

Fitzenberger, Bernd; Völter, Robert

Working Paper

Long-run effects of training programs for the unemployed in East Germany

IZA Discussion Papers, No. 2630

Provided in Cooperation with:

IZA – Institute of Labor Economics

Suggested Citation: Fitzenberger, Bernd; Völter, Robert (2007) : Long-run effects of training programs for the unemployed in East Germany, IZA Discussion Papers, No. 2630, Institute for the Study of Labor (IZA), Bonn,
<https://nbn-resolving.de/urn:nbn:de:101:1-20080416344>

This Version is available at:

<https://hdl.handle.net/10419/34449>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

IZA DP No. 2630

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

Bernd Fitzenberger
Robert Völter

February 2007

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

Bernd Fitzenberger

*Goethe University Frankfurt,
ZEW, IFS and IZA*

Robert Völter

*Goethe University Frankfurt and
CDSEM, University of Mannheim*

Discussion Paper No. 2630
February 2007

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0

Fax: +49-228-3894-180

E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of the institute. Research disseminated by IZA may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit company supported by Deutsche Post World Net. The center is associated with the University of Bonn and offers a stimulating research environment through its research networks, research support, and visitors and doctoral programs. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

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

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 reciprocity and the effects are quite similar for females and males.

JEL Classification: C14, J68, H43

Keywords: multiple treatments, training programs, East Germany

Corresponding author:

Bernd Fitzenberger
Department of Economics
Goethe University
P.O. Box 11 19 32 (PF 247)
60054 Frankfurt am Main
Germany
E-mail: fitzenberger@wiwi.uni-frankfurt.de

^{*} 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.

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.¹ Detailed administrative data have been used in

¹See 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.²

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.

²For 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.³ 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.⁴ 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

³This 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.

⁴Only 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.⁵ 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

⁵These 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 $k = 0$.⁶

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).⁷ 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

⁶Using 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.

⁷In 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)⁸

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

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,

⁸In 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.⁹ 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.

⁹Abadie 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.¹⁰

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,

¹⁰In 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 τ 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)¹¹ 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

¹¹We 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.¹² 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

¹²As 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

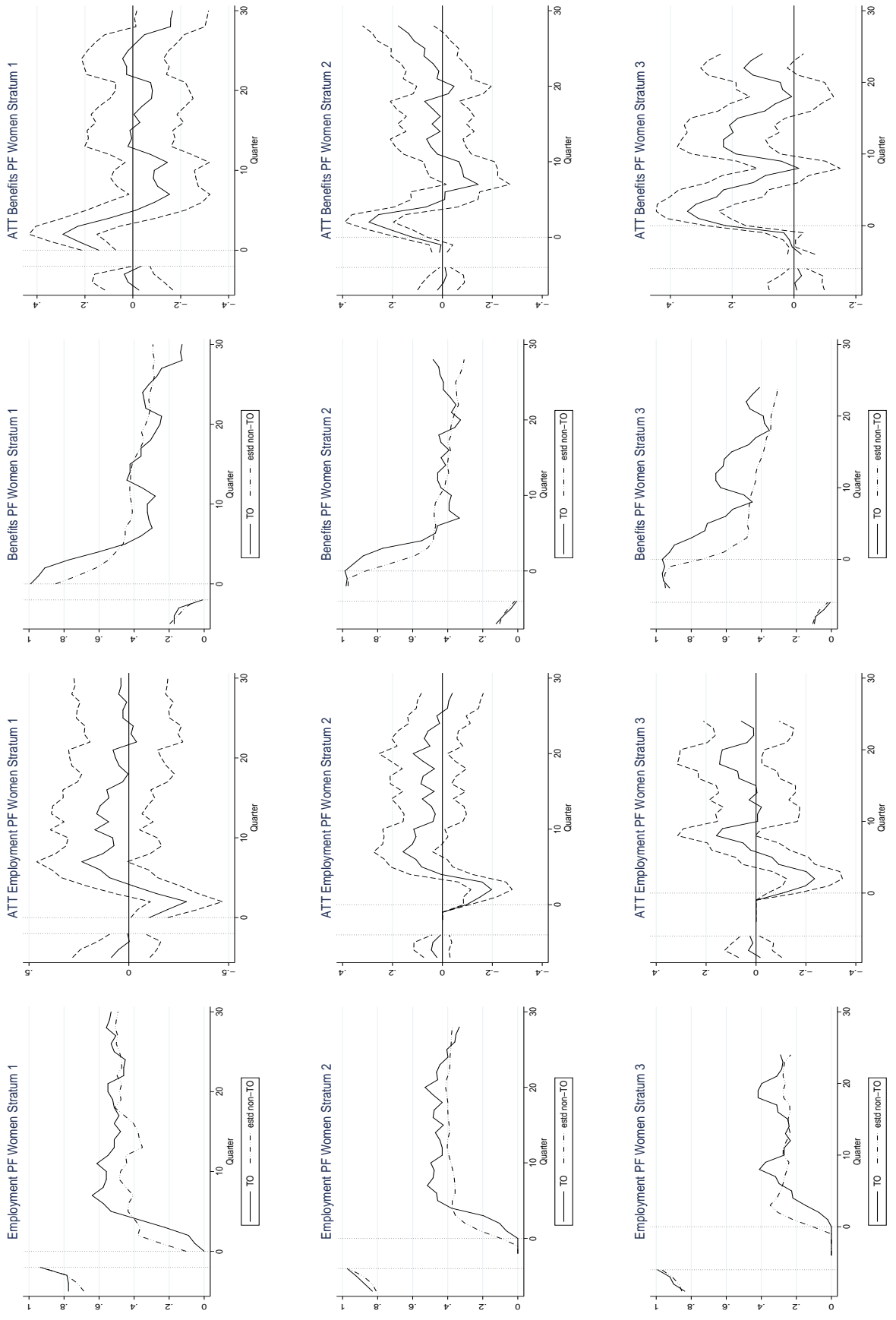


Figure 2: Practice Firm (PF) for Men

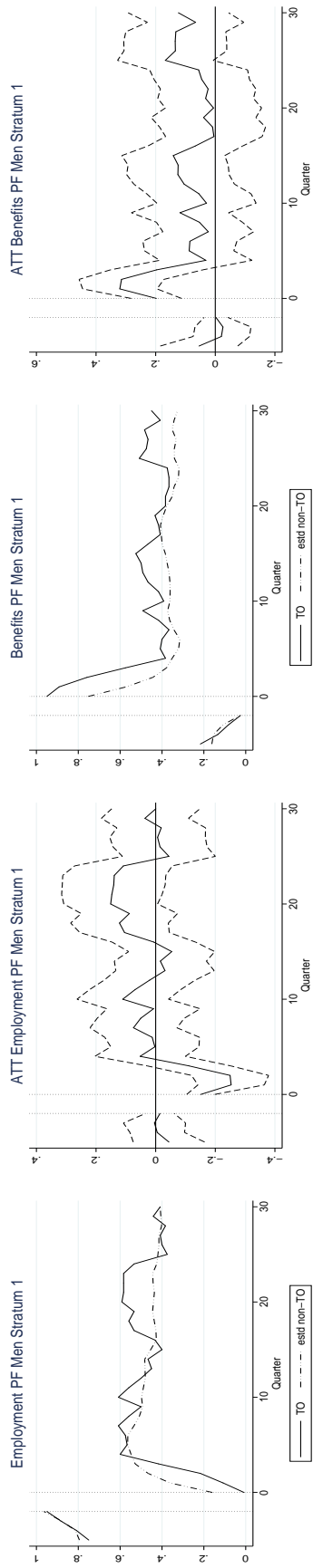


Figure 3: SPST for Women

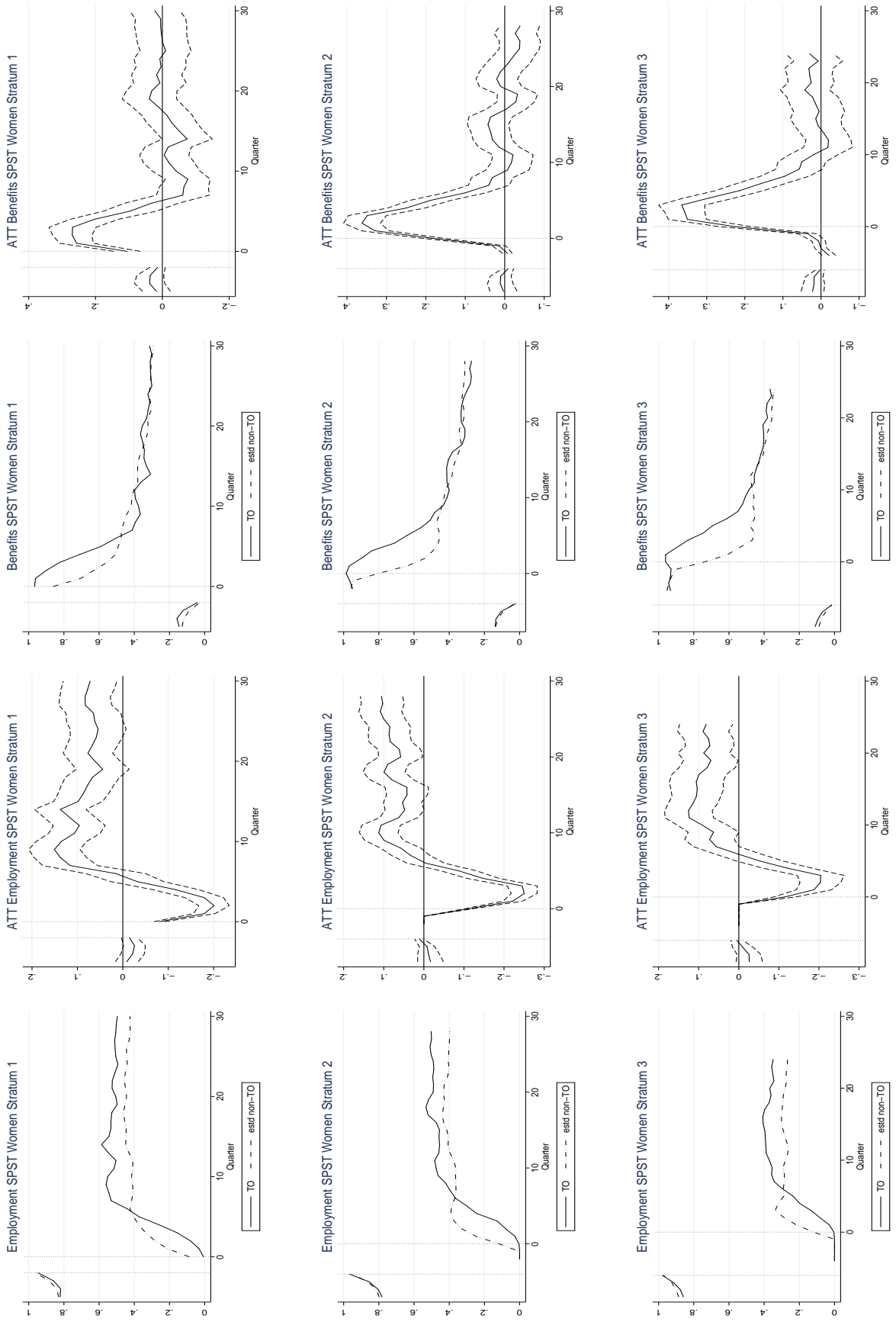


Figure 4: SPST for Men

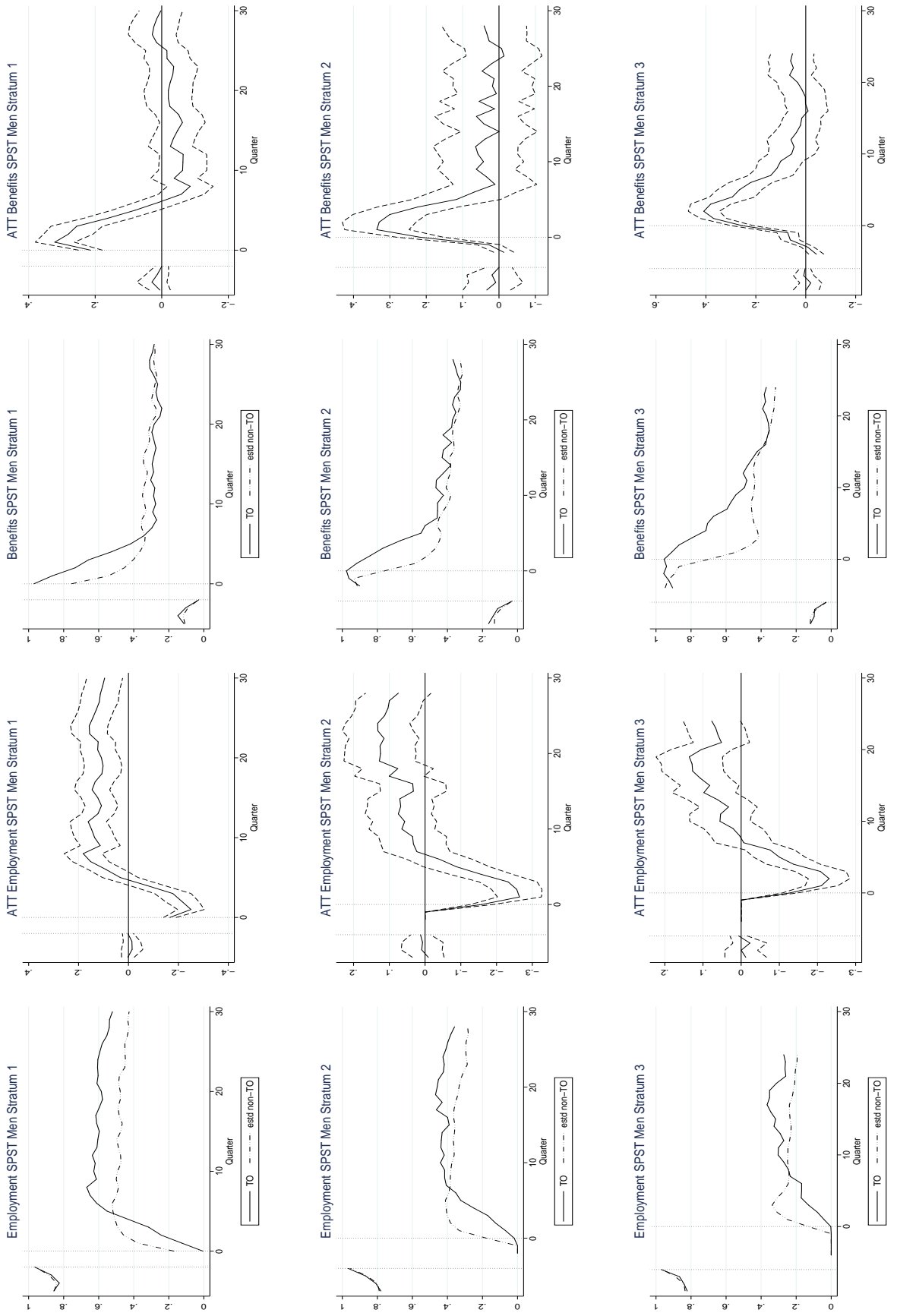


Figure 5: Retraining (RT) for Women

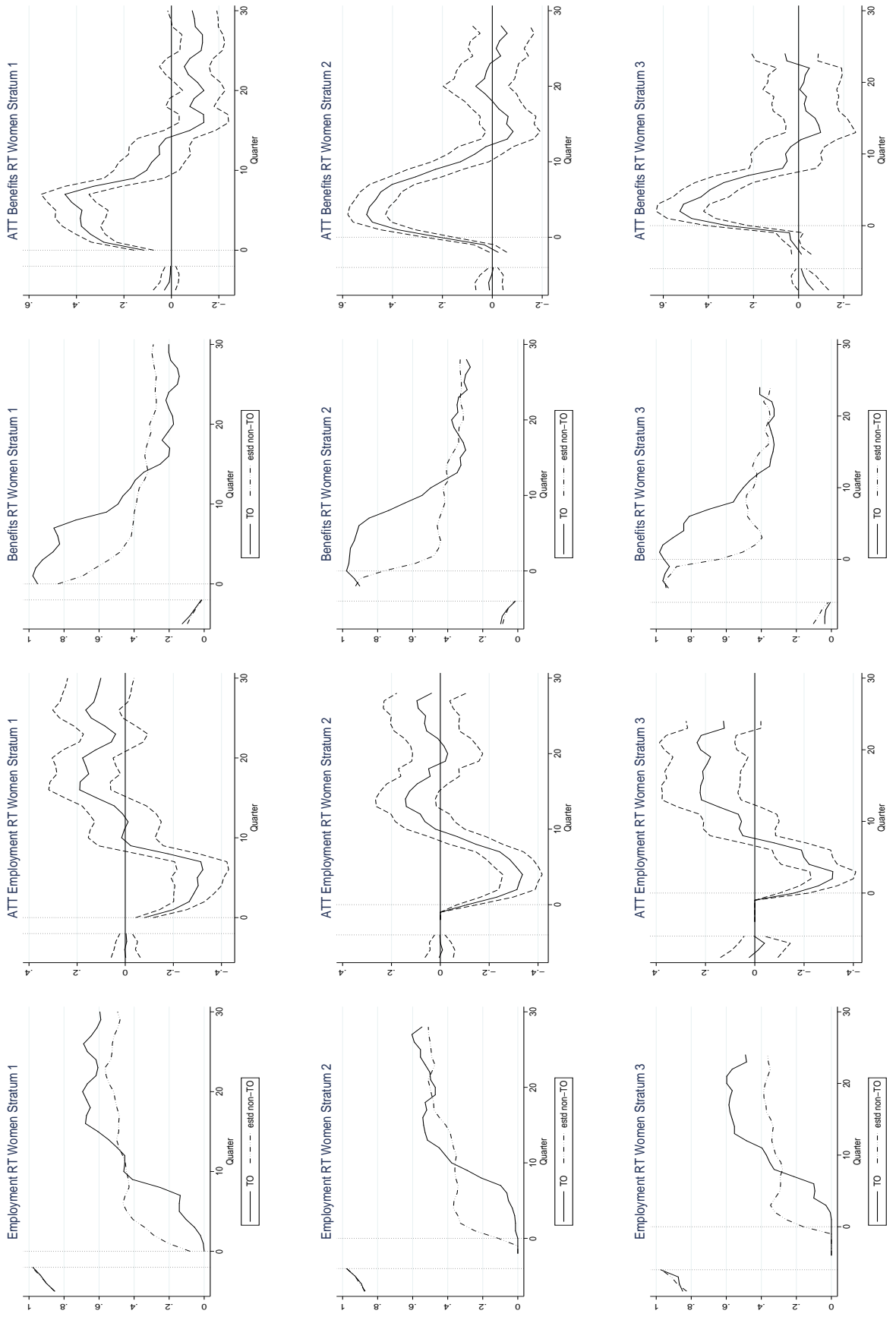


Figure 6: Retraining (RT) for Men

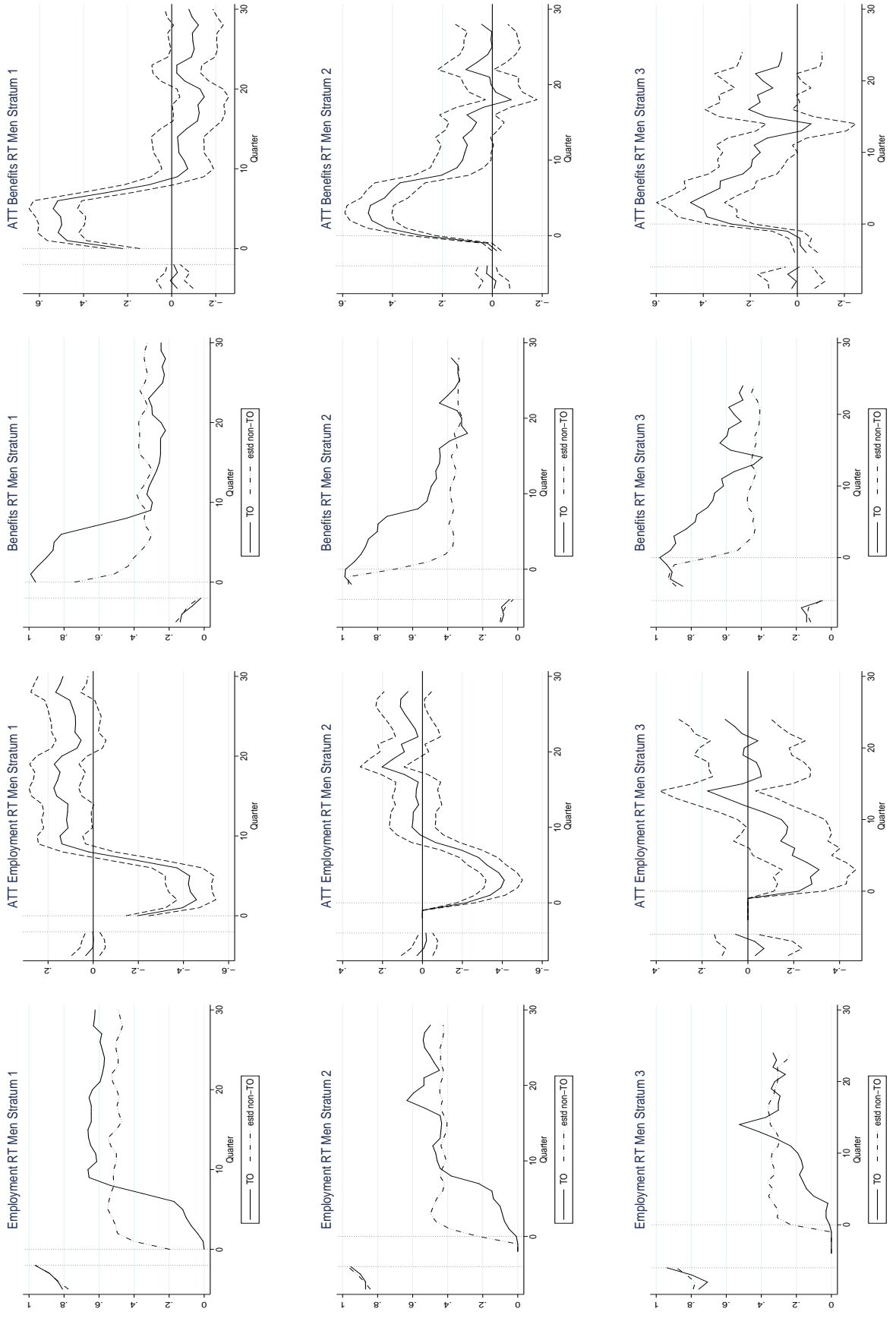


Table 6: Cumulated and Average Employment Effects

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.