The Heterogeneous Effects of Training Incidence and Duration on Labor Market Transitions[1]

BERND FITZENBERGER,[♭] ADERONKE OSIKOMINU,[♯] and MARIE PAUL[§]

[♭]Albert–Ludwigs–University Freiburg, ZEW, IZA, IFS, bernd.fitzenberger@vwl.uni-freiburg.de
[♯]Albert–Ludwigs–University Freiburg, IZA, aderonke.osikominu@vwl.uni-freiburg.de
[§]née Waller, Albert–Ludwigs–University Freiburg, marie.waller@vwl.uni-freiburg.de

This version: April 14, 2010

Abstract: This paper estimates the impact of training incidence and duration on employment transitions accounting for the endogeneity of program participation and duration. We specify a very flexible bivariate random effects probit model for employment and training participation and we use Bayesian Markov Chain Monte Carlo (MCMC) techniques for estimation. We develop a simulation approach that uses the estimated coefficients and individual specific effects from the MCMC iterations to calculate the posterior distributions of different treatment effects of interest. Our estimation results imply positive effects of training on the employment probability of the treated, lying between 12 and 21 percentage points ten quarters after program start. The effects are higher for women than for men and higher in West Germany than in East Germany. Further, we find that the effect of training versus waiting underestimates the effect of training versus no training in the medium and long run by a third. Finally, our results show that longer planned enrolment lengths of three and four quarters as opposed to just two quarters lead to an increase in employment rates in the medium and long run by four to eleven percentage points.

Keywords: evaluation, active labor market programs, dynamic non–linear panel data models, MCMC

JEL: J68, C33, C11

Correspondence:

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

Contents

1 Introduction

With rapidly changing technologies human capital development is viewed as a key element to ensure employability of individuals in the long run. Training programs are an important tool of active labor market policy in many advanced countries (see e.g. OECD, 2007, table J). Training schemes typically comprise a wide variety of programs, from job application training to basic education and advanced vocational training taking place on the job or in classrooms. In addition, there exist important heterogeneities within a given type of training.

An important dimension of heterogeneity that has received little attention in the literature is how the impact of a given type of training program varies with the length of enrolment.[2] One reason for this is that realized training duration is an endogenous variable that depends on the success of job search during training.[3] On the one hand, lucky participants who receive a suitable job offer during training may drop out early, while the unlucky ones continue until the scheduled program end or even prolong participation for lack of job opportunities. This argument suggests that dropouts are a positively selected subset of participants. On the other hand, individuals may also drop out because they are unable to follow the program (e.g. because of lack of endurance). This would suggest a negative selection of dropouts compared to completers. In any case, program drop out for reasons that are related to employment outcomes after program start raise additional endogeneity issues that are difficult to incorporate in static evaluation approaches that are commonly used in the literature.[4]

Not only the decision to continue or to drop out is dynamic but also the assignment to training often depends on the success of job search. In countries with comprehensive

---

[2]Two notable exceptions are Flores-Lagunes et al. (2007) and Kluve et al. (2007). Both studies view realized training duration as a continuous treatment and use a static matching approach to evaluate the effects of different training lengths. Kluve et al. (2007) test the robustness of their results by instrumenting realized training duration with planned duration for a subsample of their data set.

[3]See Waller (2009) for an empirical analysis of endogenous training dropouts in Germany.

[4]For general surveys of evaluation studies including training see e.g. Heckman et al. (1999), Martin and Grubb (2001), Card et al. (2009). Recent evaluation studies for training programs conducted in Germany using a static approach for program assignment include e.g. Bergemann et al. (2009) and Lechner et al. (2009). In order to apply static evaluation methods, such as those suggested by Hirano and Imbens (2004) and Imai and van Dijk (2004) for continuous treatments or Imbens (2000) and Lechner (2001) for multiple discrete treatments, one would have to assume that actual enrolment length is either exogenous or determined at the same time as participation.

systems of active labor market policies like Germany or Sweden (cf. Sianesi, 2004, on Sweden), participation in active labor market programs may take place at any point in time during unemployment. Eventually, every unemployed who does not manage to find a job on his own becomes eligible for program participation.[5]

The main contribution of this paper is to estimate the impact of training incidence and duration on employment transitions while taking into account the dynamic, endogenous nature of program participation and duration. We focus on a large scale training program in Germany that lasts eight months on average and for which enrolment lengths vary between a couple of weeks and more than one year. We specify a joint model for the transition rates into and out of employment and training using a very flexible bivariate random effects probit model. Our specification allows in a flexible way for state dependence and duration dependence in the transition rates as well as in the treatment effects and it includes interaction effects of these model components with observed covariates. Our rich administrative data allow us to integrate such flexibility into the model while performing separate estimations by gender and region (East and West Germany). Estimating a discrete time model for labor market transitions, we account for the full observed observation vector for each individual over time, irrespective of the number of spells experienced by an individual. We use Bayesian Markov Chain Monte Carlo (MCMC) techniques that allow a numerically very robust estimation of our flexible model specification.[6] Another advantage of the MCMC technique is that it provides predictions of the individual specific effects. This allows us to assess explicitly the selectivity of the treated and the nontreated individuals. We develop a simulation approach that uses the estimated individual specific effects to calculate the posterior distributions of different treatment effects of interest, such as the average treatment effect on the treated for the employment probability. Furthermore, we use our estimation results to simulate the effects of alternative policy scenarios. In particular, we examine how the impact of training changes when assigning different planned enrolment lengths.

A comparative assessment of different enrolment lengths for the same type of training is important for policy makers interested in an efficient use of active labor mar-

---

[5]See Fredriksson and Johansson (2008) for a formal analysis of the bias resulting when applying a static evaluation approach in the case of a dynamic assignment regime as described above.

[6]See Chib (2001) for a survey on MCMC methods and Buchinsky, Fougère, Kramarz, and Tchernis (2005) for a recent application in modeling discrete time transitions of employment, job mobility and wages.

ket programs. Comprehensive training schemes typically range among the most expensive active labor market programs. During training, participants generally search less intensively for a new job than comparable nonparticipants (lock-in effect). Therefore, employment effects of training are typically negative in the short run, and positive effects may unfold only some time after the completion of the program. Our analysis provides evidence to address explicitly the question whether negative short–run effects are necessary to achieve economically important positive employment effects in the long run, or whether comparable long-run effects can be obtained with shorter programs at lower costs.

An important problem in the evaluation literature is the nonrandom sorting of individuals into training programs. Our dynamic modeling approach accounts for selection into training based on observables as well as unobservables. Our identification strategy is a discrete time version of the timing-of-events-approach introduced by Abbring and van den Berg (2003). The basic idea translated to our setting is as follows. If participants are unable to anticipate the exact starting date of the program, a nonzero treatment effect can be distinguished from a nonzero correlation between the random effects in the training and the employment equation by recognizing that the treatment effect operates only after the program start while the correlation between the unobservables induces a permanent association. As we will show below, the assumption of random variation in the timing of participation during unemployment is plausible in the institutional setup underlying this study. In Germany, for the time period investigated here, participation may potentially start at any point in time during an unemployment spell. Program assignment depends on the short-term supply of training programs, with imperfect information and little control on the part of an unemployed.

The problem of dynamic program starts has been addressed in different ways so far. In one strand of the literature, sequential matching techniques mimic a sequential randomization into treatment of individuals who are unemployed for the same elapsed duration at a given date in a sequence of possible dates. The application of matching methods presumes that sufficiently rich data are available to justify that there is no remaining selection on unobservables after controlling for observable variables. This approach allows e.g. to estimate the average effect of receiving treatment at a given date of elapsed unemployment duration versus not, implying the possibility of treatment at a later time (i.e. the effect of treatment now versus

3

waiting, Sianesi, 2004). However, as this approach focuses on nonparametric iden-
tification at given elapsed durations and does not specify a model for the selection
over time, it is not possible to estimate a causal treatment effect averaged across
starting dates.[7]

A different strand of the literature uses a continuous time duration framework to
estimate treatment effects. In contrast to sequential matching, this approach models
selection into outcomes and treatment over time. Abbring and van den Berg (2003,
henceforth AvdB) show that with single spell data for unemployment duration and
time until treatment start and time constant covariates one can identify the effect of
treatment on the exit rate from unemployment under some functional form restric-
tions. In particular, one has to assume a mixed proportional structure for the hazard
rates into employment and into training and that the exact moment of treatment
cannot be anticipated at the individual level. Furthermore, the effects of treatment
on hazard rates in subsequent spells are identified. The model allows for the pres-
ence of permanent unobserved heterogeneity terms which enter the hazard rates
in a multiplicative fashion and which are independent of the observed covariates.[8]
AvdB view the duration in untreated unemployment until treatment start and until
exit to employment as two competing risks which are linked through correlation
of their permanent unobserved heterogeneity terms. Based on the proportionality
and the no-anticipation assumption, the treatment effect can be distinguished from
a correlation of the unobservables because the exit rates from unemployment after
treatment start can be contrasted with those obtained in the case where treatment
has not started. The latter are estimated through the competing risks part of the
model (see Abbring and van den Berg, 2004, page 15). Since the unobserved hetero-
geneity induces a permanent association between the two risks, its spurious selection
(or sorting) effects can be distinguished from the effects of treatment starts which
vary randomly over time across individuals.[9]

---

[7]Recent applications of this approach to the evaluation of training programs include Biewen et
al. (2007) and Dyke et al. (2006).

[8]AvdB also discuss identification based on multiple spell data where unobserved heterogeneity
terms are constant within an individual. In this case, the proportionality assumption can be relaxed
and observed covariates are not required any more. Abbring and van den Berg (2004) discuss the
link between treatment effects estimation using duration models based on the timing-of-events
approach on the one hand and cross-sectional binary treatment models as well as linear panel data
models with individual fixed effects on the other hand.

[9]Recent applications of this approach to the evaluation of training programs include Hujer et
al. (2006), Richardson and van den Berg (2008), Osikominu (2009).

Heckman and Navarro (2007) consider the semiparametric identification of dynamic treatment effects in structural dynamic discrete choice models. Similar to Abbring and van den Berg (2003), the treatment status is allowed to depend on unobserved factors in the outcome equation. Heckman and Navarro require the existence of instruments that affect choices but not outcomes for semiparametric identification of causal effects. In a dynamic context, one needs instrumental variation at each stage of the sequential selection process or variation in the impact of time–invariant instruments (see Heckman and Navarro, 2007, Theorem 1). This variation must not be fully anticipated by the agents.[10]

Our paper contributes to and extends the dynamic evaluation literature in the following way. We implement a joint model in discrete time for transitions in and out of employment and in and out of training. We do not only take into account the timing of program starts but also the endogeneity of program duration. Our model allows for selection into training based on unobservables that are assumed to be independent of observed covariates. Similar to AvdB, our approach exploits the no–anticipation assumption for identification of the causal effect of training. Similar to Heckman and Navarro (2007) our model is specified in discrete time. We model the impact of training in a very flexible way in order to avoid overly strict functional form restrictions. Based on our model estimates, we evaluate the causal effects of training incidence and duration.

Our main findings are as follows. We estimate positive effects of training on the employment probability unfolding three to four quarters after program start in all four subsamples considered. Ten quarters after program start, the effect of treatment on unconditional employment rates for the treated individuals lies between 12 and 21 percentage points (ppoints). The effects are higher for women than for men and higher in West Germany than in East Germany. To compare our results to the recent literature applying sequential matching estimators, we use our model estimates to simulate the effect of treatment start at a given date versus waiting (Sianesi, 2004). We find that the effect of treatment versus waiting underestimates the treatment effect of training versus no training in the medium and long run by a third. Finally, we use our estimates to analyze how training effects vary with planned program

---

[10]Abbring and van den Berg (2003) argue that it is often difficult to maintain exclusion restrictions in dynamic settings with forward-looking agents. Unfortunately, the data used in our analysis lack credible instruments, i.e. time–varying exogenous variables which affect the assignment process into treatment, as required by Heckman and Navarro (2007).

duration. Longer planned enrolment lengths of three and four quarters as opposed to just two quarters lead to an increase in employment rates by four to six ppoints and six to eleven ppoints, respectively, in the medium and long run. Thus, longer training programs show higher long-run employment gains, which may justify the higher costs involved.

The remainder of this paper is structured as follows. Section 2 describes the institutional setup and the data set used. Section 3 presents our evaluation framework and the MCMC estimation procedure. Section 4 discusses the results and section 5 concludes. The appendix provides further details on the data, the implementation of the estimation approach, and detailed estimation results.

2   Institutional Background and Data

2.1   Training in Germany

Training schemes have traditionally dominated active labor market policy in Germany. The legislation distinguishes three main types of training, further training (*Berufliche Weiterbildung*), retraining (*Berufliche Weiterbildung mit Abschluss in einem anerkannten Ausbildungsberuf*), and short-term training (*Trainingsmaßnahmen und Maßnahmen der Eignungsfeststellung*). Figure 1 shows the evolution of entries into the three different training programs in West and East Germany during the period 1999 to 2007. Until 2000, enrolment into further training (henceforth also referred to as long-term training) was around 260 thousand in West Germany and 170 thousand in East Germany. A policy reorientation favoring programs supposed to activate the unemployed in the short run led to a decline in further training and retraining and a sharp increase of short-term training. In 2004, participation in further training was about 100 thousand in West Germany and about 50 thousand in East Germany. The corresponding figures for short-term training were 800 thousand and 400 thousand, respectively, up from around 200 thousand in 1999. After a low point in 2005, participation recovered somewhat in 2006 and 2007.

— Insert figure 1 about here. —

6

The main goal of active labor market policy in Germany is to reintegrate unemployed individuals into employment. In this study we focus on further training programs. They are used to adjust the skills of the unemployed to changing requirements of the labor market and possibly to changed individual conditions of employability (due to health problems for example). Further training courses typically last several months to one year and are usually conducted as full-time programs. Teaching takes place in class rooms or on the job in training firms. The course curriculum may also include internships. Typical examples of further training schemes are courses on IT based accounting or on customer orientation and sales training. Similar to the much longer retraining schemes, that lead to a complete new degree within the German apprenticeship system, further training programs aim at improving the human capital and productivity of the participant. Short-term training, in contrast, primarily aims at improving job search and lasts typically about four weeks.

In order to become eligible for training, job seekers have to register personally at the local employment agency. This involves a counseling interview with a caseworker. In principle, they have in addition to fulfill a minimum work requirement and be entitled to unemployment benefits. However, there are exceptions to this rule. The most important criterion is that the training scheme has to be considered necessary by the caseworker for the unemployed to find a new job. Participation in training can occur at any time during an unemployment spell.

Before 2003, training measures were assigned by the caseworker. This was often done in agreement with the job seeker, considering his or her willingness to receive training and to work in a specific field. The final decision was subject to the discretion of the caseworker. Assignment into programs was to a large extent driven by the supply of courses that were booked in advance for a year by the employment agencies from training providers. Assignments to training often occurred at very short notice in order to fill course capacities (Schneider et al., 2006).

In 2003, the assignment procedure changed to a system where the job seeker receives a training voucher from the caseworker valid between one and three months. The voucher specifies the maximal length, the content and the objective of the eligible training program. The job seeker then chooses by himself a suitable course from a pool of certified training providers. The 2003 reform meant to make the allocation process more targeted and selective. However, potential participants were uncertain

about the actual starting date because it turned out that training providers tended to collect vouchers until a critical number of participants was reached or they shortly canceled scheduled courses if there were too few participants (Kühnlein and Klein, 2003, Schneider et al., 2006). Moreover, during the first three months of 2003, the old and new assignment system coexisted. 93% of the programs in our analysis sample start before the reform. An additional 2% starts in the first quarter of 2003, thus about 5% of the programs fall in the time when vouchers were used.

During training most participants receive a subsistence allowance of the same amount as the unemployment compensation they would receive otherwise. Participants not eligible for subsistence allowance may receive similar payments from the European Social Fund. In addition, travel and child-care costs may be covered by the employment agency.

Once a particular program or a training voucher has been assigned, participation is mandatory. Non-compliance may be sanctioned with a temporary suspension of unemployment compensation. The planned duration of the further training programs considered in this paper is eight months on average. However, not all participants who start a program complete it. In fact, according to the study in Waller (2009), one out of five participants who have started a program and attended it for at least one week drop out before having reached 80% of the planned duration. About half of the dropouts start employment soon after quitting a program. In many cases this behavior is encouraged by the employment agency because in general employment has priority over participation in active labor market programs. Exceptions from this rule are possible if completing the program is deemed necessary for a stable placement. Those dropping out for other reasons are typically not sanctioned. As opposed to dropouts, it also happens in some cases that participation in training is prolonged. Due to dropout and possible prolongment of participation the actual duration of training is endogenously determined.

2.2   Constructing a Panel Data Set

For the empirical analysis, we construct a panel data set from a rich administrative database, the Integrated Employment Biographies Sample (IEBS). The IEBS is a 2.2% random sample from a merged data file containing individual data records collected in four different administrative processes: the IAB Employment History

(*Beschäftigten-Historik*), the IAB Benefit Recipient History (*Leistungsempfänger-Historik*), the Data on Job Search Originating from the Applicants Pool Database (*Bewerberangebot*), and the Participants-in-Measures Data (*Maßnahme-Teilnehmer-Gesamtdatenbank*).[11] The data contain detailed daily information on employment subject to social security contributions, receipt of transfer payments during unemployment, job search, and participation in different active labor market programs.

We consider an inflow sample into unemployment consisting of individuals who became unemployed between the first of July 1999 and the end of December 2000, after having been continuously employed for at least 125 days. Entering unemployment is defined as the transition from non-subsidized employment to non-employment plus subsequently (not necessarily immediately) some contact with the employment agency, either through benefit receipt, program participation, or a job search spell. In order to exclude individuals eligible for specific labor market programs targeted to youths and individuals eligible for early retirement schemes, we only consider persons aged between 25 and 53 years at the beginning of their unemployment spell.

We discretize the spell information in the original data by calendar quarters. We follow a person in the sample from the quarter of his/her first inflow into unemployment over the next 16 quarters or until the end of 2004, whichever occurs first. For 76% of the individuals in the sample we observe the full sequence of 17 quarters. The sequences of the remaining individuals are shorter either because we observe less than 17 quarters from their inflow until the end of 2004, or because we censor the time path for training participants in the quarter in which they enter a long-term active labor market program other than training. We ignore participation in short-term training and do not censor employment sequences in this case.

We distinguish the two outcome states non-subsidized employment (henceforth denoted as employment) and non-employment as alternative states. We aggregate the employment information measured at a daily level into quarters as follows. First, for short gaps of a length up to 45 days between sequences of longer employment or non-employment spells we extend the longer spells through the gap. Second, we map the start of non-employment and employment spells to the quarterly employment dummy in the following way. If a transition to non-employment occurs during a calendar quarter, the employment dummy is set to zero during this quar-

---

[11]For further information on the data see Appendix A.

9

ter. It continues to equal zero in the following quarter if the elapsed duration of non-employment at the end of the quarter exceeds 90 days. From the third quarter of non-employment onwards, the employment dummy is set to zero if the share of days in non-employment exceeds one half. Third, we take care of not dropping short employment spells by defining the individual to be employed in the later quarter.

Participation in further training is coded as follows. We construct a dummy variable that equals one in the quarter in which the job seeker starts a training program and attends it for at least 27 days. In order to model the duration of the training program we apply the same rules as for the employment dummy above to the qualification dummy. Because not only the start of a program but also the program status in each following quarter is used for the estimation, it is important to use reliable information on the realized program duration. We correct the reported end dates of training programs using the correction procedures proposed in Waller (2008). Participation can already occur in the first quarter we observe for an individual.

The definition of the quarterly employment and training dummy variables mimics the timing of events. When a person starts a training program in one quarter, he is also coded to be non-employed in that quarter. While being on the program, a participants remains non-employed. When a program participant exits into employment in one quarter, even though he has been in the program at the beginning of this quarter, the training dummy changes to zero in that quarter. Consequently, our empirical analysis imposes a lag in the effect of training, such that training in one quarter is only allowed to have a causal effect on employment in future quarters.

The panel data set for the analysis is complemented by adding personal, occupational and regional information. Information on the regional labor market and the current season is updated at the beginning of each quarter. The estimations are carried out separately for males and females and West and East Germany.

2.3 Descriptive Analysis

Table 1 gives an overview of the four samples and their basic characteristics. On average we observe 13 to 15 quarters per person, with the number of non-employment quarters ranging from eight to ten. This corresponds to 1.5 to 1.9 unemployment spells and about one employment spell on average per person. One in ten to one in

five persons participate in training throughout the observation period with participation rates being higher in East Germany and among females.

— Insert table 1 about here. —

Planned and realized enrolment lengths in training vary widely. Figure 2 provides histogram plots of planned and realized program durations in the four samples. The height of the bars records the fraction of cases with a program duration corresponding to the value given on the abscissae. Realized and in particular planned durations display spikes at certain round dates like six months or one year. The share of realized durations lying below half a year is higher than that of planned durations. This indicates that some trainees drop out before the scheduled program end.

— Insert figure 2 about here. —

Figure 3 illustrates the evolution of the employment and training rates from the quarter of inflow into unemployment onwards. In the calendar quarter of the inflow, all individuals are defined as non-employed. The employment rates subsequently recover, but those of females remain at a slightly lower level than those of males. While participation rates barely reach five percent in West Germany, they peak at about eight to nine percent in East Germany.

— Insert figure 3 about here. —

Figure 4 gives a first impression of the likely order of magnitude of the treatment effects. It shows the actual employment rates and estimates of the counterfactual associated with starting a training program in a given quarter versus waiting for the treated individuals where treated and matched controls are merely aligned in time. Treatment status is a time-varying variable. This means that training participants who enrol later are counted as controls for those who enrol in an earlier quarter. The matching is performed with respect to the calendar quarter of the first inflow and the elapsed unemployment duration in the current unemployment spell. No adjustments are made for other potential sources of selection bias. West German females show the largest employment differences five to ten quarters after program

11

start, which amount to more than 15 ppoints. The initial lock-in periods character-ized by negative employment effects are substantially longer in East Germany than in West Germany.

— Insert figure 4 about here. —

3   Evaluation Framework

We estimate the effect of participating in training, measured by the sequence of quar-terly training dummy variables $\{Q_{it} : t = 0, \dots, 16\}$, on the employment dummy $E_{it'}$ in subsequent quarters ($t' > t$). We model the employment and the training decision as a two equation system with possibly dependent individual specific effects. This means that we allow for selection into and out of training based on unobservables. As both dependent variables are binary we specify a random effects probit model for each. Following Chib and Hamilton (2002) and Chib and Jacobi (2007), we estimate the model using Bayesian MCMC techniques.[12] The treatment effects are estimated using simulations for both outcomes based on the MCMC iterations. We estimate the effect of both treatment incidence and the duration of treatment on a discrete outcome variable (employment) and we allow for state and duration dependence in the outcome variable.

Estimation of discrete choice-models for labor market transitions can be viewed as a discrete time version of the timing-of-events approach by AvdB which uses a continu-ous time duration model with unobserved heterogeneity, where time until treatment start and unemployment duration constitute two competing risks. Note, however, that our approach also models the length of the training program. The goal of the timing-of-events approach by AvdB is to estimate the causal treatment effect on the hazard to leave unemployment. AvdB show that identification of the causal effect of entering a program relies on the conditional randomness of program starts, a no-anticipation condition, independent variation of the covariate effects in the two duration equations, as well as functional form assumptions involving e.g. a mixed proportional hazard model. In applications of the AvdB approach (Richardson and van den Berg, 2008), the functional form assumptions used are conceptually similar

---

[12]Chib and coauthors analyze the binary treatment case with continuous outcomes and allow separate outcome equations.

to the ones used here. Similar to AvdB, our approach relies on a selection on un-observables strategy. Our estimates allow for heterogeneity of treatment effects and we estimate both the effect of training incidence and duration. Identification of the treatment effects in our model is implied in an analogous way by the two important assumptions of the timing-of-events approach by AvdB, i.e. the non-anticipation assumption and the assumption that the random effects are uncorrelated with the observed covariates. We rule out that anticipation of participation in the future affects current or future employment. These identifying assumptions are plausible in the present context as enrolment into training largely depends on short-term indicators (cf. section 2.1). Furthermore, identification of our model is helped by using data on all transitions observed during our observations window, i.e. by using multi–spell data in terms of a continuous time duration model.

3.1 Econometric Model

Now, we describe the model more formally. Consider first the employment equation. In order to model the employment dynamics we introduce employment lags up to the order of 15 (i.e. $E_{i(t-1)}$, $E_{i(t-2)}$, ..., $E_{i(t-15)}$, where i indexes individuals and t quarters) as explanatory variables for current employment status. A lagged variable only kicks in if the inflow into unemployment has not been too recent for the corresponding lag to be available, i.e. the jth lag kicks in if $t - j \geq t_0$, where t_0 denotes the quarter of the inflow into unemployment. This way we account for the entire employment history since the inflow into unemployment, thus accounting for both state dependence and duration dependence in the most flexible way, based on our discrete time data. Furthermore, we include calendar time and a vector of observed characteristics, $\boldsymbol{x}_{it,E}$, in the employment equation. In particular, we use information on schooling and occupational qualification, age, occupation, and earnings in the previous jobs, number of days employed in the last three years before the inflow into unemployment, health, children, labor market characteristics of the residential municipality, season and year. In addition, we control in a flexible way for the elapsed number of quarters an individual is in the panel, t, and the elapsed duration in the current employment or non-employment spell, denoted by $\tau_{it,E}$. Finally, we allow for a large number of interaction effects between lagged employment status, elapsed duration, and covariates.

Following the timing-of-events assumption, we assume that participation in training in a given quarter can only in subsequent quarters have a causal effect on the employment probability. Thus, the employment equation only includes lagged training status whose impact is modelled in a flexible way. The dummy variable $Q_{i(t-1)}$ indicates if the individual attended a training program in the previous quarter. If this dummy takes a one, lagged training is depicted by a dummy if participation so far has lasted two quarters $[Q_{(t-1)} = 1] \times [Q_{(t-2)} = 1] \times [Q_{(t-3)} = 0]$, three quarters $[Q_{(t-1)} = 1] \times [Q_{(t-3)} = 1] \times [Q_{(t-4)} = 0]$, four quarters $[Q_{(t-1)} = 1] \times [Q_{(t-4)} = 1] \times [Q_{(t-5)} = 0]$ and so forth.

We distinguish explicitly the effects on the exit from nonemployment in the next quarter while being in training *from* the effect on future employment in subsequent quarters after training has ended. To do so, we add a dummy D_{it} indicating whether an individual has ever participated in training since the inflow quarter. For trainees who have already exited the program, i.e. individuals having received training before quarter $t-1$ for whom the dummy on participation in training in the last quarter $Q_{i(t-1)}$ takes a zero, D_{it} is equal to one. We account for the training history by interacting D_{it} with $\tau_{it,Q}$, which in this case indicates the completed duration, as well as a polynomial of the elapsed quarters since program start, $(t - t_{s,i})$, where $t_{s,i}$ is the quarter in which training starts for individual i. An interaction effect of $\tau_{it,Q}$ and $(t - t_{s,i})$ is also added. To allow for state dependence of the training effects in the outcome equation, the variables reflecting training history are interacted with other explanatory variables in the employment equation and the first lag of employment status $E_{i(t-1)}$.

Specifically, the employment dummy E_{it} is modelled by:

$$
(1) \qquad E_{it} = \mathbb{1}[\boldsymbol{g}\left(E_{i(t-1)}, \ldots, E_{i(t-15)}, \boldsymbol{x}_{it,E}, \tau_{it,E}, t\right) \boldsymbol{\beta}_E + \gamma_{0E} Q_{i(t-1)} +
$$

$$
D_{it}(1 - Q_{i(t-1)}) \boldsymbol{m}\left(E_{i(t-1)}, \boldsymbol{x}_{it,E}, \tau_{i(t-1),Q}, (t - t_{s,i})\right) \boldsymbol{\delta}_E
$$

$$
Q_{i(t-1)} \boldsymbol{h}\left(Q_{i(t-2)}, \ldots, Q_{i(t-16)}, \boldsymbol{x}_{it,E}, \tau_{i(t-1),Q}\right) \boldsymbol{\gamma}_{1E} + \alpha_{i,E} + \epsilon_{it,E} > 0]
$$

where $\mathbb{1}[\boldsymbol{.}]$ is the indicator function, $\boldsymbol{g}(\boldsymbol{.})$, $\boldsymbol{h}(\boldsymbol{.})$, and $\boldsymbol{m}(\boldsymbol{.})$ denote vector-valued functions, and $\boldsymbol{\beta}_E$, $\boldsymbol{\gamma}_{1E}$, $\boldsymbol{\delta}_E$ denote conformable column-vectors of coefficients. These functions may involve various interaction effects. $\alpha_{i,E}$ is the individual specific effect and $\epsilon_{it,E}$ the idiosyncratic error term. The treatment effect of training on employment is captured through the coefficients γ_{0E}, $\boldsymbol{\gamma}_{1E}$, and $\boldsymbol{\delta}_E$. It is allowed to differ by

the duration of training, the time since training start, and a number of covariates. Specification (1) allows the information determining the employment probability to accumulate endogenously over time. Therefore, the effects of lagged employment are allowed to change by allowing for interactions with time t and other covariates. We account for duration and state dependence in a flexible way. The exact specification of the functions $\boldsymbol{g}(.)$, $\boldsymbol{h}(.)$, and $\boldsymbol{m}(.)$ is guided by whether some coefficients are significant.[13]

Consider next the participation equation modeling the transition into and out of training. It is estimated simultaneously with the employment equation if the individual is not employed in the respective quarter and has not yet left a training program. Since participation can only occur during non-employment the two equation system reduces to a single equation for observations for which the employment status, E_{it}, is equal to one. Then, the treatment equation is switched off. This triangular structure provides additional identifying restrictions. We do not consider reentry into training after the completion of a first training program because this only occurs very rarely in our data and we consider this to be a different treatment.

The training equation includes a vector of observed regressors, $\boldsymbol{x}_{it,Q}$. In particular, we include variables driving the decision to enter and to stay in a program. The covariate vector comprises a dummy indicating whether the individual was enrolled in training in the previous quarter, $Q_{i(t-1)}$, a variable for the elapsed quarters in the program $\tau_{it,Q}$, a polynomial of the time until the planned end (in case the planned duration is not yet exhausted) and a dummy if the planned end is missing in the data. These variables are equal to zero if the individual was not yet enrolled in training ($Q_{i(t-1)} = 0$). Furthermore, the vector of independent variables includes variables summarizing the employment history since the inflow quarter, dummy variables indicating whether the current quarter is the inflow quarter, as well as whether a repeated transition from employment to non-employment has occurred, and a polynomial of the elapsed unemployment duration in days. Finally, information on age, schooling, vocational training, last job, number of days in employment during the last three years before the inflow, health, children and entitlement to unemployment compensation, season, and year are incorporated.

---

[13]We use the posterior variances and covariances to undertake the analogy of classical significance tests.

Specifically, the participation dummy Q_{it} is modelled by:

$$(2) \qquad Q_{it} = \quad \mathbb{1}[\boldsymbol{p}\left(Q_{i(t-1)}, \boldsymbol{x}_{it,Q}, \tau_{it,Q}\right)\boldsymbol{\beta}_Q + \alpha_{i,Q} + \epsilon_{it,Q} > 0]$$

where $\boldsymbol{p}(\boldsymbol{.})$ denotes a vector-valued function, $\boldsymbol{\beta}_Q$ the conformable coefficient vector, $\alpha_{i,Q}$ the individual specific effect, and $\epsilon_{it,Q}$ the idiosyncratic error term.

The two individual specific random effects, $\alpha_{(i,E)}$ and $\alpha_{(i,Q)}$, follow a *joint* normal distribution, $(\alpha_{i,E}, \alpha_{i,Q})' \sim \mathcal{N}(\boldsymbol{0}, \Sigma)$. The error terms $\epsilon_{it,E}$ and $\epsilon_{it,Q}$ are independent standard normals. Thus, the model includes two individual specific effects which are allowed to be correlated. Let $\boldsymbol{z}_{it,y}$, $y \in \{E, Q\}$, denote the entire vector of covariates including lagged endogenous variables and interaction terms in the employment and training equation, respectively, $\boldsymbol{\eta}'_E = (\boldsymbol{\beta}_E, \gamma_{0E}, \boldsymbol{\gamma}_{1E}, \boldsymbol{\delta}_E)'$, $\boldsymbol{\eta}_Q = \boldsymbol{\beta}_Q$. T_i is the number of quarters individual i is in the panel. Then the likelihood contribution of individual i is as follows:

$$(3) \qquad L_i = \int \prod_{t=1}^{T_i} f(E_{it}|\boldsymbol{z}_{it,E}, \alpha_{i,E}; \boldsymbol{\eta}_E) \cdot f(Q_{it}|\boldsymbol{z}_{it,Q}, \alpha_{i,Q}; \boldsymbol{\eta}_Q)^{C_{it}} dG(\alpha_{i,E}, \alpha_{i,Q})$$

where $f(y_{it}) = \Phi(\boldsymbol{z}_{it,y}\boldsymbol{\eta}_y + \alpha_{i,y})^{y_{it}} \cdot (1 - \Phi(\boldsymbol{z}_{it,y}\boldsymbol{\eta}_y + \alpha_{i,y}))^{(1-y_{it})}$, $y \in \{E, Q\}$, $\Phi(\boldsymbol{.})$ denotes the standard normal cumulative distribution function, and C_{it} is a dummy equal to one if the individual is non-employed and has not yet completed a training program. As the individual specific effects are not directly observed one would have to integrate them out with respect to the joint distribution $G(\alpha_{i,E}, \alpha_{i,Q})$ in order to estimate the model by maximum likelihood. In this paper, we use Bayesina MCMC estimation, which avoids maximization of the likelihood function. We obtain estimates of the posterior distribution of all parameters of the model including predictions of the random effects. These allow us to estimate the posterior distribution of different treatment effects of interest using a simulation approach.

3.2 MCMC Estimation of a Random Effects Probit Model

We estimate the model introduced in the previous section using Bayesian MCMC techniques (see Chib, 2001, for an overview on MCMC techniques). The values of the parameters along the MCMC iterations allow to estimate the posterior distribution of the parameters and of other model parameters of interest. From a classical perspective, the mean of the posterior distribution converges to the maximum of the

likelihood function and the variance of the posterior distribution converges to the asymptotic variance of a ML estimation. Thus, the standard deviation of the draws may be interpreted as standard errors from the classical perspective (Train, 2003). To obtain a sample from the posterior distribution, we use the Gibbs sampler. To simplify the sampling from a complex joint distribution, the Gibbs sampler forms blocks of model parameters and samples recursively from the distribution of one block conditional on the current values of the remaining parameters. The resulting sequence of simulated parameters is a Markov Chain whose invariant distribution is the desired posterior distribution. After convergence, the draws are samples from this posterior distribution.

The key idea for the estimation of probit models is to augment the data by simulating the latent variables in equations (1) and (2) as one step of the Gibbs sampler and to use standard linear estimation techniques to obtain the corresponding draws for the β–coefficients (Albert and Chib, 1993). Furthermore, one step of the Gibbs sampler involves sampling the random effects $\alpha_{i,E}, \alpha_{i,Q}$ conditional on the data and the remaining model parameters (Zeger and Karim, 1991).[14] The posterior distribution of the individual specific random effects allows to estimate the expected value of the random effects of one individual given the data and the prior distribution of parameters. These predictions allow for an assessment of selectivity effects.

We provide details of the algorithm in Appendix B. Conjugate but very diffuse priors are used. The results reported below are based on running the algorithm for 50,000 iterations. We monitor convergence by comparing the means at different stages of the chains. We discarded the first 5,000 iterations (the burn-in phase). Thus, our results are based on 45,000 draws. We implement the Gibbs sampler in Stata.

3.3   Estimation of Treatment Effects

The raw coefficient estimates are difficult to interpret because of the complex dynamic structure of the model involving many interaction effects. Therefore, we analyze directly the posterior distribution of various treatment effects of interest.

---

[14]Odejar (2002) proposes a Gibbs sampler for a model sharing important features with the one estimated in this paper. Buchinsky et al. (2005) and Horny et al. (2008) apply similar models to study employment and wage mobility but they do not make use of the predictions of the random effects.

We consider the following average effects of treatment on the treated (ATT):

Classical ATT. This is the ATT of training versus non-participation during the observation period.

Training versus Waiting. At any given quarter, those enrolling into training during this quarter are counted as treated whereas those not yet enrolling are assigned to the control group. The latter may potentially participate in a later quarter. This effect mimics the treatment parameter suggested by Sianesi (2004) and estimated in several subsequent papers for European training programs using propensity score matching (see for example Biewen et al., 2007).

Effect of a Given Planned Program Duration. Among the treated individuals, we estimate the effects of different planned program durations, allowing for the realized program length to be endogenous. In particular, we compare the effect of attending a program with a planned length of one, three, and four quarters, respectively, to attending a program with a planned duration of two quarters.

To estimate these treatment effects, we simulate draws from the posterior distribution of these treatment effects based on the sequence of MCMC iterations. To account for selection based on unobservables, we use the draws of the individual random effects $\alpha_{i,E}$ and $\alpha_{i,Q}$ from the MCMC estimation of the model. The details of the simulation procedure are given below.

First, we describe the simulation of the *Classical ATT*. For every 30th draw of the MCMC iterations (after the burn-in phase), we go through the following steps:

Step 1. For each participant, predict the treatment outcome E_{it}^1 starting with the first period after program participation ($t - t_a = 1$, where t_a denotes the last quarter of program participation).[15] In particular, go from $t - t_a = 1$ to $t - t_a = 9$ and predict the employment status for each period based on the corresponding draw from the vector of coefficients η_E, the vector of explanatory variables $z_{it,E}$, the corresponding draw of the $\alpha_{i,E}$ and a draw of the idiosyncratic error term $\epsilon_{it,E}$. The dynamic elements of $z_{it,E}$, such as lags of

---

[15]During program participation, the employment dummy is equal to zero.

employment status, are updated when moving from one quarter to the next. $\epsilon_{it,E}$ is drawn from a standard normal distribution.[16]

Step 2. For each participant, simulate the counterfactual employment outcome (i.e. the outcome if the participant had not participated in a program) E_{it}^0 for each period beginning with the quarter of program start ($t - t_s = 0$, where t_s denotes the first quarter of program participation). Again go through the dynamic process and predict the employment status for each period based on the same η_E, $\alpha_{i,E}$ and $\epsilon_{it,E}$ as before. Adapt the $z_{it,E}$ to a situation with no participation and update them while going through the process.

Step 3. To get a draw of the ATT aligned to the end of the program, average the difference of the two predictions over all treated individuals (N_1), i.e. $\frac{1}{N_1} \sum_{i=1}^{N_1} (E_{it}^1 - E_{it}^0)$ (t aligned to end of program), for each period $t > t_a$. This gives a draw from the posterior distribution of the ATT for each quarter.

Step 4. For a draw of the ATT aligned to the start of the program, average the difference of the two predictions over all treated individuals, i.e. $\frac{1}{N_1} \sum_{i=1}^{N_1} (E_{it}^1 - E_{it}^0)$ (t aligned to start of program), for each period $t \geq t_s$.

The resulting 1,500 draws provide an estimate of the posterior distribution of the *Classical ATT*. We estimate the ATT by the mean of the posterior distribution and we use the standard deviation as our estimate of estimation uncertainty.

Second, the estimation of the posterior distribution of the effect of *Training versus Waiting* proceeds in an analogous way. Step 1 remains the same. The counterfactual employment outcome E_{it}^w relates to a situation in which the participant does not start a program in the observed start quarter (t_s), so the employment status $E_{it_s}^w$ is simulated and $Q_{it_s}^w$ is set to zero. The individual may start a program later. Thus, from $t = t_s$ onwards, the counterfactual employment status E_{it}^w and the counterfactual participation status Q_{it}^w are simulated in turn, adapting the elements of $\mathbf{z}_{it,E}$ and $\mathbf{z}_{it,Q}$ that include lagged employment or participation status while going through the dynamic processes. Note that we do not allow $Q_{it}^w = 1$ if $E_{it}^w = 1$ or if individual i has already left a program. The simulation of E_{it}^w uses the respective draw of $\boldsymbol{\eta}_E$ and $\alpha_{i,E}$. The simulation of Q_{it}^w relies on the draw of $\boldsymbol{\eta}_Q$ and $\alpha_{i,Q}$. The

---

[16]One may wonder why we simulate the observed treatment outcome for the treated. Doing so, the posterior distribution of the treatment effects takes account of the randomness of the treatment outcomes.

same $\epsilon_{it,E}$ as in step 1 are used and the $\epsilon_{it,Q}$ are drawn from a standard normal distribution. In order to calculate a draw of the effect of *Training versus Waiting* aligned to the start of the program, average the difference of the two predictions for each period over all participants ($\frac{1}{N_1}\sum_{i=1}^{N_1}(E_{it}^1 - E_{it}^w)$, t aligned to start of program). Likewise, calculate $\frac{1}{N_1}\sum_{i=1}^{N_1}(E_{it}^1 - E_{it}^w)$ (t aligned to end of program) to get a draw aligned to the end of the program.

Finally, consider the *Effect of a Given Planned Program Duration*. For each 30th draw, we simulate the employment status for different planned program durations. First, we simulate the employment status and participation status for a situation in which all participants are assigned to a program with a planned length of two quarters.[17] For the quarter when a participant starts the program ($t = t_s$), Q_{it}^{p2} is set to one and $E_{it_s}^{p2}$ is set to zero, as in the original data. In the next quarter ($t - t_s = 1$), E_{it}^{p2} also remains the same as in the original data. The participation status $Q_{it_S}^{p2}$ in t_s is then simulated and in the following quarters $t - t_s > 1$, E_{it}^{p2} and Q_{it}^{p2} are simulated in turn for each period. The elements of $\boldsymbol{z}_{it,E}$ and $\boldsymbol{z}_{it,Q}$ that include lags of employment or participation status are adapted while going through the dynamic processes. Again, the values of $\boldsymbol{\eta}_E$, $\boldsymbol{\eta}_Q$, $\alpha_{i,E}$, $\alpha_{i,Q}$ of the corresponding draw are used. $\epsilon_{it,E}$ and $\epsilon_{it,Q}$ are drawn from a standard normal distribution, respectively. Similarly, the employment status and the participation status are simulated for each period for the alternative scenario in which the planned program duration is set to one quarter (E_{it}^{p1} and Q_{it}^{p1}), three quarters (E_{it}^{p3} and Q_{it}^{p3}), and four quarters (E_{it}^{p4} and Q_{it}^{p4}), respectively. The same $\epsilon_{it,E}$ and $\epsilon_{it,Q}$ as before are used. As the median planned duration of the programs in the data is slightly more than two quarters, we take E_{it}^{p2} as a benchmark and calculate the effect of a planned duration of one quarter as opposed to two quarters ($\frac{1}{N_1}\sum_{i=1}^{N_1}(E_{it}^{p1} - E_{it}^{p2})$), of three quarters versus two quarters ($\frac{1}{N_1}\sum_{i=1}^{N_1}(E_{it}^{p3} - E_{it}^{p2})$), and four quarters versus two quarters ($\frac{1}{N_1}\sum_{i=1}^{N_1}(E_{it}^{p4} - E_{it}^{p2})$) for each period $\{t : 0 \leq t - t_s \leq 9\}$. Again, we estimate the posterior distribution based on 1,500 draws from the simulations for the MCMC iterations.

---

[17]In terms of the model specification, this means that the explanatory variables in the participation equation involving the planned end date (i.e. *days/91 until planned end if enough duration left* and *days/91 until planned end if enough duration left squared*) are adapted to this scenario. As these variables are measured in days, the decision whether there is still enough planned duration left and the values these variables take depend on the day within a quarter at which the program starts. Note that it is possible in the simulation as well as in the original data that realized program participation continues beyond the planned end date. The variable *planned end missing* is set to zero.

4 Estimation Results

We estimate the impact of incidence and duration of training on the transition probabilities between employment and non-employment using the MCMC estimation approach described in the previous section. Our empirical model accounts for selection into training based on unobservables. Estimation is carried out separately for West German males, West German females, East German males, and East German females. The detailed estimation results are given in table 5 in Appendix C. The first column for each sample refers to the mean of the coefficients and the second to their standard deviation over MCMC iterations after the burn-in phase. We interpret them in an analogous way as the point estimates and standard errors of the coefficients obtained by a frequentist approach. Next, we briefly discuss the overall fit of the model and the individual level variances of the error terms. Because of the complexity of the model (it comprises about 160 parameters), we refrain from further discussing single parameters. Rather, we assess the estimated model in general and discuss the results for different treatment effects of interest.

4.1 Model Fit and Selection on Unobservables

Evidence on the fit of the model is provided in table 2 for the treated individuals from the start of the program onwards along with information on the number of observations available in each quarter. Actual and predicted employment rates of the trainees match closely in all four samples. Thus, our rich model specification does a good job in replicating the employment dynamics found in the data which suggests that our model is not grossly misspecified.

— Insert table 2 about here. —

The last panel of table 5 in Appendix C displays results on the variances and co-variances of the error terms of the employment and the training equation. The share of the variance that is due to the random effects varies between 36% and 51% for the employment equation and between 22% and 31% for the training equation. Except for the sample of West German females, the correlation between the two random effects tends to be significantly negative. The correlation coefficient is -.22 and significant at the five percent level for females in East Germany. It is -.11 to

-.12 for West and East German males and significant at the ten percent level. This suggests that those individuals who have a higher unobserved propensity to enter a program and to stay in a program tend to have a lower unobserved propensity to be employed.

4.2 Classical Treatment Effect on Employment and on Transition Rates

Figure 5 shows the average effect of training versus no training for participants on employment by quarter.[18] More precisely, we compare the average of the actual employment outcomes of trainees with the expected counterfactual outcome obtained by setting the lags of training status in the employment equation to zero. In figure 5, the average difference in the quarterly employment rates is depicted on the vertical axis, while quarters since program start are measured on the horizontal axis. The dashed lines around the estimated treatment effects are 95 percent confidence bands. We conclude that the treatment effect for a particular quarter is statistically significant if the confidence band does not contain zero. As can be seen, a participation in training reduces the employment probability during the first three to five quarters after program start. During the first two quarters $(t - t_s \leq 1)$ the employment probabilities of participants decline between seven (East German females) and 15 (East German males) ppoints compared to the situation of no participation. This lock-in effect lasts one quarter longer in the East German samples compared to the West German ones. After the first year counted from the quarter of program start, the difference in employment rates turns positive and continues to increase until the end of the observation window. Ten quarters after program start $(t - t_s = 9)$, West German females have a 21 ppoints higher employment probability than in the absence of training. The effects are of similar magnitude for East German females (17 ppoints) and somewhat smaller for the male samples (12 ppoints).

— Insert figure 5 about here. —

Based on the model simulations used to calculate the classical ATT, we also calculate transition rates from non-employment to employment and from employment

---

[18]The corresponding numbers are given in table 6 in Appendix C.

to non-employment. Table 3 depicts the difference in transition rates from non–employment to employment based on the simulated treatment outcomes as opposed to the simulated non-treatment outcomes. Table 4 depicts the difference in transition rates from employment to non-employment. When calculating the transition rates, the denominator is the number of individuals in non-employment (or employment, respectively) in the previous quarter. This number may differ between the situation where the individuals are treated and the counterfactual situation, thus the differences in transition rates can not be interpreted as conventional treatment effects because the outcome in the previous quarter is affected itself by the treament.

The results of table 3 suggest higher transition rates from non-employment to employment due to training after a lock-in period of two to three quarters. The difference in transition rates between the treatment and the non-treatment scenario reaches its maximum four quarters after treatment start, that is soon after the planned end of a typical program. Table 4 indicates much lower transition rates from employment back to non-employment, suggesting that training increases subsequent employment stability.

— Insert table 3 about here. —

— Insert table 4 about here. —

4.3 Training versus Waiting

In the recent evaluation literature, researchers often estimate the effect of treatment at a given point in time versus no treatment at that point in time, implying that the treatment may take place at some later point in time. Thus, the control group involves both individuals who never participate and those who defer participation. This effect is commonly referred to as the effect of training versus waiting (Sianesi, 2004). In order to mimic this parameter, we simulate the training status of the actual trainees imposing that they postpone participation at least one quarter beyond their observed true program start. This entails the possibility that the simulated training dummies are zero during all quarters for some individuals. Table 7 in Appendix C depicts the actual and the simulated participation rates. From the last row of table 7, it can be seen that under the simulated waiting scenario only 68.7 % of

the original participants ever enrol into training. We calculate the effect of training versus waiting as the difference between the actual employment rates of the trainees and their simulated employment rates under the waiting scenario.

— Insert figure 6 about here. —

The results are displayed in Figure 6.[19] There are negative lock-in effects of similar magnitude and length as for the classical effect of training in figure 5. In the quarter in which the program starts, participants in training have a seven to 14 ppoints lower probability to be in employment than compared to the situation of not yet starting a program. After about four to five quarters ($t - t_s = 3, 4$), the treatment effects turn positive and then increase further during the subsequent quarters. In quarter ten since program start ($t - t_s = 9$), they lie in the range of 7 to 15 ppoints. This is about a third less than in the case of a pure no-training control group.

Based on our model estimates, the estimated effect for treatment versus waiting underestimates the causal (classical) treatment effect in the medium and long run. This finding is due to the fact that control persons who obtain training in the near future also experience positive treatment effects in the medium and long run.

4.4 Treatment Effects for Different Planned Training Durations

Here, we use our model estimates to analyze how treatment effects vary with the planned program duration. Typically, participation in training decreases the exit rates from unemployment between the start and the end date of a program compared to a situation of no participation (lock-in effect). The size of the reduction in job finding probabilities may change over the course of treatment. If job finding efforts increase towards the end of the program, the time until the scheduled end date has a negative effect on exits from unemployment. In a mechanical sense, a shorter planned enrolment length should therefore be associated with a shorter and less pronounced lock-in period. However, it is unclear whether such an advantage in the short run persists over time. If, by administering shorter programs, one could decrease the lock-in effect without reducing the long-run employment gains, policy makers would be advised to shorten planned durations.

---

[19]The corresponding numbers are given in table 8 in Appendix C.

Similar to above, we simulate the training and employment histories of the subsample of training participants that result after fixing the planned program duration to a prespecified value. Specifically, we consider planned program durations of one, two, three, and four quarters. We then evaluate the effect of participating in a program scheduled over one, three, and four quarters, respectively, as opposed to two quarters, the median of planned duration. Tables 9 to 12 in Appendix C show the simulated participation and employment rates associated with different planned program durations. Note that the simulated realized program duration can be shorter or longer than the planned one. However, the tables suggest that there is a strong positive correlation between planned and realized program durations.

— Insert figure 7 about here. —

— Insert figure 8 about here. —

— Insert figure 9 about here. —

Figure 7 displays the treatment effects associated with a planned duration of one versus two quarters.[20] The gains of a shorter participation are small and only transient. In the third quarter after program start ($t - t_s = 2$), the employment rate is between one and two ppoints higher. In the medium and long run, those attending programs with a scheduled length of two quarters fare better, exhibiting employment rates that are consistently higher by three to five ppoints. A similar pattern emerges when comparing programs with a scheduled length of three and four quarters, respectively, with those planned to last two quarters, cf. figures 8 and 9.[21] Trainees attending longer programs are only slightly worse off during the additional quarters they are supposed to be in the program. After the scheduled end of the longer program, they have consistently higher employment rates than compared to the benchmark case of a six-month program. Indeed, compared to a planned duration of two quarters, the employment rates associated with attending a nine-month program are four to six ppoints higher and those associated with attending a one-year program are six to eleven ppoints higher.

---

[20]The corresponding numbers are given in table 13 in Appendix C.

[21]The corresponding numbers are given in tables 14 and 15 in Appendix C.

5 Conclusion

Training programs are an important part of active labor market policies of many advanced countries. This paper examines the effects of training incidence and duration on labor market transitions in discrete time. We estimate a dynamic random effects probit model including an employment and a participation equation based on large administrative data for Germany. The participation equation models the start of training as well as its end accounting for endogenous dropout. We control for selection on unobservables by allowing the random effects of both equations to be correlated. We account for time and duration dependence as well as for various forms of effect heterogeneity in a flexible way. Using Bayesian Markov Chain Monte Carlo (MCMC) methods, we estimate the posterior distribution of the model parameters, including the individual random effects. The analysis is separately done for West and East Germany and for males and females. Because of the complexity and the dynamic nature of our model, we simulated different treatment effects of interest using the estimated coefficients and individual specific effects based on the sequence of MCMC iterations.

Our findings imply positive effects of training on the employment probability emerging three to four quarters after program start in all four subsamples considered. Ten quarters after program start, the effect of treatment on unconditional employment rates for the treated individuals lies between 12 and 21 percentage points (ppoints). The effects are higher for women than for men and higher in West Germany than in East Germany. Our estimation approach allows us to estimate various other treatment effects of interest. Similar to the recent literature applying sequential matching techniques to estimate the effect of treatment at a given date versus not at that date but possibly later (cf. Sianesi, 2004), we simulate the effect of training versus waiting. Based on our model estimates, we find that the effect of treatment versus waiting underestimates the treatment effect of training versus no training in the medium and long run by about a third.

Finally, we use our estimates to analyze how training effects vary with planned program duration. Longer planned enrolment lengths of three and four quarters, respectively, as opposed to just two quarters lead to an increase in employment rates by four to six ppoints and six to eleven ppoints, respectively, in the medium and long run. This suggests that, on average, the higher costs of longer training

26

programs translate into higher long-run employment gains.

References

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

Abbring, J. and G.J. van den Berg (2004). "Analyzing the Effect of Dynamically Assigned Treatments using Duration Models, Binary Treatment Models, and Panel Data Models." *Empirical Economics* 29, 5-20.

Albert, J.H. and S. Chib (1993). "Bayesian Analysis of Binary and Polychotomous Response Data." *The Journal of the American Statistical Association* 88, 669-679.

Bergemann, A., B. Fitzenberger and S. Speckesser (2009). "Evaluating the Dynamic Employment Effects of Training Programs in East Germany Using Conditional Difference-in-Differences." *Journal of Applied Econometrics*, forthcoming.

Biewen, M., B. Fitzenberger, A. Osikominu and M. Waller (2007). "Which Program for Whom? Evidence on the Comparative Effectiveness of Public Sponsored Training Programs in Germany." IZA Discussion Paper, No. 2885, Bonn.

Buchinsky, M., D. Fougère, F. Kramarz and R. Tchernis (2005). "Interfirm Mobility, Wages, and the Returns to Seniority and Experience in the U.S." IZA Discussion Paper No. 1521.

Bundesanstalt für Arbeit, BA, (2001). *Arbeitsstatistik 2000 - Jahreszahlen.* Nürnberg.

Bundesagentur für Arbeit, BA, (2006). *Arbeitsstatistik 2005 - Jahreszahlen.* Nürnberg.

Bundesagentur für Arbeit, BA, (2007, 2008). *Arbeitsmarkt 2006, 2007.*" Nürnberg.

Card, D., J. Kluve, and A. Weber (2009). "Active Labor Market Policy Evaluations: A Meta-Analysis." CESifo Working Paper, No. 2570, Munich.

Chib, S. (2001). "Markov Chain Monte Carlo Methods: Computation and Inference." In J. Heckman and E. Leamer (eds.): *Handbook of Econometrics*, Volume 5, Amsterdam: Elsevier Science, 3569-3649.

Chib, S. and B.H. Hamilton (2002). "Semiparametric Bayes Analysis of Longitudinal Data Treatment Models", *Journal of Econometrics* 110, 67-89.

Chib, S. and L. Jacobi (2007). "Modeling and Calculating the Effect of Treatment at Baseline from Panel Outcomes." *Journal of Econometrics* 140, 781-801.

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

Flores-Lagunes A., A. Gonzalez, and T.C. Neumann (2007). "Estimating the Effects of Length of Exposure to a Training Program: The Case of Job Corps." Working Paper #521, Princeton University, Industrial Relations Section.

Fredriksson, P. and P. Johansson (2008). "Dynamic Treatment Assignment: The Consequences for Evaluations Using Observational Data." *Journal of Business and Economic Statistics* 26, 435-445.

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

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

Heckman, J. and S. Navarro (2007). "Dynamic Discrete Choice and Dynamic Treatment Effects." *Journal of Econometrics* 136, 341–396.

Horny, G., R. Mendes and G.J. van den Berg (2009). "Job Durations and Firm Specific Effects: MCMC Estimation with Longitudinal Employer-Employee Data." IZA Discussion Paper No. 3992, Bonn.

Hujer, R., S.L. Thomsen and C. Zeiss (2006): "The Effects of Vocational Training Programmes on the Duration of Unemployment in Eastern Germany." *AStA Advances in Statistical Analysis* 90, 299-321.

Hirano, K. and G.W. Imbens (2004). "The Propensity Score with Continuous Treatments." In A. Gelman and X.L. Meng (eds.), *Applied Bayesian Modeling and Causal Inference from Incomplete-Data Perspectives*. West Sussex: Wiley, 73-84.

Imai, K. and D.A. van Dijk (2004). "Causal Inference With General Treatment Regimes: Generalizing the Propensity Score." *Journal of the American Statistical Association* 99, 854-866.

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

Kluve, J., H. Schneider, A. Uhlendorff and Z. Zhao (2007). "Evaluating Continuous Training Programs Using the Generalized Propensity Score." IZA Discussion Paper No. 3255, Bonn.

Kühnlein, G. and B. Klein (2003). "Bildungsgutscheine - mehr Eigenverantwortung, mehr Markt, mehr Effizienz? - Erfahrungen bei der Neuausrichtung der beruflichen Weiterbildung." Arbeitspapier Band 74, Hans-Böckler-Stiftung.

Lechner, M. (2001). "Identification and Estimation of Causal Effects of Multiple Treatments Under the Conditional Independence Assumption." In M. Lechner and F. Pfeiffer (eds.), *Econometric Evaluation of Labour Market Policies.* Heidelberg, New York: Physica, 43-58.

Lechner, M., R. Miquel and C. Wunsch (2009). "Long-Run Effects of Public Sector Sponsored Training in West Germany." *Journal of the European Economic Association*, forthcoming.

Martin, J.P. and D. Grubb (2001). "What Works for Whom: A Review of OECD Countries' Experiences with Active Labour Market Policies." *Swedish Economic Policy Review* 8, 96.

OECD (2007). *OECD Employment Outlook.* Paris.

Odejar, A. (2002). "Bayesian Analysis of Sample Selection and Endogenous Switching Regression Models with Random Coefficients Via MCMC Methods." SFB 386 Discussion Paper No. 291.

Osikominu, A. (2008). "Is Short Training Short-Lived and Long Training Long-Lasting? A Multi-State Duration Analysis of the Dynamic Effects of Training Schemes for the Unemployed." Discussion Paper, Albert-Ludwigs-University Freiburg.

Plaßmann, G. (2002). "Der Einfluss der Arbeitslosenversicherung auf die Arbeitslosigkeit in Deutschland." *Beiträge zur Arbeitsmarkt- und Berufsforschung, BeitrAB* 255.

Richardson, K. and G.J. van den Berg (2008). "Duration dependence versus unobserved heterogeneity in treatment effects: Swedish labor market training and the transition rate to employment." IFAU Working Paper, 2008:7, Uppsala.

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

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

Train, K. (2003). *Discrete Choice Methods with Simulation.* Cambridge: Cambridge University Press.

Waller, M. (2008). "On the Importance of Correcting Reported End Dates of Labor Market Programs." *Schmollers Jahrbuch* 128, 213-236.

Waller, M. (2009). Many Dropouts? Never Mind! - Employment Prospects of Dropouts from Training Programs, Freiburg University, not published.

Zeger, S. and M. Karim (1991). "Generalized linear models with random effects: a Gibbs sampling approach." *Journal of the American Statistical Association*, 86, 79-86.

Zimmermann, R., S. Kaimer, D. Oberschachtsiek (2007). "Dokumentation des Scientific Use Files der Integrierten Erwerbsbiographien (IEBS-SUF V1) Version 1.0." FDZ Datenreport 1/2007, Institut für Arbeitsmarkt- und Berufsforschung (IAB), Nürnberg.

Figures

Figure 1: Entries into Training Programs in West and East Germany (in 1000)
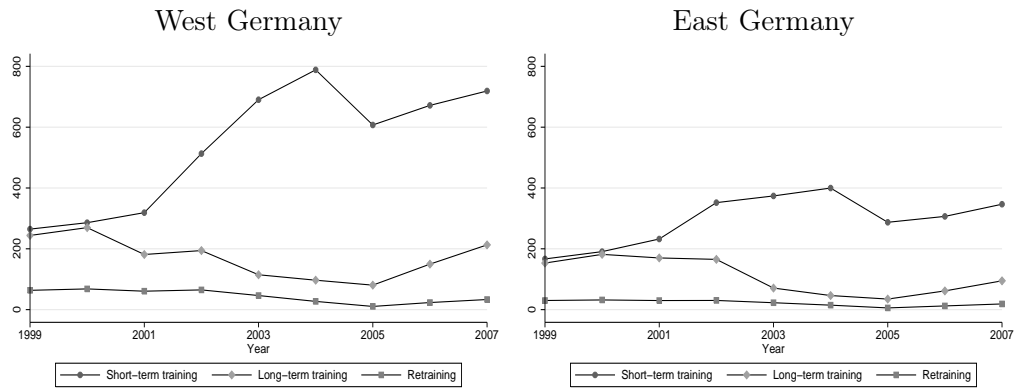
Source: BA (2001, 2006, 2007, 2008); own calculations.

Figure 2: Planned and Realized Training Durations
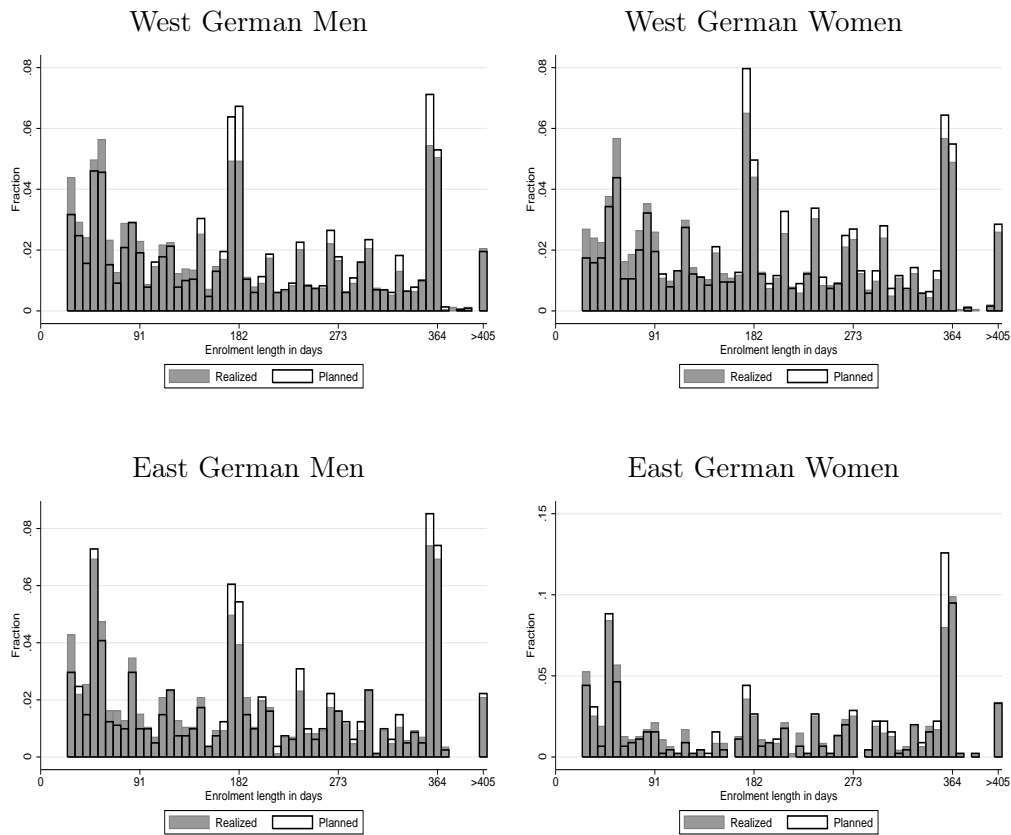
Figure 3: Employment and Participation Rates over Time
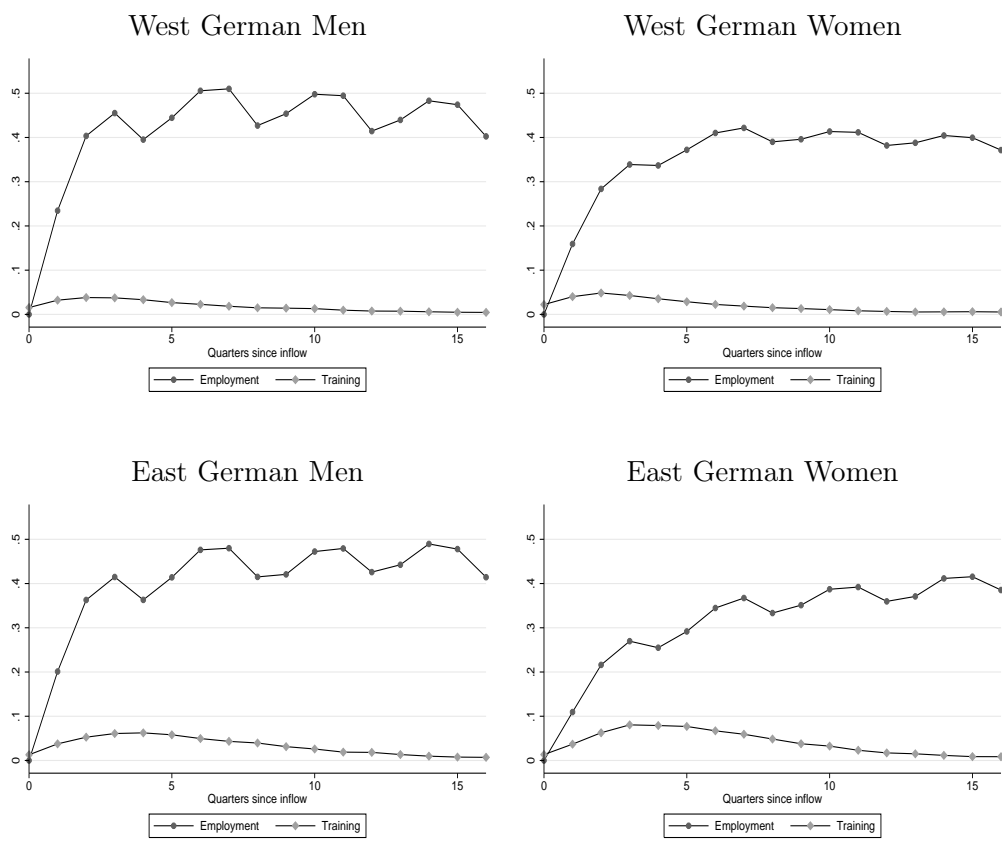
Figure 4: Raw Treatment and Nontreatment Employment Rates

West German Men

West German Women

East German Men

East German Women
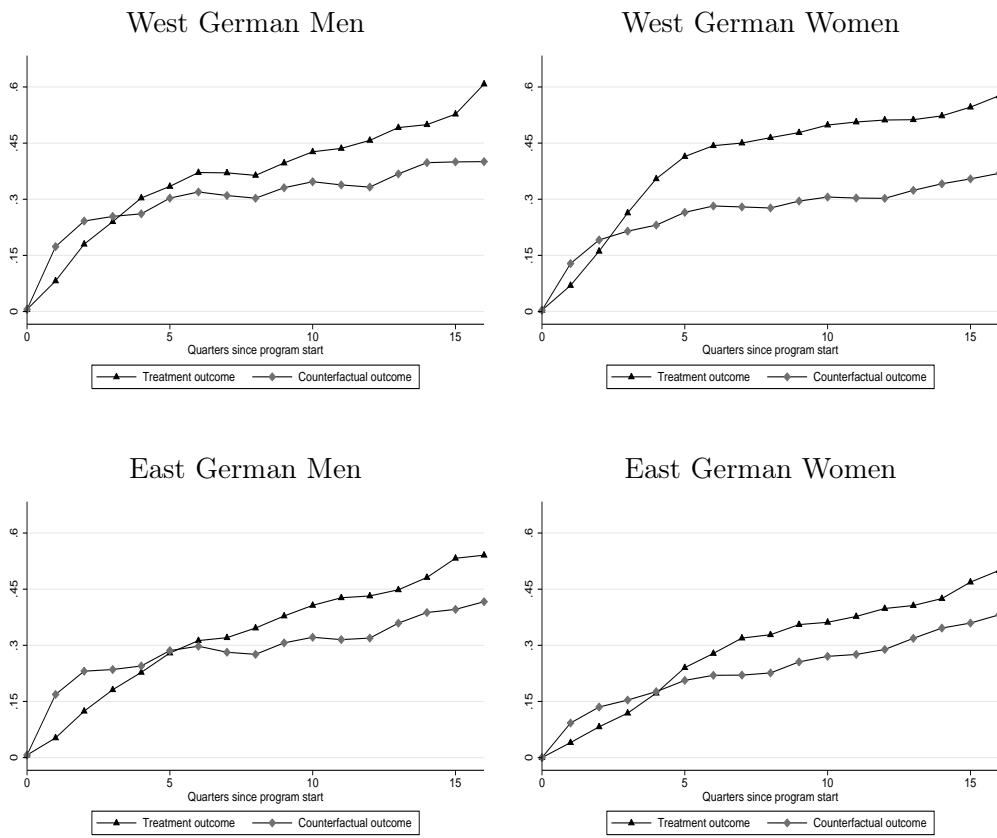
Note: Raw estimates of the treatment effect on the treated, where treated and controls are aligned in the time dimension only. In particular, treated and nontreated individuals are matched on the calendar quarter of their first inflow and elapsed unemployment duration in the current spell. No adjustments are made for other potential sources of selection bias.

Figure 5: Classical Average Treatment Effect on the Treated

West German Men

West German Women

East German Men

East German Women
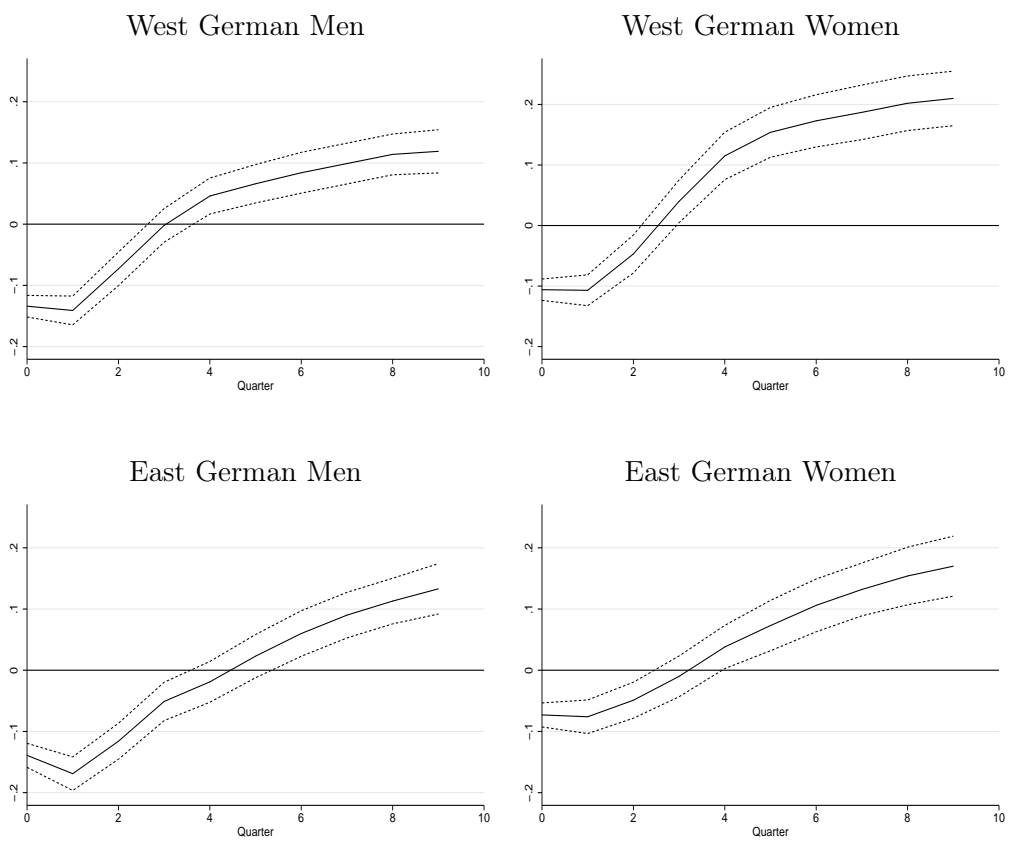
Figure 6: ATT of Training versus Waiting

West German Men

West German Women

East German Men

East German Women
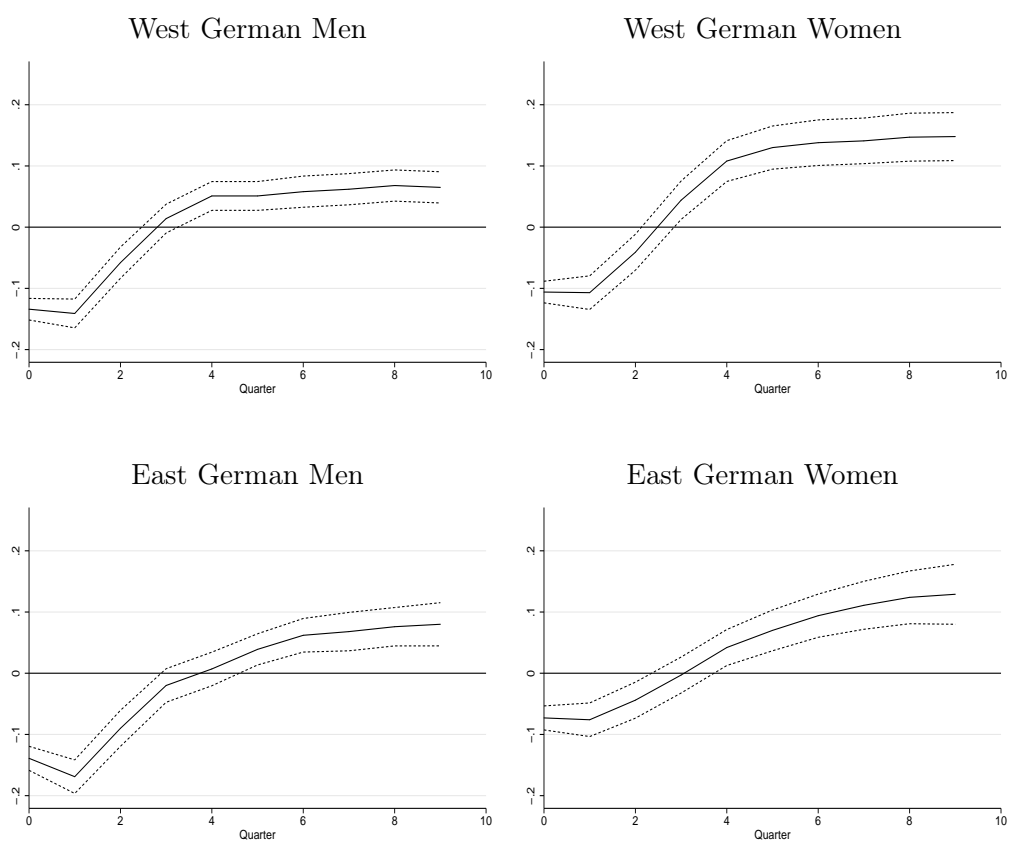
Figure 7: ATT of Attending a Program Scheduled for One versus Two Quarters

West German Men

West German Women

East German Men

East German Women
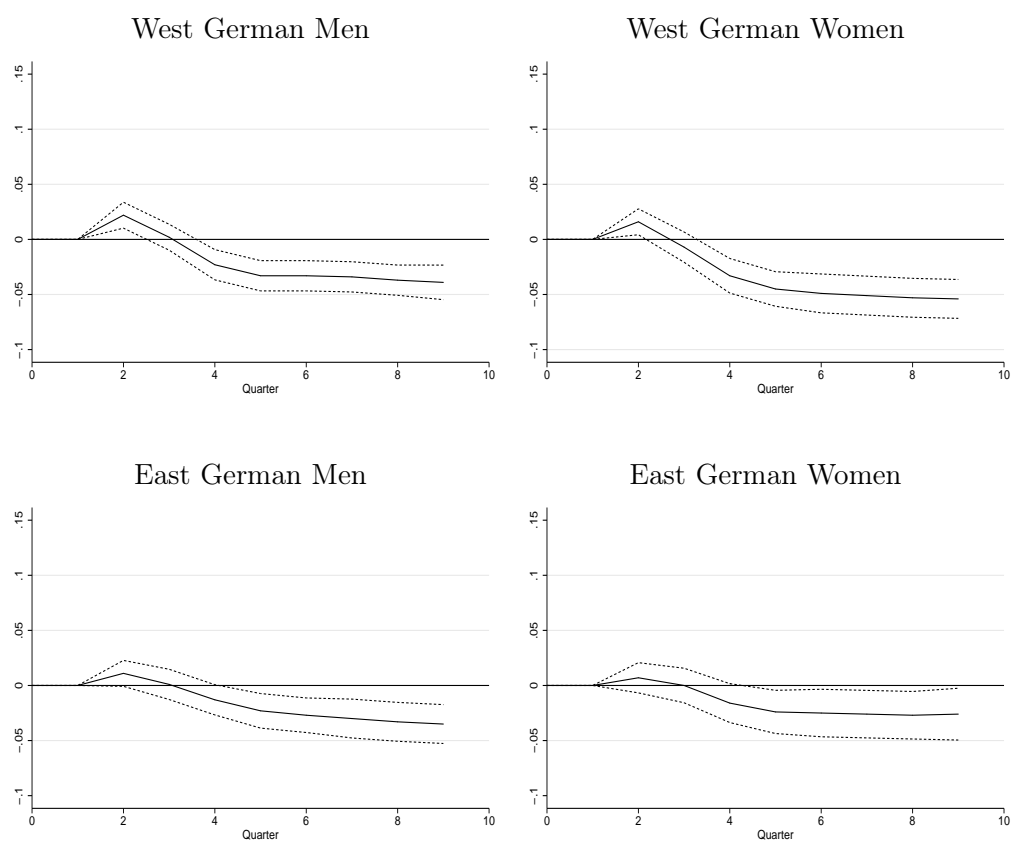
Figure 8: ATT of Attending a Program Scheduled for Three versus Two Quarters

West German Men

West German Women

East German Men

East German Women
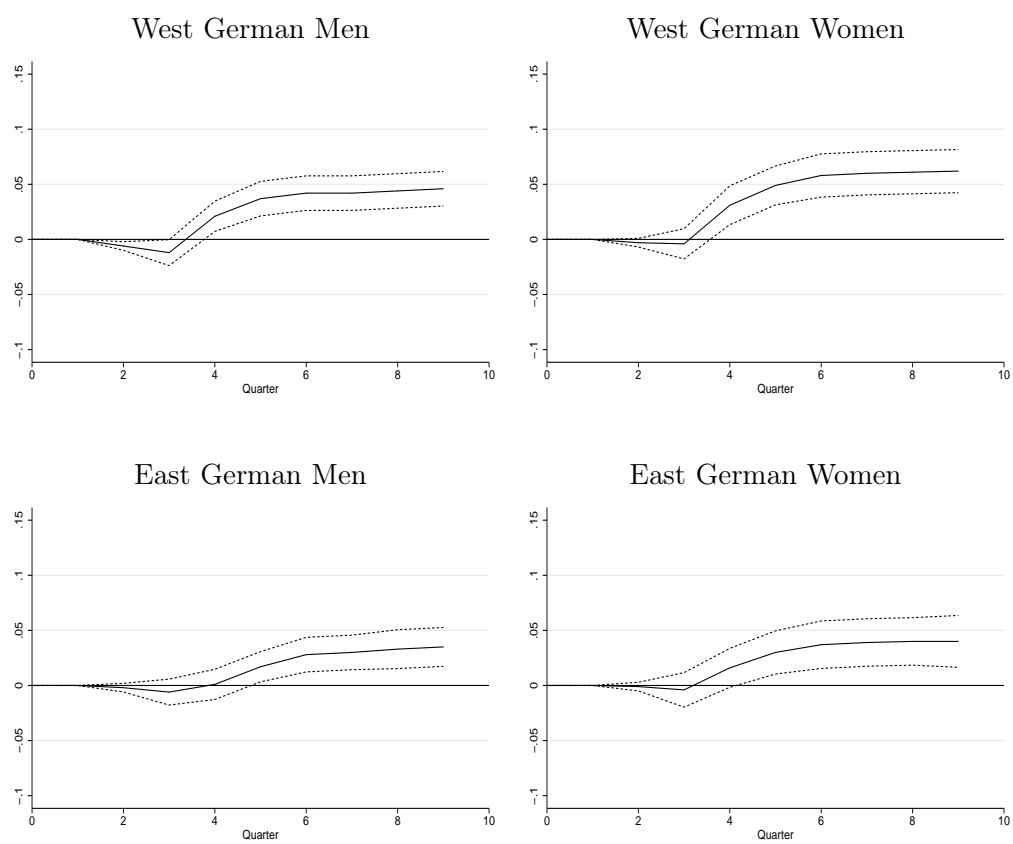
Figure 9: ATT of Attending a Program Scheduled for Four versus Two Quarters

West German Men

West German Women

East German Men

East German Women
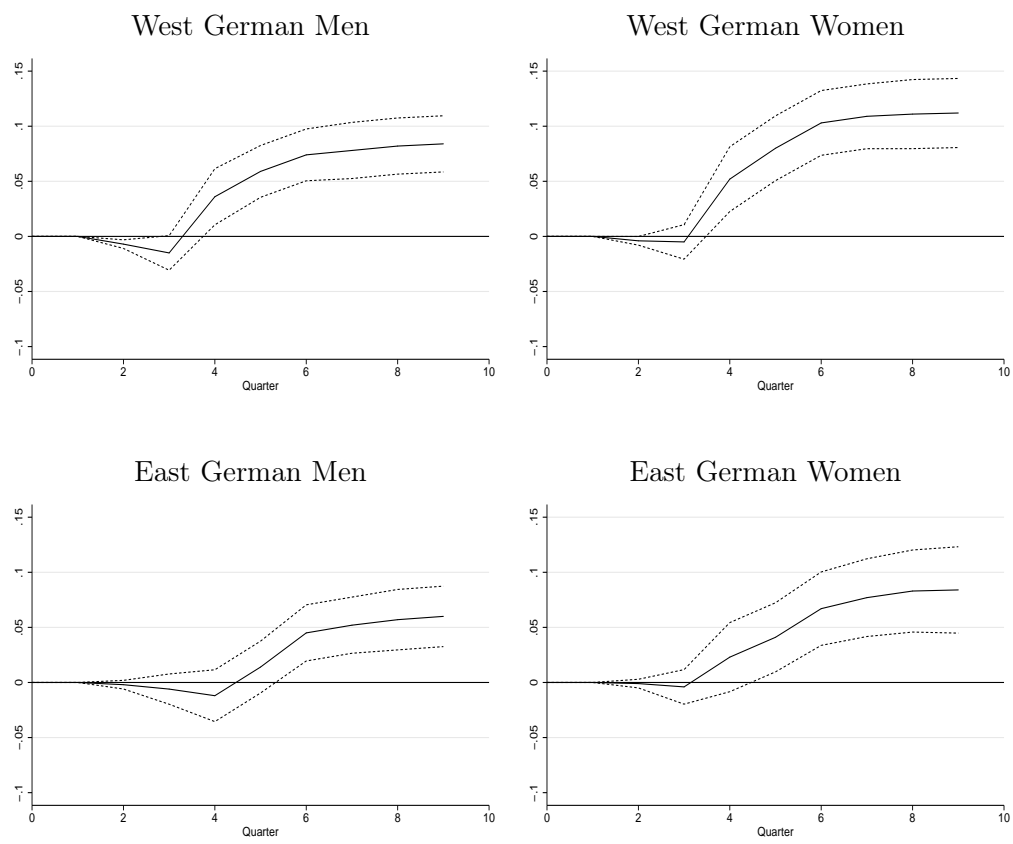
Tables

Table 1: Descriptive Statistics

	Male, West	Female, West	Male, East	Female, East
Individuals	16,317	12,328	8,737	4,869
Quarters per Person	14.7	15.3	13.2	13.5
Quarters Employed p. P.	6.0	5.2	5.1	4.0
Quarters Unemployed p. P.	8.7	10.1	8.2	9.5
Quarters in Training p. P.	0.27	0.31	0.45	0.56
Employment Spells p. P.	1.21	0.85	1.06	0.74
Unemployment Spells p. P.	1.89	1.53	1.78	1.48
Training Spells p. P.	0.11	0.12	0.16	0.19

Table 2: Employment Rate and Number of Participants Still Observed Aligned to Start of Program

$t - t_s$	Male West $\bar{E}_t$	$\hat{E}_t$	N_t	Female West $\bar{E}_t$	$\hat{E}_t$	N_t	Male East $\bar{E}_t$	$\hat{E}_t$	N_t	Female East $\bar{E}_t$	$\hat{E}_t$	N_t
0	0.000	0.000*	1740	0.000	0.000*	1431	0.000	0.000*	1300	0.000	0.000*	848
1	0.079	0.079*	1740	0.070	0.070*	1431	0.048	0.048*	1300	0.039	0.039*	848
2	0.179	0.177	1721	0.162	0.165	1411	0.120	0.123	1290	0.081	0.078	840
3	0.239	0.244	1696	0.264	0.267	1385	0.178	0.177	1282	0.118	0.121	834
4	0.302	0.301	1664	0.354	0.355	1366	0.226	0.218	1265	0.172	0.185	825
5	0.334	0.346	1623	0.415	0.412	1338	0.278	0.284	1229	0.240	0.239	816
6	0.371	0.370	1577	0.442	0.440	1316	0.311	0.320	1201	0.279	0.274	795
7	0.371	0.377	1526	0.450	0.454	1291	0.318	0.335	1159	0.320	0.300	765
8	0.364	0.387	1465	0.464	0.468	1253	0.345	0.355	1106	0.329	0.330	741
9	0.396	0.405	1393	0.478	0.481	1213	0.378	0.391	1039	0.356	0.355	710

Notes: t_s denotes the quarter of program start, and $\bar{E}_t$ the sample mean of the employment dummy in quarter t. $\hat{E}_t$ is the mean of the employment dummy as predicted using the simulation strategy (prediction of treatment outcomes).

* Observed value is taken for the first period.

Table 3: Differences Between Treated and Untreated in Transition Rates out of Non-Employment

$t - t_s$	Male West		Female West		Male East		Female East	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
0	-0.134	0.009	-0.106	0.009	-0.139	0.010	-0.073	0.010
1	-0.058	0.010	-0.034	0.010	-0.083	0.011	-0.027	0.010
2	0.011	0.009	0.026	0.010	-0.016	0.010	-0.006	0.009
3	0.039	0.008	0.071	0.009	0.009	0.009	0.013	0.008
4	0.049	0.009	0.089	0.010	0.002	0.010	0.040	0.010
5	0.023	0.009	0.064	0.011	0.028	0.011	0.034	0.011
6	0.017	0.010	0.043	0.011	0.027	0.011	0.038	0.012
7	0.013	0.009	0.034	0.011	0.023	0.010	0.031	0.011
8	0.012	0.009	0.035	0.011	0.024	0.011	0.034	0.013
9	0.006	0.010	0.028	0.011	0.029	0.013	0.028	0.013

Table 4: Differences Between Treated and Untreated in Transition Rates out of Employment

$t - t_s$	Male West		Female West		Male East		Female East	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
2	-0.085	0.037	-0.089	0.040	-0.125	0.052	-0.170	0.073
3	-0.072	0.028	-0.095	0.028	-0.106	0.038	-0.143	0.059
4	-0.048	0.024	-0.071	0.025	-0.074	0.031	-0.096	0.049
5	-0.056	0.020	-0.058	0.021	-0.073	0.026	-0.095	0.041
6	-0.073	0.019	-0.058	0.020	-0.091	0.025	-0.101	0.039
7	-0.088	0.019	-0.066	0.020	-0.106	0.026	-0.107	0.036
8	-0.091	0.020	-0.068	0.020	-0.103	0.025	-0.099	0.036
9	-0.085	0.019	-0.062	0.019	-0.092	0.025	-0.091	0.034

Appendix

A Detailed Information on the Data

This study uses data from the IEBS Version 4.02. A German description of the IEBS Version 3.01 can be found in Zimmermann et al. (2007). Information in English can be found on the website of the Research Data Center of the Federal Employment Agency (http://fdz.iab.de/en.aspx). The website also describes the conditions under which researchers may obtain access to the IEBS.

The first of the four administrative data sources included in the IEBS, the IAB Employment History, consists of social insurance register data for employees subject to contributions to the public social security system. It covers the time period from 1990 to 2004. The main feature of these data is detailed daily information on the employment status of each recorded individual. For each employment spell, in addition to start and end dates, data from the Employment History contain information on personal as well as job and firm characteristics such as wage, industry or occupation.

The IAB Benefit Recipient History, the second data source, includes daily spells of unemployment benefit, unemployment assistance and subsistence allowance payments the individuals received between January 1990 and June 2005. In addition to the sort of the payment and the start and end dates of periods of transfer receipt the spells contain further information like sanctions, periods of disqualification from benefit receipt and personal characteristics. Furthermore, the information in the Employment and the Benefit Recipient History allows one to calculate the individual entitlement periods to unemployment benefits.[22]

The third data source included in the IEBS is the so-called Data on Job Search Originating from the Applicants Pool Database, which contains rich information on individuals searching for jobs. It contains all the records starting January 2000 to June 2005 and partly also those beginning before 2000 if the person in question keeps the same client number throughout. The database includes a rich variety of information on personal characteristics (in particular education, family status and health condition), information related to placement fields (e.g. qualification and experience in the target profession), and regional information.

The Participants-in-Measures Data, the fourth data source, contains diverse information on participation in public sector sponsored labor market programs, for example training programs, job-creation measures, integration subsidies, business start-up allowances covering the period January 2000 to July 2005. Comparing the

---

[22]For the calculation of the claims, the present study relies on Plaßmann (2002) that contains a summary of the different regulations.

entries into different programs in 1999 with the figures for later years shows that information on programs starting in 1999 seems to be already complete for most active labor market programs. Furthermore, this database allows to distinguish subsidized employment in the context of active labor market policy from regular employment. Similar to the other sources, information comes in the form of spells indicating the start and end dates at the daily level, the type of the program as well as additional information on the program such as the planned end date or if the program ends with a certificate.

B   Algorithm for MCMC Estimation

The posterior distribution combines the likelihood (see section 3.1) and the priors. We set the following independent priors: the prior distributions of the coefficients $\boldsymbol{\eta}_E$ are given by independent normal priors with distribution $\mathcal{N}(b_{E,0}, B_{E,0})$. $\mathcal{N}(\bullet)$ denotes the normal distribution. Setting very large values for the variance $B_{E,0}$, we use extremely diffuse priors. The same is done for the elements of the coefficient vector $\boldsymbol{\eta}_Q$, whose prior distributions are given by $\mathcal{N}(b_{Q,0}, B_{Q,0})$. The prior distribution of the random effects is $\mathcal{N}(\mathbf{0}, \Sigma)$. The hyperparameter Σ^{-1} follows the prior distribution $\mathcal{W}^{-1}(H_0, h_0)$, where H_0 is the inverse scale matrix and h_0 denotes the degrees of freedom. $\mathcal{W}^{-1}$ denotes the inverse Wishart distribution. In order to set a diffuse prior, we choose a small value for h_0. The diagonal elements of H_0 are set to the individual level variances of separate Maximum Likelihood estimations of the two equations multiplied by h_0, and the off-diagonal elements are set to zero.

- Step 0: Set starting values for the coefficient vectors $\boldsymbol{\eta}_E$ and $\boldsymbol{\eta}_Q$, the random effects $(\alpha_{i,E}, \alpha_{i,Q})$, and the variance covariance matrix of the random effects Σ.

- Step 1a: Sample E_{it}^* from truncated $\mathcal{N}(\boldsymbol{z}_{it,E}\boldsymbol{\eta}_E + \alpha_{i,E}, 1)$ with support $[0, \infty]$ if $E_{it} = 1$ and with support $[-\infty, 0]$ if $E_{it} = 0$.

- Step 1b: Sample Q_{it}^* from truncated $\mathcal{N}(\boldsymbol{z}_{it,Q}\boldsymbol{\eta}_Q + \alpha_{i,Q}, 1)$ with support $[0, \infty]$ if $Q_{it} = 1$ and with support $[-\infty, 0]$ if $Q_{it} = 0$ (provided the training equation is to be estimated).

- Step 2: Sample $(\alpha_{i,E}, \alpha_{i,Q})'$ from its bivariate normal conditional posterior distribution $\mathcal{N}(\boldsymbol{\mu}, V_{\alpha_i})$, where $\boldsymbol{\mu} = V_{\alpha_i} \cdot \begin{pmatrix} T_{i,E} & 0 \\ 0 & T_{i,Q} \end{pmatrix} \cdot \begin{pmatrix} (\bar{E}_i^* - \bar{\boldsymbol{z}}_{i,E}\boldsymbol{\eta}_E) \\ (\bar{Q}_i^* - \bar{\boldsymbol{z}}_{i,Q}\boldsymbol{\eta}_Q) \end{pmatrix}$ and $V_{\alpha_i} = \left(\Sigma^{-1} + \begin{pmatrix} T_{i,E} & 0 \\ 0 & T_{i,Q} \end{pmatrix} \right)^{-1}$, a bar over a variable denotes its mean across time, $T_{i,E}$ the number of observations for person i, and $T_{i,Q}$ the number of observations for person i for which the training equation is to be estimated.

- Step 3a: Sample the $\boldsymbol{\eta}_E$ vector from its multivariate normal conditional posterior distribution $\mathcal{N}(M_E, V_E)$, where $M_E = V_E(B_{E,0}^{-1}b_{E,0} + \sum_{i=1}^{N} \sum_{t=1}^{T_{i,E}} \boldsymbol{z}_{it,E}'(E_{it}^* - \alpha_{i,E}))$ and $V_E = (B_{E,0}^{-1} + \sum_{i=1}^{N} \sum_{t=1}^{T_{i,E}} \boldsymbol{z}_{it,E}'\boldsymbol{z}_{it,E})^{-1}$. N is the number of persons in the data.

- Step 3b: If the training equation is to be estimated, sample the $\boldsymbol{\eta}_Q$ vector from its multivariate normal conditional posterior distribution $\mathcal{N}(M_Q, V_Q)$, where $M_Q = V_Q(B_{Q,0}^{-1}b_{Q,0} + \sum_{i=1}^{N} \sum_{t=1}^{T_{i,Q}} \boldsymbol{z}_{it,Q}'(Q_{it}^* - \alpha_{i,Q}))$ and $V_Q = (B_{Q,0}^{-1} + \sum_{i=1}^{N} \sum_{t=1}^{T_{i,Q}} \boldsymbol{z}_{it,Q}'\boldsymbol{z}_{it,Q})^{-1}$.

- Step 4: Sample Σ^{-1} from its conditional posterior distribution

$$\mathcal{W}^{-1}\left(\begin{pmatrix} \sum_{i=1}^{N} \alpha_{i,E}^2 & \sum_{i=1}^{N} \alpha_{i,E}\alpha_{i,Q} \\ \sum_{i=1}^{N} \alpha_{i,E}\alpha_{i,Q} & \sum_{i=1}^{N} \alpha_{i,Q}^2 \end{pmatrix} + H_0,\ N + h_0\right).$$ Go to Step 1. Always use the current parameter values.

C   Detailed Estimation Results

Table 5: Means and Standard Deviations of Parameters
from MCMC Estimation

Name	Male West Mean	Male West SD	Female West Mean	Female West SD	Male East Mean	Male East SD	Female East Mean	Female East SD
Employment Equation								
Q_{t-1}	-0.478	0.064	-0.300	0.079	-0.761	0.080	-0.351	0.112
$[Q_{t-1}=1] \times \ \ldots$								
$\ldots [Q_{t-2}=1] \times [Q_{t-3}=0]$	0.320	0.078	0.323	0.088	0.474	0.097	0.120	0.147
$\ldots [Q_{t-3}=1] \times [Q_{t-4}=0]$	0.681	0.091	0.844	0.096	0.776	0.108	0.526	0.143
$\ldots [Q_{t-4}=1] \times [Q_{t-5}=0]$	1.222	0.111	1.418	0.120	0.834	0.124	0.996	0.148
$\ldots \sum_{j=5}^{6}[Q_{t-j}=1 \times Q_{t-7}=0]$	0.666	0.137	1.012	0.132	1.090	0.121	0.711	0.171
$\ldots \sum_{j=7}^{16}[Q_{t-j}=1]$	0.596	0.131	0.373	0.084	0.919	0.158	0.216	0.293
$\ldots \tau_{t,Q} \times$ unskilled	0.013	0.032	-0.035	0.036	0.079	0.060	-0.075	0.097
$\ldots \tau_{t,Q} \times$ high school	-0.041	0.031	-0.004	0.033	-0.039	0.037	0.013	0.047
$\ldots \tau_{t,Q} \times$ health probl.	0.052	0.044	0.090	0.049	-0.077	0.090	0.089	0.108
$\ldots \tau_{t,Q} \times$ age ≥ 50	0.043	0.051	-0.127	0.058	-0.021	0.045	-0.142	0.062
$[Q_{t-1}=0] \times [E_{t-1}=0] \times \ \ldots$								
$\ldots D_t$	0.125	0.134	-0.008	0.162	-0.599	0.161	-0.602	0.236
$\ldots D_t \times (t-t_s)$	-0.095	0.027	-0.026	0.031	0.083	0.035	0.109	0.050
$\ldots D_t \times (t-t_s)^2$	0.008	0.002	0.004	0.002	-0.003	0.002	-0.001	0.003
$\ldots D_t \times \tau_{t,Q}$	0.362	0.068	0.432	0.076	0.303	0.077	0.307	0.098
$\ldots D_t \times \tau_{t,Q}^2$	0.001	0.009	-0.002	0.011	-0.012	0.011	0.018	0.014
$\ldots D_t \times (t-t_s) \times \tau_{t,Q}$	-0.029	0.005	-0.030	0.006	-0.009	0.006	-0.033	0.008

<continued on next page>

Means and Standard deviations of Parameters from
MCMC Estimation <continued>

Name	Male West Mean	SD	Female West Mean	SD	Male East Mean	SD	Female East Mean	SD
$\dots D_t \times$ unskilled	-0.205	0.111	-0.093	0.139	0.238	0.208	0.394	0.341
$\dots D_t \times$ high school	-0.228	0.150	-0.322	0.183	0.514	0.206	0.539	0.234
$\dots D_t \times$ health problems	-0.275	0.168	-0.074	0.229	-0.295	0.316	0.607	0.334
$\dots D_t \times$ age ≥ 50	0.099	0.177	0.004	0.191	-0.283	0.192	-0.616	0.262
$\dots D_t \times \tau_{t,Q} \times$ unskilled	0.035	0.045	0.011	0.054	-0.078	0.080	-0.293	0.131
$\dots D_t \times \tau_{t,Q} \times$ high school	0.047	0.049	0.062	0.063	-0.199	0.058	-0.114	0.066
$\dots D_t \times \tau_{t,Q} \times$ health problems	0.045	0.065	-0.040	0.095	-0.080	0.100	-0.039	0.112
$\dots D_t \times \tau_{t,Q} \times$ age ≥ 50	-0.038	0.071	-0.080	0.075	0.030	0.058	0.105	0.072
$[Q_{t-1}=0] \times [E_{t-1}=1] \times \dots$								
$\dots D_t$	0.107	0.167	0.349	0.213	-0.018	0.215	0.650	0.348
$\dots D_t \times (t - t_s)$	0.025	0.033	-0.044	0.039	0.051	0.043	0.041	0.065
$\dots D_t \times (t - t_s)^2$	-0.3^{-4}	0.002	0.002	0.002	-0.002	0.003	-0.003	0.004
$\dots D_t \times \tau_{t,Q}$	0.203	0.082	0.220	0.090	0.123	0.097	-0.170	0.169
$\dots D_t \times \tau_{t,Q}^2$	-0.011	0.011	-0.004	0.011	0.001	0.013	0.040	0.026
$\dots D_t \times (t - t_s) \times \tau_{t,Q}$	-0.006	0.006	-0.001	0.007	-0.003	0.007	0.001	0.010
$\dots D_t \times$ unskilled	-0.169	0.133	0.172	0.162	0.502	0.299	0.723	0.536
$\dots D_t \times$ high school	-0.140	0.173	0.146	0.202	0.133	0.221	-0.152	0.277
$\dots D_t \times$ health problems	-0.442	0.200	-0.118	0.260	1.523	0.650	-0.478	0.495
$\dots D_t \times$ age ≥ 50	-0.692	0.226	-0.166	0.215	-0.255	0.234	-0.316	0.342
$\dots D_t \times \tau_{t,Q} \times$ unskilled	0.006	0.055	-0.081	0.062	-0.146	0.114	-0.288	0.205
$\dots D_t \times \tau_{t,Q} \times$ high school	-0.028	0.055	-0.077	0.065	-0.086	0.065	0.006	0.084
$\dots D_t \times \tau_{t,Q} \times$ health prob.	0.206	0.075	0.096	0.093	-0.987	0.322	0.521	0.165
$\dots D_t \times \tau_{t,Q} \times$ age ≥ 50	0.206	0.094	0.022	0.084	0.058	0.083	0.094	0.108
E_{t-1}	1.913	0.129	2.027	0.152	1.721	0.204	2.342	0.280
E_{t-2}	-0.220	0.032	-0.316	0.044	-0.267	0.046	-0.359	0.075
E_{t-3}	-0.278	0.034	-0.293	0.047	-0.272	0.049	-0.279	0.081
E_{t-4}	0.425	0.014	0.277	0.021	0.268	0.021	0.290	0.035
$\sum_{j=5}^{8} E_{t-j}$	-0.178	0.007	-0.212	0.009	-0.141	0.010	-0.250	0.015
$\sum_{j=9}^{12} E_{t-j}$	-0.123	0.008	-0.128	0.010	-0.070	0.011	-0.138	0.018
$\sum_{j=13}^{16} E_{t-j}$	-0.220	0.012	-0.229	0.017	-0.197	0.018	-0.284	0.031
$t > 1$	0.327	0.022	0.396	0.028	0.348	0.031	0.430	0.048
$t > 2$	0.041	0.021	0.078	0.026	0.046	0.030	0.091	0.046
$t > 3$	-0.154	0.021	-0.052	0.027	-0.114	0.030	-0.146	0.047
$t > 4$	0.286	0.021	0.233	0.027	0.294	0.030	0.297	0.048
$t > 5$	0.035	0.020	0.076	0.025	0.004	0.028	0.120	0.044
$t > 9$	0.207	0.019	0.192	0.025	0.187	0.027	0.231	0.042
$t > 13$	0.244	0.021	0.200	0.026	0.164	0.030	0.246	0.045
$E_{t-1} \times t$	0.079	0.045	0.014	0.064	-0.068	0.065	-0.070	0.109
$E_{t-2} \times t$	-0.003	0.048	-0.001	0.068	0.076	0.069	-0.114	0.113

<continued on next page>

Name	Male West Mean	SD	Female West Mean	SD	Male East Mean	SD	Female East Mean	SD
$E_{t-3} \times t$	0.482	0.049	0.412	0.070	0.407	0.073	0.462	0.118
$\tau_{t,E} \times [E_{t-1} = 0]$	0.101	0.019	-0.121	0.020	0.028	0.029	-0.116	0.034
$\tau_{t,E} \times [E_{t-1} = 1]$	0.002	0.023	-0.039	0.027	0.088	0.036	-0.105	0.051
$\tau_{t,E}^2 \times [E_{t-1} = 0]$	0.003	0.001	0.006	0.001	0.004	0.001	0.003	0.001
$\tau_{t,E}^2 \times [E_{t-1} = 1]$	0.005	0.001	0.001	0.001	0.000	0.001	0.001	0.002
last job: assisting workers	-0.096	0.023	0.017	0.044	-0.149	0.032	-0.096	0.079
last job: jobs in service	-0.097	0.035	0.057	0.038	-0.081	0.057	-0.077	0.065
last job: office or business job	-0.077	0.038	0.082	0.042	-0.199	0.068	-0.039	0.073
last job: technician or related	-0.032	0.039	0.054	0.047	-0.068	0.058	-0.050	0.084
last job: academic or managers	-0.050	0.044	0.092	0.049	-0.109	0.066	-0.043	0.090
share last wages censored	0.860	0.136	0.303	0.199	0.647	0.262	1.105	0.378
log last average real wage	0.202	0.042	0.148	0.054	0.168	0.140	0.242	0.118
log last average real wage squared	0.017	0.006	-0.007	0.008	0.022	0.020	-0.009	0.017
last job: whitecollar job	-0.077	0.032	-0.037	0.035	-0.128	0.050	-0.017	0.060
last job: seasonal worker	0.198	0.031	0.215	0.038	0.202	0.040	0.305	0.062
last job: parttime worker	-0.069	0.039	-0.107	0.031	-0.044	0.067	-0.009	0.055
days/91 employed last 3 years	0.125	0.014	0.128	0.019	0.061	0.021	0.085	0.034
days/91 employed last 3 years squ.	-0.006	0.001	-0.008	0.001	-0.001	0.001	-0.005	0.002
age/100	-0.650	0.136	0.846	0.183	-0.720	0.187	-0.571	0.327
no vocational degree	-0.104	0.021	0.072	0.028	-0.014	0.042	-0.220	0.071
no schooling degree	-0.047	0.022	-0.193	0.038	-0.200	0.046	-0.074	0.094
high school (Abitur)	-0.275	0.034	-0.039	0.036	-0.147	0.054	-0.154	0.072
health problems	-1.467	0.039	-1.280	0.051	-1.328	0.069	-1.825	0.133
at least one child	0.084	0.014	0.259	0.019	-0.050	0.020	-0.044	0.032
region with bad conditions	-0.042	0.069	0.094	0.100	-0.379	0.053	-1.021	0.087
urban region with high unempl.	-0.145	0.023	-0.060	0.032	-0.494	0.057	-1.010	0.095
unemployment rate in community	-0.001	0.002	-0.5^{-4}	0.003	-0.010	0.003	0.008	0.004
winter (Jan.-Mar.)	0.033	0.015	0.017	0.018	-0.089	0.022	-0.032	0.033
spring (Apr.-Jun.)	0.502	0.013	0.215	0.016	0.483	0.018	0.290	0.028
summer (Jul.-Sept.)	0.362	0.012	0.110	0.015	0.420	0.017	0.190	0.026
year 1999 or 2000	0.477	0.057	0.208	0.072	0.365	0.081	-0.108	0.130
year 2001	0.331	0.044	0.165	0.056	0.257	0.063	-0.075	0.101
year 2002	0.211	0.033	0.120	0.042	0.120	0.048	-0.069	0.075
year 2003	0.201	0.022	0.086	0.028	0.139	0.032	-0.029	0.049
age/100 $\times E_{t-1}$	-0.591	0.158	-0.994	0.212	0.023	0.230	-1.512	0.383
low skilled $\times E_{t-1}$	-0.099	0.026	-0.354	0.037	-0.105	0.057	-0.037	0.095
high school (Abitur) $\times E_{t-1}$	0.577	0.043	0.168	0.046	0.540	0.070	0.439	0.090
health problems $\times E_{t-1}$	-0.076	0.054	-0.190	0.073	0.033	0.099	-0.001	0.199
share last wages censored $\times E_{t-1}$	-0.201	0.158	-0.828	0.219	-0.925	0.225	-1.709	0.386
log last average real wage $\times E_{t-1}$	-0.147	0.028	-0.091	0.030	-0.166	0.045	-0.158	0.061
age/100 $\times \tau_{t,E}$	-0.282	0.022	-0.055	0.024	-0.213	0.033	-0.021	0.042
low skilled $\times \tau_{t,E}$	-0.003	0.004	-0.014	0.004	-0.015	0.008	-0.008	0.011

Means and Standard deviations of Parameters from
MCMC Estimation <continued>

Name	Male West Mean	SD	Female West Mean	SD	Male East Mean	SD	Female East Mean	SD
high skilled $\times \tau_{t,E}$	0.032	0.005	0.012	0.005	0.025	0.008	0.016	0.010
health problems $\times \tau_{t,E}$	0.043	0.007	0.015	0.008	0.038	0.011	0.066	0.016
share last wages censored $\times \tau_{t,E}$	-0.122	0.020	-0.051	0.025	-0.057	0.027	-0.065	0.048
log last average real wage $\times \tau_{t,E}$	-0.029	0.003	-0.012	0.003	-0.021	0.006	-0.004	0.006
age/100 $\times \tau_{t,E} \times E_{t-1}$	0.274	0.034	0.245	0.040	0.128	0.050	0.262	0.077
low skilled $\times \tau_{t,E} \times E_{t-1}$	0.027	0.006	0.043	0.007	0.054	0.013	0.009	0.020
high school (Abitur) $\times \tau_{t,E} \times E_{t-1}$	-0.029	0.008	-0.024	0.008	-0.021	0.013	-0.066	0.016
health problems $\times \tau_{t,E} \times E_{t-1}$	-0.012	0.011	0.025	0.014	-0.013	0.019	-0.064	0.042
share last wages cens. $\times \tau_{t,E} \times E_{t-1}$	0.126	0.032	0.140	0.044	0.118	0.047	0.355	0.091
log last av. real wage $\times \tau_{t,E} \times E_{t-1}$	0.028	0.006	0.020	0.006	0.020	0.009	0.032	0.012
constant	-2.667	0.141	-2.665	0.180	-1.900	0.304	-1.393	0.340
Participation Equation								
$\tau_{t,Q}$	0.038	0.048	0.071	0.060	0.060	0.056	0.230	0.080
Q_{t-1}	0.663	0.138	0.688	0.157	0.429	0.162	-0.534	0.247
planned end missing	1.458	0.108	0.892	0.151	1.482	0.185	0.942	0.329
days/91 to pl. end if enough dur. left	1.852	0.061	2.009	0.070	2.206	0.075	2.523	0.100
days/91 to pl. end if . . . squ.	-0.175	0.007	-0.168	0.008	-0.206	0.010	-0.214	0.011
inflow quarter	0.035	0.044	0.012	0.047	-0.211	0.058	-0.178	0.079
days/91 elapsed unempl. duration	0.061	0.014	-0.036	0.016	0.092	0.019	0.046	0.021
days/91 elapsed unempl. duration sq.	-0.008	0.001	0.001	0.001	-0.010	0.001	-0.006	0.002
days inflow to end quarter if $t=0$	-0.009	0.001	-0.009	0.001	-0.009	0.001	-0.009	0.002
repeated inflow	-0.031	0.043	0.022	0.062	-0.005	0.056	-0.057	0.088
winter (Jan.-Mar.)	0.206	0.032	0.301	0.037	0.254	0.041	0.339	0.051
spring (Apr.-Jun.)	0.134	0.032	0.176	0.036	0.194	0.040	0.375	0.048
summer (Jul.-Sept.)	0.117	0.031	0.147	0.035	0.124	0.041	0.207	0.050
year 1999 or 2000	0.559	0.108	0.590	0.124	0.706	0.156	0.415	0.187
year 2001	0.379	0.099	0.382	0.111	0.639	0.144	0.452	0.173
year 2002	0.338	0.093	0.232	0.099	0.601	0.137	0.401	0.158
year 2003	0.111	0.094	0.033	0.095	0.293	0.137	0.029	0.152
younger than 30	0.037	0.039	-0.169	0.050	0.038	0.057	-0.041	0.074
30-34 years old	-0.022	0.035	-0.154	0.042	0.001	0.049	-0.043	0.061
40-44 years old	0.021	0.037	0.083	0.044	-0.018	0.049	0.074	0.059
45-49 years old	-0.067	0.042	0.057	0.047	-0.051	0.050	-0.015	0.064
50 years or more	-0.263	0.047	-0.218	0.055	-0.165	0.058	-0.079	0.068
no schooling degree	0.028	0.040	-0.263	0.074	-0.194	0.081	-0.234	0.126
high school (Abitur)	0.224	0.036	0.045	0.037	0.160	0.054	0.036	0.052
no vocational degree	-0.007	0.028	-0.039	0.036	-0.065	0.054	-0.136	0.070
last job: office or business jobs	0.183	0.046	0.368	0.039	0.038	0.077	0.128	0.046
last job: technician or related	0.164	0.047	0.029	0.048	0.086	0.069	0.159	0.062
last job: whitecollar job	0.132	0.038	0.070	0.043	0.209	0.055	0.265	0.055
last job: seasonal worker	-0.231	0.052	-0.088	0.057	-0.226	0.065	-0.133	0.067
last job: parttime worker	0.078	0.053	-0.042	0.043	0.002	0.088	0.140	0.055

Means and Standard deviations of Parameters from
MCMC Estimation <continued>

Name	Male West Mean	SD	Female West Mean	SD	Male East Mean	SD	Female East Mean	SD
log last average real wage	0.068	0.019	0.013	0.025	0.056	0.027	0.093	0.034
health problems	0.112	0.039	0.201	0.050	-0.067	0.061	0.051	0.075
at least one child	0.135	0.027	0.197	0.034	0.150	0.034	0.266	0.045
days/91 employed last 3 years	-0.008	0.022	-0.020	0.027	-0.038	0.031	-0.081	0.037
days/91 employed last 3 years squ.	0.001	0.001	0.002	0.002	0.003	0.002	0.006	0.002
entitled to unempl. compensation	0.169	0.046	0.112	0.061	0.122	0.063	0.300	0.086
unemployment rate in community	0.005	0.004	0.005	0.005	0.012	0.005	-0.002	0.006
younger than 30 $\times \tau_{t,Q}$	-0.325	0.165	-0.143	0.212	-0.322	0.250	0.924	0.348
30-34 years old $\times \tau_{t,Q}$	-0.066	0.146	-0.079	0.193	-0.282	0.200	0.894	0.293
40-44 years old $\times \tau_{t,Q}$	0.382	0.165	0.107	0.180	0.350	0.220	0.824	0.305
45-49 years old $\times \tau_{t,Q}$	0.264	0.172	-0.363	0.184	0.171	0.230	0.804	0.306
50 years or more $\times \tau_{t,Q}$	0.446	0.206	-0.104	0.237	0.466	0.261	1.242	0.337
no vocat. degree $\times \tau_{t,Q}$	-0.183	0.103	-0.292	0.127	-0.669	0.218	-0.142	0.268
younger than 30 $\times Q_{t-1}$	0.138	0.079	0.041	0.093	-0.025	0.110	-0.242	0.123
30-34 years old $\times Q_{t-1}$	0.053	0.060	0.068	0.085	0.002	0.079	-0.172	0.104
40-44 years old $\times Q_{t-1}$	-0.197	0.074	-0.075	0.079	-0.167	0.084	-0.203	0.112
45-49 years old $\times Q_{t-1}$	-0.033	0.072	0.099	0.077	-0.143	0.088	-0.192	0.110
50 years or more $\times Q_{t-1}$	-0.116	0.103	-0.059	0.121	-0.229	0.097	-0.318	0.117
no vocational degree $\times Q_{t-1}$	0.054	0.046	0.061	0.055	0.247	0.099	0.158	0.099
constant	-3.477	0.186	-3.223	0.219	-3.483	0.261	-3.424	0.306
Individual Level Variances and Covariances								
$\mathrm{Var}(\alpha_E)$	0.617	0.028	0.880	0.047	0.553	0.039	1.047	0.087
$\mathrm{Var}(\alpha_Q)$	0.286	0.065	0.451	0.089	0.356	0.073	0.292	0.072
$\mathrm{Cov}(\alpha_E, \alpha_Q)$	-0.046	0.026	-0.003	0.042	-0.053	0.032	-0.122	0.057
$\mathrm{Var}(\alpha_E)/(\mathrm{Var}(\alpha_E)+1)$	0.381	0.011	0.468	0.013	0.356	0.016	0.511	0.021
$\mathrm{Var}(\alpha_Q)/(\mathrm{Var}(\alpha_Q)+1)$	0.221	0.039	0.308	0.042	0.261	0.039	0.224	0.042
$\mathrm{Corr}(\alpha_E + \epsilon_{t,E}, \alpha_Q + \epsilon_{t,Q})$	-0.032	0.018	-0.002	0.025	-0.036	0.021	-0.075	0.034
$\mathrm{Corr}(\alpha_E, \alpha_Q)$	-0.109	0.060	-0.004	0.067	-0.120	0.069	-0.220	0.096

Notes: $t = 0, \ldots, 16$ indexes the quarters since the inflow. E_t indicates the employment status and Q_t the training status in period t. $\tau_{t,E}$ and $\tau_{t,Q}$ indicate the elapsed duration in employment/unemployment and training, respectively. D_t is a dummy equal to one if a participation in training occurred during any previous quarter since the inflow. $(t - t_s)$ denotes the elapsed time since the beginning of the program. α_E (α_Q) denotes the individual specific effect in the employment (qualification) equation, $\epsilon_{t,E}$ ($\epsilon_{t,Q}$) the idiosyncratic error term in the employment (qualification) equation.

Table 6: Classical ATT Aligned to Program Start

	Male West		Female West		Male East		Female East	
$t - t_s$	Mean	SD	Mean	SD	Mean	SD	Mean	SD
0	-0.134	0.009	-0.106	0.009	-0.139	0.010	-0.073	0.010
1	-0.141	0.012	-0.107	0.013	-0.169	0.014	-0.076	0.014
2	-0.073	0.014	-0.047	0.016	-0.116	0.015	-0.049	0.015
3	-0.002	0.014	0.040	0.018	-0.051	0.016	-0.010	0.017
4	0.046	0.015	0.115	0.020	-0.019	0.017	0.038	0.018
5	0.066	0.016	0.154	0.021	0.023	0.018	0.073	0.021
6	0.084	0.017	0.173	0.022	0.060	0.019	0.106	0.022
7	0.099	0.017	0.187	0.023	0.090	0.019	0.132	0.022
8	0.114	0.017	0.202	0.023	0.113	0.019	0.154	0.024
9	0.119	0.018	0.210	0.023	0.133	0.021	0.170	0.025

t_s denotes the quarter in which the program starts.

Table 7: Predicted Participation Rate of Participants if Postponing Participation (Aligned to Start of Program)

	Male West		Female West		Male East		Female East	
$t - t_s$	$\hat{Q}_t$	$\bar{Q}_t$	$\hat{Q}_t$	$\bar{Q}_t$	$\hat{Q}_t$	$\bar{Q}_t$	$\hat{Q}_t$	$\bar{Q}_t$
0	0.000	1.000	0.000	1.000	0.000	1.000	0.000	1.000
1	0.270	0.650	0.175	0.708	0.337	0.729	0.220	0.742
2	0.317	0.413	0.192	0.464	0.374	0.533	0.221	0.619
3	0.292	0.226	0.165	0.229	0.333	0.349	0.189	0.432
4	0.241	0.074	0.132	0.064	0.269	0.142	0.151	0.195
5	0.177	0.020	0.095	0.025	0.197	0.024	0.107	0.029
6	0.132	0.020	0.072	0.021	0.141	0.014	0.073	0.021
7	0.099	0.019	0.053	0.018	0.097	0.010	0.052	0.018
8	0.078	0.006	0.043	0.011	0.071	0.006	0.039	0.012
9	0.060	0.001	0.035	0.008	0.050	0.000	0.028	0.002
Total	0.687	1	0.523	1	0.767	1	0.626	1

t_s denotes the quarter in which the program starts. $\bar{Q}_t$ is the mean of the participation dummy of participants as observed in data, $\hat{Q}_t$ the mean as predicted under the waiting scenario. The row labeled "Total" gives the share of those who ever enrol into a program.

Table 8: ATT of Training versus Waiting Aligned to Program Start

	Male West		Female West		Male East		Female East	
$t - t_s$	Mean	SD	Mean	SD	Mean	SD	Mean	SD
0	-0.134	0.009	-0.106	0.009	-0.139	0.010	-0.073	0.010
1	-0.141	0.012	-0.107	0.014	-0.169	0.014	-0.076	0.014
2	-0.058	0.013	-0.041	0.015	-0.090	0.015	-0.044	0.015
3	0.014	0.012	0.044	0.016	-0.020	0.014	-0.003	0.015
4	0.051	0.012	0.108	0.017	0.007	0.014	0.042	0.015
5	0.051	0.012	0.130	0.018	0.039	0.013	0.070	0.017
6	0.058	0.013	0.138	0.019	0.062	0.014	0.094	0.018
7	0.062	0.013	0.141	0.019	0.068	0.016	0.111	0.020
8	0.068	0.013	0.147	0.020	0.076	0.016	0.124	0.022
9	0.065	0.013	0.148	0.020	0.080	0.018	0.129	0.025

t_s denotes the quarter in which the program starts.

Table 9: Predicted Participation and Employment Rates for Different Planned Program Durations: Male, West

	One quarter		Two quarters		Three quarters		Four quarters	
$t - t_s$	$\hat{Q}_t$	$\hat{E}_t$	$\hat{Q}_t$	$\hat{E}_t$	$\hat{Q}_t$	$\hat{E}_t$	$\hat{Q}_t$	$\hat{E}_t$
0	1.000	0.000	1.000	0.000	1.000	0.000	1.000	0.000
1	0.399	0.079	0.793	0.079	0.904	0.079	0.918	0.079
2	0.086	0.187	0.286	0.165	0.706	0.159	0.812	0.158
3	0.024	0.246	0.061	0.244	0.246	0.232	0.617	0.229
4	0.007	0.267	0.017	0.290	0.050	0.311	0.198	0.326
5	0.003	0.296	0.006	0.329	0.015	0.365	0.043	0.387
6	0.001	0.316	0.002	0.348	0.005	0.390	0.013	0.422
7	0.001	0.323	0.001	0.358	0.002	0.399	0.005	0.435
8	0.000	0.328	0.001	0.365	0.001	0.410	0.003	0.448
9	0.000	0.344	0.000	0.383	0.001	0.429	0.001	0.467

t_s denotes the quarter in which the program starts. $\hat{Q}_t$ and $\hat{E}_t$ are the simulated means of the participation and employment probability, respectively.

Table 10: Predicted Participation and Employment Rates for Different Planned Program Durations: Female, West

	One quarter		Two quarters		Three quarters		Four quarters	
$t - t_s$	$\hat{Q}_t$	$\hat{E}_t$	$\hat{Q}_t$	$\hat{E}_t$	$\hat{Q}_t$	$\hat{E}_t$	$\hat{Q}_t$	$\hat{E}_t$
0	1.000	0.000	1.000	0.000	1.000	0.000	1.000	0.000
1	0.416	0.070	0.825	0.070	0.922	0.070	0.931	0.070
2	0.099	0.166	0.311	0.151	0.748	0.147	0.836	0.147
3	0.030	0.246	0.071	0.253	0.269	0.249	0.647	0.248
4	0.011	0.292	0.021	0.325	0.059	0.357	0.219	0.377
5	0.005	0.328	0.008	0.372	0.019	0.422	0.051	0.452
6	0.002	0.350	0.004	0.398	0.007	0.456	0.016	0.501
7	0.001	0.364	0.002	0.415	0.004	0.475	0.007	0.524
8	0.001	0.375	0.001	0.428	0.002	0.490	0.005	0.540
9	0.001	0.386	0.001	0.440	0.002	0.501	0.003	0.552

t_s denotes the quarter in which the program starts. $\hat{Q}_t$ and $\hat{E}_t$ are the simulated means of the participation and employment probability, respectively.

Table 11: Predicted Participation and Employment Rates for Different Planned Program Durations: Male, East

	One quarter		Two quarters		Three quarters		Four quarters	
$t - t_s$	$\hat{Q}_t$	$\hat{E}_t$	$\hat{Q}_t$	$\hat{E}_t$	$\hat{Q}_t$	$\hat{E}_t$	$\hat{Q}_t$	$\hat{E}_t$
0	1.000	0.000	1.000	0.000	1.000	0.000	1.000	0.000
1	0.455	0.048	0.869	0.048	0.946	0.048	0.952	0.048
2	0.096	0.127	0.333	0.117	0.798	0.115	0.873	0.115
3	0.027	0.178	0.068	0.177	0.295	0.171	0.723	0.171
4	0.009	0.214	0.020	0.227	0.062	0.229	0.264	0.215
5	0.004	0.256	0.007	0.279	0.018	0.296	0.052	0.294
6	0.001	0.281	0.003	0.308	0.006	0.336	0.013	0.353
7	0.001	0.291	0.001	0.321	0.003	0.352	0.006	0.374
8	0.000	0.308	0.001	0.341	0.001	0.374	0.003	0.398
9	0.000	0.339	0.000	0.373	0.001	0.408	0.001	0.434

t_s denotes the quarter in which the program starts. $\hat{Q}_t$ and $\hat{E}_t$ are the simulated means of the participation and employment probability, respectively.

Table 12: Predicted Participation and Employment Rates for Different Planned Program Durations: Female, East

$t - t_s$	One quarter $\hat{Q}_t$	One quarter $\hat{E}_t$	Two quarters $\hat{Q}_t$	Two quarters $\hat{E}_t$	Three quarters $\hat{Q}_t$	Three quarters $\hat{E}_t$	Four quarters $\hat{Q}_t$	Four quarters $\hat{E}_t$
0	1.000	0.000	1.000	0.000	1.000	0.000	1.000	0.000
1	0.443	0.039	0.899	0.039	0.959	0.039	0.961	0.039
2	0.075	0.083	0.352	0.076	0.866	0.074	0.919	0.074
3	0.019	0.123	0.062	0.123	0.345	0.119	0.823	0.119
4	0.006	0.159	0.016	0.175	0.064	0.191	0.323	0.198
5	0.002	0.195	0.006	0.219	0.018	0.249	0.064	0.260
6	0.001	0.217	0.002	0.242	0.007	0.279	0.019	0.309
7	0.001	0.237	0.001	0.263	0.003	0.302	0.008	0.340
8	0.000	0.263	0.001	0.290	0.002	0.330	0.005	0.373
9	0.000	0.287	0.001	0.314	0.002	0.354	0.003	0.398

t_s denotes the quarter in which the program starts. $\hat{Q}_t$ and $\hat{E}_t$ are the simulated means of the participation and employment probability, respectively.

Table 13: ATT of Planned Program Duration of One Quarter versus Two Quarters

$t - t_s$	Male West Mean	Male West SD	Female West Mean	Female West SD	Male East Mean	Male East SD	Female East Mean	Female East SD
0	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
2	0.022	0.006	0.016	0.006	0.011	0.006	0.007	0.007
3	0.002	0.006	-0.007	0.007	0.001	0.007	-0.000	0.008
4	-0.023	0.007	-0.033	0.008	-0.013	0.007	-0.016	0.009
5	-0.033	0.007	-0.045	0.008	-0.023	0.008	-0.024	0.010
6	-0.033	0.007	-0.049	0.009	-0.027	0.008	-0.025	0.011
7	-0.034	0.007	-0.051	0.009	-0.030	0.009	-0.026	0.011
8	-0.037	0.007	-0.053	0.009	-0.033	0.009	-0.027	0.011
9	-0.039	0.008	-0.054	0.009	-0.035	0.009	-0.026	0.012

t_s denotes the quarter in which the program starts.

Table 14: ATT of Planned Program Duration of Three Quarters versus Two Quarters

$t - t_s$	Male West Mean	SD	Female West Mean	SD	Male East Mean	SD	Female East Mean	SD
0	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
2	-0.006	0.002	-0.003	0.002	-0.002	0.002	-0.001	0.002
3	-0.012	0.006	-0.004	0.007	-0.006	0.006	-0.004	0.008
4	0.021	0.007	0.031	0.009	0.001	0.007	0.016	0.009
5	0.037	0.008	0.049	0.009	0.017	0.007	0.030	0.010
6	0.042	0.008	0.058	0.010	0.028	0.008	0.037	0.011
7	0.042	0.008	0.060	0.010	0.030	0.008	0.039	0.011
8	0.044	0.008	0.061	0.010	0.033	0.009	0.040	0.011
9	0.046	0.008	0.062	0.010	0.035	0.009	0.040	0.012

t_s denotes the quarter in which the program starts.

Table 15: ATT of Planned Program Duration of Four Quarters versus Two Quarters

$t - t_s$	Male West Mean	SD	Female West Mean	SD	Male East Mean	SD	Female East Mean	SD
0	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
2	-0.007	0.002	-0.004	0.002	-0.002	0.002	-0.001	0.002
3	-0.015	0.008	-0.005	0.008	-0.006	0.007	-0.004	0.008
4	0.036	0.013	0.052	0.015	-0.012	0.012	0.023	0.016
5	0.059	0.012	0.080	0.015	0.014	0.012	0.041	0.016
6	0.074	0.012	0.103	0.015	0.045	0.013	0.067	0.017
7	0.078	0.013	0.109	0.015	0.052	0.013	0.077	0.018
8	0.082	0.013	0.111	0.016	0.057	0.014	0.083	0.019
9	0.084	0.013	0.112	0.016	0.060	0.014	0.084	0.020

t_s denotes the quarter in which the program starts.