs for taking part in this program versus not taking part in any program (“waiting”) for training starting during the three time intervals for elapsed unemployment duration, i.e. 1–2 quarters (stratum 1), 3–4 quarters (stratum 2), and 5–8 quarters (stratum 3). The additional appendix reports our preferred specifications, which are obtained after extensive specification search, summary statistics of the covariates used, detailed results of the balancing tests, and figures on common support.

The covariates considered are all defined for the beginning of unemployment and are thus time-invariant for an individual during the unemployment spell. Personal characteristics considered are age, marital status and formal education (with/without vocational training degree, tertiary education degree). In addition, we use information about the last employer, namely industrial sector and firm size, and a number

of characteristics of the previous job such as employment status and information on earnings in the previous job. Regarding the employment and program participation history, we consider the employment history and participation in any ALMP program in the year before the beginning of unemployment. Differences in regional labor market conditions as well as supply of programs are the reason to include regional variables in the specification. We use the federal state and the population density at the district level. Finally, we also use the calendar month of the beginning of the unemployment period.

Our specification search starts by using as many as possible of the covariates mentioned above without interactions. The specification search is mainly led by the following two criteria: (i) single and joint significance, and (ii) balance of the covariates according to the regression based balancing test in Smith and Todd (2005). In general, insignificant covariates are dropped. We also test for the significance of interaction effects, in particular interactions with age. In order to achieve balance of covariates, we test different functional forms and interaction effects. In a few cases, we keep insignificant covariates or interactions, when they help to achieve balance. As we find the balancing test to be somewhat sensitive to small cell sizes we occasionally aggregate small groups that have similar coefficients.

The results for the Probit estimates show that the final specifications vary considerably between men and women and the three time intervals for a given program. Age effects are significant in most cases. In particular, participants in retraining are younger than individuals in other groups.

Our chosen specifications for the propensity score pass the regression based balancing test (no rejection) of Smith and Todd (2005) for a sufficiently large number of covariates. We graphically examine the common support requirement for estimating the average treatment effect on the treated (ATT). Overall, we are satisfied with the overlap of support in all cases and proceed without restricting the samples.<sup>10</sup>

## 5.2 Estimated Treatment Effects

We estimate the effects of the three types of training programs PF, SPST, and RT, separately for males and females. The two outcome variables considered are the individual quarterly employment rate (ER) and benefit reciprocity rate (BR: UB, UA,

---

<sup>10</sup>In four cases (out of 16) we have to drop one and in one case two treated individuals from the treatment effect estimations due to numerical problems.

or IM; see section 2.3). We match participants in treatment  $k$  and nonparticipants in any treatment, who are still unemployed in the quarter before treatment starts and have received benefits the same number of months until the quarter before treatment starts, by their similarity in the estimated propensity scores and the starting month of the unemployment spell. The ATT is then estimated separately for quarters  $\tau$  since the beginning of program  $k$  for stratum 1, 2, and 3.

Figures 1–6 display the estimated treatment effects  $\hat{\theta}(k; \tau)$  on the horizontal axis against quarter  $\tau \geq 0$  since the beginning of treatment or quarter  $\tau < 0$  before the beginning of treatment. The time axis is divided into three parts by two vertical lines, which denote the last quarter before the unemployment spell starts and the treatment start  $\tau = 0$ , respectively. The left part shows the four quarters before unemployment starts, the middle part the gap between the beginning of the unemployment spell and the beginning of treatment and the right part the time since treatment start. Each figure contains a panel of three times four graphs (except PF for males, with only stratum 1 in figure 2), where each row represents represents one stratum of elapsed duration of unemployment. The first and third column show the evolution of average outcomes for treated individuals (solid line) and their estimated nontreatment counterfactual (dashed line). The differences of these lines are displayed in the second and fourth column (solid line), respectively, as the estimated treatment effects together with pointwise 95%–confidence bands (dashed lines). To summarize the graphical evidence in a systematic way, tables 6 and 7 provide cumulated treatment effects ( $\sum_{\tau=0}^{L-1} \hat{\theta}(k; \tau)$ ) over the first  $L = 8, 16,$  and  $24$  quarters since beginning of treatment and average treatment effects during quarter 4 to 23 and 8 to 23 [ $1/(24 - l) \sum_{\tau=l}^{23} \hat{\theta}(k; \tau)$  for  $l = 4, 8$ ]. These aggregated effects are calculated as sums or averages of the effects depicted graphically.

The treatment PF (figures 1–2) basically shows statistically significant negative lock-in effects on ER during the first six quarters (the solid line in the first columns lies below the dashed line)<sup>11</sup> and no significant positive ER effects afterwards. The BR effects are almost symmetric, with positive BR effects during the lock-in period and mostly no significant BR effects afterwards, except for stratum 3 for women where the BR effect seems to be quite volatile and often significantly positive in the medium- and long-run. The results are quite similar in stratum 1 for both genders. The graphical evidence is confirmed in tables 6 and 7. We restrict our discussion of the aggregated effects to the cumulated effects over 24 quarters and to the average effects during quarter 8 to 23. None of the aggregated ER effects is significant. For

---

<sup>11</sup>We discuss lock-in effects for the time it takes for the treated individuals to catch up with the nontreated individuals.

BR, we find no significant aggregated effects on women for stratum 1 and 2. For Men in stratum 1 the cumulated effect on BR is significantly positive, but the average effect is insignificant. For stratum 3, we find both effects to be significantly positive. Thus, the treatment PF shows no positive employment effects, but it increases the benefit reciprocity rate for women starting treatment later in their unemployment spell.

The evidence for SPST in figures 3–4 is much more positive and confirms the results in FS. After strong negative lock-in effects during a period of almost two years, we find positive and mostly significant medium- and long-run employment effects of around 10 percentage points (pp), which typically persist until the end of the observation period. The effects on BR are similar to PF, i.e. treatment increases BR in the short run, and the medium- and long-run effects are not significantly different from zero. The cumulated ER increases lie between 0 and 1.5 quarters. They are significant for stratum 1 and insignificant for the later strata. The average ER effects are highly significant and amount to about 10 pp in all cases. All cumulated BR effects are positive and significantly so for strata 2 and 3. The average BR effects are never significant. The effects for both genders are very similar.

For RT, the evidence in figures 5–6 is more mixed. As to be expected, we find the longest (typically lasting 10 quarters) and deepest lock-in effects for this treatment, with stratum 1 for men showing the strongest decline. The medium- and long-run ER effects are only significantly positive for males in stratum 1 and females in stratum 3. For women in stratum 1 the effects are sometimes significantly positive. The three other cases basically show insignificant ER effects in the medium- and long-run, although they are positive in most periods. Again, we find positive BR effects during the lock-in period and typically insignificant BR effects in the medium- and long-run for strata 2 and 3. For stratum 1 we see a medium- and long-run reduction, but which is only sometimes significant. Almost all of the cumulated ER effects are insignificantly negative, stratum 2 for men shows a significantly negative effect and stratum 3 for women an insignificantly positive one. Confirming the graphical evidence, the average ER effects are significant only for males in stratum 1 (around 12 pp) and females in stratum 3 (around 16 pp). All cumulated BR effects are significantly positive. The average BR effects are only significant for males in stratum 2 and 3.

No case in 1–6 shows significant differences in outcomes before the beginning of the unemployment spell. Since we include the employment history in the propensity score estimation, this is not a pre-program test of the CIA. But the results

show that our matching approach balances well the employment history of treated and nontreated individuals. Note furthermore that lock-effects last fairly long in comparison to results for West Germany, see Lechner et al. (2005a), LMW, FS, and Fitzenberger et al. (2006). A likely reason is that search frictions in the labor market are higher in East Germany compared to West Germany.

Overall, our results do not confirm the gender differences in the treatment effects as found in LMW. Neither for SPST, which comprises most of the long training as in LMW, and nor for RT, we find that employment effects are higher for females compared to males and that males show zero or negative long-run effects.<sup>12</sup> To explore reasons for the differences in results, we first would like to reexamine the evidence on gender differences in the content of training as reported in LMW, which the authors identify as a potential reason for the gender differences in the treatment effects. Programs are characterized by the target profession of training. This information is contained in table 5 stratified by gender, program, and stratum. Large differences show up between genders as also documented in LMW. PF for women mainly train in office professions (38%–48%) and in broader programs (20–27%), which can not be related to a specific profession. For female participants in SPST these fields are also the most important with 20–30% for office professions and 13–31% for broader programs. RT for women train mainly in service professions (17–28%), office professions (12–25%) and health professions (10–22%). For males, the programs PF and RT are dominated by target professions in *construction*, which have a share of at least 40%, and even 56% for men in RT in stratum 3. Metal professions are second most important for PF and RT in stratum 1 and 2 with about 25%. RT in stratum 3 trains only 12% in metal professions. SPST for men is concentrated in service professions (13–22%) and technical professions (13–19 %) for all strata. In strata 1 and 2 metal professions are most important with 27 and 23% and construction is also important with 13 and 17%. In the third stratum broad programs are most important with 32%. Thus, our data show similar gender differences in the content of training as reported by LMW.

Now, we explore further possible explanations of the differences in the estimated treatment effects for RT. We focus on RT because SPST differs from long training as defined in LMW and target professions in construction have a fairly small share in SPST. First, the differences to LMW are not due to the fact that LMW use a static

---

<sup>12</sup>As one exception, we find positive effects of RT for females and not for males in stratum 3. However, the number of treated males in stratum 3 is very small and the results in LMW correspond mainly to stratum 1 and 2 because the construction of the treatment sample in LMW oversamples early treatments, see discussion below.



evaluation approach, while we estimate the effects of treatment versus waiting. To investigate this, we reestimate the treatment effects in stratum 1 excluding the future participants in any training program from the control group (around 10% of the male and around 20% of the female controls are excluded, see additional appendix). The results for males basically do not change while the estimated treatment effects for females are reduced to some extent (these results are available upon request). Thus, the difference in evaluation approach should work in the opposite direction and can not explain the differences in the results. Second, since LMW suggest that males do not show positive long-run employment effects from RT because of the large share of target professions in construction, we estimate the treatment effects of RT for males separately with target profession in construction and in nonconstruction. We exclude the cases where the target profession is missing. The results (see additional appendix for details) clearly show that the employment effects for target profession construction are by no means smaller than for target profession nonconstruction. In fact, the point estimates for stratum 1 and 2 even suggest that in most cases the medium- and long-run employment effects are higher for target professions in construction (these differences are, however, not significant). Third, the differences in the sample construction (see table in additional appendix for a juxtaposition) between our paper and LMW show that LMW oversample early treatments. This should work in the opposite direction of the differences in the results, because in stratum 1 men but not women show positive employment effects for RT (see footnote 12). There are a number of further differences in the construction of the sample which, however, seem unlikely to explain the differences in results.

Concluding, we can not replicate the gender differences in results reported in LMW and we can not confirm differences in treatment effects by target profession as suggested by LMW. We have explored possible reasons to rationalize these differences but, unfortunately, the reason for these differences in results remains an open question.

## 6 Conclusions

Using a dynamic multiple treatment framework, this study analyzes the effects of three exclusive training programs for inflows into unemployment for the two years 1993/94. We evaluate medium- and long-run treatment effects both for employment and benefit reciprocity up to 24-30 quarters after the beginning of the treatment depending on the starting date of the treatment and we distinguish by gender. Our

results imply positive medium- and long-run employment effects for the largest program, Provision of Specific Professional Skills and Techniques (SPST), a program which involves sizeable off-the-job class room training. In contrast, practice firms show no positive employment effects and this holds also for retraining (the longest program) in four out of six cases. Furthermore, we do not find any of the three programs to reduce significantly the benefit recipiency rate in the medium and long run, in the short run all programs show the lock-in effect with an increase in the benefit recipiency rate, thus providing evidence for 'benefit churning' as in Kluve et al. (2004). The fact that we see increased ER and constant BR in the long run for SPST means that nonparticipation in the labor market went down. This suggests that such programs prevent its participants from leaving the labor force. Overall, the treatment effects are quite similar for females and males, thus, we can not confirm the gender differences found in Lechner et al. (2005b). Our evidence confirms the necessity to analyze long-term effects of sizeable training programs because all programs show strong negative lock-in effects in the short run. The positive assessment of SPST compared to practice firms is in contrast to the conventional wisdom in most of the literature. As a final caveat, an overall assessment of the microeconomic effects is not possible since various necessary information for a comprehensive cost-benefit-analysis are lacking in our data set.

## 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.
- Bender, S., A. Bergemann, B. Fitzenberger, M. Lechner, R. Miquel, S. Speckesser, and C. Wunsch (2005). "Über die Wirksamkeit von Fortbildungs- und Umschulungsmaßnahmen", *Beiträge zur Arbeitsmarkt- und Berufsforschung*, IAB, Nürnberg.
- Bender, S., A. Haas, and C. Klose (2000). "IAB employment subsample 1975-1995", *Schmollers Jahrbuch (Journal of Applied Social Science Studies)* 120, 649-662.
- Bergemann, A. B. Fitzenberger, and S. Speckesser (2004). "Evaluating the Dynamic Employment Effects of Training Programs in East Germany Using Conditional Difference-in-Differences." ZEW Discussion Paper, Mannheim.
- BA ([Bundesanstalt für Arbeit] 1993, 1997, 2001). *Berufliche Weiterbildung*. Nürnberg: Bundesanstalt für Arbeit (various issues).

- Bundesanstalt für Arbeit (2003), *Geschäftsbericht 2002*, Einundfünfzigster Geschäftsbericht der Bundesanstalt für Arbeit. Nürnberg: Bundesanstalt für Arbeit.
- Fay, R. (1996) “Enhancing the Effectiveness of Active Labour Market Policies: Evidence from Programme Evaluations in OECD countries.” Labour Market and Social Policy Occasional Papers, 18, OECD, Paris.
- Fitzenberger, B., A. Osikominu, and R. Völter (2006). “Get Training or Wait? Long-Run Employment Effects of Training Programs for the Unemployed in West Germany.” IZA Discussion Paper No. 2121.
- Fitzenberger, B. and S. Speckesser (2007) “Employment Effects of the Provision of Specific Professional Skills and Techniques in Germany.” *Empirical Economics* (forthcoming).
- 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) “Microeconomic Evaluation of the Active Labor Market Policy in Switzerland.” *Economic Journal*, 112(482), 854–893.
- Heckman, J. H. Ichimura, and P. Todd (1998) “Matching as an Econometric Evaluation Estimator.” *Review of Economic Studies*, 65, 261–294.
- Heckman, J. 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.
- Horowitz, J. (2001) “The Bootstrap.” In: J.J. Heckman, E. Leamer, editors, *Handbook of Econometrics*, Volume 5, Elsevier, Amsterdam, 3159-3228.
- Hujer, R., S. Thomsen und C. Zeiss (2006). “The Effects of Vocational Training Programmes on the Duration of Unemployment in Eastern Germany” *Allgemeines Statistisches Archiv* 90(2), 299-321.
- Imbens, G. (2000). “The Role of the Propensity Score in Estimating Dose-Response Functions” *Biometrika* 87, 706-710.
- Kluve, J. H. Lehmann, and C. Schmidt (2004). “Disentangling Treatment effects of labor market histories: the role of employment histories.” Discussion Paper, RWI, Essen.
- Kluve, J., and C. Schmidt (2002). “Can Training and Employment Subsidies Combat European Unemployment?” *Economic Policy*, 35, 411-448.
- Kraus, F., P. Puhani, and V. Steiner (1999). “Employment effects of publicly financed training programs - The East German Experience.” *Jahrbücher für Nationalökonomie und Statistik*, 219, 216–248.
- Lechner, M. (2000). “An evaluation of public sector sponsored continuous vocational training programs in East Germany.” *Journal of Human Resources*, 35, 347–375.

- Lechner, M. (2001). "Identification and Estimation of Causal Effects of Multiple Treatments under the Conditional Independence Assumption." In: M. Lechner and F. Pfeifer (eds.) (2000), *Econometric Evaluation of Active Labor Market Politics in Europe*, Heidelberg: Physica-Verlag.
- Lechner, M., 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.
- Lubyova, M. and J.C. van Ours (1999). "Effects of Active Labour Market programs on the transition rate from unemployment into regular jobs in the Slovak republic." *Journal of Comparative Economics*, 27, 90–112.
- Martin, J.P. and Grubb, D. (2001) "What works and for whom: A review of OECD countries' experiences with active labour market policies." *Swedish Economic Policy Review*, 8, 9–56.
- OECD (2005) "Labour Market Programmes and Activation Strategies: Evaluating the Impacts." Chapter 4 of *Employment Outlook*, OECD, Paris.
- Puhani, P. (1999) "Evaluating active labour market policies - empirical evidence for Poland during transition." *ZEW Economic Studies*, 5, Physica, Heidelberg.
- 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. Steinwede und 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. (2003) "Differential Effects of Swedish Active Labour Market Programs for Unemployed Adults in the 1990s." Discussion Paper, Institute for Fiscal Studies, London.
- 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.
- Smith, J.A. and P. Todd (2005). "Rejoinder." *Journal of Econometrics* 125, 365–375.
- Wunsch, C. (2006). "Labour Market Policy in Germany: Institutions, Instruments and Reforms since Unification." Discussion Paper, University of St. Gallen.

# Appendix

## Descriptive Statistics and Description of Data

Table 1: Participation in First Training Program for the Inflow Samples into Unemployment

Training Program	Frequency	Percent of inflow sample	Percent among treated
Women			
Practice Firm	145	2.4	9.4
SPST	1,210	19.7	78.1
Retraining	195	3.2	12.6
No training program above	4,585	74.7	–
Total inflow sample	6,135	100	100
Men			
Practice Firm	73	1.2	8.7
SPST	528	8.9	63.2
Retraining	234	4.0	28.0
No training program above	5,076	85.9	–
Total inflow sample	5,911	100	100

Remark: Programs that start before a new job is found are considered. We exclude training programs which start together with a job (like integration subsidies) or which involve a very small number of participants since they are not targeted on inflows into unemployment (as career advancement and German language courses).

Table 2: Number of Training Spells and Length of Unemployment before Program Start

	Women	Men
Practice Firm		
1–2 quarters	37	40
3–4 quarters	51	15
5–8 quarters	48	14
>8 quarters	9	4
Total	145	73
SPST		
1–2 quarters	254	200
3–4 quarters	374	141
5–8 quarters	435	144
>8 quarters	147	43
Total	1,210	528
Retraining		
1–2 quarters	61 (61)	113 (107)
3–4 quarters	76 (75)	82 (79)
5–8 quarters	53 (53)	35 (33)
>8 quarters	5 (5)	4 (4)
Total	195 (194)	234 (223)

Remark: The time intervals indicate the quarter of program start relative to the beginning of the unemployment spell. The numbers in parenthesis for RT are participants who are less than 51 years old when entering unemployment.

Table 3: Elapsed Duration of Unemployment in Months at Beginning of Training Spell

	Women	Men
Practice Firm		
Average	10.9	8.1
25%-Quantile	5	2
Median	10	5
75%-Quantile	14	11
SPST		
Average	12.8	10.4
25%-Quantile	6	4
Median	11	7.5
75%-Quantile	18	15
Retraining		
Average	8.9	6.8
25%-Quantile	4	3
Median	8	6
75%-Quantile	12	10

Table 4: Realized Duration of Training Spells in months

	Women	Men
Practice Firm		
Average	6.5	6.1
25%-Quantile	6	4
Median	6	6
75%-Quantile	7	8
SPST		
Average	9.1	8.8
25%-Quantile	6	4
Median	10	9
75%-Quantile	12	12
Retraining		
Average	18.7	17.3
25%-Quantile	15	12
Median	21	21
75%-Quantile	22	22

Remark: The duration of the training spell is defined as the number of months of continuous training. No interruptions are allowed. If in any month we do not identify the program we assume the program has ended the month before.



Table 5: Program fields of (target) profession

		Program field (see below)												
	Stratum	1	2	3	4	5	6	7	8	9	10	11	12	missing
		Women												
PF	1	14	5	5	3	0	0	8	0	38	0	0	27	0
PF	2	6	4	6	2	4	2	4	4	43	0	6	20	0
PF	3	4	2	2	2	2	6	4	0	48	0	8	21	0
SPST	1	5	0	4	0	3	9	1	18	30	6	11	13	26
SPST	2	4	2	5	0	4	11	2	10	20	8	14	20	30
SPST	3	3	2	2	1	2	11	0	7	21	6	13	31	31
RT	1	6	0	4	2	0	10	2	8	12	20	28	8	18
RT	2	9	1	4	6	4	1	0	1	16	22	27	6	12
RT	3	12	2	4	8	4	8	0	0	25	10	17	8	9
		Men												
PF	1	2	25	0	42	2	0	0	2	2	0	5	18	0
SPST	1	3	27	1	13	13	3	8	4	1	0	22	5	29
SPST	2	2	23	1	17	19	4	5	6	1	2	13	5	34
SPST	3	2	10	1	7	15	0	7	4	1	0	21	32	43
RT	1	3	27	0	43	2	1	7	0	0	2	12	3	17
RT	2	1	26	1	40	6	4	7	0	1	6	6	1	12
RT	3	0	12	0	56	3	0	3	0	6	3	15	3	3
Total		4	9	3	10	5	6	3	6	16	5	14	17	25

Remark: The table shows the distribution of the fields of profession for the programs by stratum and gender in percent of the nonmissing information. The fields are the following: 1 agriculture, basic materials, leather, textiles 2 metal 3 food 4 construction 5 technical 6 retail sales 7 transport 8 accounting 9 office 10 health 11 services 12 broader program. The last column gives the share of missing information.

## Estimated Effects of Further Training Measures

Figures 1–6 display the estimated treatment effects  $\hat{\theta}(k; \tau)$  on the horizontal axis against quarter  $\tau \geq 0$  since the beginning of treatment or quarter  $\tau < 0$  before the beginning of treatment. The time axis is divided into three parts by two vertical lines, which denote the last quarter before the unemployment spell starts and the treatment start  $\tau = 0$ , respectively. The left part shows the four quarters before unemployment starts, the middle part the gap between the beginning of the unemployment spell and the beginning of treatment and the right part the time since treatment start.

Figure 1: Practice Firm (PF) for Women

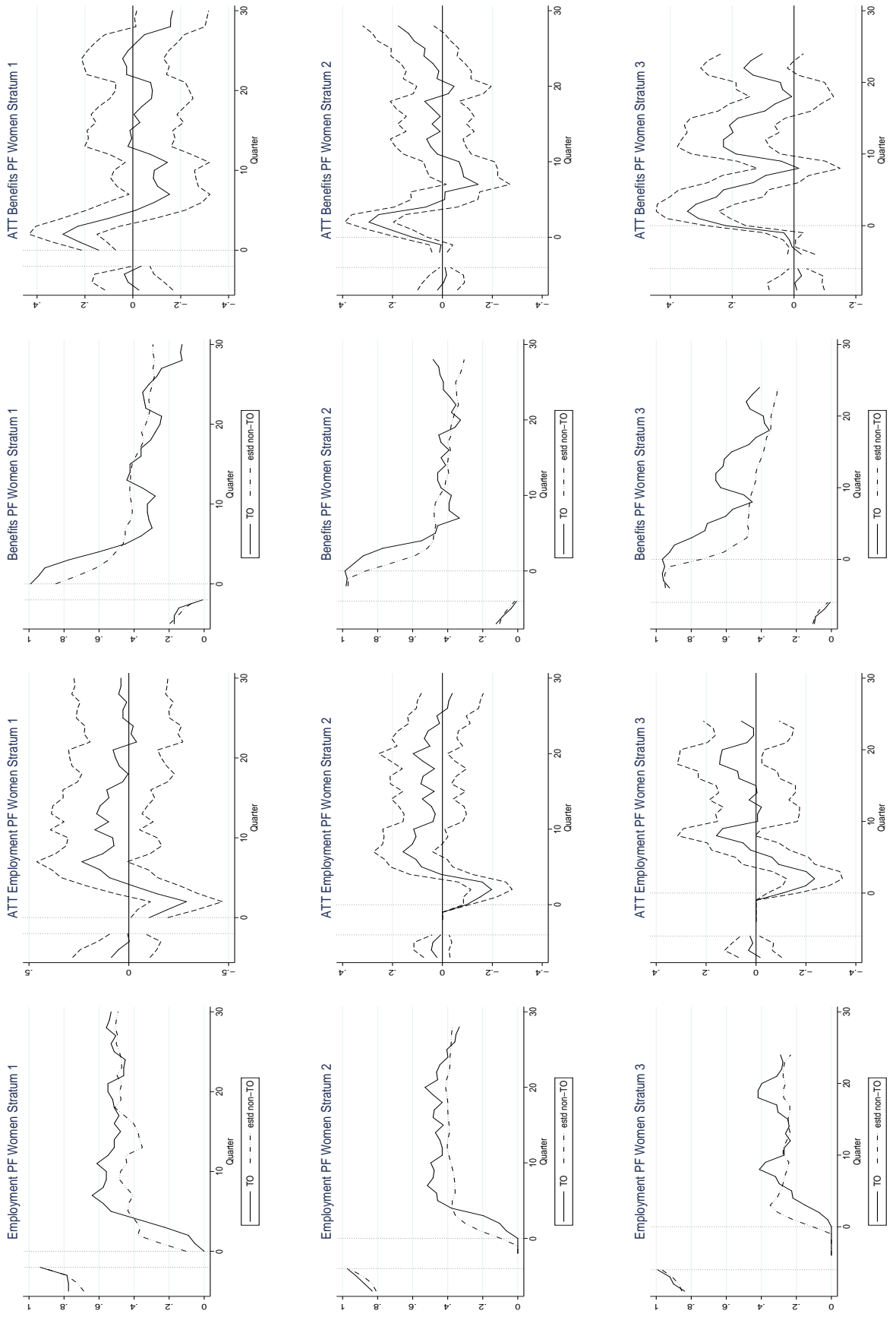


Figure 2: Practice Firm (PF) for Men

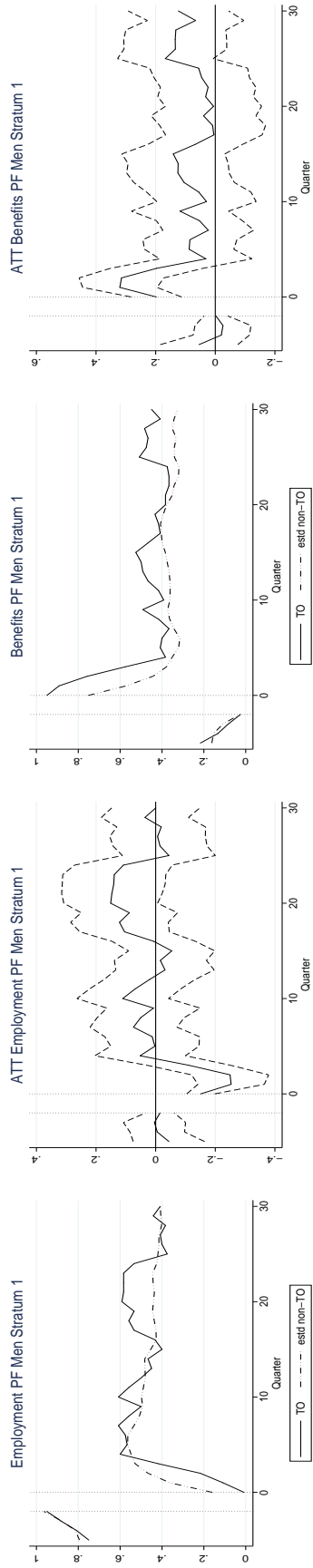


Figure 3: SPST for Women

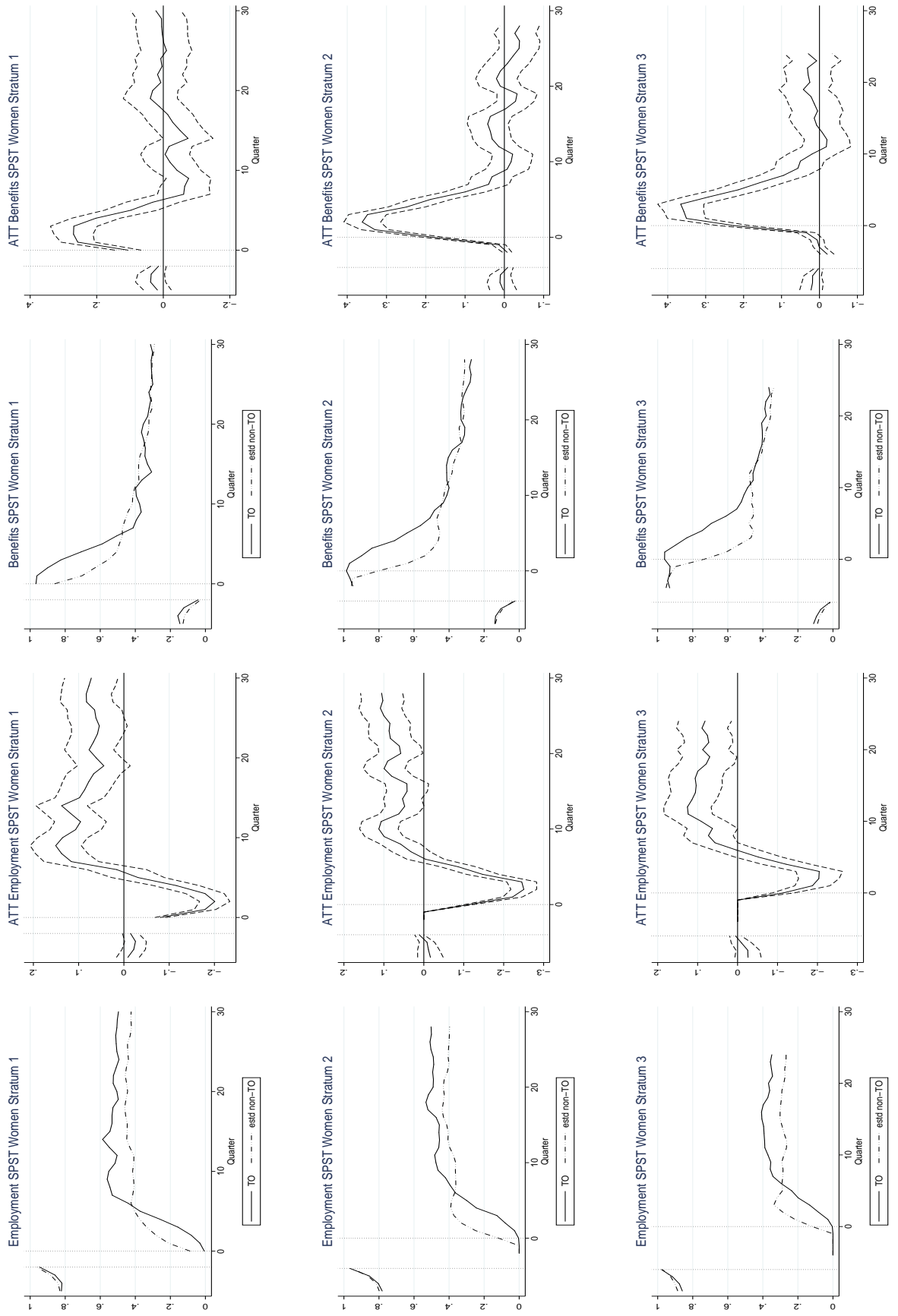


Figure 4: SPST for Men

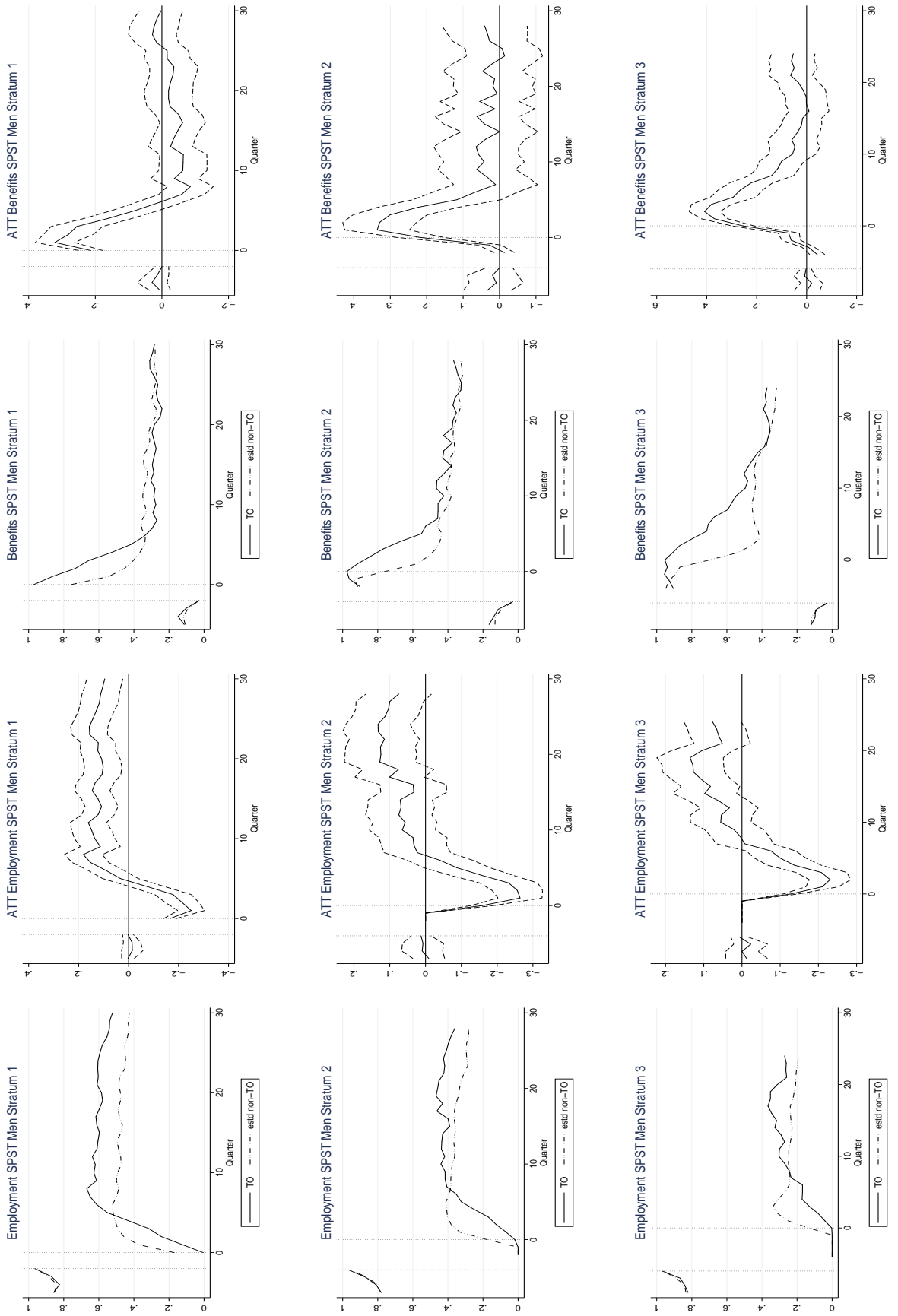


Figure 5: Retraining (RT) for Women

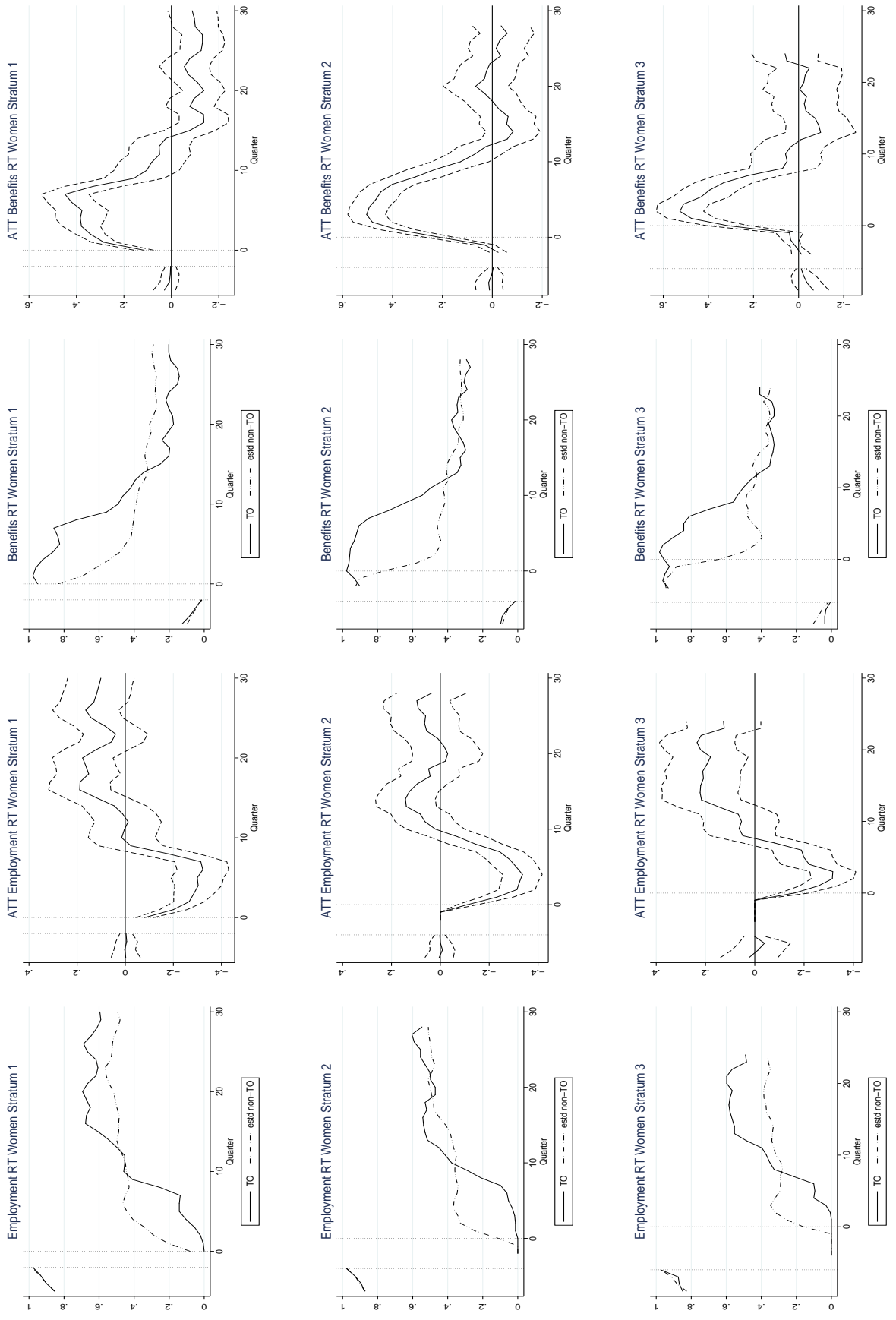


Figure 6: Retraining (RT) for Men

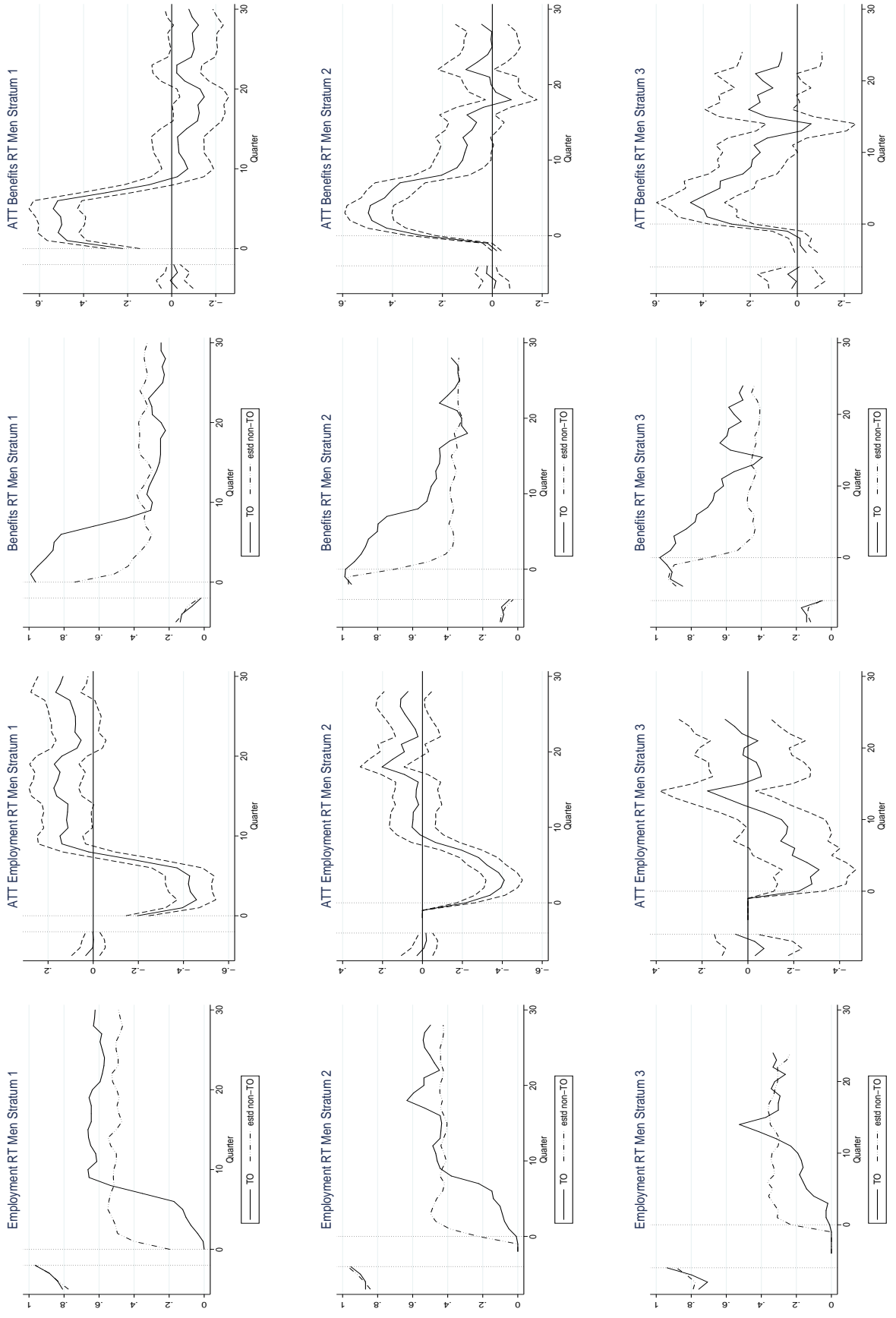


Table 6: Cumulated and Average Employment Effects

Stratum	Cumulated Effects			Average Effects		
	Q0-Q7	Q0-Q15	Q0-Q23	Q4-Q23	Q8-Q23	Q8-Q23
Women						
PF 1	-0.27 (0.60)	0.70 (1.21)	0.98 (1.81)	0.085 (0.085)	0.078 (0.088)	0.078 (0.088)
PF 2	-0.25 (0.27)	0.31 (0.58)	0.88 (0.86)	0.074 (0.040)*	0.071 (0.044)	0.071 (0.044)
PF 3	-0.82 (0.37)**	-0.54 (0.77)	0.08 (1.19)	0.041 (0.056)	0.057 (0.061)	0.057 (0.061)
SPST 1	-0.65 (0.13)***	0.32 (0.26)	0.86 (0.38)**	0.075 (0.018)***	0.095 (0.020)***	0.095 (0.020)***
SPST 2	-1.04 (0.10)***	-0.45 (0.22)**	0.15 (0.37)	0.049 (0.017)***	0.074 (0.019)***	0.074 (0.019)***
SPST 3	-0.82 (0.17)***	-0.02 (0.33)	0.66 (0.51)	0.068 (0.024)***	0.092 (0.025)***	0.092 (0.025)***
RT 1	-2.07 (0.25)***	-2.09 (0.54)***	-1.00 (0.84)	-0.008 (0.041)	0.067 (0.049)	0.067 (0.049)
RT 2	-2.15 (0.26)***	-1.80 (0.57)***	-1.61 (0.91)*	-0.032 (0.043)	0.034 (0.048)	0.034 (0.048)
RT 3	-1.74 (0.31)***	-0.70 (0.69)	0.89 (1.03)	0.097 (0.049)**	0.164 (0.055)***	0.164 (0.055)***
Men						
PF 1	-0.62 (0.40)	-0.47 (0.71)	0.43 (1.05)	0.060 (0.048)	0.066 (0.052)	0.066 (0.052)
SPST 1	-0.61 (0.19)***	0.48 (0.39)	1.48 (0.56)***	0.114 (0.026)***	0.131 (0.027)***	0.131 (0.027)***
SPST 2	-1.17 (0.24)***	-0.74 (0.49)	0.10 (0.74)	0.050 (0.035)	0.079 (0.036)**	0.079 (0.036)**
SPST 3	-1.10 (0.18)***	-0.69 (0.38)*	0.09 (0.59)	0.044 (0.028)	0.074 (0.030)***	0.074 (0.030)***
RT 1	-2.88 (0.30)***	-1.96 (0.56)***	-0.96 (0.82)	0.026 (0.037)	0.120 (0.040)***	0.120 (0.040)***
RT 2	-2.54 (0.28)***	-2.36 (0.54)***	-1.64 (0.81)**	-0.014 (0.037)	0.056 (0.042)	0.056 (0.042)
RT 3	-1.88 (0.53)***	-2.13 (1.06)**	-2.20 (1.61)	-0.056 (0.073)	-0.020 (0.077)	-0.020 (0.077)

Remark: The cumulated (average) effects are the sum (average) of the quarter specific average treatment effect on the treated over the respective quarters. \*, \*\* and \*\*\* denote significance at the 10%, 5% and 1% level, respectively, and QiQj denotes quarter  $i$  to quarter  $j$  since beginning of treatment.



Table 7: Cumulated and Average Effects on Benefit Reciprocity

Stratum	Cumulated Effects			Average Effects		
	Q0-Q7	Q0-Q15	Q0-Q23	Q4-Q23	Q8-Q23	Q8-Q23
Women						
PF 1	0.73 (0.48)	0.27 (0.85)	0.02 (1.21)	-0.043 (0.057)	-0.045 (0.059)	-0.045 (0.059)
PF 2	0.78 (0.30)***	0.73 (0.62)	0.86 (0.86)	-0.001 (0.040)	0.005 (0.045)	0.005 (0.045)
PF 3	1.89 (0.36)***	3.15 (0.70)***	3.83 (1.03)***	0.132 (0.048)***	0.121 (0.053)**	0.121 (0.053)**
SPST 1	1.17 (0.20)***	0.82 (0.34)**	0.89 (0.51)*	-0.000 (0.024)	-0.017 (0.026)	-0.017 (0.026)
SPST 2	1.80 (0.13)***	1.91 (0.24)***	1.91 (0.35)***	0.035 (0.016)**	0.007 (0.018)	0.007 (0.018)
SPST 3	2.06 (0.18)***	2.16 (0.33)***	2.34 (0.49)***	0.052 (0.023)**	0.018 (0.024)	0.018 (0.024)
RT 1	2.75 (0.26)***	3.47 (0.51)***	2.65 (0.71)***	0.077 (0.034)**	-0.006 (0.040)	-0.006 (0.040)
RT 2	3.38 (0.25)***	3.94 (0.50)***	4.01 (0.76)***	0.122 (0.035)***	0.039 (0.040)	0.039 (0.040)
RT 3	3.21 (0.36)***	3.15 (0.74)***	3.01 (1.08)***	0.059 (0.050)	-0.013 (0.054)	-0.013 (0.054)
Men						
PF 1	1.26 (0.39)***	2.01 (0.69)***	2.24 (1.03)**	0.060 (0.048)	0.061 (0.054)	0.061 (0.054)
SPST 1	1.26 (0.19)***	0.84 (0.35)**	0.57 (0.51)	-0.025 (0.023)	-0.043 (0.025)*	-0.043 (0.025)*
SPST 2	1.60 (0.30)***	1.94 (0.58)***	2.17 (0.84)**	0.050 (0.040)	0.036 (0.041)	0.036 (0.041)
SPST 3	2.34 (0.23)***	2.79 (0.46)***	3.01 (0.63)***	0.080 (0.029)***	0.042 (0.030)	0.042 (0.030)
RT 1	3.53 (0.34)***	3.32 (0.60)***	2.57 (0.81)***	0.044 (0.036)	-0.060 (0.038)	-0.060 (0.038)
RT 2	3.38 (0.28)***	4.32 (0.48)***	4.55 (0.69)***	0.143 (0.032)***	0.073 (0.035)**	0.073 (0.035)**
RT 3	2.80 (0.43)***	3.71 (0.71)***	4.82 (1.12)***	0.165 (0.051)***	0.127 (0.058)**	0.127 (0.058)**

Remark: The cumulated (average) effects are the sum (average) of the quarter specific average treatment effect on the treated over the respective quarters. \*, \*\* and \*\*\* denote significance at the 10%, 5% and 1% level, respectively, and QiQj denotes quarter  $i$  to quarter  $j$  since beginning of treatment.

# Additional Appendix to “Long-Run Effects of Training Programs for the Unemployed in East Germany”

## A Estimation Results for the Propensity Score

### A.1 Sample Sizes

#### A.1.1 Sample Sizes by Stratum

Women				
	Stratum 1	Stratum 2	Stratum 3	
Waiting	5783 (4652)	3855 (2996)	2294 (1671)	
PF	37	51	48	
SPST	254	374	435	
RT	61 (61)	76 (75)	53 (53)	
Men				
	Stratum 1	Stratum 2	Stratum 3	
Waiting	5558 (4444)	2705 (2046)	1381 (997)	
PF	40	15	14	
SPST	200	141	144	
RT	113 (107)	82 (79)	35 (33)	

Remark: Numbers in Parentheses exclude the 51–55 year old. We use this further restricted sample to evaluate RT. We do not evaluate PF for males in stratum 2 and 3 due to the small sample size.

### A.1.2 Sample Sizes by Quarter

Quarter of unempl.	Inflows	Outflows	Job	PF	SPST	RT	controls	share trt later	alternative controls	alt share trt later
Women										
1	6135	621	12	119	33	5783	0.207	5971	0.232	
2	5350	806	25	135	28	5162	0.232	5162	0.232	
3	4356	582	28	206	42	3855	0.181	4080	0.226	
4	3498	443	23	168	34	3273	0.213	3273	0.213	
5	2830	408	21	157	31	2294	0.070	2621	0.186	
6	2213	241	10	115	11	1886	0.085	2077	0.169	
7	1836	155	10	99	7	1645	0.098	1720	0.137	
8	1565	127	7	64	4	1490	0.108	1490	0.108	
9+	1363	1202	9	147	5		mean		mean	
							0.175		0.206	
Men										
1	5911	1315	20	80	57	5558	0.087	5754	0.118	
2	4439	1300	20	120	56	4243	0.114	4243	0.114	
3	2943	729	10	84	52	2705	0.090	2797	0.120	
4	2068	402	5	57	30	1976	0.123	1976	0.123	
5	1574	272	4	52	23	1381	0.037	1495	0.110	
6	1223	176	2	34	6	1109	0.046	1181	0.104	
7	1005	102	3	29	3	933	0.055	970	0.091	
8	868	68	5	29	3	831	0.061	831	0.061	
9+	763	712	4	43	4		mean		mean	
							0.088		0.113	

Remark: The table shows quarter by quarter of elapsed unemployment duration the number of those who are still unemployed at the beginning of the quarter (inflows) and the number of those who during the quarter start a job (job) or a treatment (PF, SPST, RT). Controls are all those who are still unemployed at the beginning of the quarter but do not start a treatment during the stratum. The share of the controls who start a treatment during a later stratum is also given. An alternative definition of control persons (not pursued in the paper) would take as controls all those, who are still unemployed at the beginning of the quarter but do not start a treatment during the *quarter*. This would lead to a slightly higher share of controls who receive treatment later. The means are weighted means. The table considers the sample age 25–55 at the beginning of unemployment. The restricted version age 25–50 for RT is available upon request from the authors. The number for outflows in jobs in quarter 9+ include those, who never again start a job.

## A.2 Variable Definitions

Table 8: Variable Definitions

Label	Definition
Personal Attributes	
aXYY	Age at start of unemployment $\geq XX$ and $\leq YY$
age	Age at start of unemployment
married	Married
qual.l	No vocational training degree or education information missing
qual.m	Vocational training degree
qual.h	University/College degree
Last Employment	
BER1	Apprentice
BER2	Blue Collar Worker
BER3	White Collar Worker
BER4	Worker at home with low hours or BER missing
BER5	Part-time working
pearn	Daily earnings $\geq 15$ Euro per day in 1995 Euro
earnlow	Daily earnings $< 15$ Euro per day in 1995 Euro
earncens	Earnings censored at social security taxation threshold
earn	Daily earnings
logearn	$\log(\text{earn})$ if $\text{pearn}=1$ and $\text{earncens}=0$ , otherwise zero
Last Employer	
industry1	Agriculture
industry2	Basic materials
industry3	Metal, vehicles, electronics
industry4	Light industry
industry5	Construction
industry6	Production oriented services, trade, banking
industry7	Consumer oriented services, organization and social services
frmsize1	Firm Size (employment) missing or $\leq 10$
frmsize2	Firm Size (employment) $> 10$ and $\leq 200$
frmsize3	Firm Size (employment) $> 200$ and $\leq 500$
<continued on next page>	

Table 8: Variable Definitions &lt;continued&gt;

Label	Definition
frmsize4	Firm Size (employment) > 500
Employment and Program History	
preexM	Employed M (M=6, 12) month before unemployment starts
preex12cum	Number of months employed in the last 12 months before unemployment starts, standardized
pretx1	Participation in any ALMP program reported in our data in the year before unemployment starts
Regional Information	
state1	Mecklenburg-Vorpommern
state2	Brandenburg
state3	Sachsen-Anhalt
state4	Sachsen
state5	Thüringen
popdens	population density (standardized)
Calendar Time of Entry into Unemployment	
uentry	First unemployment month (months counted from January 1993)
Interaction of Variables / Functional Form	
._sq	squared
._	interaction
All variables are defined at the time of entry into unemployment and constant during the unemployment spell.	

### A.3 Summary Statistics

The following six tables document the mean values of the variables in the three strata for women and men. The means are shown for the dynamic control group and the participants in PF, SPST and RT, respectively. Since we restrict the age for the evaluation of RT to lie between 25 and 50, we show the means for the dynamic control group also for this more restricted group and for RT only for this age group.

Table 9: **Women Stratum 1**

Variable	control	control 25–50	PF	SPST	RT 25–50
age	40.088	36.886	40.757	39.819	33.689
married	.59	.576	.514	.61	.738
qual l	.108	.098	.081	.039	.066
qual m	.843	.851	.892	.89	.869
qual h	.049	.05	.027	.071	.066
BER1	.003	.004	0	0	0
BER2	.428	.425	.432	.252	.443
BER3	.391	.399	.405	.535	.393
BER4	.002	.002	0	0	0
BER5	.176	.171	.162	.213	.164
pearn	.966	.967	.973	.996	1
earncens	.004	.004	0	.004	0
logearn	3.447	3.445	3.521	3.647	3.657
industry1	.066	.064	.027	.047	.016
industry2	.047	.044	.027	.043	.115
industry3	.064	.064	.189	.075	.066
industry4	.074	.072	.081	.071	.033
industry5	.036	.038	.081	.035	.049
industry6	.264	.269	.162	.303	.18
industry7	.45	.45	.432	.425	.541
frmsize1	.229	.235	.081	.193	.18
frmsize2	.447	.447	.676	.433	.344
frmsize3	.152	.148	.189	.177	.213
frmsize4	.173	.169	.054	.197	.262
preex6	.842	.835	.784	.846	.934
preex12	.764	.747	.757	.827	.836
preex12cum	10.153	10.052	9.811	10.354	11.033
pretx1	.077	.083	.054	.031	.066
state1	.139	.144	.081	.154	.131
state2	.167	.169	0	.118	.246
state3	.209	.207	.162	.205	.115
state4	.304	.299	.216	.362	.377

*Continued on next page...*

... table 9 continued

Variable	control	control 25–50	PF	SPST	RT 25–50
state5	.18	.179	.541	.161	.131
popdens	468.023	463.781	422.692	544.084	379.189
uentry	10.721	10.847	8.892	13	10.033
N	5783	4652	37	254	61

Table 10: **Women Stratum 2**

Variable	control	control 25–50	PF	SPST	RT 25–50
age	40.777	37.183	41.176	38.623	33.12
married	.587	.568	.667	.61	.627
qual l	.133	.122	.039	.078	.08
qual m	.828	.838	.863	.834	.893
qual h	.039	.04	.098	.088	.027
BER1	.003	.003	0	.005	0
BER2	.432	.433	.353	.294	.427
BER3	.373	.378	.49	.559	.413
BER4	.002	.002	0	0	0
BER5	.191	.185	.157	.142	.16
pearn	.963	.963	1	.979	1
earncens	.003	.003	.02	.003	.013
logearn	3.427	3.419	3.562	3.593	3.609
industry1	.058	.058	0	.037	.053
industry2	.051	.048	.078	.029	.053
industry3	.061	.058	.098	.091	.08
industry4	.073	.073	.02	.056	.027
industry5	.038	.041	0	.027	.027
industry6	.261	.263	.294	.342	.32
industry7	.457	.459	.51	.417	.44
frmsize1	.222	.228	.176	.203	.173
frmsize2	.439	.439	.353	.428	.44
frmsize3	.156	.154	.235	.144	.16

Continued on next page...

... table 10 continued

Variable	control	control 25–50	PF	SPST	RT 25–50
frmsize4	.182	.179	.235	.225	.227
preex6	.859	.85	.922	.832	.92
preex12	.793	.77	.784	.786	.867
preex12cum	10.374	10.242	10.843	10.257	11.027
pretx1	.072	.079	.059	.053	.08
state1	.135	.138	.157	.139	.227
state2	.172	.174	.02	.155	.147
state3	.216	.215	.333	.233	.147
state4	.292	.288	.275	.329	.36
state5	.185	.186	.216	.144	.12
popdens	466.479	461.806	520.108	561.464	392.724
uentry	10.321	10.433	8.157	13.024	11.733
N	3855	2996	51	374	75

Table 11: **Women Stratum 3**

Variable	control	control 25–50	PF	SPST	RT 25–50
age	41.668	37.309	43.604	40.609	33.453
married	.589	.57	.542	.591	.547
qual l	.159	.147	.083	.099	.075
qual m	.805	.814	.896	.857	.906
qual h	.036	.038	.021	.044	.019
BER1	.003	.003	.021	0	0
BER2	.449	.454	.292	.368	.528
BER3	.341	.342	.583	.446	.434
BER4	.001	.002	0	.002	0
BER5	.206	.199	.104	.184	.038
pearn	.959	.959	.938	.982	1
earncens	.004	.003	0	.002	0
logearn	3.395	3.388	3.452	3.562	3.674
industry1	.06	.06	.042	.048	.038

*Continued on next page...*



... table 11 continued

Variable	control	control 25–50	PF	SPST	RT 25–50
industry2	.051	.046	.125	.064	.075
industry3	.063	.059	.104	.057	.094
industry4	.075	.073	.042	.094	.075
industry5	.041	.045	0	.037	.038
industry6	.246	.247	.313	.308	.189
industry7	.464	.47	.375	.391	.491
frmsize1	.221	.23	.188	.152	.075
frmsize2	.436	.437	.438	.453	.396
frmsize3	.158	.151	.063	.182	.17
frmsize4	.185	.181	.313	.214	.358
preex6	.866	.855	.917	.906	.887
preex12	.804	.774	.813	.853	.849
preex12cum	10.448	10.284	10.979	10.906	10.755
pretx1	.06	.068	.083	.078	.057
state1	.121	.121	.229	.103	.189
state2	.167	.168	.021	.2	.094
state3	.228	.228	.229	.172	.208
state4	.289	.284	.333	.368	.226
state5	.195	.199	.188	.156	.283
popdens	453.54	451.059	642.06	559.151	534.315
uentry	10.242	10.373	9.458	9.874	6.755
N	2294	1671	48	435	53

Table 12: Men Stratum 1

Variable	control	control 25–50	PF	SPST	RT 25–50
age	40.124	36.84	36.975	40.035	33.449
married	.466	.435	.475	.56	.421
qual l	.094	.089	.125	.06	.075
qual m	.842	.851	.825	.75	.897
qual h	.063	.059	.05	.19	.028

Continued on next page...

... table 12 continued

Variable	control	control 25–50	PF	SPST	RT 25–50
BER1	.001	.002	.025	0	0
BER2	.794	.804	.85	.6	.86
BER3	.167	.161	.075	.34	.112
BER4	0	0	0	.005	0
BER5	.037	.033	.05	.055	.028
pearn	.991	.991	1	.995	.991
earncens	.015	.014	0	.03	0
logearn	3.647	3.649	3.72	3.711	3.727
industry1	.079	.074	.05	.05	.084
industry2	.072	.075	.075	.115	.103
industry3	.116	.113	.1	.165	.196
industry4	.056	.056	.05	.05	.065
industry5	.207	.227	.175	.115	.121
industry6	.221	.234	.15	.29	.262
industry7	.248	.222	.4	.215	.168
frmsize1	.255	.27	.175	.19	.131
frmsize2	.501	.503	.425	.49	.533
frmsize3	.12	.109	.1	.115	.121
frmsize4	.124	.118	.3	.205	.215
preex6	.845	.835	.875	.845	.869
preex12	.776	.761	.75	.87	.776
preex12cum	10.223	10.097	10.175	10.56	10.533
pretx1	.064	.069	.025	.03	.075
state1	.14	.141	.2	.2	.168
state2	.147	.152	.05	.13	.14
state3	.209	.213	.4	.165	.243
state4	.323	.319	.15	.35	.346
state5	.18	.175	.2	.155	.103
popdens	460.115	463.21	340.144	604.944	499.402
uentry	10.6	10.657	9.425	12.01	10.533
N	5558	4444	40	200	107

Table 13: **Men Stratum 2**

Variable	control	control 25–50	PF	SPST	RT 25–50
age	41.316	37.473	38.933	40.099	34.443
married	.44	.396	.333	.44	.405
qual l	.12	.118	.133	.078	.114
qual m	.805	.812	.867	.801	.797
qual h	.075	.07	0	.121	.089
BER1	.002	.002	0	0	0
BER2	.761	.771	1	.645	.823
BER3	.197	.19	0	.319	.165
BER4	0	0	0	0	0
BER5	.04	.037	0	.035	.013
pearn	.988	.988	1	.993	1
earncens	.019	.016	0	.021	.013
logearn	3.61	3.615	3.67	3.666	3.715
industry1	.06	.053	.2	.035	.051
industry2	.078	.081	0	.057	.089
industry3	.115	.107	.067	.128	.165
industry4	.053	.055	.133	.043	.076
industry5	.174	.195	.133	.149	.139
industry6	.23	.242	.133	.312	.241
industry7	.29	.266	.333	.277	.241
frmsize1	.229	.247	.133	.199	.19
frmsize2	.487	.489	.6	.447	.57
frmsize3	.131	.118	.2	.163	.139
frmsize4	.153	.146	.067	.191	.101
preex6	.839	.82	.867	.844	.886
preex12	.783	.761	.733	.787	.861
preex12cum	10.216	10.005	10.067	10.277	10.772
pretx1	.062	.07	.067	.043	.038
state1	.134	.135	.333	.113	.19
state2	.152	.157	0	.106	.114
state3	.232	.236	.333	.284	.177
state4	.308	.304	.133	.355	.304

*Continued on next page...*

... table 13 continued

Variable	control	control 25–50	PF	SPST	RT 25–50
state5	.174	.167	.2	.142	.215
popdens	487.763	491.47	361.654	646.669	410.726
uentry	10.461	10.605	13.267	13.376	9.405
N	2705	2046	15	141	79

Table 14: Men Stratum 3

Variable	control	control 25–50	PF	SPST	RT 25–50
age	42.142	37.872	41.929	42.201	34.364
married	.433	.391	.5	.41	.333
qual l	.125	.123	.071	.153	.242
qual m	.791	.794	.929	.674	.727
qual h	.084	.082	0	.174	.03
BER1	.001	.002	0	.007	0
BER2	.736	.741	.714	.611	.788
BER3	.218	.217	.286	.333	.182
BER4	0	0	0	0	0
BER5	.045	.04	0	.049	.03
pearn	.986	.987	1	.986	.97
earncens	.025	.022	0	.056	0
logearn	3.572	3.579	3.651	3.503	3.669
industry1	.056	.045	.071	.021	.03
industry2	.078	.077	.071	.069	.091
industry3	.122	.113	.143	.125	.091
industry4	.053	.06	.071	.035	.091
industry5	.152	.165	.071	.167	.182
industry6	.227	.237	.357	.236	.273
industry7	.311	.302	.214	.347	.242
frmsize1	.236	.259	.429	.16	.182
frmsize2	.461	.458	.5	.528	.455
frmsize3	.133	.116	0	.125	.182

Continued on next page...

... table 14 continued

Variable	control	control 25–50	PF	SPST	RT 25–50
frmsize4	.169	.166	.071	.188	.182
preex6	.857	.839	.857	.861	.788
preex12	.789	.767	.786	.819	.788
preex12cum	10.319	10.129	9.857	10.493	9.667
pretx1	.052	.055	.071	.063	.121
state1	.127	.128	.357	.146	.394
state2	.165	.176	0	.153	.061
state3	.217	.221	.214	.229	.333
state4	.322	.312	.071	.313	.091
state5	.169	.163	.357	.16	.121
popdens	508.681	505.011	311.301	555.184	460.436
uentry	10.345	10.395	10.5	10.833	8.758
N	1381	997	14	144	33

#### A.4 Results of Propensity Score Estimations and Balancing Tests

Remark: The propensity score tables show the estimated coefficients of the probit regressions of the conditional probability to participate in the program mentioned in the header against the alternative of not taking part in any program in the stratum. The estimations are carried out separately for each time window of elapsed unemployment duration (Stratum 1, 2, and 3). Standard errors are in parentheses. \*, \*\*, \*\*\* means significant at the 10%, 5%, 1% level, respectively, in a two-sided test. Each probit table is followed by a table indicating how many regressors pass the Smith/Todd (2005) balancing test at different significance levels using a cubic and a quartic of the propensity score, respectively. Graphs with the densities of the propensity scores are in the next subsection.

Table 15: **Propensity Score Estimates Women Practice Firm**

COEFFICIENT	Stratum 1		Stratum 2		Stratum 3	
age	0.0669	(0.073)	0.0972	(0.065)	0.0615	(0.075)
age_sq	-0.000798	(0.00089)	-0.00118	(0.00080)	-0.000643	(0.00089)
married	-0.221*	(0.13)			-0.169	(0.13)
qual_l	0.254	(0.43)				
qual_m	0.255	(0.39)				
BER2	-0.0267	(0.13)			-0.0419	(0.20)
entglow	0.292	(0.55)				
logearn_sq	0.0298	(0.029)	0.0213	(0.018)		
frmsize2	0.521***	(0.17)				
frmsize3	0.457**	(0.21)				
state3	0.399*	(0.24)			0.796**	(0.36)
state4	0.337	(0.22)	0.707**	(0.33)	0.840**	(0.36)
state5	0.901***	(0.21)			0.779**	(0.37)
BER3_married			0.193	(0.12)		
preex12			-0.0887	(0.14)	-0.0445	(0.16)
state1			0.801**	(0.34)	1.084***	(0.37)
state35			0.870***	(0.32)		
popdens			0.0436	(0.11)		
popdens_sq			-0.0213	(0.064)		
BER3					0.333*	(0.19)
pearn					-1.260*	(0.74)
logearn					0.283	(0.19)
uentry_sq					-0.000311	(0.00040)
Constant	-5.174***	(1.53)	-5.103***	(1.36)	-3.966**	(1.56)
Observations	5820		3906		2342	
PseudoR <sup>2</sup>	0.0977		0.0397		0.0685	

Table 17: **Propensity Score Estimates Women SPST**

COEFFICIENT	Stratum 1		Stratum 2		Stratum 3	
a3034	0.0240	(0.11)	0.123	(0.098)	0.247**	(0.11)
a3539	0.482*	(0.26)	0.107	(0.10)	-0.0693	(0.14)
a4044	0.523**	(0.26)	-0.0184	(0.10)	0.0334	(0.14)
a4549	0.124	(0.30)	-0.223*	(0.12)	0.0154	(0.14)
a5055	0.154	(0.30)	-0.310***	(0.10)	-0.288**	(0.13)

married	-0.00535	(0.063)	0.0481	(0.059)	-0.225**	(0.11)
qual_l	-0.178*	(0.10)	-0.0439	(0.084)	-0.199**	(0.085)
BER2	-0.313***	(0.089)	-0.0569	(0.088)	-0.0701	(0.086)
BER3	-0.0116	(0.087)	0.250***	(0.089)	0.0820	(0.094)
logearn_sq	0.0529***	(0.012)				
earncens	0.510	(0.51)	0.842	(0.68)	0.822	(0.70)
state1	-0.0397	(0.094)			-0.184*	(0.10)
state2	-0.248**	(0.098)			0.0280	(0.091)
state3	-0.0977	(0.085)			-0.282***	(0.088)
state5	-0.152*	(0.091)			-0.214**	(0.092)
uentry	-0.0546***	(0.015)	0.113***	(0.017)		
uentry_sq	0.00324***	(0.00063)	-0.00339***	(0.00069)		
preex12	0.652***	(0.21)				
preex12_a3544	-0.607**	(0.29)				
preex12_a4555	-0.748**	(0.35)				
preex12cum_sq	-0.0601	(0.071)	0.0410***	(0.015)	0.0333	(0.041)
preex12cum_sq_a3544	0.0320	(0.095)				
preex12cum_sq_a4555	0.318***	(0.10)				
preex12cum	-0.244*	(0.15)			0.168**	(0.076)
preex12cum_a3544	0.132	(0.20)				
preex12cum_a4555	0.636***	(0.24)				
pretx1	-0.453***	(0.16)	-0.286**	(0.13)	0.333***	(0.12)
qual_h			0.227*	(0.13)	-0.276	(0.17)
pearn			-1.080***	(0.40)	-0.765*	(0.43)
logearn			0.350***	(0.10)	0.307***	(0.11)
industry3			0.364**	(0.15)	-0.153	(0.15)
industry4			0.0907	(0.16)	0.117	(0.14)
industry5			-0.149	(0.20)	-0.102	(0.18)
industry67			0.112	(0.11)		
industry6					0.0405	(0.11)
industry7					-0.175*	(0.10)
frmsize2					0.247***	(0.087)
frmsize3					0.259**	(0.11)
frmsize4					0.179*	(0.11)
popdens					0.0718	(0.063)
popdens_sq					0.00989	(0.037)
married_a3544					0.318**	(0.16)
married_a4555					0.271*	(0.15)

Constant	-2.601***	(0.25)	-2.445***	(0.24)	-1.346***	(0.25)
Observations	6037		4229		2729	
PseudoR <sup>2</sup>	0.0742		0.0807		0.0550	

Table 19: Propensity Score Estimates Women RT

COEFFICIENT	Stratum 1		Stratum 2		Stratum 3	
a2529	0.597***	(0.18)			1.106***	(0.37)
a3034	0.404**	(0.18)			1.206***	(0.36)
a3539	0.0882	(0.21)			0.942**	(0.37)
a4044	0.0421	(0.21)			0.818**	(0.38)
married	0.377***	(0.12)	0.172	(0.11)	-0.0573	(0.14)
qual_l	-0.159	(0.17)	-0.116	(0.16)	-0.731**	(0.29)
qual_h	-0.0540	(0.25)	-0.307	(0.32)		
BER2	0.0616	(0.15)			0.592**	(0.29)
BER3	-0.100	(0.17)			0.542*	(0.30)
earn	0.0106**	(0.0042)	0.0110***	(0.0038)		
preex12cum	0.142**	(0.064)	0.263*	(0.14)		
state1	0.211	(0.18)				
state24	0.375***	(0.12)				
uentry	-0.0989***	(0.025)	0.133***	(0.033)	-0.0368***	(0.013)
uentry_sq	0.00431***	(0.0011)	-0.00510***	(0.0014)		
age			0.120	(0.10)		
age_sq			-0.00243*	(0.0015)		
preex12cum_sq			0.159*	(0.082)		
preex12cum_sq_a2539			-0.112*	(0.066)		
state2			-0.304*	(0.18)	-0.575**	(0.26)
state3			-0.417**	(0.18)	-0.352	(0.22)
state4			-0.103	(0.15)	-0.448**	(0.22)
state5			-0.446**	(0.19)	-0.150	(0.21)
logearn_sq					0.0538**	(0.026)
frmsize4					0.376	(0.23)
frmsize4_uentry					0.00685	(0.022)
Constant	-3.083***	(0.29)	-4.104**	(1.79)	-3.394***	(0.56)
Observations	4713		3071		1724	
PseudoR <sup>2</sup>	0.0950		0.117		0.162	



Table 16: Balancing Tests Women Practice Firm

Stratum	Degree of Polynomial	P-values			Regressors
		>.10	>.05	>.01	
1	3	10	11	13	13
1	4	12	12	13	13
2	3	9	9	10	10
2	4	9	10	10	10
3	3	8	10	13	13
3	4	7	11	13	13

Table 18: Balancing Tests Women SPST

Stratum	Degree of Polynomial	P-values			Regressors
		>.10	>.05	>.01	
1	3	23	24	26	27
1	4	23	24	27	27
2	3	20	20	21	21
2	4	19	21	21	21
3	3	29	29	32	32
3	4	30	31	32	32

Table 20: Balancing Tests Women RT

Stratum	Degree of Polynomial	P-values			Regressors
		>.10	>.05	>.01	
1	3	13	14	14	15
1	4	9	10	14	15
2	3	14	14	15	15
2	4	11	13	15	15
3	3	14	14	16	16
3	4	14	15	16	16

Table 21: **Propensity Score Estimates Men Practice Firm**

COEFFICIENT	Stratum 1	
age	0.0515	(0.067)
age_sq	-0.000821	(0.00085)
BER3	-0.231	(0.21)
earn	-0.00313	(0.0050)
frmsize4	0.419***	(0.14)
state1	0.447**	(0.18)
state3	0.490***	(0.16)
state5	0.363**	(0.18)
uentry	0.0181	(0.030)
uentry_sq	-0.00119	(0.0013)
Constant	-3.367***	(1.28)
Observations	5598	
PseudoR <sup>2</sup>	0.0663	

Table 22: Balancing Tests Men Practice Firm

Stratum	Degree of Polynominal	P-values			Regressors
		>.10	>.05	>.01	
1	3	9	9	9	10
1	4	7	8	10	10

Table 23: **Propensity Score Estimates Men SPST**

COEFFICIENT	Stratum 1		Stratum 2		Stratum 3	
a2529	0.225	(0.15)				
a3034_4044	0.476***	(0.10)				
a3539	0.318**	(0.15)			0.254	(0.19)
a4549	0.307**	(0.12)	-0.334	(0.26)	0.0794	(0.22)
married	0.133*	(0.069)	-0.0324	(0.087)	-0.105	(0.097)
qual_l	-0.0892	(0.11)	0.0145	(0.11)	0.292	(0.22)
qual_h	0.470***	(0.12)	-0.104	(0.17)	0.438***	(0.16)
BER3	0.152	(0.096)	0.429*	(0.24)	0.0447	(0.22)
logearn	0.128	(0.093)	0.0582	(0.096)		
earncons	0.541	(0.43)	0.0931	(0.49)		
industry6	0.0362	(0.088)				
industry57	-0.208***	(0.078)				

frmsize2	0.163*	(0.088)			0.275**	(0.12)
frmsize3	0.148	(0.12)			0.154	(0.17)
frmsize4	0.356***	(0.11)			0.193	(0.15)
preex12	0.183	(0.12)	-0.136	(0.12)		
preex12cum_sq	0.0195	(0.026)			0.0434	(0.033)
preex12_a4055	0.124	(0.10)				
popdens	0.203***	(0.064)				
popdens_sq	-0.104**	(0.042)				
uentry	-0.0176	(0.016)	0.0954***	(0.026)		
uentry_sq	0.00145**	(0.00068)	-0.00260***	(0.00100)		
a3039			0.000506	(0.14)		
a4044			0.233	(0.15)	0.266	(0.20)
a5055			-0.674**	(0.27)	0.227	(0.19)
BER2			0.0651	(0.22)	-0.133	(0.21)
preex12_a4555			0.529**	(0.26)		
a3034					0.139	(0.18)
qual_l_a3544					-0.547*	(0.33)
qual_l_a4555					-0.291	(0.29)
preex12cum_sq_a4055					-0.123**	(0.061)
Constant	-3.019***	(0.39)	-2.551***	(0.42)	-1.603***	(0.27)
Observations	5758		2846		1525	
PseudoR <sup>2</sup>	0.0807		0.0567		0.0353	

Table 24: Balancing Tests Men SPST

Stratum	Degree of Polynomial	P-values			Regressors	
		>.10	>.05	>.01		
1	3	17	20	22	22	
1	4	17	20	22	22	
2	3	14	14	15	15	
2	4	11	13	15	15	
3	3	15	15	17	17	
3	4	13	14	17	17	

Table 25: Propensity Score Estimates Men RT

COEFFICIENT	Stratum 1	Stratum 2	Stratum 3
a2529	0.970*** (0.24)	0.843*** (0.22)	
a3034	1.034*** (0.24)	0.683*** (0.22)	
a3539	0.840*** (0.24)	0.684*** (0.22)	

a4044	0.651***	(0.25)	0.499**	(0.22)		
married	0.0561	(0.090)	0.0817	(0.11)	0.00675	(0.18)
qual_l	-0.174	(0.14)	-0.193	(0.27)	-0.179	(0.23)
qual_h	-0.604*	(0.35)			-0.244	(0.43)
logearn	0.0752	(0.092)			0.106	(0.15)
frmsize2	0.325***	(0.12)				
frmsize3	0.377**	(0.16)				
frmsize4	0.511***	(0.15)				
preex12	-0.303*	(0.17)			0.213	(0.22)
preex12cum	0.352***	(0.13)				
preex12cum_sq	0.124**	(0.056)	-0.0654*	(0.038)		
state1	0.0417	(0.13)			1.031***	(0.27)
state2	-0.0132	(0.14)			0.0906	(0.36)
state3	0.0538	(0.12)			0.716***	(0.27)
state5	-0.290**	(0.15)			0.380	(0.31)
popdens	0.250***	(0.089)				
popdens_sq	-0.192***	(0.066)				
uentry	0.00237	(0.0062)	0.0428	(0.029)		
qual_m			-0.260	(0.24)		
BER2			0.322	(0.41)		
BER3			0.140	(0.43)		
logearn_sq			0.0196	(0.019)		
uentry_sq			-0.00240*	(0.0013)		
age					-0.0502	(0.13)
age_sq					0.000270	(0.0017)
pretx1					0.530*	(0.30)
Constant	-3.085***	(0.44)	-2.761***	(0.56)	-1.474	(2.35)
Observations	4551		2125		1030	
PseudoR <sup>2</sup>	0.0814		0.0509		0.118	

Table 26: Balancing Tests Men RT

Stratum	Degree of Polynomial	P-values			Regressors
		>.10	>.05	>.01	
1	3	17	18	21	21
1	4	19	21	21	21
2	3	11	13	13	13
2	4	11	11	13	13
3	3	10	11	12	12
3	4	7	9	12	12

## A.5 Common Support

Figure 7: Densities of Propensity Scores for Women

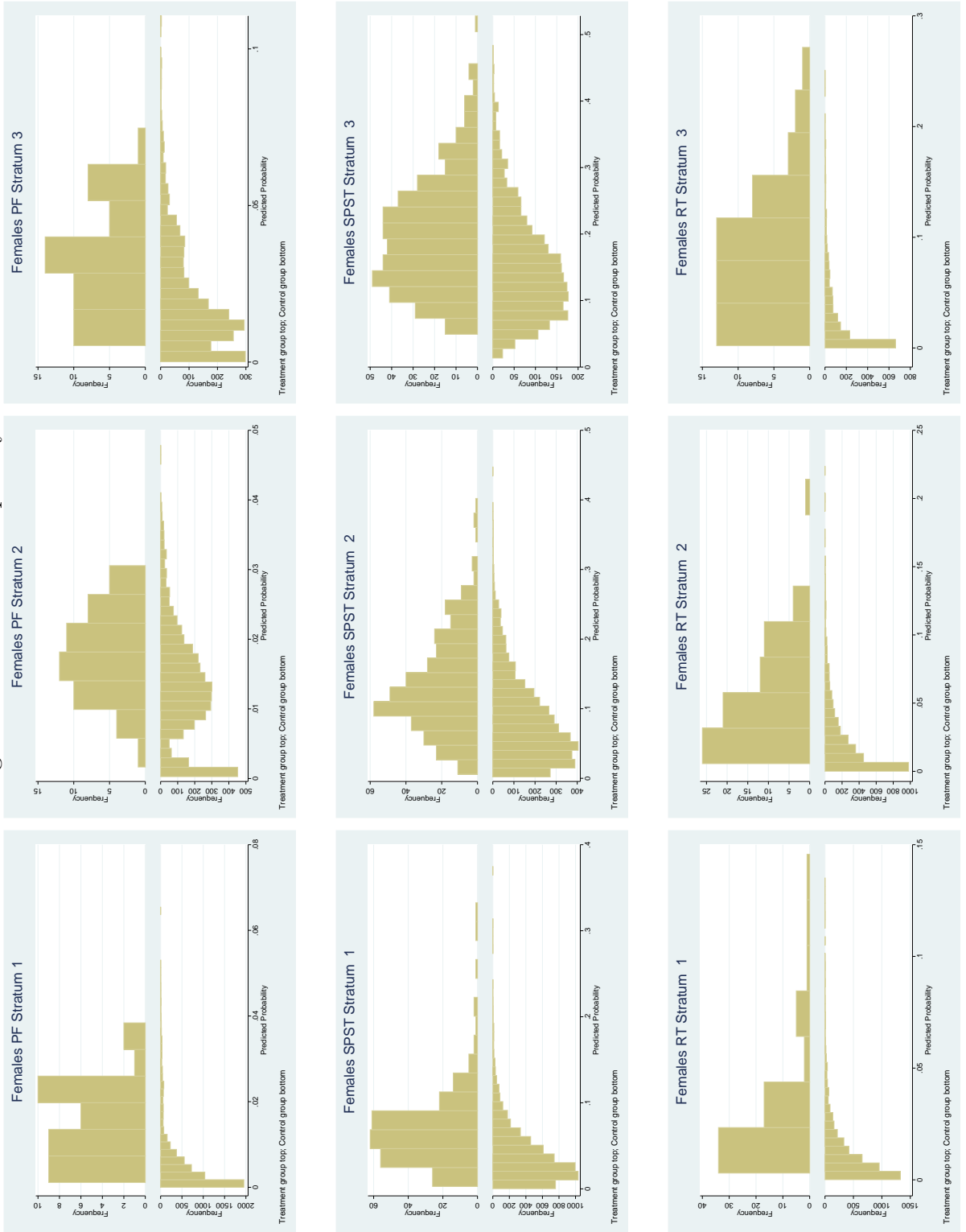
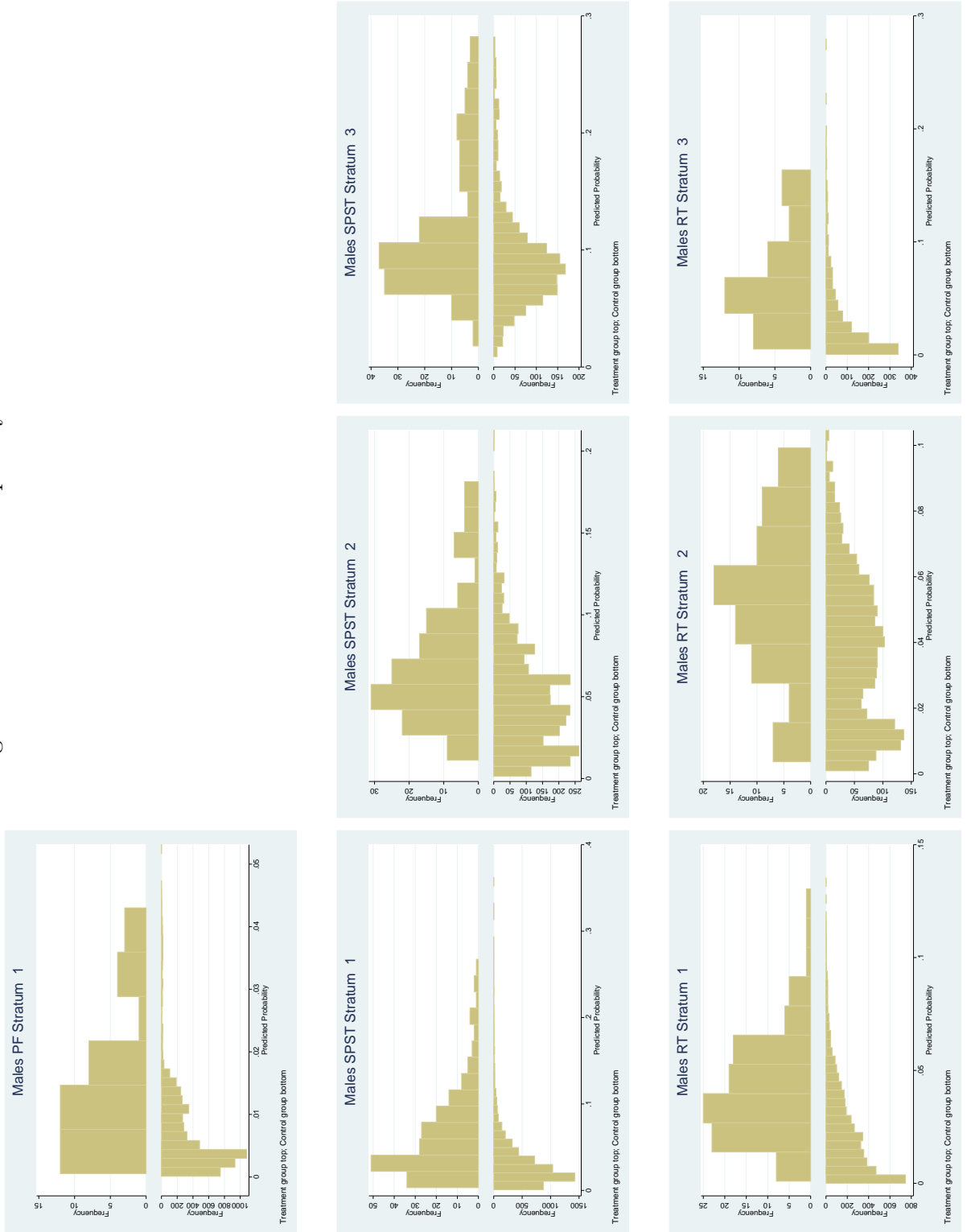


Figure 8: Densities of Propensity Scores for Men



## B Information about the data

### B.1 Other types of further training

In this study we are interested in active labor market programs for unemployed who have previously been employed and who have not already found a new job. However, we also want to give a short overview about other programs regulated by the labor promotion act (AFG) which we do not evaluate.

#### *German Courses*

The German Courses are intended for newly arrived immigrants. So the participants typically have not been employed in Germany before the German Course and hence are not part of the focus group of this study, the previously employed unemployed.

#### *Career Advancement*

These programs are typical programs directed at the employed, which were more important when the labor promotion act was introduced in 1969. By providing additional human capital the participant's risk of becoming unemployed should be lowered. Prime examples are courses in which the participants with a vocational training degree obtain additional certificates which allow them to independently run craftsman's establishments and to train trainees in the dual system of vocational training.

#### *Wage subsidies*

Wage subsidies are paid for the employment of formerly long-term unemployed and are intended to decrease the competitive disadvantage of these recruits for the period of familiarization with the skill requirement of the job. Even if the target group of wage subsidies are also unemployed we do not evaluate them because they require a job for which the wage subsidy is paid. This means provision of wage subsidies is already conditional on employment which is the success criteria for the other programs.

#### *Any program which starts together with a job*

For the same reasons why we do not evaluate wage subsidies we also do not evaluate any program which starts together with employment. Because we want to evaluate the program's effect on employment we do not consider programs which start together with employment.



## B.2 Construction of the monthly panel

The IABS employment and LED benefit payment data are daily register data whereas the FuU training data gives monthly information about program participation. This study uses the merged data as described in Bender et al. (2005). From the merged data we construct a monthly panel. If the original daily data contain more than one spell overlapping a specific month we take the information from the spell with the largest overlap as the spell defining the monthly information.

The defining condition to be part of our inflow sample into unemployment is a transition from an employment month to a nonemployment month, in which the last employment month was between December 1992 and November 1993 and thus the first unemployment month was between January 1993 and December 1994. In order to divide nonemployment (to be precise: not employed subject to social security contributions) into unemployment and other states (like labor market leavers, transition into self employment, employment as civil servant) we additionally require a month with benefit payments from the employment office within the first twelve month of nonemployment or indication of participation in any labor market program in one of our data to be part of the inflow sample in unemployment.

Later on we aggregate the information further from monthly to quarterly information. Whereas the monthly employment information is binary the quarterly employment information can take the values 0, 1/3, 2/3 or 1.

We identify program participation if a person starts a program while being in the defining unemployment spell. The participant must not be employed in the first month of the program. Otherwise we would consider such a program as a program which starts together with a job which we do not evaluate. In this case we would treat such a person as being employed. The exact identification of the program types will be explained in the following.

## B.3 Identifying program participation

We identify participation in a further training program from a combination of FuU training data information, the benefit payment information and the employment status information. In principle, every participant in a further training program should be recorded in the FuU training data and we would not need the benefit

payment data for identification of participation. There are two reasons to use the benefit payment data as well. First we find the training data to be incomplete, many recipients of training related benefits are not contained in the training data.<sup>13</sup> Only using the benefit payment data identifies these participants. Second, quite often the type of training in the training data is given very unspecific as “Other adjustment of working skills”. The benefit payment data can give more information about these programs. Finally we need the employment status to identify participation because we only evaluate programs which start while being unemployed.

In the remaining part of this section we describe how we aggregate the benefit payment information and the training data information. The next section contains the exact coding plan. We disclose in detail which combination of information from benefit payment and training data we identify as PF, SPST or RT.<sup>14</sup>

### **Benefit payment information from the LED-data**

The merged data we use contain three variables with benefit payment information from the original LED data, (“parallel original benefit information 1-3” [*Leistungsart im Original 1-3*] L1LA1, L2LA1, L3LA1). The main variable is L1LA1. If there are two parallel payment informations in the original data L1LA2 also contains information and only if there is a third parallel payment spell L3LA1 is also filled. In general we use L1LA1. Only if L1LA1 is not informative about program participation and L2LA1 is we use L2LA1 and only if L1LA1 and L2LA1 are not informative but L3LA1 we use L3LA1. The benefit payment information is given in time varying three-digit codes (for the coding plan see Bender et al. 2005). We extracted the program related information from the benefit payment information as given in table 27. The main distinction regarding program participation is the distinction between no benefits at all or unemployment benefits/assistance on the one hand and program related maintenance benefits on the other hand. There are five types of program related benefits. Most important for us are the more general maintenance benefits while in further training and the more specific maintenance benefits while in retraining.

---

<sup>13</sup>Remember the purpose of the training data was only internal documentation. This might explain its incompleteness.

<sup>14</sup>More details about the benefit payment data and training data can be found in Speckesser (2004), Fitzenberger and Speckesser (2005) and Bender et al. (2005).

Table 27: Aggregated types of benefit payment

German Abbreviation	Description
ALG	unemployment benefits
ALHi	unemployment assistance
UHG §41a	maintenance payment while in specific short term measure
UHG Fortbildung	maintenance payment while in further training
UHG Umschulung	maintenance payment while in retraining
UHG Darlehen	maintenance payment as a loan
UHG Deutsch	maintenance payment while in a German course

The original benefit payment information is given in three variables L1LA1, L2LA1 and L3LA1 with time varying three-digit codes.

### Type of training from FuU-data

In this evaluation study one of the most important advantages compared to survey data is the information about the precise type of training. It allows us to identify homogeneous treatments for the evaluation. In the merging process, up to two parallel FuU-spells were merged to one spell of the IABS data because in many cases the FuU-data provided more than one parallel spell. These two parallel spells provide two variables indicating the type of course (*Maßnahmart* [FMASART1, FMASART2]).

**Aggregating the training type information** Since type of treatment (*Maßnahmart*) is often coded as “other adjustment” (FMASART1=12 [*Sonstige Anpassungen*]) in the FuU-data, we increase the precision of information about the type of treatment by relying on the second parallel information about the type of training: The second FuU-spell is used if the first FuU-spell is coded as “other adjustment” (*”Sonstige Anpassungen”*) and a second spell includes a code different from 12. Such combined information of FMASART1 and FMASART2 is referred to as FMASART\* in the following.

### Combining the information

When using information from different sources, the sources may give differing information. If the training data indicated training participation and the benefit payment data did not or vice versa we relied on the source which indicated training for the following reasons. If somebody receives training related benefits it is more likely that the employment agency forgot to fill in the training data record than the agency

wrongly induced payment of benefits. And if somebody is contained in the training data but does not receive maintenance benefits he either receives no benefits, which is possible while being in training, or receives unemployment benefits/assistance and the payment is just wrongly labelled.

If both training and benefit payment data indicate program participation but differ in the type of program we generally use the training data information. An example: the benefit payment indicates maintenance payments for further training and the training data indicates Retraining. We use Retraining from the training data. The only exception is unspecific program information from the training data “other adjustment”. If in such cases the benefit payment data give specific information like Retraining we use the information from benefit payment data. All possible combinations of training and benefit payment information which we use to identify participation in one of the three programs are given in the following section.

## B.4 Coding plan for the treatment information

This section gives the exact coding plans for identification of Practice Firm, SPST and Retraining. In general we identify program participation as start of a program in an unemployment spell before another employment begins. This means that we only identify a start of a program if the employment status in the first month of the program indicates no employment (BTYP $\neq$ 1).

### Practice Firm

Practice Firm is a consolidation of the program types Practice enterprise and Practice studio from the FuU training data. There is no specific benefit payment type related to Practice Firms, rather the participants shall receive the general maintenance payment for further training. Since the training data are more reliable than the benefit payment data regarding type of the program we identify Practice Firm whenever FMASART shows the codes 11 or 12 independently of the payment information.

Program code	Label	Label in German
10	Practice enterprise	Übungsfirma
11	Practice studio	Übungswerkstatt

In table 28 we show how often which combination of benefit payment information and program type information identifies *Practice Firm* in the two inflow samples.

Table 28: Identification of *Practice Firm* with program type and benefit payment type: Frequencies

Program	Type of payment				Total	
	no benefits	UB/UA	maintenance benefits for short term training	benefits for further retraining		
Practice enterprise	0	0	0	106	0	106
Practice studio	0	0	0	110	2	112
Total	0	0	0	216	2	218

Women and Men together. BTYP $\neq$ 1 as an additional requirement.

### Provision of specific professional skills and techniques

We identify SPST in the following cases.

- (a) Identification from training data and benefit payment data

We identify SPST if the training data indicates the general program “Other adjustment” and the benefit payment information is no benefit payments, unemployment benefits, unemployment assistance or maintenance payments while in retraining.

Program code	Label	Label in German
12	Other adjustment of working skills	sonst. Anpassung der berufl. Kenntnisse

- (b) Reliance on benefit payment data

We identify SPST if the program information from the training data is missing and the benefit payment information is maintenance payments while in further training.

Program code	Label	Label in German
-9	missing	fehlende Angabe

- (c) Additional program from training data

We also identify SPST when another program of little quantitative importance but SPST-comparable content is recorded in the training data independent of the benefit payment information.

Program code	Label	Label in German
31	Further education of trainers and multidisciplinary qualification	Heran-/Fortbildung v. Aus- bildungskräften/ berufs- feldübergreifende Qualifikation

(d) Additional combination

Finally we identify SPST if the training data indicate the unspecific “other career advancement” and the benefit payment information indicates further training.

Program code	Label	Label in German
28	Other promotion	sonstiger Aufstieg (< 97)

In table 29 we show how often which combination of benefit payment information and program type information identifies *SPST* in the two inflow samples.

Table 29: Identification of *SPST* with program type and benefit payment type: Frequencies

Program	Type of payment			Total
	no benefits	UB/UA	maintenance benefits for further training	
missing	0	0	549	549
Other adjustment of working skills	6	10	1158	1174
Other promotion	0	0	6	6
Further education of trainers and multidisciplinary qualification	0	0	9	9
Total	6	10	1722	1738

Women and Men together. BTYP≠1 as an additional requirement.

## Retraining

Retraining or longer “Qualification for the first labor market via the education system” is taking part in a new vocational training and obtaining a new vocational training degree according to the German dual education system. Additionally, but quantitatively of little importance we see the make up of a missed examination “Certification” as comparable to retraining because the result is the same. Furthermore and also only of marginal importance we see participation in the programs “Technican” or “Master of Business administration (not comparable to an american style

MBA)” while not receiving maintenance benefits as a loan as Retraining. Conventionally these two programs are considered as career advancement programs which we do not evaluate. Benefits as a loan would underline their character as career advancements.

(a) Identification from training data

We identify the following two programs as Retraining independent of the benefit payment information.

Program code	Label	Label in German
29	Certification	berufl. Abschlussprüfung
32	Retraining	Umschulung

(b) Reliance on benefit payment data

If the training data is uninformative and maintenance benefits for Retraining are paid we identify Retraining.

Program code	Label	Label in German
-9	missing	fehlende Angabe
12	Other adjustment of working skills	sonst. Anpassung der berufl. Kenntnisse

(c) Other programs from training data

Two other programs are identified from the training data. They typically also take two years full time and require an existing vocational training degree, hence are somewhat comparable to retraining in a narrower definition. Not identified if maintenance benefits are paid as a loan.

Program code	Label	Label in German
26	Technician	Techniker (<97)
27	Master of business administration	Betriebswirt (<97)

In table 30 we show how often which combination of benefit payment information and program type information identifies *Retraining* in the two inflow samples.

Table 30: Identification of *Retraining* with program type and benefit payment type: Frequencies

Program	Type of payment			Total		
	no benefits	UB/UA	maintenance benefits			
			further training	retraining	loan	
missing	0	0	0	55	0	55
Other adjustment of working skills	0	0	0	13	0	13
Technician	0	0	0	0	0	0
Master of business administration	0	0	0	0	0	0
Certification	0	0	1	0	0	1
Retraining	2	2	219	137	0	360
Total	2	2	220	205	0	429

Women and Men together. BTYP≠1 as an additional requirement.



## B.5 Sample construction in comparison to LMW

Table 31: Overview sample construction

	this paper	LMW
Inflow sample	starts unemployment spell in 93/94	starts unemployment spell in 93/94
Treatment group	starts a program within 24 months after beginning of unemployment spell	starts a program between beginning of unemployment spell and the end of 94
Control group	dynamic control group: does not start a program in the stratum of unemployment under consideration	static control group: does not start a program between beginning of unemployment spell and the end of 94
Treatment identification	training spell in the training participation data or income maintenance spell in the benefit payment data indicating program participation	training spell in the training participation data
Age restriction	25–55 years (25–50 in case of RT) in the year of entry into unemployment	20–53 years in the year of the (simulated) program start
Benefit payment restriction	controls have to receive unemployment benefits at least once during the first 12 months of their unemployment spell	recipience of benefits in the month before program start for participants and in the month before as well as in the month of the simulated program start for controls
Other restrictions	without East Berlin	last employment before defining unemployment spell not as trainee, home worker, apprentice, or in part-time with less than half of the usual ours; no foreigners
Sample size RT	Women: 189 (=61+75+53), Men: 219 (=107+79+33), numbers in parenthesis differentiated by strata	Women: 190, Men: 255
Sample size SPST (this paper) and short and long training (LMW)	Women: 1063 (=254+374+435), Men: 485 (=200+141+144)	Women: 557 (=209 (short) + 348 (long)), Men: 302 (=112+190)
Sample size non-participants	Women: 4585, Men: 5076 (not directly comparable to the dynamic control groups used in the paper)	Women: 2914, Men: 1690

# C Heterogeneous Treatment Effects by Target Profession

## C.1 Retraining for Men

In this section we show heterogeneous treatment effects of retraining on men. We contrast retraining (RT) with target profession in construction with RT with other target profession (non-construction).

The effects are estimated in the same way as the non-disaggregated effects in the paper. We used the same propensity score specifications and bandwidth as in the paper.

Table 32: Sample sizes: Retraining for men by target profession

	Stratum 1	Stratum 2	Stratum 3
Construction	40	29	19
Non-construction	50	40	13
missing	17	10	1
Total	107	79	33

Remark: The participants are classified according to the field in which they are retrained. This information is only available from the training participation data and hence is missing if participation is identified from the benefit payment data.

Figure 9: Retraining (RT) for Men stratum 1 – Construction (top) vs. Non-construction (bottom)

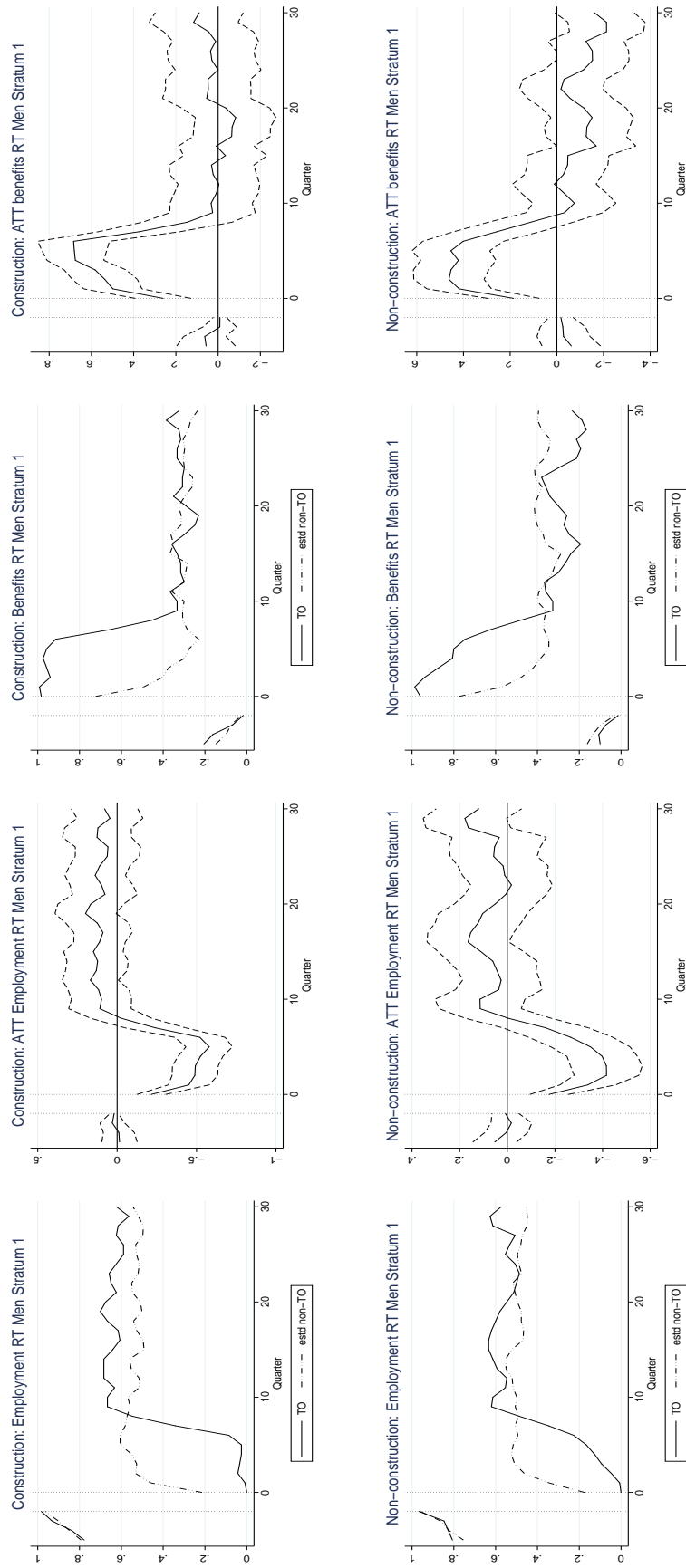


Figure 10: Retraining (RT) for Men stratum 2 – Construction (top) vs. Non-construction (bottom)

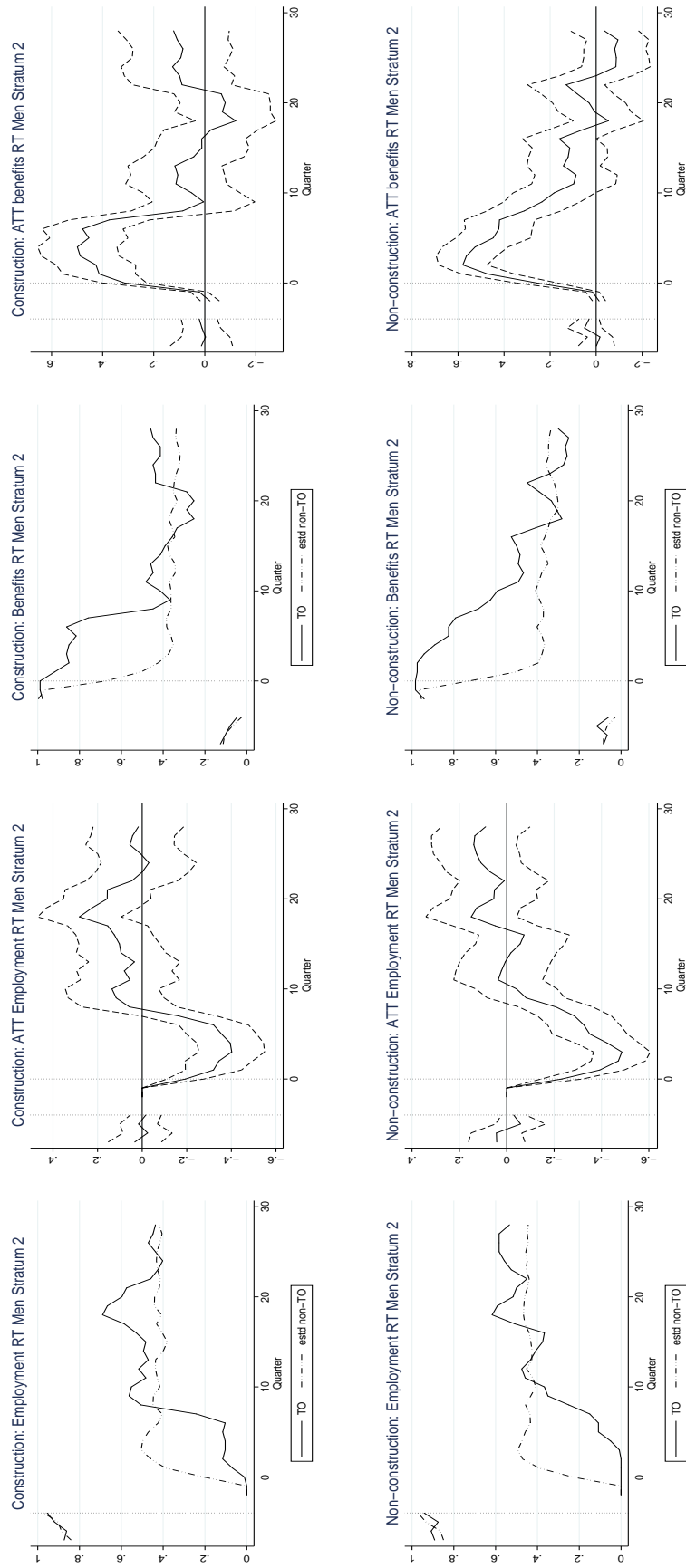


Figure 11: Retraining (RT) for Men stratum 3 – Construction (top) vs. Non-construction (bottom)

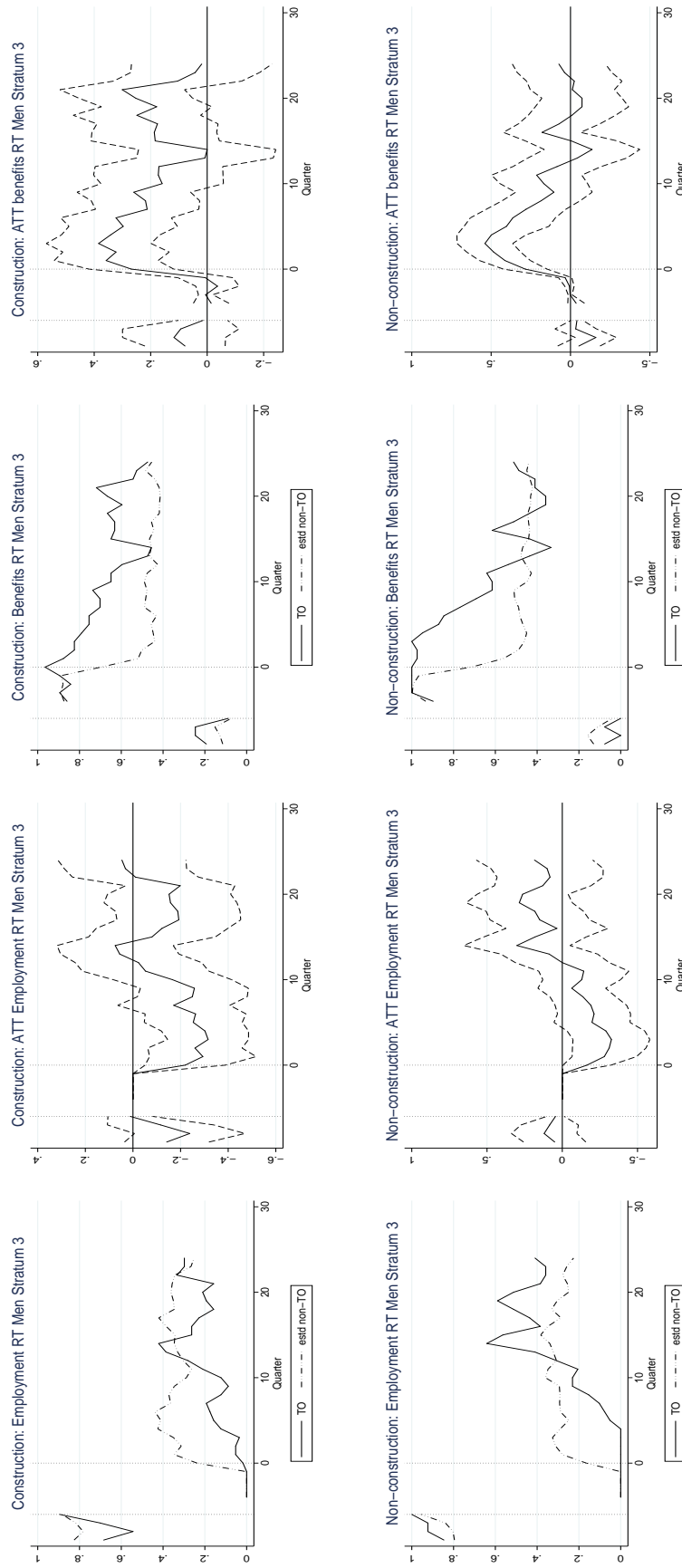


Table 33: Cumulated and Average Employment Effects RT for Men – Construction vs. Non-construction

Stratum	Target Profession	Cumulated Effects			Average Effects		
		Q0-Q7	Q0-Q15	Q0-Q23	Q4-Q23	Q8-Q23	
Employment							
1	construction	-3.51 (0.44)***	-2.64 (0.91)***	-1.64 (1.37)	0.001 (0.064)	0.117 (0.071)*	
1	non-construction	-2.52 (0.46)***	-2.02 (0.90)**	-1.42 (1.32)	-0.004 (0.061)	0.069 (0.066)	
2	construction	-2.49 (0.47)***	-1.82 (0.91)**	-0.68 (1.34)	0.029 (0.063)	0.114 (0.071)	
2	non-construction	-2.97 (0.41)***	-3.31 (0.91)***	-2.87 (1.39)**	-0.064 (0.064)	0.006 (0.070)	
3	construction	-2.08 (0.56)***	-2.78 (1.15)**	-3.77 (1.81)**	-0.134 (0.084)	-0.106 (0.096)	
3	non-construction	-1.97 (0.77)**	-1.87 (1.45)	-0.64 (2.34)	0.022 (0.106)	0.083 (0.112)	
Benefit Reciprocity							
1	construction	4.29 (0.41)***	4.53 (0.88)***	4.43 (1.27)***	0.128 (0.061)**	0.009 (0.069)	
1	non-construction	3.05 (0.46)***	2.91 (0.88)***	2.12 (1.24)*	0.030 (0.056)	-0.058 (0.061)	
2	construction	3.45 (0.47)***	3.98 (0.80)***	3.83 (1.15)***	0.109 (0.054)**	0.024 (0.062)	
2	non-construction	3.69 (0.42)***	4.98 (0.75)***	5.40 (1.11)***	0.176 (0.052)***	0.107 (0.057)*	
3	construction	2.51 (0.54)***	3.68 (0.84)***	5.17 (1.36)***	0.192 (0.063)***	0.167 (0.077)**	
3	non-construction	3.27 (0.67)***	3.84 (1.27)***	3.94 (1.84)**	0.111 (0.087)	0.042 (0.090)	

Remark: The cumulated (average) effects are the sum (average) of the quarter specific average treatment effect on the treated over the respective quarters. \*, \*\* and \*\*\* denote significance at the 10%, 5% and 1% level, respectively, and QiQj denotes quarter  $i$  to quarter  $j$  since beginning of treatment.