

Discussion Paper No. 06-065

**The Effects of Short-Term Training
Measures on the Individual Unemployment
Duration in West Germany**

Reinhard Hujer, Stephan L. Thomsen
and Christopher Zeiss

ZEW

Zentrum für Europäische
Wirtschaftsforschung GmbH

Centre for European
Economic Research

Discussion Paper No. 06-065

The Effects of Short-Term Training Measures on the Individual Unemployment Duration in West Germany

Reinhard Hujer, Stephan L. Thomsen
and Christopher Zeiss

Download this ZEW Discussion Paper from our ftp server:

<ftp://ftp.zew.de/pub/zew-docs/dp/dp06065.pdf>

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

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

Non-Technical Summary

With about 800 thousand newly promoted individuals in West and about 1.2 million in Germany in 2004, short-term training measures (*Maßnahmen der Eignungsfeststellung und Trainingsmaßnahmen*, TM) are the most important intervention of German active labor market policy. This is the first study analyzing the effects of these programs on the individual unemployment duration in West Germany. Since participation in TM should improve the search process for employment, we measure the effects of programs on the duration until individuals become employed. By applying a multivariate mixed proportional hazards model, we are able to consider information of the timing of treatment in the unemployment spell as well as observable and unobservable factors to control for selectivity. Moreover, we allow treatment effects to vary over time and take account of heterogeneity in the effects due to individual differences.

Based on three inflow samples into unemployment in West Germany of 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-run. 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 the beginning of programs. Effects start to decrease after that time. More than 12 months after participation, program effects have vanished completely. From the analysis of heterogeneity due to individual characteristics, gender differences in impacts could be established. Although low qualified persons 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 improve the employment chances of job seekers clearly.

THE EFFECTS OF SHORT-TERM TRAINING MEASURES ON THE INDIVIDUAL UNEMPLOYMENT DURATION IN WEST 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 11, 2006

Abstract

Short-term training measures are the most important intervention of German active labor market policy in terms of persons promoted. However, evidence on the impacts of programs is missing. This study analyzes the effects of these programs on the individual unemployment duration in West Germany. By applying a multivariate mixed proportional hazards model, we are able to consider information of the timing of treatment in the unemployment spell as well as observable and unobservable factors to control for selectivity. Moreover, we allow treatment effects to vary over time and take account of heterogeneity in the effects due to individual differences.

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

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 the annual budget – about 19.5 billion Euro (36 percent) in 2004 – with the purpose to improve the employment chances of about 2.5 million persons participating in the different active labor market policy (ALMP) programs.¹ The most important program are short-term training measures (*Maßnahmen der Eignungsfeststellung und Trainingsmaßnahmen*, TM) with about 1.2 million newly promoted individuals in 2004 of which 788,533 joined programs in the western part. TM exceed other programs in West Germany by far, e.g., the second most important program have been bridging allowances for 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 persons 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 intervention that either attempt to support the job placement on part of the employment agencies as well as the self-contained job search of the participants, or should adjust the qualification of the participants to the demands of the market. Therefore, TM generally aim to improve the prospects of job search, i.e., the job search process. For the empirical analysis it is useful to measure the impact of TM on the search process in terms of the duration of unemployment until 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 the individual joins the TM in the unemployment spell. Standard evaluation literature usually deals only with binary information if an individual has received a treatment or not, see e.g., 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 microeconomic 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 allows to control for observable and unobservable factors to identify the treatment effect in presence of selectivity, which is a major issue for 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 of three inflow samples into unemployment from June, August and

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.

October 2000, where observations are followed until December 2003. The study is restricted to West Germany, since labor market and economic situation of West and East Germany are clearly different even more than a decade after German Unification 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 of 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 of Short-Term Training Measures in Germany

TM were introduced with the enactment 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 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 are 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*). 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 twelve weeks at maximum. TM are accomplished at service providers (*Bildungsträger*) and firms ensuring 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 of the first, about 19 percent of the second, and about 28 percent of the third module. Combinations amounted to 18 percent of all promotions. 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 funded 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 must not be supported if it should lead to a recruitment at a former employer of the individual (during the last four years for more than three months subject to compulsory insurance), 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 engage the participant without promotion in TM or if placement of suitable experts is possible.

Caseworkers possess a lot of discretion in the allocation of participants. Hence, it is interesting to know the determinants of 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 the integration chances, but also previous knowledge as well as motivation of job seekers. The results indicate that caseworkers assess the preceding unemployment duration 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 persons (with university degree) are regarded more rarely.

The rising importance of TM within ALMP in West (and East) Germany becomes obvious from table 1 presenting the number of entries into the three most important ALMP programs as well as the unemployment rates for the years 2000 to 2004. Whereas the East German economy has been plagued by unemployment rates of 17.1 (2000) to 18.4 percent (2003), the analogue figures for West Germany were 7.2 (2001) to 8.5 percent (2004). The development of the ALMP mix reflects this regional difference as well. In West Germany, the focus is on programs that aim at adjusting the qualification of the individuals to the demands of the market. The emphasis in East Germany is on employment programs relieving the tense situation of the market. In both regions, but with a stronger emphasis in the West, the number of TM has increased significantly. In 2000, TM have been the second most important program with 285 (201) thousand persons promoted in West (East) 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 decrease of the more traditional programs and reflects the reforms

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).

of German ALMP in 1998 and the following years.² The main reason for that reform was the high and persistent unemployment associated with a tense budgetary situation of 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. Their importance decreased as both are long in duration (for up to three years) and expensive.³ TM are clearly shorter and program costs are much lower than for other measures. In 2004 (2003), the FEA spent 496 (577) million Euro on TM; the average costs per participant and month amounted to 538 Euro (Bundesagentur für Arbeit, 2005b).

Despite those facts, empirical evidence on the effectiveness is rare for Germany. Reasons for that maybe lack of appropriate data, heterogeneity of programs, and that TM are sometimes used as preparative measures for continuative 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. In particular, they compare short-term training measures, further training, and retraining programs (medium- and long-term training), and find positive effects for both gender in West Germany. The study by Stephan, Rässler, and Schewe (2006) is intended as an illustrative example for the evaluation of several ALMP programs on the basis of a new scientific data base of 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 of the third module accomplished at firms reduce the probability of unemployment as well as the remaining unemployment duration. TM aiming to assess the willingness and the ability to work have negative effects on those outcomes.

On the European level, programs directly comparable to German TM are difficult to find since they

² Since 1998, the legal basis for ALMP in Germany was 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.

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

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 on the microeconomic level. The comprehensive survey tables in Kluve (2006) show rather mixed results of the effects that are negative in 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 locking-in effects reported frequently. 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 participant’s unemployment duration; programs started later have the opposite effect. Unfortunately, a sufficient explanation for that large drop is missing. 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 job placement on part of the employment agency as well as the self-contained job search of the participants. Second, programs are used to adjust the qualification of job seekers to 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 individuals in terms of an optimal stopping problem in a dynamic and uncertain environment.⁴ 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 simplicity, one can think of 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 he is assumed to know 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 accept or to reject the wage offer without possibility of recall. If the job seeker accepts a wage offer the search process stops and he becomes employed at wage w forever.⁵ 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

⁴ See Mortensen (1986) and Mortensen and Pissarides (1999) for a detailed discussion of the search model.

⁵ In the simple model, job-to-job transitions are excluded.

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

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

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:

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

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 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 how a participation in TM affects the duration of unemployment. According to the institutional set-up of TM, we can deduce two channels how programs affect the job search of the 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 in particular TM in form of the first (*Eignungsfeststellung*) and the second module (*Überprüfung der Verfügbarkeit/Bewerbertraining*) improve the search behavior of the participants. Programs in terms of 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 terms of the second module may increase the efficiency and intensity of the self-contained job search by counsels with respect to the possibilities of job search or courses with respect to the application process. Improving the search behavior either by increasing the 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 a participation in a training measure increases the number of job offers that arrive in the small interval dt . The impact of a participation in a TM on the 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

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

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 a participation in a TM which increases the search efficiency, directly lowers the unemployment duration on the one hand, but on the other hand makes the workers more selective with respect to the 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, a participation in a TM may improve job-relevant skills and therefore increases the job opportunities of the participants. In particular, TM in form of the third module have the objective to teach fundamental skills that are required to be placed into an employment or a vocational education. In the case a training measure increases the skills of the participant this is equivalent to an increased productivity. This allows participants to apply for job’s, which are on average associated with higher wages. In the following analysis we therefore assume that a participation in a training measure shifts the mean of the wage offers distribution $F(w)$ to a higher level.⁶ Following Mortensen (1986) 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

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

we find that a marginal increase in the mean of the distribution $F(w)$ decreases the probability to obtain 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

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

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):

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

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

⁶ Mortensen (1986) also considers changes in the variance of $F(w)$.

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 of the policy analysis, which we have not addressed so far, is the question if policy changes are anticipated by the individuals. Individuals anticipating a future participation will adjust their optimal search strategy at the point in time the information of a 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, and we could only estimate the composite effect of TM on the hazard rate into employment. Hence, the empirical analysis does distinguish between effects on the offer arrival rate and on the wage offer distribution. Moreover, it does not differentiate between the three modules and combinations of modules either.

3. Econometric Model

In this section we present the econometric model for the estimation of the treatment effect of TM on the unemployment duration. 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 caseworkers decision 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. At the point in time an individual enters unemployment, we measure the duration until the individual enters employment (T_e) and the duration until he/she joins a TM (T_p). 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 (x) 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 . 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 in 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 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., $\mu(t - t_p, x)^{I(t>t_p)} = \mu_1^{I(t_p < t \leq t_p + c)} \mu_2^{I(t > t_p + c)}$, 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 allows to analyze the development of the treatment effect 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 is related to the observable and unobservable factors. Technically, selectivity means that those individuals who are observed to receive a treatment at t_p are a non-random subset with respect to t_e . Whereas any selectivity conditional on observable is captured by the systematic part in equation (7) possible selection on unobservable 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 . In the case selectivity cannot fully be captured by the observable characteristics, we would observe a dependence of the unobserved heterogeneity terms and the indicator function for the treatment effect appears as an endogenous time-varying regressor.

A further important aspect of the model is the 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 a useful information since it allows to distinguish between time-invariant selection effects embodied by the 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 of how long the elapsed duration of unemployment is. In contrast, in case of a selection effect we would observe a correlation between the points in time of the transitions into employment and program. E.g., a positive selection effect results in a pattern where a quick transition into 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 to impose exclusion restrictions on the observable variables as it is the case in selection models. Such exclusion restrictions on x are often hardly to justify from a theoretical point of view, since 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 that case, the hazard rate at t of an individual that anticipates a future treatment at time t_p , will be different from the hazard rate of an individual that obtains an alternative treatment at time t_p^* for $t \leq \min\{t_p, t_p^*\}$.⁷ 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 of the treatment event is usual not available. Therefore, we assume that either participations in TM are not anticipated, or if they are 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 the individuals act on the determinants of T_p . That is, 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

⁷ The alternative treatment at t_p^* includes the no-treatment case, see Abbring and van den Berg (2003).

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

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

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

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)

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

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. Then four combinations with an associated probability are 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

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

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

4. Data and Descriptive Statistics

4.1. Data

The empirical analysis is based on three samples of inflows into unemployment in West Germany in months June, August and October 2000. The labor market status are observed until December 2003. The data were merged from several data sets for administrative purposes of the FEA. The main source of information is the job seekers data base (*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. In addition, there are some subjective ones as well, like 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 (see section 2.1), and 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 all persons who are registered in the German social security system proving the individual pension claims. These are all persons employed compulsorily to social security.⁸ 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 with day as 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 that kind of proxy includes periods out of the labor force or receiving social benefits as well. For that reason, labor force movements as well as episodes of employment not subject to social security are not identified in the data. Therefore, the non-employment duration has to be assumed to be an upward biased proxy of the true unemployment duration. Fitzenberger and Wilke (2006) analyze unemployment durations in Germany dealing with similar problems. They suggest to use a lower (times of permanent income transfers) and an upper bound (non-employment duration) for the unemployment duration. Their results indicate that both bounds do not differ too strong if early retired older persons are excluded. For this reason and limitations of our data, we refrain from estimating the effects for the lower bound.

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 2000.⁹ From this sample, we exclude all individuals 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 persons became unemployed from employment or entered the labor force for the first time. For that reason, unemployment entry dates in the sample correspond to unemployment entry in economical sense. Furthermore, we restrict the sample for homogeneity reasons to domestic people who are neither disabled nor affected by other health restraints. Moreover, to avoid influences related to professional training we exclude persons 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 reduce the bias of the unemployment proxy as well (see Fitzenberger and Wilke, 2006). 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 un-

⁸ Self-employed and pensioners are not included.

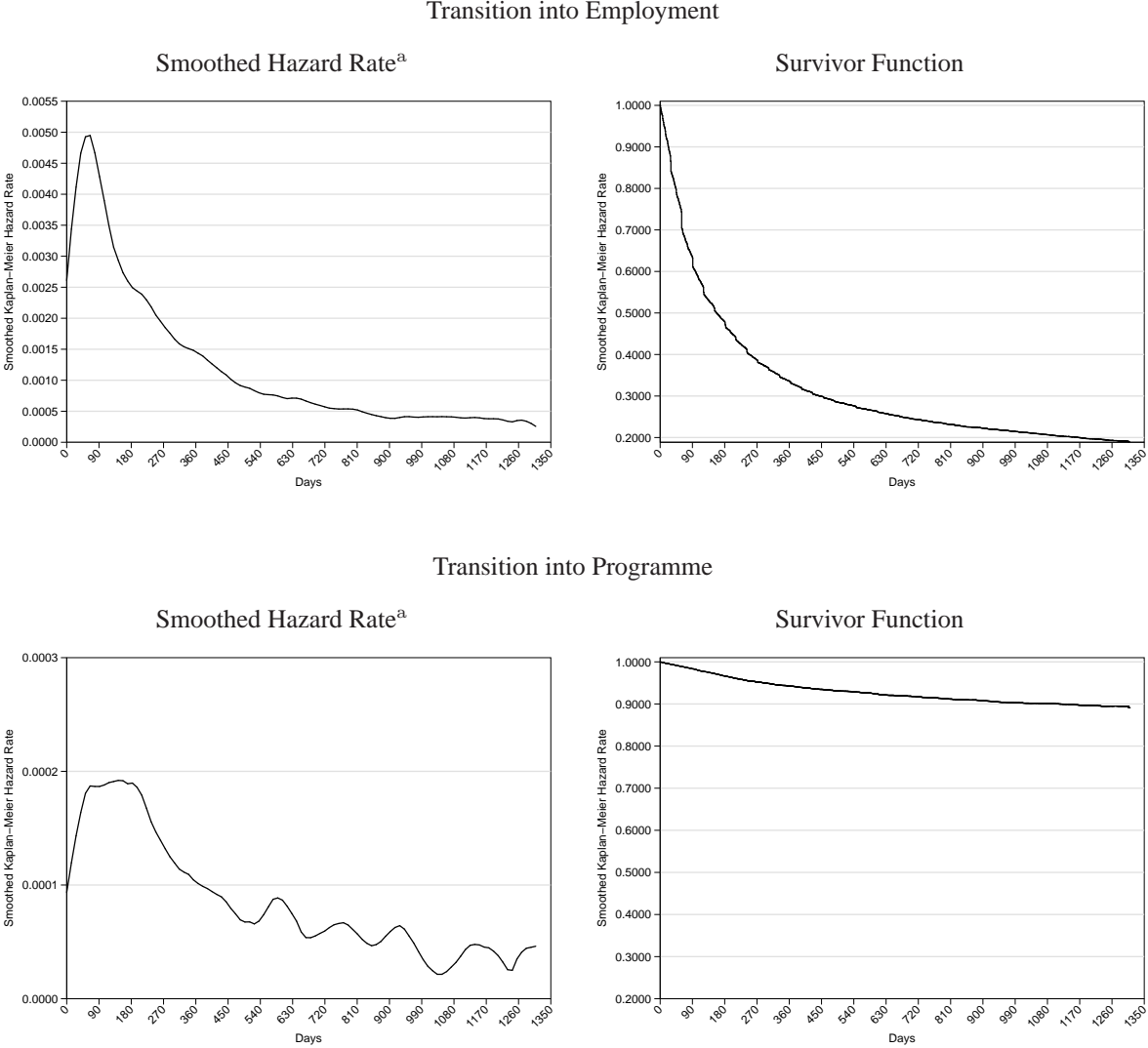
⁹ We consider differences due to the starting dates of the unemployment spell in calendar time by including dummy variables in the empirical analysis

employment 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 program. For the transition into employment we find a quite typical picture. In particular during the first three months, job seekers experience the highest probability to leave for employment. After that time, the chances of finding a job decrease strongly. The corresponding survivor function implies that the probability being still not employed after three months is almost 60 percent; after three years, this probability decreases to about 20 percent. The transition rate into TM

FIG. 1: NON-PARAMETRIC ESTIMATES



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

establishes a slightly different picture. Job-seekers have the highest chances to enter a TM within the

first six to seven months after the start 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.

TAB. 2: DESCRIPTIVE STATISTICS FOR COVARIATES¹

	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 (<i>Realschulabschluss</i>)	20.74	23.57	20.63
Advanced Technical College (<i>Fachhochschulreife</i>)	5.85	5.42	5.87
A-Level (<i>Abitur</i>)	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.

Table 2 presents means and frequencies of the observable covariates used in the analysis to point out equalities and differences. The results of Kurtz (2003) indicate that important determinants for the decisions of caseworkers to promote job seekers by TM are the placement chances after participation, the compensation of missing occupational qualification as well as previous knowledge and motivation. In the empirical analysis, we approximate missing 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 persons without graduation. It becomes obvious, that participants do not differ much in these variables from other job seekers. However, the ratio of participants owning an O-level degree (*Realschulabschluss*) is larger (23.57 part. /20.63 non-part. percent) and that of persons with an A-level degree (*Abitur*) is smaller compared to that of non-participants (10.83/13.10 percent). Analogously, participants do less often own a technical school or university degree.

The life cycle position of the individual is an important determinant for the labor market performance as well. 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 of 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 do more often apply for technical professions than the average job seekers (5.20/3.62 percent). However, none of the covariates seems to determine participation or non-participation clearly. We are also not able to approximate the motivation of the 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). The main interest is 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

TABLE 3: ESTIMATION RESULTS (BASIC MODEL)¹

Variable	Transition Rate into Employment		Transition Rate into Training- Programme	
	Coeff.	t-Value	Coeff.	t-Value
Baseline Hazard				
$\lambda_{90} \geq Y < 180$; $\lambda_{180} \geq S < 360$	0.3292	9.47	-0.0529	-0.498
$\lambda_{180} \geq Y < 360$; $\lambda_{360} \geq S < 540$	0.8828	10.35	-0.4016	-2.283
$\lambda_{360} \geq Y < 540$; $\lambda_{540} \geq S < 720$	0.5902	7.72	-0.3532	-1.525
$\lambda_{540} \geq Y < 720$; $\lambda_{720} \geq S < 900$	0.1723	2.21	-0.3882	-1.350
$\lambda_{720} \geq Y < 900$; $\lambda_S \geq 900$	-0.2004	-2.43	-0.5738	-1.684
$\lambda_{900} \geq Y < 1080$	-0.3794	-4.40		
$\lambda_Y \geq 1080$	-0.4673	-5.09		
Unobserved Heterogeneity (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 (<i>Realschulabschluss</i>)	0.0643	1.60	0.3573	2.283
– Advanced Technical College (<i>Fachhochschulreife</i>)	-0.0061	-0.11	0.1931	0.883
– A-Level (<i>Abitur</i>)	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 categorical 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*).

1.48. That is, 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 enhance the search process of the participating individuals clearly, i.e., participation reduces the time persons seek for employment.

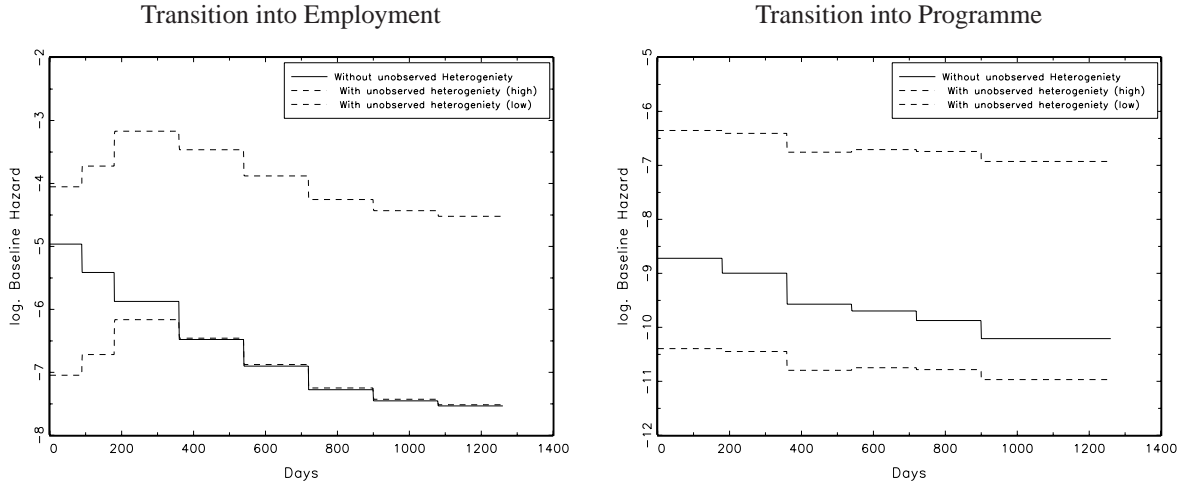
The observable covariates affect the transition rate into employment in different directions. It increases with the *number of children* and for *single parents* and *married* persons as well. As these variables are indicators on the responsibility of the job seeker for closely related persons, the results show that those persons are more successful in finding jobs. A possible reason maybe the higher willingness to actively seek for employment, but also the higher necessity to work. In addition, it should be noted that with respect to gender women require a shorter time to find a new occupation. Moreover, the transition rate into employment increases with qualification. *Skilled employees* are better off in finding a job compared to the unskilled. Owning a graduation from school provides a significant estimate for *CSE* only; for the other groups no differences could be attested. Persons who seek for jobs in the *manufacturing industry*, *technical occupations* or in *service professions* do also experience transitions into employment more often than the reference group (*agriculture, fishery, mining and miscellaneous other occupations*).

A quite common finding in the empirical literature is that older unemployed workers face barriers to employment. The estimate for *age* is in line with that. In addition, persons who are not open to 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 persons who became unemployed in June 2000. Becoming unemployed in August is worse and the lowest transition rate is established for the October entries.

The estimates of the influence of the observable covariates on the transition rate into programs show a mixed picture. Due to statistical insignificance of most parameters, deriving clear rules of selection into programs is hampered. However, persons with *children* are preferred for participation in TM. The positive parameter of *off-the-job training* shows that those persons are favored by caseworkers compared to persons without vocational education. An increase of the transition rate could also be established for persons owning an *O-level* degree (compared to persons without graduation). In association with the large number of TM that are accomplished in Germany, the positive parameter for persons applying for *technical occupations* reflects (at least to some extent) the structural change of the German economy. Hence, in particular for those people the contents of TM may be useful to adjust their qualification to the demands of the market. Finally, the dummy for the unemployment entry shows that persons who became unemployed in August 2000 have increased participation chances. One reason could be that some TM are used as preparative courses for professional training in the apprenticeship system (starting in September).

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

FIG. 2: ESTIMATED BASELINE HAZARDS



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, 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 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).¹⁰ 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 is decreasing 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 we ignore unobserved heterogeneity 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 av-

¹⁰ 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.

erage treatment effects that are 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 captures the dynamic accumulation of the treatment effect over the unemployment spell. However, considering these effects requires to be explicit with respect 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

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

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.,

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

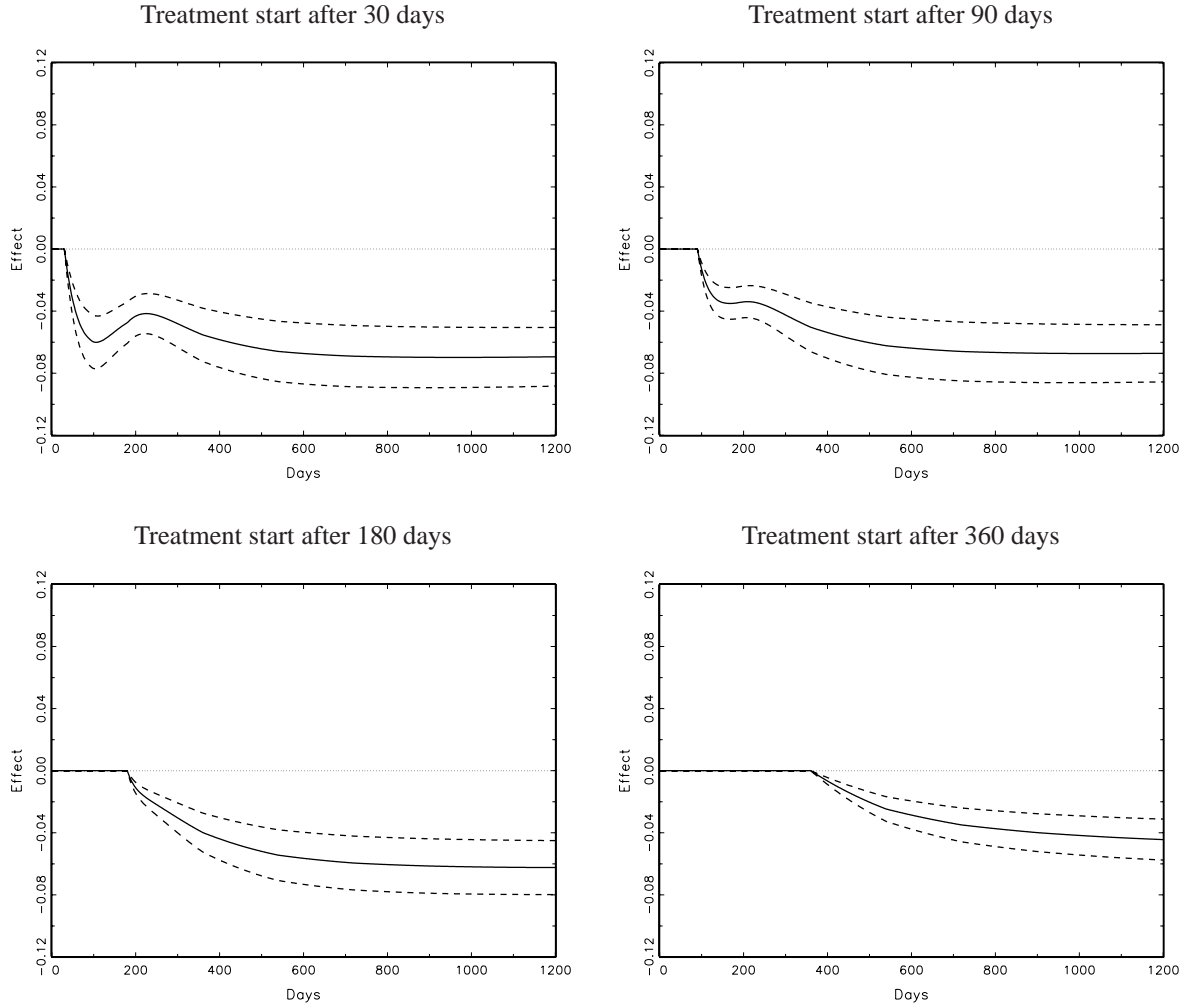
To calculate the effect of a participation in a TM, we predict the survivor function for the empirical model using the estimated parameters and means of the observable and unobservable covariates. The effects on the survivor functions are calculated for hypothetical program starts after 30, 90, 180 and 360 days of unemployment that are compared to the no-treatment case.

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 no-treatment case. That is, the probability to be still unemployment at time t is significantly lowered. What becomes obvious from the figures is that impacts of TM are stronger when programs start earlier compared to starting 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 is 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

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.

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 the newly received job application advices of the participants are not associated with instantaneous employment, but perhaps better 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.

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

TAB. 4: TIME VARYING TREATMENT EFFECT

Effect	$c = 90$		$c = 180$		$c = 360$	
	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	

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 we refrain from presentation.¹¹

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 is shifted 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. A 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 the finding. Here, a positive effect of TM is visible 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 have completely vanished one year after participation. This finding implicates 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, like 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 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 to change 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, when effects are assumed to shift after 360 days. During the first year after 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,

¹¹ The results are available on request by the authors.

