THE EFFECTS OF SHORT-TERM TRAINING MEASURES ON INDIVIDUAL UNEMPLOYMENT DURATION IN WESTERN GERMANY*

Reinhard Hujer[†], Stephan L. Thomsen[‡] and Christopher Zeiss[§]

[†] *J.W. Goethe-University, Frankfurt/Main, IZA, Bonn, ZEW, Mannheim*
[‡] *ZEW, Mannheim*
[§] *J.W. Goethe-University, Frankfurt/Main*

This version: September 26, 2006

Abstract

This paper analyzes the effects of short-term training measures on individual unemployment duration in western Germany. We take account of information on the timing of treatment in the unemployment spell and of observable and unobservable factors to control for selectivity applying a multivariate mixed proportional hazards model (MMPH). Moreover, we suggest an extension of the model to estimate treatment effects varying over time. In addition, to shed more light the effects, we provide estimates in terms of average treatment effects that are common in the evaluation literature. The results show that training measures significantly reduce unemployment duration.

Keywords: Training Measures, Active Labor Market Policy, Western Germany, Time-Varying Treatment Effects, Multivariate Mixed Proportional Hazards

JEL Classification: J64, J24, I28, C41, C14

*The authors thank Ralf Wilke for valuable comments, and Christian Brinkmann, Steffen Kaimer and the Institute for Employment Research (IAB) in Nuremberg for support in the preparation of the data. The usual disclaimer applies.

[†]Reinhard Hujer is Professor of Statistics and Econometrics at the J.W. Goethe-University of Frankfurt/Main and Research Fellow of IZA, Bonn and ZEW, Mannheim.

[‡]Stephan L. Thomsen is Research Associate at the Centre for European Economic Research (ZEW), L 7,1, D-68161 Mannheim, e-mail: thomsen@zew.de.

[§]Christopher Zeiss is Research Assistant at the Chair of Statistics and Econometrics, J.W. Goethe-University, PO Box 11 19 32, D-60054 Frankfurt/Main, e-mail: zeiss@wiwi.uni-frankfurt.de.

1 Introduction

The Federal Employment Agency (*Bundesagentur für Arbeit*, FEA) spends a significant share of its annual budget – about 19.5 billion euros (36 percent) in 2004 – on attempting to improve the employment opportunities of about 2.5 million people participating in a number of different active labor market policy (ALMP) programs.[1] The most important of these programs are short-term training measures (*Maßnahmen der Eignungsfestellung und Trainingsmaßnahmen*, TM) which supported around 1.2 million individuals in 2004 of whom 788,533 joined programs in the western part of Germany. The importance of TM far exceeds that of any other programs in western Germany, e.g., the second and the third most important programs have been bridging allowances for the self-employed (*Überbrückungsgeld bei Aufnahme einer selbständigen Tätigkeit*) with about 137,400 participants and vocational training programs (*Förderung der beruflichen Weiterbildung*) with about 124,000 individuals newly promoted.

The main purpose of TM is the integration of unemployed individuals and people threatened by unemployment into employment by supporting them with a set of different courses and activities. This set comprises, e.g., aptitude tests, courses teaching presentation techniques for job applicants, as well as traditional training courses providing specific skills and techniques. TM are a labor supply-side oriented form of treatment which either attempt to support job placements made by the employment agencies and the job search activities undertaken by participants themselves, or which attempt to bring participants' skills into line with market demands. This means that TM generally aim to improve the outcomes of job searches, i.e., of the job search process. For the purpose of empirical analysis, it is useful to measure the impact of TM on the search process in terms of the duration of unemployment prior to a transition into employment or equivalently in the corresponding hazard rate. A further aspect to be considered in this context relates to the timing of treatment, i.e., the point of time at which an unemployed person joins the TM. Standard evaluation literature usually deals only with binary information: i.e. whether an individual has been subject to treatment or not (refer for example to Heckman, LaLonde, and Smith, 1999). In contrast, recent empirical literature points out the importance of information on the timing of treatment events. Abbring and van den Berg (2003) show that the timing of events conveys useful information for the identification of the treatment effect. In addition, Fredriksson and Johansson (2004) highlight the dynamic assignment of treatments and its serious implications for the validity of the conditional independence assumption commonly invoked in microeconometric evaluation studies.

These findings are important for evaluating TM as well. We therefore apply a multivariate mixed proportional hazards model (MMPH) to estimate the effects that uses the timing of treatment as identifying information. The model enables observable and unobservable factors to be controlled to identify the treatment effect in the presence of selectivity, which is a major issue in all non-experimental evaluations. We focus on the impacts of TM on the search process for employment. In addition, we estimate the effects on the survivor function and the expected unemployment duration to allow interpretation of treatment effects in the vein of average treatment effects that are common to many evaluation studies. Our empirical analysis is based on data from three inflow samples into unemployment from June, Au-

[1] Besides the goal of improving the employment chances there are a number of further purposes of German ALMP, e.g., the improvement of the balance between labor demand and supply or gender equality. All figures in this section are taken from Bundesagentur für Arbeit (2005a) except noted otherwise.

gust and October 2000, where observations are followed until December 2003. The study is restricted to western Germany, as the labor market and economic situation of western and eastern Germany differ sharply even now, more than a decade after German reunification in 1990.

A further important aspect of the following analysis is the development of the treatment effect over time. A time-varying treatment effect may arise if, e.g., it takes some time for the effects to develop and affect the search process, or if after a certain amount of time, other effects, e.g., discouraged worker effects etc., overlay the program effect. To account for that we estimate an extended version of the model where treatment effects are allowed to vary over time. Moreover, effects may be heterogeneous due to individual characteristics, i.e., programs are more effective for some sub-groups of the labor market than for others. We regard this type of effect heterogeneity in a third model, and estimate the effects for selected sub-groups.

The paper is structured as follows: The first part of section two provides some stylized facts about the programs in Germany, the second part discusses theoretical impacts of TM on the search process for employment within the prototypical search model by Mortensen (1986). Section three presents the econometric model. The data used in the analysis and selected descriptive statistics are introduced in the fourth section. The empirical estimates of the impacts of TM are presented in section five. The final section concludes.

2 Short-Term Training Measures

2.1 Stylized Facts about Short-Term Training Measures in Germany

TM were introduced with the enaction of Social Code III (*Sozialgesetzbuch III*) in 1997/1998, see §§48-52. They replaced the former short-term qualification measures (*kurzzeitige Qualifizierungsmaßnahmen*), training measures for unemployment assistance/benefit[2] recipients and employment counseling measures (*Maßnahmen der Arbeitsberatung*). The primary purpose of TM is to improve the integration prospects of the participating individuals. For this reason, programs consist of three different types of measures (modules) that can be accomplished separately or in combination and allow a flexible implementation in line with the specific needs of the job seekers and the options of the local employment agencies as well.

The first module involves aptitude tests (*Eignungsfeststellungen*) that last for up to four weeks. These tests are used to assess the suitability of job seekers in terms of skills, capability and labor market opportunities for employment or training. The measures of the second module of TM aim at improving the applicant's presentation and job search abilities (*Überprüfung der Verfügbarkeit/Bewerbertraining*).

[2] It may be worth noting that unemployed persons in Germany receive(d) two different kinds of payments conditional on the unemployment duration and the individual's contribution period to unemployment insurance (UI). Until 2004, persons generally received unemployment benefits (*Arbeitslosengeld*, UB) amounting to 60 percent (67 with dependent children) of the last net earnings for at least six months if they had contributed to UI for at least twelve months during three years before unemployment. The maximum duration of UB were up to 32 months depending on the contribution period and the individual's age. For unemployed people who had exhausted the UB entitlement unemployment assistance (*Arbeitslosenhilfe*, UA) was paid amounting to 53 (57) percent of the last average net earnings from insured employment. UA could have been paid until retirement age. In 2005, UA were pooled with social assistance in the so-called unemployment benefits II (*Arbeitslosengeld II*).

The activities support the individual's efforts to find work or efforts by the employment agency to place him/her, especially through job-application training, counseling on job search possibilities or measures assessing the unemployed person's willingness and ability to work (work-tests). Measures of the second module are promoted for up to two weeks. The last module contains practical training of the participants (for up to eight weeks) providing necessary skills and techniques required for placement in employment or vocational training (*Vermittlung notwendiger Kenntnisse und Fertigkeiten*). The courses cover specific working techniques (e.g., business administration), computer courses and language courses. Combinations of modules, e.g., a job aptitude test followed by a computer course, could be granted for a maximum of twelve weeks. TM are provided by service providers (*Bildungsträger*) and firms ensuring and this ensures that activities are closely related to the market. Referring to the official statistics of the FEA, in 2005 about 34 percent of the participants joined programs in the first, about 19 percent in the second, and about 28 percent in the third module. Combinations amounted to 18 percent of all support action. Furthermore, more than 95 percent of the participating individuals complete the TM; the main reason may be the short duration of programs.

Financial support is provided by FEA and covers course costs, examination fees, travel grants as well as child care. In addition, participants receive unemployment insurance (UI) payments or maintenance allowances if not entitled to UI. Decisions about support of courses and placement of job seekers are made by the employment agencies. Support is authorized on recommendation or with the approval of the agency only and activities are often initiated by caseworkers. However, TM may be initiated by job seekers, service providers or firms as well. A program is not eligible for support if it is intended to facilitate the re-recruitment (in a socially insured position for more than three months within a period of four years) of the unemployed person by their former employer or if the employer has offered a job to the unemployed person before the current unemployment spell. Moreover, to avoid deadweight-losses, support is denied if the service provider could be expected to take on the participant without support action in TM or if placement of suitable experts is possible.

Caseworkers possess a great deal of discretion in the allocation of participants and it is consequently interesting to know on what basis they reach their decisions. According to Kurtz (2003) who has interviewed a number of caseworkers about their preferences/ objectives/ reasons for offering TM, the most important factors are the placement chances of the individual after participation, the compensation of missing (professional) qualification, the improvement of integration chances, but also previous knowledge as well as motivation of job seekers. The results indicate that caseworkers regard preceding unemployment duration as being of minor importance for placement. Similar to the majority of ALMP programs, TM are offered to job seekers facing barriers to employment in particular, e.g., long-term unemployed. Higher educated people (with university degree) are less likely to be regarded as suitable TM candidates.

The growing importance of TM within ALMP in western (and eastern) Germany is clearly demonstrated in Table 1 which presents the number of entries into the three most important ALMP programs as well as the unemployment rates for the years 2000 to 2004. While the economy in eastern Germany has been plagued by unemployment rates of 17.1 (2000) to 18.4 percent (2003), the analogous figures for western Germany were 7.2 (2001) to 8.5 percent (2004). The development of the ALMP mix reflects

TAB. 1: ENTRIES INTO SELECTED ALMP PROGRAMS AND UNEMPLOYMENT RATES IN 2000-2004

	2000	2001	2002	2003	2004
Germany					
Short-term Training Measures	485,339	551,176	864,961	1,064,293	1,188,369
Vocational Training Programmes	522,939	441,907	454,699	254,718	185,041
Job Creation Schemes	265,563	194,633	162,737	146,824	153,021
Unemployment Rate (in percent)	9.6	9.4	9.8	10.5	10.6
East Germany					
Short-term Training Measures	200,712	232,261	351,867	373,930	399,836
Vocational Training Programmes	213,654	188,423	195,533	93,676	61,089
Job Creation Schemes	181,395	130,147	119,869	115,300	112,921
Unemployment Rate (in percent)	17.1	17.3	17.7	18.5	18.4
West Germany					
Short-term Training Measures	284,627	318,915	513,094	690,363	788,533
Vocational Training Programmes	337,880	261,199	259,166	161,042	123,952
Job Creation Schemes	78,684	61,890	42,862	31,515	40,079
Unemployment Rate (in percent)	7.5	7.2	7.7	8.4	8.5

Source: Bundesanstalt für Arbeit (2003; 2005a).

this regional difference as well. In western Germany, the focus is on programs that aim to adjust the qualification of the individuals to meet the demands of the market. The emphasis in eastern Germany is on employment programs designed to relieve the tense situation of the market. In both regions - but particularly so in the west, the number of TM has increased significantly. In 2000, TM have been the second most important program with 285 (201) thousand people promoted in western (eastern) Germany behind vocational training programs. Five years later, TM are the largest program with 789 (400) thousand participants (2004). This strong rise of TM has been accompanied by a tailing off in the use of more traditional programs and reflects the reforms of German ALMP in 1998 and the following years.[3] The main reason for this reform was the high and persistent unemployment associated with tense budgetary pressures on the FEA. Until the end of the 1990s, vocational training programs and job creation schemes (*Arbeitsbeschaffungsmaßnahmen*) have been the most important ALMP programs in Germany. They have become less important as both are long term in nature (up to three years) and expensive.[4] TM are clearly shorter and program costs are much lower than for other measures. In 2004 (2003), the FEA spent 496 (577) million euros on TM; the average costs per participant and month amounted to 538 euros (Bundesagentur für Arbeit, 2005b).

Despite these facts, empirical evidence on effectiveness is scarce in Germany. Possible reasons for this may be the lack of appropriate data,heterogeneity of programs, or that TM are sometimes used as preparative measures for continuing participation in other ALMP programs. Biewen, Fitzenberger, Osikominu, and Waller (2006) estimate the impacts of different ALMP programs on the employment rate using propensity score matching that takes account of the unemployment duration. In particular, they compare short-term training measures, further training, and retraining programs (medium and long-term

[3] Since 1998, the legal basis for ALMP in Germany has been amended twice. In 2002, new instruments and a more 'activating' labor market policy were introduced; from 2004 onwards the four laws Modern Services on the Labor Market (Hartz-reforms) have been enacted to reach the goals of Lisbon treaty from March 2000.

[4] In comparison, the spending of the FEA on vocational training programs (job creation schemes) amounted to 3,616 (1,212) million euros in 2004. Costs per participant and month in 2004 were 1,573 (1,179) euros.

training), and find positive effects for both gender in western Germany. The second study demonstrating empirical evidence on TM in Germany is Stephan, Rässler, and Schewe (2006). It is intended to provide an illustrative example for the evaluation of several ALMP programs on the basis of a new scientific data base from the FEA. Based on a 10 percent sample of all data, they estimate the impact of TM on the probability of remaining unemployed and the virtual duration of unemployment. The results indicate that only TM in the third module undertaken in firms reduce the probability of unemployment as well as the remaining unemployment duration. TM aiming to asses the willingness and the ability to work have negative effects on those outcomes. However, Stephan, Rässler, and Schewe mention that their estimates have to be interpreted with care. Moreover, a drawback of these studies is that only observable characteristics are controlled to identify treatment effects. Both analyses are based on the conditional independence assumption, i.e., unobserved influences are ruled out by assumption. They do neither account explicitly for further unobserved influences nor is the sensitivity of the estimates tested against possible unobserved heterogeneity. In addition, Biewen, Fitzenberger, Osikominu, and Waller have to discretize the duration of unemployment to estimate treatment effects with matching. Although this takes the timing of treatment more seriously than in Stephan, Rässler, and Schewe (2006), it is still regarded only crudely in the estimation as no comparisons between different points of time are possible without imposing further assumptions.

It is difficult to find programs which are directly comparable to German TM at the European level as they are designed as a mixture of 'traditional' vocational training and job search assistance programs. There is a broad variety of studies evaluating effects of vocational training programs on different outcomes, particularly at the microeconomic level. The comprehensive survey tables in Kluve (2006) show rather mixed results of the effects that are negative in a few cases and often insignificant or modestly positive. However, since vocational training programs are on average clearly longer than TM, effects are not directly comparable due to frequently reported locking-in effects. In contrast, job search assistance programs seem to be more similar to German TM. As the concept of these programs is relatively new, empirical evidence is less frequent. Examples are the studies by Weber and Hofer (2004a; 2004b) evaluating the effects of job search assistance programs for Austria. Their findings show that programs started during the first year of unemployment reduce participants' unemployment duration; programs started later have the opposite effect. Unfortunately, no adequate explanation has been provided for this large drop. Crépon, Dejemeppe, and Gurgand (2005) analyze the effects of intensive counseling schemes in France with respect to the duration as well as the recurrence of unemployment. Their results indicate positive effects on both, i.e., a reduced duration and a lower recurrence of unemployment for participants.

2.2 Impact of TM on the Search Process

Choosing a suitable outcome variable to measure program effects is an important issue for evaluation. As seen above, in order to improve the prospects for integration into employment, TM focus on two objectives. First, they attempt to improve the success of the employment agency's job placement activities as well as the job seeking behavior of participants. Second, programs are used to modify the qualification of job seekers to meet the demands of the market. Therefore, TM should be expected to accelerate the job search period of the participants, i.e., they should reduce the unemployment duration. For a precise discussion of the impacts of TM on the unemployment duration, a consideration of a formal theoretical model is useful. To do so, we embed our discussion in the standard search model proposed by Mortensen (1986).

The prototype model explains the search behavior of unemployed persons in terms of an optimal stopping problem in a dynamic and uncertain environment.[5] The model specifies job search as a sequential sampling process where an unemployed job seeker sequentially draws a sample from a wage offer distribution. For the sake of simplicity, one could consider a job seeker who sequentially applies for randomly selected jobs which are characterized by a wage offer (w). Due to market imperfections, the job seeker cannot observe the exact wage an offered job pays, but it is assumed he knows the distribution of the wage offers. The wage offer distribution is characterized by the cumulative distribution function $F(w)$ for $0 < w < \infty$. The job seeker sequentially decides to irreversibly accept or to reject the wage offer. If the job seeker accepts a wage offer, the search process stops and he becomes employed at wage w forever.[6] Otherwise, the search process continues. The worker's decision problem involves a choice of strategy for searching and the selection of a criterion that determines when an offered wage is acceptable (Mortensen, 1986).

In the model unemployed individuals aim at maximizing their expected present income over an infinite horizon, with the subjective rate of discount given by r. Wage offers arrive at random intervals following a Poisson-process with arrival rate λ, and during the period of search unemployed job seekers receive unemployment benefits b net of search cost a per unit time. The basic version of the model is assumed to be stationary, i.e., the parameters λ, $F(w)$, b, a and r are constant and independent of time. Mortensen (1986) shows that the optimal strategy can be characterized by a reservation wage w^* that is determined by the fundamental equation

$$(\lambda + r)w^* = \lambda E(w) + \lambda \int_0^{w^*} F(w)dw + r(b - a). \tag{1}$$

In the empirical analysis, the variable of interest is the duration of unemployment until a transition into employment or equivalently the hazard rate, i.e., the rate at which job seekers escape from unemployment. Assuming that the reservation wage is stationary, the hazard rate results from the rate at which wage offers arrive times the probability that this offer is acceptable:

$$\theta = \lambda[1 - F(w^*)]. \tag{2}$$

Under the stationarity assumption, the hazard rate is constant over time which is not reasonable for the empirical analysis. In particular, analyzing the effect of policy changes implies that the relevant parameters are not stationary. In the case in which parameters are non-stationary, but changes are not anticipated, the hazard rate simply generalizes to a time dependent hazard $\theta(t) = \lambda(t)[1 - F(w^*; t)]$ (see van den Berg, 2001).

Having introduced a simple search model framework, the question arises as to how participation in TM affects the duration of unemployment. According to the institutional set-up of TM, we can deduce

[5] See Mortensen (1986) and Mortensen and Pissarides (1999) for a detailed discussion of the search model.

[6] In the simple model, job-to-job transitions are excluded.

two channels through which programs affect the job search of participants. First, TM that improve or support the job placement on part of the employment agency or the self-contained job search of the participants, can be expected to improve the search behavior of the participants by increasing the intensity as well as the efficiency of the search efforts. Second, TM that teach job relevant skills may improve the job opportunities of the participants by allowing them to apply for jobs which are on average associated with higher wages. In the following we will discuss both channels.

Considering the first channel, we assume that TM in form of the first (*Eignungsfeststellung*) and the second module (*Überprüfung der Verfügbarkeit / Bewerbertraining*) , in particular, improve the search behavior of the participants. Programs in the first module may increase the efficiency of the job-placement process, since they support caseworkers in order to select more suitable job offers. Analogously, TM in the second module may increase the efficiency and intensity of the self-contained job search by advisors in terms of job search opportunities or courses relating to the application process. Improving search behavior either by increasing search intensity or efficiency means that TM affect the participants such that the number of jobs offers that participants receive increases. In what follows we assume that participation in a training measure increases the number of job offers that arrive in the small interval dt. The impact of participation in a TM on search behavior is therefore represented by a change of the offer arrival rate λ. The impact of an increased arrival rate on the unemployment duration is given by

$$\frac{\partial\theta}{\partial\lambda} = [1 - F(w^*)] - \lambda f(w^*)\frac{\partial w^*}{\partial\lambda}. \tag{3}$$

The first term is the direct increase of the hazard rate due to an increased offer arrival rate λ. This positive effect is counteracted by a negative effect due to the reservation wage represented by the second term. From eq. (1) we find that $\frac{\partial w^*}{\partial\lambda} > 0$, i.e., a higher arrival rate increases the reservation wage which induces a negative indirect effect on the hazard rate. The net effect is obtained from the sum of the positive direct and the negative indirect effects, where a sufficient condition for a positive net effect on the hazard rate is a 'log-concave' wage offer density function (Mortensen, 1986). The model shows that participation in a TM increases search efficiency and directly lowers unemployment duration on the one hand whilst, on the other hand, making workers more selective with respect to wage offers. However, note that the positive effect on the offer arrival rate may also be counteracted by a locking-in effect. Locking-in effects arise if individuals reduce their search activity during the period they actually participate in the program. An overall positive effect on the search efficiency therefore requires that a positive after-program effect dominates a negative locking-in effect.

In addition to the effect on the search behavior, participation in a TM may improve job-relevant skills and therefore improve the job opportunities of the participants. In particular, TM in the form of the third module are intended to teach fundamental skills that are required for placement in employment or vocational education. If a training measure increases the skills of the participant, this is equivalent to increased productivity. This allows participants to apply for jobs which are on average associated with higher wages. In the following analysis we therefore assume that participation in a training measure shifts the mean of the wage offers distribution $F(w)$ to a higher level.[7] Following Mortensen (1986)

[7] Mortensen (1986) also considers changes in the variance of $F(w)$.

we define a translation G of the wage offer distribution as $G(w+\mu) = F(w)$, where the mean of G is exactly μ units larger, but all other higher moments around the mean are the same. From

$$\lim_{\mu\to 0}\{[G(w) - F(w)]/\mu\} = \lim_{\mu\to 0}\{[G(w) - G(w+\mu)]/\mu\} = -f(w), \tag{4}$$

we find that a marginal increase in the mean of the distribution $F(w)$ decreases the probability of obtaining a wage offer less or equal to w, provided that $\partial F(w)/\partial w = f(w)$ exists. Rewriting eq. (1) associated with the translation we get

$$(\lambda + r)w^*(\mu) = \lambda\mu + \lambda E_F(w) + \lambda \int_0^{w^*(\mu)} F(w-\mu)dw + r(b-a), \tag{5}$$

where $w^*(\mu)$ is the reservation wage associated with the wage offer distribution $G(w)$. Differentiating with respect to μ gives $\partial w^*(\mu)/\partial\mu = \theta(\mu)/[r+\theta(\mu)]$. With $0 < \theta(\mu)/[r+\theta(\mu)] < 1$, an increase in the mean of the wage offer distribution increases the reservation wage by an amount less than the increase in the mean (Mortensen, 1986). To obtain the effect of an increase of the mean of $F(w)$ on the unemployment duration, we derive from eq. (2):

$$\frac{\partial\theta(\mu)}{\partial\mu} = \lambda\left\{f[w^*(\mu)-\mu]\left[1 - \frac{\partial w^*(\mu)}{\partial\mu}\right]\right\} > 0. \tag{6}$$

An increased mean of the wage offer distribution increases the hazard rate since the reservation wage increases by less than the mean of the wage offer distribution. Therefore, for the given higher mean the workers are less selective with respect to the wage offers. However, the effect on the reservation wage will be very small if the hazard rate is large compared to the interest rate.

An important issue in the policy analysis which we have not addressed so far is whether policy changes are anticipated by individuals. Individuals anticipating future participation will adjust their optimal search strategy at the point in time the information of participation arrives. Van den Berg (1990) shows that a shift in future time paths of the structural parameters induce searchers to be more selective in their search process if that shift increases expected discounted lifetime income. Furthermore, he notes that the signs of the derivatives with respect to the structural parameters are in accordance with signs of the derivatives in the stationary model.

The theoretical analysis in this section shows that all three modules of TM can affect the search process – and therefore the individual unemployment duration – of the participants. However, the empirical analysis in the following section is restricted to a reduced form approach due to data limitations that allow no distinction between the modules and combinations of modules either. Hence, we could only estimate the composite effect of TM on the hazard rate into employment, and cannot distinguish between effects on the offer arrival rate and on the wage offer distribution.

3 Econometric Model

In this section we present the econometric model for the estimation of the treatment effect of TM on unemployment duration. The major task of an econometric analysis in the non-experimental setting is to distinguish the causal treatment effect from possible selection effects with respect to the programme assignment. Generally, the assignment into a TM depends on the decision of caseworkers and the agreement of the potential participant. Therefore, the decision whether to join a TM most likely depends on the expected labor market performance of the potential participant. In other words, the assignment into programme is likely to be endogenous in a model that explains the unemployment duration. Therefore, the following empirical analysis is based on a multivariate duration framework introduced by Abbring and van den Berg (2003) that enables us to identify the treatment effect.

In the following we consider the population of inflows into unemployment. The duration until the individual enters employment (T_e) and the duration until he/she joins a TM (T_p) are measured from the point of time an individual enters unemployment. T_e and T_p are assumed to be non-negative and continuous random variables with realizations denoted as t_e and t_p. The durations T_e and T_p are assumed to vary with time-invariant observable characteristics (x) and time-invariant unobservable characteristics (v_e, v_p). For the observable characteristics, we do not impose any exclusion restrictions, i.e., the observable characteristics are assumed to be the same for both durations. With respect to the unobservable characteristics we assume that v_e captures the unobserved heterogeneity of T_e and v_p captures the unobserved heterogeneity of T_p.

The fundamental assumption of the following model is that any dependence between T_e and T_p conditional on x and (v_e, v_p) stems from the causal effect of T_p on T_e. Then, the joint distribution $T_e, T_p|x, v$ is the product of the conditional distributions $T_e|T_p, x, v$ and $T_p|x, v$. Assuming further that $T_e, T_p|x, v$ is absolutely continuous we can specify the conditional distributions in terms of their hazard rates (Abbring and van den Berg, 2004). Both hazard rates are specified as mixed proportional hazard (MPH) models,

$$\theta_e(t|t_p, x, v_e) = \lambda_e(t)\exp(x'\beta_e)v_e\mu(t-t_p, x)^{I(t>t_p)}, \quad (7)$$

$$\theta_p(t|x, v_p) = \lambda_p(t)\exp(x'\beta_p)v_p. \quad (8)$$

The hazard rate for the transition into employment (eq. 7) at time t consists of a baseline hazard $\lambda_e(t)$, a systematic part $\exp(x'\beta_e)$ and the unobserved heterogeneity term v_e. A basic feature of the MPH specification is that duration dependence and individual heterogeneity enter the hazard multiplicatively (see Lancaster, 1979). The duration dependence, i.e., the shape of the hazard over time, is represented by the baseline hazard. Individual heterogeneity is regarded by the systematic part and the unobserved heterogeneity term. It is common to MPH models to specify the systematic part such that $\theta_e(t|t_p, x, v_e)$ and $\theta_p(t|x, v_p)$ are multiplicative in each element of x. The transition rate from unemployment into TM (eq. 8) is specified analogously with baseline hazard $\lambda_p(t)$, systematic part $\exp(x'\beta_p)$ and unobserved heterogeneity term v_p.

The treatment effect $\mu(t-t_p, x)^{I(t>t_p)}$ represents the causal effect of t_p on the hazard rate $\theta_e(t|t_p, x, v_e)$, where $I(t > t_p)$ is an indicator function taking the value 1 if $t > t_p$. The treatment effect can be interpreted as a shift of the hazard rate by $\mu(t - t_p, x)$ that is directly associated with the expected remaining unemployment duration. In that sense, a positive treatment effect will shorten the expected remaining unemployment duration. Hence, in the general specification, the treatment effect is allowed to depend on the time since treatment has started ($t - t_p$) and on the observable characteristics x as well.

In the empirical analysis, we consider three (computational manageable) specifications of the treatment effect $\mu(t-t_p,x)^{I(t>t_p)}$. The first specifies the effect as a permanent and constant shift of the hazard rate at the moment the treatment starts (basic model). In this specification the effect is defined as $\mu(t-t_p,x)^{I(t>t_p)} = \mu^{I(t>t_p)}$. This specification serves as a reference for two extensions with respect to the specification of the treatment effect. The first extension allows for a time-varying treatment effect, where the effect that is modelled as a piecewise-constant with two intervals, i.e., $\mu(t-t_p,x)^{I(t>t_p)} = \mu_1^{I(t_p<t\leqslant t_p+c)}\mu_2^{I(t>t_p+c)}$, and c is an exogenous constant. In this specification, the hazard rate shifts by μ_1 at the moment the individual starts to participate, and after a duration of length c the hazard is shifted by μ_2. This extended specification enables the development of the treatment effect to be analyzed over time. A time-varying treatment effect might arise if, e.g., it takes some time for the effects to develop and affect the search process, or after a certain amount of time other effects, e.g., discouraged worker effects etc., overlay the program effect. Moreover, program effects may also differ by individual characteristics, i.e., programs are more effective for some subgroups of the labor market than for others. We take account of effect heterogeneity due to individual characteristics in a second extension, where we specify the treatment effect as a time-invariant effect that is allowed to vary with the observable characteristics, i.e., $\mu(t-t_p,x)^{I(t>t_p)} = \mu(x)$.

The basic assumption of the empirical model is that any selectivity relates to observable and unobservable factors. Technically, selectivity means that those individuals observed to receive a treatment at t_p are a non-random subset with respect to t_e. Whereas any selectivity conditional on observable characteristics is captured by the systematic part in equation (7), possible selection on unobservable factors is captured by a dependence of v_e and v_p. Generally, we assume that (v_e, v_p) is a random vector with distribution function $G(v_e, v_p)$ independent of x. If selectivity cannot be fully captured by the observable characteristics, we would observe a dependence of the unobserved heterogeneity terms. Then, the indicator function for the treatment effect appears as an endogenous time-varying regressor.

A further important aspect of the model is consideration of the information on the timing of the treatment within the unemployment spell. As Abbring and van den Berg (2003) demonstrate, this additional information conveys useful information on the treatment effect in the presence of selectivity. The timing of treatment is useful information as it enables a distinction to be made between time-invariant selection effects embodied by observable and unobservable characteristics and a causal treatment effect that becomes effective at the moment the treatment starts. If we consider the timing of treatment, a positive causal treatment effect leads to a pattern where a transition into employment is typically realized very quickly after a transition into treatment, no matter how long the elapsed duration of unemployment is. In contrast, in the case of a selection effect we would observe a correlation between the points in time of the transitions into employment and the program. E.g., a positive selection effect results in a pattern where a quick transition into the program is followed by a quick transition into employment, i.e., both transitions occur very rapidly after the unemployment spell has started. Thus, the main difference between a treatment and a selectivity effect is that the treatment affects the transition rate into employment only after it has been realized whereas selectivity affects the transition rate everywhere. Furthermore, the inclusion of the timing of events as identifying information avoids imposing exclusion restrictions on the observable variables as is the case in selection models. Such exclusion restrictions on x can seldom be justified from a theoretical point of view as the information that is available to the researcher is usually available to the individual under consideration as well.

Identification of the treatment effect requires that individuals do not anticipate future treatments. Anticipatory effects are present if, for example, those individuals who are informed about a future TM reduce their search activity in order to wait for the program. In this case, the hazard rate at t of an individual who anticipates a future treatment at time t_p, will be different from the hazard rate of an individual who obtains an alternative treatment at time t_p^* for $t \leqslant \min\{t_p, t_p^*\}$.[8] Due to the anticipatory effect, the information on the timing of the event would not be sufficient for identification since a causal change of the hazard occurs at the moment the information shock of the treatment arrives. However, information on the timing when people start to anticipate future participation is usually not available. Therefore, we assume that either participation in TM is not anticipated, or if it is anticipated individuals do not act on this information. In this context, it has to be noted that the assumption of no anticipatory effects does not rule out that individuals act on the determinants of T_p. In other words, individuals are allowed to adjust their optimal behavior to the determinants of the treatment process, but not to the realizations of t_p.

Abbring and van den Berg (2003) prove that with assumptions similar to those made in standard univariate MPH models, the bivariate model in eqs. (7) and (8) and the treatment effect in particular are identified. The identification is nonparametric, since no parametric assumptions with respect to the baseline hazard and the unobserved heterogeneity distribution are required (Abbring and van den Berg, 2003). In order to build the likelihood function for the estimation of the model, we have to consider censored observations. Let δ_e and δ_p be censoring indicators, with $\delta_e = 1$ ($\delta_p = 1$) if T_e (T_p) is right censored, the individual likelihood-contributions are given by

$$\ell_e(t|t_p, x, v_e) = f_e(t|t_p, x, v_e)^{\delta_e} \exp[-\int_0^t \theta_e(u|t_p, x, v_e)du]^{1-\delta_e}, \tag{9}$$

$$\ell_p(t|x, v_p) = f_p(t|x, v_p)^{\delta_p} \exp[-\int_0^t \theta_p(u|x, v_p)du]^{1-\delta_p}. \tag{10}$$

With the assumption that $T_e|t_p, x, v_e$ is independent from $T_p|x, v_p$ we can write (see van den Berg, 2001)

$$\ell_{e,p}(t|x) = \int_0^\infty \int_0^\infty \ell_e(t|t_p, x, v_e)\ell_p(t|x, v_p)dG(v_e, v_p). \tag{11}$$

Following Heckman and Singer (1984), the arbitrary distribution function $G(v_e, v_p)$ can be approximated by a discrete distribution with a finite number of mass points. For the unobserved heterogeneity distribution we assume two possible values for v_e and v_p each. Four combinations with an associated probability are then possible. This specification is rather flexible and computationally feasible (Richardson and van den Berg, 2001). The estimation is accomplished by maximum likelihood where the joint unobserved heterogeneity distribution adds seven unknown parameters to the model. For the estimation by maximum likelihood it is helpful to utilize a logistic specification for the probability, and the four probabilities are

$$\pi_{j,k} = \frac{q_{j,k}}{\sum_{m=1}^2 \sum_{n=1}^2 q_{m,n}}, \tag{12}$$

and $q_{j,k}$ are free parameters to be estimated.

[8] The alternative treatment at t_p^* includes the non-treatment case, see Abbring and van den Berg (2003).

4 Data and Descriptive Statistics

4.1 Data

The empirical analysis is based on three samples of inflows into unemployment in western Germany in the months June, August and October 2000. The labor market status is observed until December 2003. The data are merged from several data sets for the administrative purposes of the FEA. The main source of information is the job seekers database (*Bewerberangebotsdatei*, BewA) that contains all registered job seekers in Germany, and comprises a large set of characteristics surveyed by caseworkers at the local employment agencies. Those characteristics cover information on the socio-demographic background of the individuals (e.g., age, marital status, gender), qualification details and placement restraints (e.g., schooling or health restrictions), and the date of entry into unemployment. The majority of the attributes in BewA are objective facts. The data also include some subjective aspects such as the assessment of the individual's qualification by the responsible caseworker (*level of qualification*).

Additional information on programs is derived from an excerpt of the program participants' master data set (*Maßnahme-Teilnehmer-Grunddatei*, MTG). This data set consolidates details on all ALMP programs funded by the FEA. These data allow us to identify episodes of participation in TM and other ALMP programs. Unfortunately, we cannot distinguish between different modules of TM and therefore analyze the effect of TM as a whole.

The outcome of interest (transition into employment) is extracted from the employment statistics register (*Beschäftigtenstatistik*, BSt). BSt includes everybody who is registered in the German social security system as having documented individual pension claims. These are all persons in socially insured employment.[9] There are a number of wage subsidy programs that are recorded as employment in BSt. Hence, we merge information of MTG to distinguish spells of employment and programs in the observation period. For the employment periods we observe the associated record dates (usually at the end of the month) and for the program spells the exact entry and exit dates. The duration of unemployment until the first transition into employment, T_e, and until the first transition into TM, T_p, are calculated from this information using 'day' as a unit of time. Unfortunately, we are not able to observe the unemployment duration in terms of registered unemployment at the FEA. Instead, the time from entry into unemployment until employment (non-employment duration) serves as a proxy for the real unemployment duration of the individuals. In addition to registered unemployment this kind of proxy includes periods out of the labor force or receipt of social benefits as well. Since labor force movements as well as episodes of employment not subject to social security are not identified in the data, using the non-employment duration is expected being an upward biased proxy of the true unemployment duration. Due to data limitations, we have to rely on this proxy. However, Fitzenberger and Wilke (2006) analyze unemployment durations in Germany dealing with similar problems. They use a lower (times of permanent income transfers) and an upper bound (non-employment duration) to proxy the true unemployment duration. Their results indicate that neither bounds differ too strongly if early retired older people are excluded. For this reason, using the non-employment duration seems to be reasonable for our question.

[9] Self-employed and pensioners are not included.

If an individual joins an alternative ALMP program before he/she becomes employed, we consider the unemployment spell to be censored at the point in time this transition occurs. In addition, both durations are censored if no transition within the observation window can be observed. Since the available data cover transitions from unemployment into employment only, we do not account for job-to-job transitions.

The initial sample contains 76,697 individuals with 23,630 individuals entering unemployment in June, 31,217 in August and 21,850 in October 2000.[10] We exclude all individuals from this sample who either joined alternative ALMP programs in the period from January 2000 up to their unemployment entry or exhibited failures in the data. This exclusion should ensure (to a limited extent) that people became unemployed or entered the labor force for the first time. For this reason, unemployment entry dates in the sample correspond to unemployment entry in the economic sense. Furthermore, we restrict the sample for homogeneity reasons to German citizens who are neither disabled nor affected by other health restraints. Moreover, to avoid influences related to professional training we exclude people younger than 25 years. Older individuals (above 55 years) are not considered in order to rule out selection due to early retirement. This exclusion should also reduce the possible bias of proxy for the unemployment. By imposing these restrictions, we are left with 35,706 individuals for analysis. We observe 1,366 of the individuals to enter a TM, i.e., 3.8 percent of the unemployment spells until a transition into program are non-censored. With respect to the unemployment spells until a transition into employment we observe 25,651 (72 percent) non-censored spells.

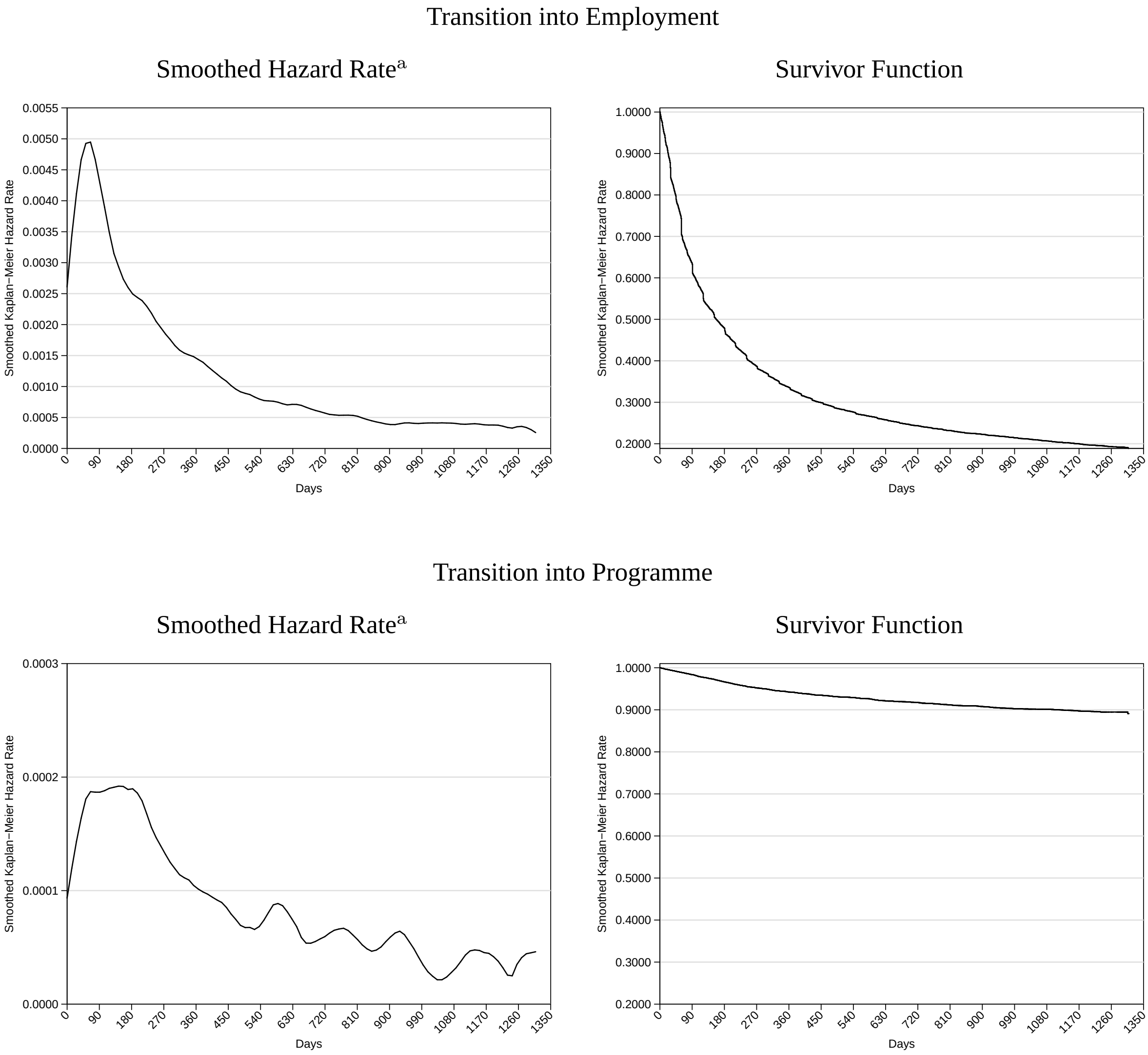
4.2 Descriptive Statistics

Figure 1 presents Kaplan-Meier estimates of the hazard rates and survivor functions for the transition into employment and the transition into a program. For the first transition, we find a quite typical picture. In particular during the first three months, job seekers experience the highest probability of leaving and taking up employment. The chances of finding a job decrease strongly thereafter. The corresponding survivor function implies that the probability of still being not employed after three months is almost 60 percent; after three years, this probability decreases to about 20 percent. The transition rate into TM establishes a slightly different picture. Job-seekers have the highest chances of entering a TM within the first six to seven months of unemployment. Afterwards, the hazard rate decreases clearly. It has to be noted that the hazard rate for the transition into TM is throughout significantly lower than the hazard rate for the transition into employment. Hence, the corresponding survivor function shows that an individual is still not assigned to TM with a probability of 90 percent even after three years.

Based on the results of the non-parametric estimates, we choose the number and limits of the intervals for the piecewise-constant baseline hazard rates of our model. Since the Kaplan-Meier estimates provide some differences in the development of both hazard rates over time, we regard eight intervals for the transition rate into employment and six for the transition rate into program. The interval limits of the hazard rate into employment are 90, 180, 360, 540, 720, 900 and 1,080 days. The analogue limits for the hazard rate into program are 180, 360, 540, 720 and 900 days, i.e., intervals last for six months.

[10] We consider differences due to the starting dates of the unemployment spell in calendar time by including dummy variables in the empirical analysis

FIG. 1: NON-PARAMETRIC ESTIMATES



[a] The bandwidth used in the kernel smooth to plot the estimated hazard function was set to 30.

Table 2 presents means and frequencies of the observable covariates considered in the estimation to point out equalities and differences. The results of Kurtz (2003) indicate that important determinants for the decisions of caseworkers to support job seekers by TM are the placement chances after participation, compensation for lack of occupational qualification as well as previous knowledge and motivation. In the empirical analysis, we approximate lack of occupational qualification as well as previous knowledge of the job seekers by using information on *occupational experience*, *vocational education*, *level of qualification* and *schooling*. The categorial variables have to be interpreted with respect to the following references: *vocational education* refers to missing education. For the assessment of the individual's qualification by the caseworker (*level of qualification*) we use individuals with or without technical knowledge. The *schooling* categories are in reference to people who have no school qualifications. It becomes obvious that participants do not differ much in these variables from other job seekers. However, the

TAB. 2: DESCRIPTIVE STATISTICS FOR COVARIATES[1]

	Total	Particip.	Non-Particip.
Observations	35,706	1,366	34,340
Frequencies (in %)			
Women	47.40	48.02	47.38
Applicant for Full Time Job	79.01	77.45	79.07
Occupational Experience (Yes)	92.54	92.75	92.53
Vocational Education[2]			
In-Firm Training	48.13	51.36	48.00
Off-the-Job Training	1.36	1.90	1.34
Vocational School	1.93	1.90	1.93
Technical School	4.47	3.37	4.52
University	5.17	4.03	5.22
Advanced Technical College	1.88	1.46	1.89
Level of Qualification[3]			
University Level	6.11	4.32	6.18
Advanced Technical College Level	2.64	1.90	2.67
Technical School Level	2.95	2.64	2.65
Skilled Employee	44.39	47.29	44.28
Schooling[4]			
CSE[5]	48.74	48.98	48.73
O-Level (*Realschulabschluss*)	20.74	23.57	20.63
Advanced Technical College (*Fachhochschulreife*)	5.85	5.42	5.87
A-Level (*Abitur*)	13.01	10.83	13.10
Family Status[6]			
Single Parent	6.21	6.59	6.19
Married	49.18	48.68	49.20
Desired Occupational Group[7]			
Manufacturing Industry	33.10	31.26	33.17
Technical Occupation	3.68	5.20	3.62
Service Professions	60.04	59.96	60.04
Means			
Age	36.92	37.33	36.90
No. of Children	0.67	0.73	0.67

[1] All statistics are calculated at start of the unemployment spell.
[2] Reference Category: missing education.
[3] Reference Category: with and without technical knowledge.
[4] Reference Category: without graduation.
[5] Certificate of secondary education (*Hauptschulabschluss*).
[6] Reference Category: singles/not married.
[7] Reference Category: agriculture, mining, fishery and miscellaneous occupations.

ratio of participants with O-level school qualifications (*Realschulabschluss*) is larger (23.57 part. /20.63 non-part. percent) and that of people with A-level qualifications (*Abitur*) is smaller compared to that of non-participants (10.83/13.10 percent). Analogously, fewer participants have a technical school or university degree.

The life cycle position of the individual is also an important determinant of labor market performance. To capture its influence, we account for a number of sociodemographic attributes in the estimation. The *age*, *gender* (women), *marital status* and the *number of children* of the job seeker are considered. Moreover, we incorporate the labor market attachment and occupational group of the individual by using information on *application for full time job only* and *desired occupational group*. For the sake of completeness, it should be noted that the dummy variables for the *family status* are in reference to singles/ not married individuals and the dummy variables for the *desired occupational group* refer to individuals who want to work in the agriculture and fishery industry, the mining industry and miscellaneous professions. The figures in Table 2 indicate that participating and non-participating people are on average not very different in the life cycle position, labor market attachment and occupational group. One obvious difference is that participants in TM more often apply for technical professions than average job seekers (5.20/3.62 percent). However, none of the covariates seems to determine participation or non-participation clearly. We are not able either to approximate the motivation of job seekers from the set of variables. Hence, it is part of the unobserved heterogeneity we consider.

5 Empirical Evidence

5.1 Impacts of TM – Basic Model

We start the discussion of the effects of TM with the results of the basic model where the treatment effect is specified as a constant and permanent shift of the hazard rate (see Table 3). We are mainly interested in parameter μ, i.e., the causal impact of participation in a TM on the hazard rate into employment. The result establishes a clear positive treatment effect of $\exp(0.3915) = 1.48$ which could be interpreted as follows: At the point in time an individual enters a TM, the hazard rate into employment shifts by 1.48. In other words, the hazard rate of a participant, at any point in time after he/she has entered a TM, is 48 percent higher compared to an individual who has not entered a TM so far. Hence, TM clearly enhance the search process of the participating individuals, i.e., participation reduces the time people are seeking employment.

The observable covariates affect the transition rate into employment in different directions. It increases with the *number of children* and for *single parents* and *married* people as well. As these variables are indicators of the responsibility the job seeker has for closely related people, the results show that these people are more successful in finding jobs. One possible reason may be a greater willingness to actively seek employment as well as the greater need for work. In addition, it should be noted that, in terms of gender, women require less time to find a new job. Moreover, the transition rate into employment increases with qualification. *Skilled employees* are better at finding a job compared to the unskilled. School qualifications provide a significant estimate for *CSE* only; for the other groups no differences could be found. People who seek work in the *manufacturing industry, technical occupations* or in *service professions* also experience transitions into employment more often than the reference group (*agriculture, fishery, mining and miscellaneous other occupations*).

One fairly common finding in the empirical literature is that older unemployed workers face barriers to employment. This is confirmed by the estimate for *age*. In addition, people who are not willing to take on part-time employment have a lower transition rate into employment (*applicant for a full time job only*). As we pool data from three months of inflows into unemployment for the analysis, possible differences due to seasonal figures have to be considered. The dummies indicate that the transition rate to employment is highest in the sample for people who became unemployed in June 2000. Becoming unemployed in August is worse and the lowest transition rate is established for the October entries.

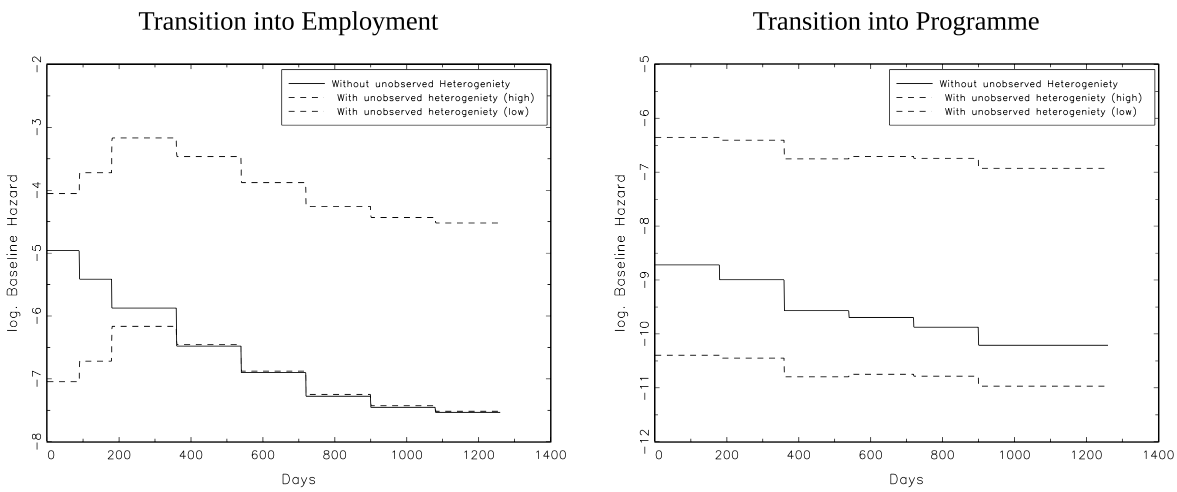
TAB. 3: ESTIMATION RESULTS (BASIC MODEL)[1]

Variable	Transition Rate into Employment		Transition Rate into Training-Programme	
	Coeff.	t-Value	Coeff.	t-Value
Baseline Hazard				
$\lambda_{90 \geqslant Y < 180}$; $\lambda_{180 \geqslant S < 360}$	0.3292	9.47	-0.0529	-0.498
$\lambda_{180 \geqslant Y < 360}$; $\lambda_{360 \geqslant S < 540}$	0.8828	10.35	-0.4016	-2.283
$\lambda_{360 \geqslant Y < 540}$; $\lambda_{540 \geqslant S < 720}$	0.5902	7.72	-0.3532	-1.525
$\lambda_{540 \geqslant Y < 720}$; $\lambda_{720 \geqslant S < 900}$	0.1723	2.21	-0.3882	-1.350
$\lambda_{720 \geqslant Y < 900}$; $\lambda_{S \geqslant 900}$	-0.2004	-2.43	-0.5738	-1.684
$\lambda_{900 \geqslant Y < 1080}$	-0.3794	-4.40		
$\lambda_{Y \geqslant 1080}$	-0.4673	-5.09		
Unobserved Heterogenity (v_u, v_p)	2.9934	40.94	-4.0393	-8.733
Constant	-7.0471	-64.31	-6.3548	-13.843
Age	-0.0173	-14.71	-0.0015	-0.297
Women	0.0901	4.03	0.0410	0.355
Applicant for Full Time Job only	-0.0703	-2.65	0.0042	0.026
Occupational Experience (Yes)	-0.0466	-1.41	-0.1342	-0.884
No. of Children	0.0234	2.16	0.1098	2.266
Vocational Education				
− In-Firm Training	0.0282	1.03	0.1218	0.985
− Off-the-Job Training	-0.0052	-0.06	0.6648	2.227
− Vocational School	-0.0057	-0.08	-0.0254	-0.081
− Technical School	0.0763	1.45	-0.2677	-1.036
− University	-0.0195	-0.27	0.0591	0.175
− Advanced Technical College	-0.0611	-0.65	-0.0981	-0.253
Level of Qualification				
− University Level	-0.0467	-0.74	-0.4892	-1.614
− Advanced Technical College Level	-0.0723	-0.92	-0.3995	-1.110
− Technical School Level	0.0469	0.74	0.0097	0.021
− Skilled Employee	0.0558	2.14	0.1895	1.456
School Education				
− CSE[2]	0.1108	3.28	0.1546	1.168
− O-Level (*Realschulabschluss*)	0.0643	1.60	0.3573	2.283
− Advanced Technical College (*Fachhochschulreife*)	-0.0061	-0.11	0.1931	0.883
− A-Level (*Abitur*)	0.0036	0.07	0.1435	0.698
Family Status				
− Single Parent	0.1367	3.20	-0.0072	-0.028
− Married	0.1278	5.63	-0.1606	-1.517
Occupational Group				
− Manufacturing Industry	0.1895	3.45	-0.1107	-0.492
− Technical Occupation	0.2402	3.24	0.8380	2.778
− Service Professions	0.2392	4.39	0.1134	0.500
Entry into the Sample				
− Entry in August	-0.0630	-2.98	0.2697	2.837
− Entry in October	-0.1723	-7.18	0.1718	1.709
Treatment Effect (μ)	0.3915	6.95		
q_1	2.3651	7.75		
q_2	-0.7747	-2.78		
q_3	2.4279	8.19		
π_1	0.0427			
π_2	0.4541			
π_3	0.0197			
π_4	0.4836			
Log-Likelihood	-186,602.27			

[1] Reference categories for categorial variables: Vocational education, *missing education*; level of qualification, *with and without technical knowledge*; schooling, *without graduation*; family status, *singles/not married*; desired occupational group, *agriculture, mining, fishery and miscellaneous occupations*.
[2] Certificate of Secondary Education (*Hauptschulabschluss*).

The estimates of the influence of the observable covariates on the transition rate into programs show a mixed picture. The statistical insignificance of most parameters makes it very difficult to derive clear rules of selection into programs. However, preference is given to people with *children* for participation in TM. The positive parameter of *off-the-job training* shows that caseworkers prefer this group of people to people without any vocational education. An increase in the transition rate was also established for people with *O-level* qualifications (compared to people with no school qualifications). In association with the large number of TM that are carried out in Germany, the positive parameter for people applying for *technical occupation*s reflects (at least to some extent) structural changes in the German economy. This means that the contents of TM may be particularly useful in adjusting the skills of this group of people to meet the demands of the market. Finally, the dummy for the unemployment entry shows that people who became unemployed in August 2000 have increased participation chances. One reason could be that some TM are used as preparation courses for professional training in the apprenticeship system (starting in September).

FIG. 2: ESTIMATED BASELINE HAZARDS



To test the sensitivity of our results with respect to the unobserved heterogeneity distribution, we have estimated a model that accounts for selection on observables only (see Table A.1 in the appendix). With $\mu = \exp(0.1881) = 1.21$, the estimated treatment effect is smaller. Therefore, ignoring the unobserved influences in the selection process leads to a downward biased estimate of the treatment effect. Comparison of the estimates of the observable covariates shows that the inclusion of unobserved heterogeneity reduces the significance of most of the parameters. The largest differences result for the estimated piecewise-constant duration dependence. The graphs of Figure 2 compare the logarithms of the estimated duration dependence for the models with and without unobserved heterogeneity (baseline hazard rates). Starting with the model without unobserved heterogeneity (solid line in the graph), we find that the hazard rates into employment as well as into programs are decreasing functions. Hence, the model establishes a negative duration dependence. This finding is similar to the Kaplan-Meier estimates from above (see Figure 1). In contrast, the hazard rates for the model considering unobserved influences (dashed lines in the graph) provide a different picture. For the hazard rate into employment, the graph show a positive duration dependence during the first three intervals (0-89, 90-179, 180-359 days).[11] For the remaining period until the end of the observation window, the function is decreasing and we find a negative duration dependence similar to the non-mixed model.

A similar picture could be revealed for the transition rate into programs. In the model accounting for unobserved heterogeneity, the function decreases during intervals one to three (0-179, 180-359, 360-539 days), but increases during the fourth interval (540-719). Afterwards, it decreases again until the end of the observation period. The findings point towards a dynamic sorting process captured by unobserved heterogeneity. A stronger duration dependence is a typical finding when unobserved heterogeneity is not considered, see e.g., Lancaster (1990). Hence, taking account of unobserved heterogeneity primarily affects the shape of the baseline hazard rates and the treatment effect. If unobserved heterogeneity is ignored the dynamic sorting processes due to unobserved characteristics would be assigned misleadingly to duration dependence (treatment effect or baseline hazard).

To shed more light on the treatment effect, we additionally calculate the effect of participation on the survivor function and the expected unemployment duration. These effects are comparable to average treatment effects that are the subject of many evaluation studies, see e.g., Heckman, LaLonde, and Smith (1999). In contrast to the effect on the hazard rate, the effect on the survivor function and the expected unemployment duration capture the dynamic accumulation of the treatment effect over the unemployment spell. However, considering these effects requires explicitness with regard to the timing of treatment. Consider the average treatment effect of a treatment at time s compared to a treatment at a time k for $k \neq s$ in terms of the survivor function $\bar{F}_e(t|t_p, x, v_e)$ at time t. In the terminology of Holland (1986), we would refer to s as the treatment and to k as the control. The causal effect of the treatment s relative to the control k for individual i is then given by the difference of the survivor functions

$$\Delta(t)_{sk} = \bar{F}_e(t|s, x, v_e) - \bar{F}_e(t|k, x, v_e). \tag{13}$$

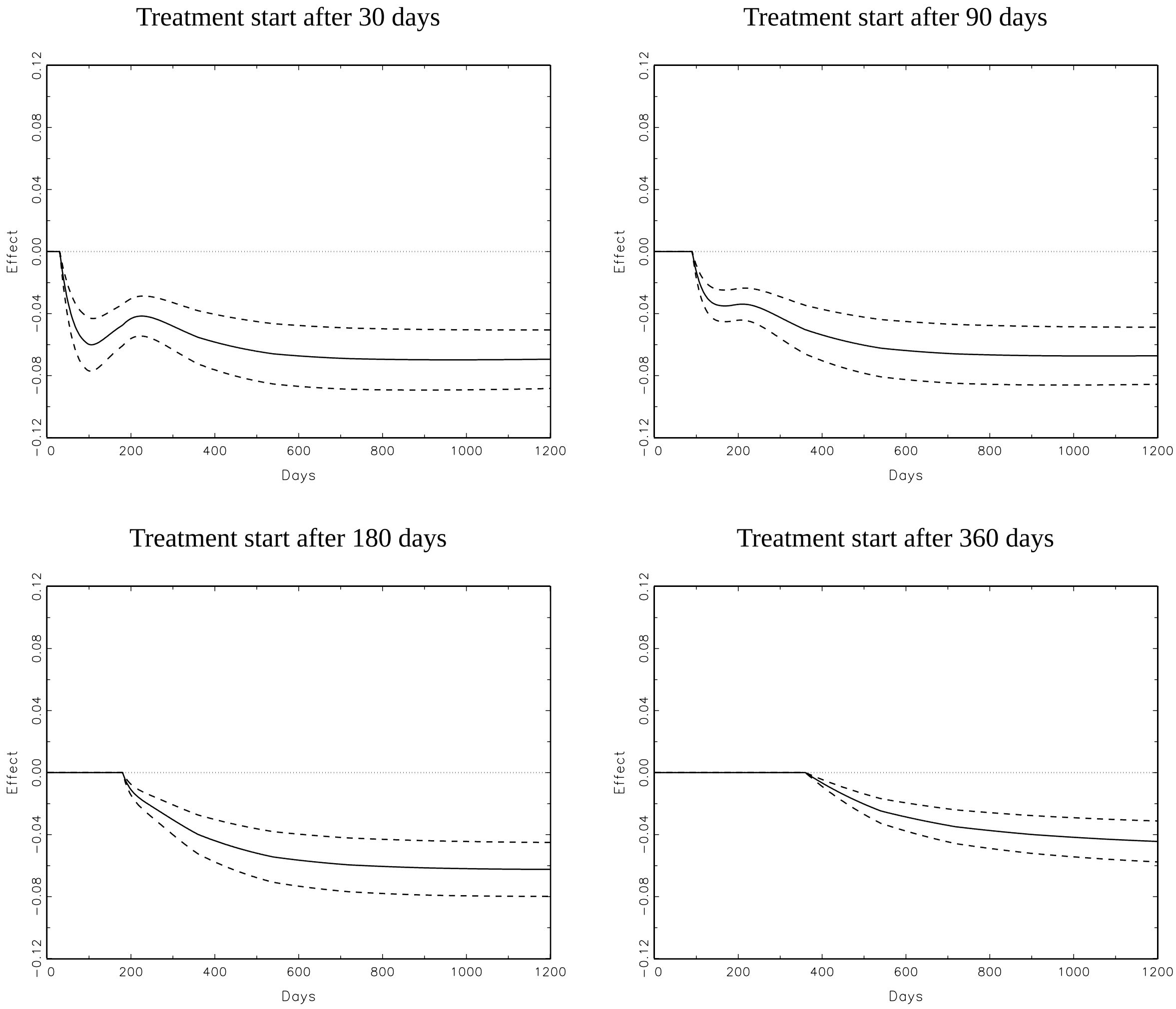
It is important to note that treatments are characterized by the time when they occur in this set-up, i.e., they could relate to the same type of course. The effect in terms of the survivor function implies a time path of the treatment effect which is determined by the effect of a treatment on the hazard rate. As Fredriksson and Johansson (2004) mention, this estimator is more fundamental than the effect in terms of the expected unemployment duration since the difference in the survivor functions integrates to the difference in the expected durations, i.e.,

$$\int_0^\infty \Delta(t)_{sk} dt = E[T_e|s] - E[T_e|k]. \tag{14}$$

To calculate the effect of participation in TM, we predict the survivor function for the empirical model using the estimated parameters and means of the observable and unobservable covariates. The effects on the survivor functions are calculated for hypothetical starts of programs after 30, 90, 180 and 360 days of unemployment that are compared to the non-treatment case.

[11] We have tested a set of different specifications for the numbers and lengths of the intervals for the baseline hazards. The final specification was chosen by two objectives: First, it provides the maximum of the likelihood function, and second, it fits well to the non-parametric estimates from Figure 1.

FIG. 3: EFFECT ON THE PREDICTED SURVIVOR FUNCTION [a]



[a] Solid line represents the treatment effect on the predicted survivor function and the dashed lines represent the 95% confidence band. Confidence bands are calculated by Delta-Method.

Figure 3 shows the treatment effect on the predicted survivor functions for the basic model with unobserved heterogeneity. Since the effect on the hazard rate is significantly positive, the effect on the predicted survivor function turns out to be significantly negative. Hence, for the period after the program start the predicted survivor function is generally below the survivor function for the non-treatment case. That is, the probability of still being unemployed at time t is significantly lowered. What these graphs clearly show is that impacts of TM are stronger when programs start earlier rather than later. Note, that this pattern primarily results from the specification of the treatment effect as a constant and permanent shift of the hazard rate. Furthermore, the impact is particularly strong early in the unemployment spell due to the multiplicative specification of the hazard rate and the shape of the baseline hazard. Moreover, we are able to derive the effect on the expected unemployment durations from the predicted survivor functions. The following results are obtained: We find a similar reduction of the expected unemployment duration for treatments starting after 30 and after 90 days with 40 and 39 percent respectively. However, if TM are started after six months or even one year of unemployment, the reduction of the unemployment duration is not as strong with only 36 and 30 percent.

5.2 Impacts of TM – Effect Heterogeneity

Up to now, the treatment effect of TM has been modelled as a permanent and constant shift of the hazard rate occurring at the moment the individual joins the program. However, it is reasonable to expect treatment effects to vary over time. On the one hand, program effects may need some time to unfold. This could be the case if participation in a TM is associated with a certificate handed out after the end of the course, e.g., for a computer course. Program effects may also be delayed since participants' newly obtained job application advices are not associated with instantaneous employment but may be associated with improved perspectives. On the other hand, effects may vanish after a certain amount of time if, for example, participants are informed about available jobs they could apply for and this information becomes obsolete over time. The effect of 'being informed' consequently decreases.

TAB. 4: TIME VARYING TREATMENT EFFECT

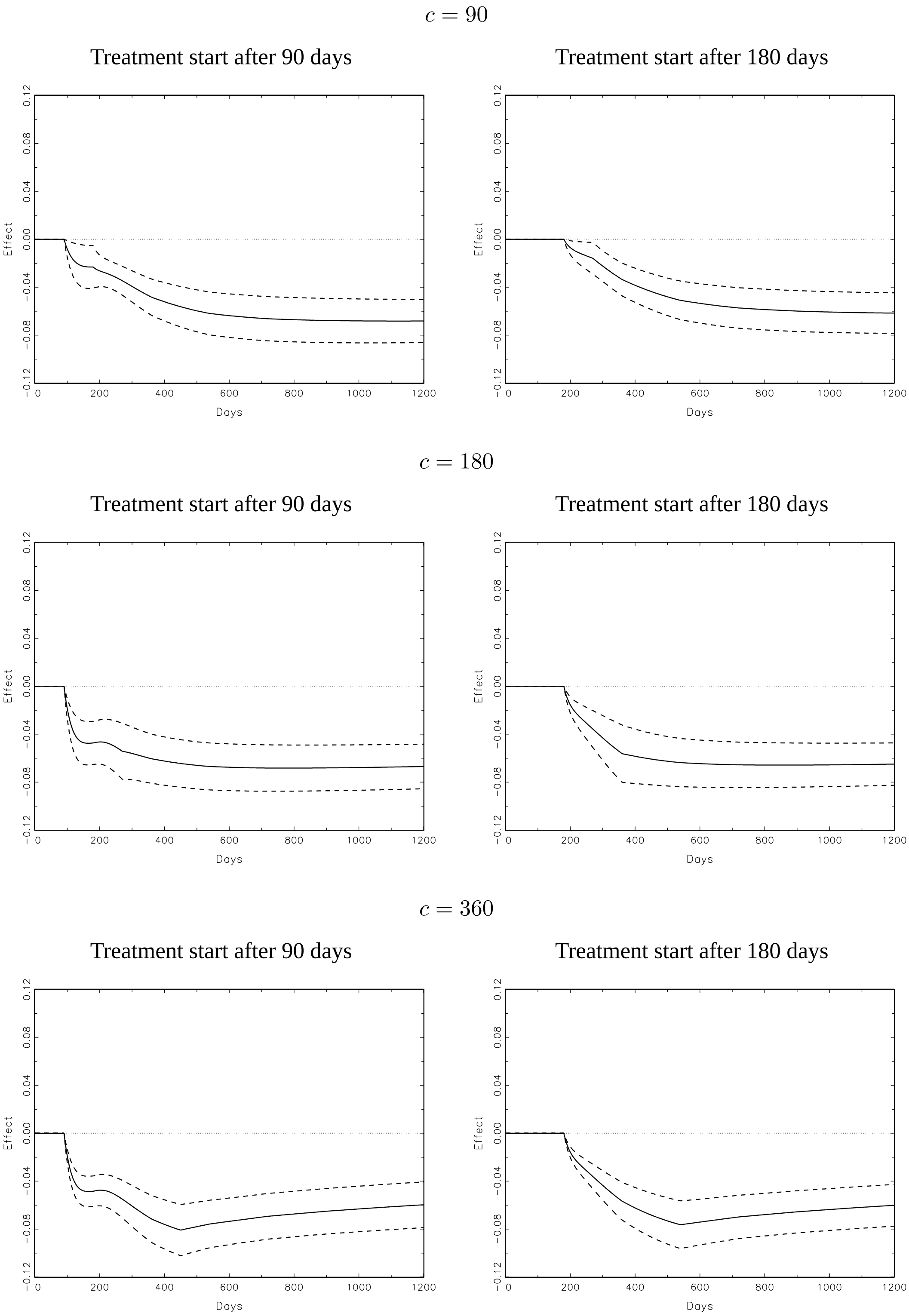
	$c = 90$		$c = 180$		$c = 360$	
Effect	Coeff.	t-Value	Coeff.	t-Value	Coeff.	t-Value
μ_1	0.2578	2.50	0.5297	5.37	0.5381	8.03
μ_2	0.4104	7.23	0.3412	5.26	0.1152	1.15
Log-Likelihood	-186,601.20		-186,600.92		-186,595.62	

In order to analyze the dynamic development of the treatment effect, we estimate an extended model where the treatment effect is allowed to vary over time. As presented in section 3, we specify the treatment effect as a piecewise-constant function $t - t_p$, where μ_1 is the treatment effect for period $[t_p, t_p + c)$ and μ_2 for period $[t_p + c, \infty)$. The specifications of baseline hazard, systematic part and unobserved heterogeneity are the same as in the basic model. We estimate three different models, with c set to 90, 180 and 360 days, i.e., the treatment effect is assumed to shift at these points of time. The results are given in Table 4. The estimates for baseline hazard, systematic part and unobserved heterogeneity are similar to that of the basic model and are not presented here.[12]

For the first two models, where the treatment effect is assumed to switch after 90 and 180 days, we find a positive effect on the hazard rate into employment for μ_1 and μ_2. For the first model, the hazard rate shifts by 30 percent during the first 90 days after the start of the TM and by 50 percent afterwards. The estimates of the second model imply that the shift of the hazard is even stronger during the first 180 days with 70 percent. For the remaining period, the effect is lower with an associated shift of 40 percent. This result suggests that the treatment effect increases within the first 6 months after the start of programs, and starts to decrease slightly afterwards. One possible explanation is that participants need some time to put the learned skills into practice. Taking a look at the model with $c = 360$ supports this finding. In this case a positive TM effect is apparent for μ_1 only, i.e., during the first year after start of programs with about 71 percent. Hence, as there is no effect of TM afterwards, i.e., program effects

[12] The results are available on request by the authors.

FIG. 4: EFFECT ON THE PREDICTED SURVIVOR FUNCTION FOR THE EXTENDED MODEL[a]



[a] Solid line represents the treatment effect on the predicted survivor function and the dashed lines represent the 95% confidence band. Confidence bands are calculated by Delta-Method.

have completely vanished one year after participation. This finding suggests two conclusions: First, the positive effects of TM last for a limited period only. Participants who do not find employment during this period will lose the gains afterwards. Second, a possible reason for the variation of the treatment effect over time is the content of the program. The set-up of TM provides necessary skills, techniques, but also incentives for job seekers to apply for jobs. Apparently, after a certain amount of time negative effects of unemployment, such as discouraged worker effects, stigmatization etc., overlay the positive treatment effects.

In analogy to the basic model we estimate the treatment effect on the predicted survivor functions for different starting dates of the treatment (after 90 and 180 days) for the extended model with time-varying impacts (figure 4). The pictures show some interesting features of the treatment effects when these are allowed to vary over time. Assuming the impact of TM to shift after 90 days ($c = 90$) reveals an almost similar effect on the survivor function as in the basic model. In contrast, if we assume the treatment effect changes after 180 days post program start, the picture is clearly different compared to the basic model. In particular during the first 180 days after program start we find a more pronounced positive effect of TM than in the basic model. Again, we could establish stronger effects if programs are started early in the unemployment spell. The strongest differences are observable for the case in which effects are assumed to shift after 360 days. During the first year after the start of TM the effect on the survivor function increases steadily, so after one year it turns out to be considerably stronger than in the basic model. However, it subsequently decreases and is almost identical to that of the basic model three years later. These results support the above findings. The effect on the predicted survivor functions points towards a treatment effect during the first year after program start only.

Finally, we wish to analyze whether treatment effects are heterogeneous due to individual characteristics. In particular, we analyze to what extent low qualified men with some work experience are affected by TM. In addition, we compare the effects with groups that differ in individual characteristics. More specifically, we estimate the effects for low qualified men who lack any work experience as well as for high qualified men with work experience (university or advanced technical college level). Finally, we compare the results for men to those for low qualified women with work experience. To do so, we use another extension of the model where the impacts of TM are allowed to vary with observable characteristics. The treatment effect is specified as a permanent and constant shift of the hazard rate similar to the basic model. Again, for the baseline hazards, systematic part and unobserved heterogeneity we use the specifications of the basic model and do not report the estimates here.[13] Table 5 shows the results for the treatment effects.

The effect for low qualified men with work experience is $\exp(0.4854) = 1.62$ and higher than the average (basic model). Unfortunately, for higher educated people and for people without occupational experience no differences could be found. However, for low qualified women with work experience, we estimate a treatment effect of $\exp(0.3099) = 1.36$. Although this group benefits from participation, the increase of the hazard rate is not as strong as for comparable men. Nevertheless, the hazard rate into employment for low qualified, but experienced men (women) shifts by about 62 (36) percent as a result of participation. Hence, TM are clearly successful in improving the search efficiency for employment.

[13] The results are available on request from the authors.

TAB. 5: EFFECT HETEROGENEITY DUE TO INDIVIDUAL CHARACTERISTICS

Effect	Coeff.	t-Value
Main Effect	0.4854	6.37
Women	-0.1755	-1.91
High Qualification	-0.0546	-0.29
Without Occupational Experience	0.1628	0.87
Log-Likelihood	-186,600.03	

6 Conclusion

TM are the largest single most important form of intervention undertaken on behalf of the unemployed in the context of German active labor market policy. Programs aim at improving the search efficiency for employment by offering a diversity of courses and counseling activities. Based on data from administrative processes of the FEA we have analyzed the empirical effects of these programs. An important aspect for the evaluation of program effects is information on the timing of the treatment event during the unemployment spell. To take account of this as well as of observable and unobservable influences, we use a multivariate mixed proportional hazards model for estimation as suggested by Abbring and van den Berg (2003). In addition, we extend the model for analysis of heterogeneity in the effects: First, treatment effects are permitted to vary over time, i.e., we explicitly regard the possibility that program effects develop or degenerate over time. Second, we consider differences in the effects due to individual characteristics. To shed more light on program impacts, we calculate the effects on survivor functions and the expected unemployment duration as well.

Based on three inflow samples into unemployment in western Germany for June, August and October 2000 that are followed up to December 2003, the estimates show that participation in TM clearly reduces the time individuals search for employment. Hence, programs are effective in shortening the unemployment duration of job seekers. The positive effects of TM affect the search process immediately from the start of the programs. The results show that TM are particularly successful in reducing the unemployment duration in the short to mid-term. Considering the dynamics of the effects from the results of the extended model indicates that impacts of TM on the transition into employment are strongest during months 3 to 6 after programs begin. Effects subsequently tail off. More than 12 months after participation, program effects have vanished completely. The analysis of heterogeneity due to individual characteristics revealed gender differences in impact. Although low qualified people with some work experience benefit from programs, the impacts are larger for men than for women. In summary, the results show that TM are successful in reducing the unemployment duration of participating individuals and substantially improve the employment chances of job seekers.

The empirical estimates of TM for western Germany are quite positive compared to the results of many ALMP programs in Germany and other countries. However, recommending an unrestricted use of programs in the future requires further research in several directions. Unfortunately, the available data do not yet enable us to conduct such research. First, one shortcoming is that we could not distinguish between the different modules and combinations of modules in the analysis. These may, of course, have an impact on program effects. Second, despite the positive effects on the transition rate into employment the time horizon of the analysis is too short to study the recurrence of unemployment. Third, the analysis is limited to western Germany. However, the eastern German labor market is plagued by higher unemployment. It is therefore essential that impacts on eastern Germany are also studied. Finally, as some TM are used as preparation measures for other ALMP programs, TM should be analyzed with respect to this purpose. Despite these ongoing research questions, the results presented in this study offer important first evidence on the effects of TM and show that reforms in labor market policy in recent years in Germany is now bearing fruit.

References

ABBRING, J. H., AND G. J. VAN DEN BERG (2003): "The Nonparametric Identification of Treatment Effects in Duration Models," *Econometrica*, 71(5), 1491–1517.

——— (2004): "Analyzing the Effect of Dynamically Assigned Treatments Using Duration Models, Binary Treatment Models, and Panel Data Models," *Empirical Economics*, 29(1), 5–20.

BIEWEN, M., B. FITZENBERGER, A. OSIKOMINU, AND M. WALLER (2006): "Employment Effects of Short and Medium Term Further Training Programmes in Germany in the Early 2000s," Working Paper, Goethe University Frankfurt.

BUNDESAGENTUR FÜR ARBEIT (2005a): "Arbeitsmarkt 2004 – Amtliche Nachrichten der Bundesanstalt für Arbeit," 53. Jahrgang, Nürnberg.

——— (2005b): "Daten zu den Eingliederungsbilanzen 2004," 53. Jahrgang, Nürnberg.

BUNDESANSTALT FÜR ARBEIT (2003): "Arbeitsmarkt 2002 – Amtliche Nachrichten der Bundesanstalt für Arbeit," 51. Jahrgang, Nürnberg.

CRÉPON, B., M. DEJEMEPPE, AND M. GURGAND (2005): "Counselling the Unemployed: Does It Lower Unemployment Duration and Recurrence?," Working Paper no. 2005-27, Paris-Jourdan Sciences Economiques.

FITZENBERGER, B., AND R. WILKE (2006): "Censored Box-Cox Quantile Regression for Unemployment Duration in Germany," mimeo, Centre for European Economic Research.

FREDRIKSSON, P., AND P. JOHANSSON (2004): "Dynamic Treatment Assignment - The Consequences for Evaluations Using Observational Data," Discussion Paper No. 1062, IZA, Bonn.

HECKMAN, J. J., R. J. LALONDE, AND J. A. SMITH (1999): "The Economics and Econometrics of Active Labor Market Programs," in *Handbook of Labor Economics*, ed. by O. Ashenfelter, and D. Card, vol. 3A, chap. 31, pp. 1865–2097. Elsevier, Amsterdam.

HECKMAN, J. J., AND B. SINGER (1984): "The Identifiability of the Proportional Hazard Model," *Review of Economic Studies*, 39-6257, 231–241.

HOLLAND, P. W. (1986): "Statistics and Causal Inference," *Journal of the American Statistical Association*, 81, 945–960.

KLUVE, J. (2006): "The Effectiveness of European Active Labor Market Policy," Discussion Paper No. 2018, IZA, Bonn.

KURTZ, B. (2003): "Trainingsmaßnahmen – Was verbirgt sich dahinter?," Werkstattbericht Nr. 8, IAB.

LANCASTER, T. (1979): "Econometric Methods for the Duration of Unemployment," *Econometrica*, 47, 939–956.

——— (1990): *The Econometric Analysis of Transition Data*. Cambridge University Press, Cambridge, New York, Oakleigh.

MORTENSEN, D. (1986): "Job Search and labor market analysis," in *Handbook of Labor Economics Vol.II*, ed. by O. Ashenfelter, and D. Card, pp. 849–919. Elsevier, Amsterdam.

MORTENSEN, D., AND C. PISSARIDES (1999): "Unemployment Responses to 'Skill-Biased' Technology Shocks: The Role of Labour Market Policy," *The Economic Journal*, 109, 242–265.

RICHARDSON, K., AND G. J. VAN DEN BERG (2001): "The Effect of Vocational Employment Training on the Individual Transition Rate from Unemployment to Work," *Swedish Economic Policy Review*, 8(2), 175–213.

STEPHAN, G., S. RÄSSLER, AND T. SCHEWE (2006): "Wirkungsanalyse in der Bundesagentur für Arbeit - Konzeption, Datenbasis und ausgewählte Befunde," Discussion Paper No. 07/2006, IAB.

VAN DEN BERG, G. J. (1990): "Nonstationary Job Search Theory," *Review of Economic Studies*, 57, 255–277.

VAN DEN BERG, G. J. (2001): "Duration Models: Specification, Identification, and Multiple Durations," in *Handbook of Econometrics*, ed. by J. J. Heckman, and E. Leamer, vol. 5, chap. 55, pp. 3381–3462. North-Holland, Amsterdam.

WEBER, A., AND H. HOFER (2004a): "Are Job Search Programmes a Promising Tool? A Microeconometric Evaluation for Austria," Discussion Paper No. 1075, IZA, Bonn.

——— (2004b): "Employment Effects of Early Interventions on Job Search Programmes," Discussion Paper No. 1076, IZA, Bonn.

A Tables

TAB. A.1: ESTIMATION RESULTS (WITHOUT UNOBSERVED HETEROGENEITY)[1]

Variable	Transition Rate into Employment Coeff.	t-Value	Transition Rate into Training-Programme Coeff.	t-Value
Baseline Hazard				
$\lambda_{90\geqslant Y<180}$; $\lambda_{180\geqslant S<360}$	-0.4496	-26.97	-0.2771	-3.99
$\lambda_{180\geqslant Y<360}$; $\lambda_{360\geqslant S<540}$	-0.9084	-50.84	-0.8501	-8.36
$\lambda_{360\geqslant Y<540}$; $\lambda_{540\geqslant S<720}$	-1.5131	-56.84	-0.9750	-8.24
$\lambda_{540\geqslant Y<720}$; $\lambda_{720\geqslant S<900}$	-1.9369	-54.50	-1.1542	-8.40
$\lambda_{720\geqslant Y<900}$; $\lambda_{S\geqslant 900}$	-2.3105	-50.92	-1.4882	-11.49
$\lambda_{900\geqslant Y<1080}$	-2.4894	-47.97		
$\lambda_{Y\geqslant 1080}$	-2.5684	-42.51		
Constant	-4.9643	-86.41	-8.7210	-37.06
Age	-0.0171	-21.11	-0.0028	-0.83
Women	0.0495	3.20	0.0067	0.10
Applicant for Full Time Job only	0.0814	4.35	0.0370	0.47
Occupational Experience (Yes)	-0.0374	-1.56	-0.0580	-0.55
No. of Children	0.0082	1.09	0.1036	3.38
Vocational Education				
– In-Firm Training	0.0741	3.86	0.1302	1.58
– Off-the-Job Training	0.0722	1.30	0.4538	2.16
– Vocational School	0.0786	1.65	0.0235	0.12
– Technical School	0.1388	3.78	-0.1595	-0.91
– University	-0.0005	-0.01	0.1635	0.70
– Advanced Technical College	-0.0011	-0.02	0.0123	0.04
Level of Qualification				
– University Level	-0.0446	-1.00	-0.4983	-2.31
– Advanced Technical College Level	-0.0094	-0.17	-0.4384	-1.62
– Technical School Level	0.0740	1.65	-0.0278	-0.14
– Skilled Employee	0.0654	3.56	0.0798	1.02
School Education				
– CSE[2]	0.0948	4.28	0.1188	1.23
– O-Level (*Realschulabschluss*)	0.0652	2.45	0.2348	2.06
– Advanced Technical College (*Fachhochschulreife*)	0.0562	1.53	0.1148	0.72
– A-Level (*Abitur*)	0.0530	1.61	0.0277	0.19
Family Status				
– Single Parent	0.1294	4.38	0.0410	0.32
– Married	0.0869	5.49	-0.0996	-1.45
Occupational Group				
– Manufacturing Industry	0.0810	2.15	-0.1408	-1.02
– Technical Occupation	0.1354	2.64	0.5850	3.15
– Service Professions	0.1407	3.76	-0.0019	-0.02
Entry into the Sample				
– Entry in August	-0.0665	-4.44	0.1646	2.46
– Entry in October	-0.0789	-4.90	0.1059	1.48
Treatment Effect (μ)	0.1881	4.90		
Log-Likelihood	-186,973.44			

[1] Reference categories for categorial variables: Vocational education, *missing education*; level of qualification, *with and without technical knowledge*; schooling, *without graduation*; family status, *singles/not married*; desired occupational group, *agriculture, mining, fishery and miscellaneous occupations*.
[2] Certificate of Secondary Education (*Hauptschulabschluss*).