The Effects of 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*

Working Paper

May 12, 2006

Comments welcome!

Abstract

With about 800 thousand newly promoted individuals in West and about 1.2 million in Germany in 2004, training measures (TM) are the most important intervention of German active labour market policy. This is the first study provides providing empirical evidence on the effects of these programmes in West Germany. Since participation in TM should improve the search process for employment, we measure the effects of programmes 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. The estimates show that TM clearly reduce the time individuals search for employment. The analysis of the variation over time indicates that effects are strongest during months 3 to 6 after the start of the programmes and decrease afterwards. More than 12 months after, effects have vanished completely. Moreover, programmes affect the search efficiency for low qualified men with some work experience more strongly than for comparable women. The results point out that TM in West Germany reduce the unemployment duration significantly.

Keywords: Training Measures, Active Labour Market Policy, West Germany, Multivariate Mixed Proportional Hazards, Time-Varying Treatment Effects, Evaluation, Unobserved Heterogeneity

JEL Classification: J64, J24, I28, J68

[*]The autors thank 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 the IZA, Bonn and ZEW, Mannheim, e-mail: hujer@wiwi.uni-frankfurt.de.

[‡]Stephan L. Thomsen is Research Assistant 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 labour market policy (ALMP) programmes.[1] The most important programme are training measures (*Maßnahmen der Eignungsfestellung und Trainingsmaßnahmen*, TM) with about 1.2 million newly promoted individuals in 2004 of which 788,533 joined programmes in the western part. Hence, TM exceed other programmes in West Germany by far, e.g., the second most important programme have been bridging allowances for self-employed (*Überbrückungsgeld bei Aufnahme einer selbständigen Tätigkeit*) with about 137,400 participants and vocational training programmes (*Förderung der beruflichen Weiterbildung*) with about 124,000 individuals newly promoted (Bundesagentur für Arbeit, 2005a). Although programmes are used on this large extent, there is no empirical evidence on the effects for the participating individuals. The lack of empirical studies seems to be surprising, as the number of evaluation studies for German ALMP programmes increased over the last years[2]. But, lack of appropriate data prevented evaluation of TM. Fortunately, we can base our empirical analysis on unique and very informative data for administrative processes of the FEA that are merged for scientific purposes. Hence, this is the first study analysing the microeconomic effects of the most important ALMP programme in West Germany.

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. In that sense, TM are a labour supply side oriented intervention and should improve the job placement process at the employment agencies. In words of economic theory, TM are expected to affect the search process for employment positively. To analyse these effects empirically, it is useful to measure changes in 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, in recent empirical literature the importance of information on the timing of treatment events has been emphasised. Abbring and van den Berg (2003) have shown that the timing of events conveys useful information for the identification of the treatment effect. In addition, Fredriksson and Johansson (2004) point out that the dynamic assignment of treatments has serious implications for the validity of the conditional independence assumption usually invoked to estimate treatment effects.

Therefore, for estimation of the effects we apply a multivariate mixed proportional hazards model (MMPH) 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 evaluate the effects of TM on the search process for employment based on three inflow samples into unemployment from June, August and October 2000, where observations are followed until December 2003. We restrict our analysis to programmes accomplished in West Germany,

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

[2] See Wunsch (2005) for a recent overview.

since labour market and economic situation of West and East Germany are clearly different even more than a decade after German Unification in 1990.

Whereas recent applications of the MMPH assume the treatment effect to be a constant and permanent shift of the hazard rate (see e.g., Hujer, Thomsen, and Zeiss (2006)), treatment effects should be modelled varying over time. This is due to the fact that, e.g., it may take some time for the effects to develop and affect the search process. In this sense, treatment effects would increase over time. Alternatively, it could be expected that after a certain amount of time other effects, e.g., discouraged worker effects etc. overlay the programme effect. We consider these possible differences by estimating an extended version of the model, where treatment effects are allowed to vary over time. Moreover, programme effects may also differ by individual characteristics, i.e., programmes are more effective for some subgroups of the labour market than for others. We take account of effect heterogeneity due to individual characteristics in a third model, and estimate the effects for selected subgroups.

The paper is organised as follows: the first part of section two provides some stylised facts on TM in Germany, the second part discusses the theoretical impacts of the programmes on the search process for employment within the prototypical search model by Mortensen (1986). The econometric model is discussed in section three. The fourth section introduces the data used in the analysis and provides selected descriptive statistics. The empirical estimates of the impacts of TM are presented in section five. The last part of section five predicts programme effects for different starting points in the unemployment spell. The final section concludes.

2. Training Measures

2.1. Stylised Facts on 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 recipients and employment counselling measures (*Maßnahmen der Arbeitsberatung*). The primary purpose of TM is to improve the integration prospects of the participating individuals. For this reason, programmes 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.

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 labour 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*). These activities should support the individual's efforts to find work or efforts by the employment agency to place him/her, especially through job-application training, counselling on job-search possibilities or measures assessing the unemployed person's willingness and ability to work. Measures of the second module are promoted for up to two weeks. The third module contains a practical training (for up to eight weeks) providing necessary skills and techniques required to be placed in employment or vocational training (*Vermittlung notwendiger Kenntnisse und Fertigkeiten*). These are 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.

To give an idea of the sizes of each of the modules, it is useful to mention the shares the FEA notes on its website (figures refer to 2005). About 34 percent of the participants join programmes in the first module, about 28 percent of the third, and about 19 percent of the second modules. Combinations amount 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 programmes.

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 authorised on recommendation or with the approval of the agency only and activities are often initiated by caseworkers. However, TM may also be initiated by job-seekers, service providers or firms. A programme must not be supported if it should lead to a recruitment at an employer who had already employed the person during the last four years for more than three months subject to compulsory insurance, or if he/she has offered an employment to the unemployed person before the start of the unemployment spell. Moreover, support is denied if the employer could be expected to engage the unemployed person without promotion in TM or if placement of suitable experts is possible.[3]

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 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 programmes, TM are offered to job-seekers with barriers to employment in particular, e.g., long-term unemployed. Higher educated persons (with university degree) are regarded more rarely.

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

	2000	2001	2002	2003	2004
Germany					
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					
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					
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).

The rising importance of TM within ALMP in West (and East) Germany becomes obvious from ta-

[3] Those precautions are imposed to avoid deadweight losses, see Layard, Nickell, and Jackman (1991).

ble 1 that presents the number of entries into the three most important ALMP programmes as well as the unemployment rates for the years 2000 to 2004. Whereas the East German economy was plagued by unemployment rates between 17.1 (2000) and 18.4 percent (2003), the analogue figures for West Germany were between 7.2 (2001) and 8.5 percent (2004). This difference is reflected in the ALMP mix, too. In West Germany, the focus is on programmes that aim at adjusting the qualification of the individuals to the demands of the market. The emphasis in East Germany is on employment programmes 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 programme with 285 (201) thousand persons promoted in West (East) Germany behind vocational training programmes. Five years later, TM are the most important measures with 789 (400) thousand participants (2004). This strong rise of TM was accompanied by a decrease of the more traditional programmes and reflects the reforms of German ALMP in 1998 and the following years.[4] The main reasons for that reform were the high and persistent unemployment and the tense budgetary situation of the FEA. Until the end of the 1990s, vocational training programmes and job creation schemes (*Arbeitsbeschaffungsmaßnahmen*) have been the most important ALMP programmes in Germany. Their importance decreased as both are long in duration (for up to three years) and expensive. TM are clearly shorter and programme 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 amounted to 538 Euro (Bundesagentur für Arbeit, 2005b).[5]

2.2. Impact of Training Measures on the Search Process

Choosing a suitable outcome variable to measure programme 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 process on part of the employment agency as well as the self-contained job-search. Second, programmes 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, that is they should reduce the unemployment duration. For a precise discussion of the impacts of TM on the unemployment duration, consideration of a formal theoretical model seems reasonable. To do so, we embed our discussion in the standard search model proposed by Mortensen (1986).

The prototype model explains the search behaviour of unemployed individuals in terms of an optimal stopping problem in a dynamic and uncertain environment.[6] 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 characterised 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 characterised by the cumulative distribution function $F(w)$ for $0 < w < \infty$. Under these circumstances, 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

[4] Since 1998, the legal basis for ALMP in Germany was amended twice. In 2002, new instruments and a more ‘activating’ labour market policy were introduced; from 2004 onwards the four laws ‘modern services on the labour market’ have been enacted to reach the goals of Lisbon treaty from March 2000.

[5] In comparison, the spending of the FEA for vocational training programmes (job creation schemes) amounted to 3,616 (1,212) million Euro in 2004.

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

w forever.[7] Otherwise, the search process continues if he rejects the offer. 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).

Unemployed individuals aim at maximising their expected present income over an infinite horizon (van den Berg, 2001). Wage offers arrive at random intervals following a Poisson-process with arrival rate λ. During the period of search, unemployed job-seekers receive unemployment benefits b net of search cost a per unit time. The subjective rate of discount is denoted with r. 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 time independent. Let U denote the expected present value of search. With the stationarity and infinite horizon assumption, the optimal strategy is given by the following asset equation [see Mortensen (1986)],

$$rU = b - a + \lambda \int_0^\infty \max\left\{\frac{w}{r} - U, 0\right\} dF(w). \tag{1}$$

Eq. (1) prices the asset value of search by requiring that the opportunity cost of holding it rU are equal to the current income flow $(b - a)$ plus the expected capital gain if a wage offer arrives (Mortensen and Pissarides, 1999). The realised capital gain when an offer arrives depends whether the offer is accepted or rejected. If the offer is accepted the excess value is $\frac{w}{r} - U$, and 0 if it is rejected. The optimal strategy can be characterised by a reservation wage w^*, with $w^* = rU$. Making use of the reservation strategy we can rewrite the asset eq. (1) as

$$(\lambda + r)w^* = \lambda E(w) + \lambda \int_0^{w^*} F(w)dw + r(b - a). \tag{2}$$

Eq. (2) implies that the reservation wage increases with the unemployment benefits and the interest rate, but decreases with the cost of search. Furthermore, it shows that the reservation wage depends on the offer arrival rate and the distribution of the wage offers. The distribution of wage offers summarises the employment opportunities given job availability, and job availability is indicated by the offer arrival rate (Mortensen, 1986).

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:

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

Eq. (3) shows that the hazard rate increases with the offer arrival λ rate and decreases with the reservation wage w^*. Under the stationarity assumption, the hazard rate is constant over time which is not reasonable for the empirical analysis. In particular, analysing 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 generalises to a time dependent hazard $\theta(t) = \lambda(t)[1-F(w^*; t)]$, see van den Berg (2001). However, an important question when considering policy changes is whether these changes are anticipated by individuals. Individuals anticipating a future participation will adjust their optimal search strategy at the point 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 the shift increases

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

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.

Having introduced a simple search model framework, the question arises how participation in TM affects the duration of unemployment. According to the institutional set-up of TM, we expect two effects to arise. First, TM attempt to improve the search effectiveness of the participants either by enhancing the placement process on part of the employment agency or by enhancing the self-contained job-search. Efficiency of job search may be increased by, e.g., aptitude tests that allow caseworkers at the employment agency to offer more suitable jobs. Job-application training, counselling on job-search possibilities and motivational training may increase the activity and efficiency of job search as well. In particular a motivational training can be expected to counteract the discouraged worker effect, and thus to maintain the search activity of the unemployed persons.[8] In the case a participation in a TM increases the search effectiveness, this results in an increase in the number of job-offers that arrive in the small interval δt. Thus, the impact of TM on the search efficiency can be represented by a change of the offer arrival rate λ.

The impact of an increased arrival rate on the unemployment duration is given by

$$\frac{\partial \theta}{\partial \lambda} = [1 - F(w^*)] - \lambda f(w^*) \frac{\partial w^*}{\partial \lambda}. \tag{4}$$

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. (2) 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 positiv 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 programme. An overall positive effect on the search efficiency therefore requires that a positive after-programme effect dominates a negative locking-in effect.

The second objective of TM is adjusting the qualification of the job-seekers by enhancing job-relevant skills and techniques (e.g. computer and language courses). Such increase of the skills of the participant is equivalent to an increased productivity and allows him or her to apply for jobs that are associated with higher wages on average. Therefor, we assume that participation in TM shifts the mean of the wage offers distribution $F(w)$ to a higher level in the following. According to 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

$$\lim_{\mu \to 0} \{[G(w) - F(w)]/\mu\} = \lim_{\mu \to 0} \{[G(w) - G(w + \mu)]/\mu\} = -f(w), \tag{5}$$

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. (2) associated with the translation we get

$$(\lambda + r)w^*(\mu) = \lambda\mu + \lambda E_F(w) + \lambda \int_0^{w^*(\mu)} F(w - \mu)dw + r(b - a), \tag{6}$$

[8] For or more detailed discussion see Calmfors, Forslund, and Hemström (2001).

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. (3):

$$\frac{\partial\theta(\mu)}{\partial\mu} = \lambda\left\{f[w^*(\mu)-\mu]\left[1-\frac{\partial w^*(\mu)}{\partial\mu}\right]\right\} > 0. \tag{7}$$

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.

3. Econometric Model

In the structural model of the preceding section, we have argued that a participation may affect the job offer arrival rate or the wage offer distribution. Furthermore, we have shown that both effects influence the hazard rate into employment which is directly associated with the expected unemployment duration. However, the data (see below) provide information on the individual unemployment duration only and the empirical analysis is restricted to a reduced form approach estimating the composite effect on the hazard rate into employment. Hence, in contrast to our theoretical discussion, the empirical analysis cannot distinguish between the effect on the offer arrival rate and the effect on the wage offer distribution. A structural analysis primarily fails due to the lack of information about the wage distribution.

The outcome of interest in the empirical analysis is the duration of unemployment until the first transition into employment.[9] The evaluation of the impact of TM on the transition into employment is done with a bivariate duration model as suggested by Abbring and van den Berg (2003). We normalise the point in time when an individual enters unemployment to zero and 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 realisations denoted as t_e and t_p. We consider the population of inflows into unemployment and the conditional distribution functions defined below are associated to this population. The durations T_e and T_p are assumed to vary with time-invariant observable characteristics (x) and time-invariant unobservable characteristics (v). The observable characteristics (x) are assumed to be the same for both distributions, i.e., no exclusion restrictions on x are imposed. With respect to the unobserved covatiates we assume that v is an $\mathbb{R}^2_+$-valued random vector (v_e, v_p) with distribution function $G(v_e, v_p)$ independent of x. It is further assumed that $T_e \perp v_p|t_p, x, v_e$ and $T_p \perp v_e|x, v_p$, i.e., 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, v) stems from causal effects 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 hazard rates (Abbring and van den Berg, 2004). Both hazard rates are specified as mixed proportional hazard (MPH) models,

$$\theta_e(t|t_p, x, v_e) \quad = \quad \lambda_e(t)\exp(x'\beta_e)v_e\mu(t-t_p, x)^{I(t>t_p)}, \tag{8}$$

[9] We define employment as all employment compulsory to social insurance, but without further subsidies.

$$\theta_p(t|x, v_p) \quad = \quad \lambda_p(t)\exp(x'\beta_p)v_p. \tag{9}$$

The hazard rate for the transition into employment (eq. 8) 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. 9) 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., $\mu(t - t_p, x)^{I(t>t_p)} = \mu_1^{I(t_p<t\leqslant t_p+c)}\mu_2^{I(t<t_p+c)}$, 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 allows to analyse 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 programme effect. Moreover, programme effects may also differ by individual characteristics, i.e., programmes are more effective for some subgroups of the labour 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. 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. When this assumption holds, the conditional durations $T_e|x, v_e$ and $T_p|x, v_p$ are only dependent through the term $\exp[\mu(t - t_p, x)I(t > t_p)]$. Therefore, this parameter can be given a causal interpretation as the treatment effect (Abbring and van den Berg, 2003). It is useful to mention that if selectivity results from a dependence of the unobserved heterogeneity terms, the indicator function for the treatment effect can be interpreted as an endogenous time-varying regressor.

An important advantage 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 a time-invariant selection effect embodied by a dependence between v_e and v_p, 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 realised 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 programme. E.g., a positive selection effect results in a pattern where a quick transition into programme 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 realised whereas selectivity affects the transition rate everywhere. 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 programme. 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 $t \leqslant \min\{t_p, t_p^*\}$.[10] 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. We could not identify the moment individuals are informed about a future treatment in the data. However, the duration between informing the participant and the actual starting date is short, and we rule out anticipatory effects of TM. 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 behaviour to the determinants of the treatment process, but not to the realisations 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. (8) and (9) 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

$$\ell_e(t|t_p, x, v_e) = f_e(t|t_p, x, v_e)^{\delta_e} \exp[-\int_0^t \theta_e(u|t_p, x, v_e)du]^{1-\delta_e}, \tag{10}$$

$$\ell_p(t|x, v_p) = f_p(t|x, v_p)^{\delta_p} \exp[-\int_0^t \theta_p(u|x, v_p)du]^{1-\delta_p}. \tag{11}$$

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)]

$$\ell_{e,p}(t|x) = \int_0^\infty \int_0^\infty \ell_e(t|t_p, x, v_e)\ell_p(t|x, v_p)dG(v_e, v_p). \tag{12}$$

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

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

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 utilise a logistic specification for the probability, and the four probabilities are

$$\pi_{j,k} = \frac{q_{j,k}}{\sum_{m=1}^{2} \sum_{n=1}^{2} q_{m,n}}, \tag{13}$$

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 labour market status are observed until December 2003. The data were merged from several datasets 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. The characteristics included cover information on the sociodemographic background of the individuals (e.g., age, marital status, gender), qualification details and placement restraints (e.g., schooling or health restrictions), a short labour market history (e.g., duration of last job before unemployment, number of placement propositions by the caseworker) and the date of entry into unemployment. The majority of characteristics included in BewA are objective attributes, but there are also some subjective ones, like the assessment of the individual's qualification by the responsible caseworker (*level of qualification*).

Additional information on programmes is derived from an excerpt of the programme participants' master data set (*Maßnahme-Teilnehmer-Grunddatei*, MTG). This dataset consolidates details on all ALMP programmes funded by FEA. These data allow us to identify episodes of participation in TM and other ALMP programmes. Unfortunately, we cannot distinguish between the different modules of TM (see section 2.1). Hence, we analyse the effect of TM as a whole.

The outcome of interest (transition into employment) is extracted from the employment statistics register (*Beschäftigtenstatistik*, BSt). The BSt incloses all persons who are registered in the German social security system proving the individual pension claims. These are all persons employed compulsory to social security.[11] Since several wage subsidy programmes are included, we merge this information with MTG to identify all spells of employment and programmes in the observation period. For the employment periods we observe the associated record dates (usually at the end of the month) and for the programme 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. We have to mention that 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. For that reason, labour force movements as well as episodes of employment not subject to social security are not identified in the

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

data. If an individual joins an alternative ALMP programme before he/she becomes employed, we consider the unemployment spell to be censored at the point in time when this transition occurs. In addition, both durations are censored if no transition within the observation window can be observed. It has to be noted that no job-to-job transitions are considered in the empirical analysis, since the available data cover transitions from unemployment into employment only.

The initial sample contains 76,697 individuals with 23,630 individuals of June, 31,217 individuals of August and 21,850 individuals of October.[12] From this sample, we exclude all individuals who either joined alternative ALMP programmes in the period from January 2000 up to their unemployment entry or exhibited failures in the data. 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. 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 programme 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 rate and survivor function for the transition into employment and the transition into programme. 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 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 significantly lower than the hazard rate for the transition into employment at all points of time. 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 programme. 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 programme are 180, 360, 540, 720 and 900 days, i.e., intervals last for six months.

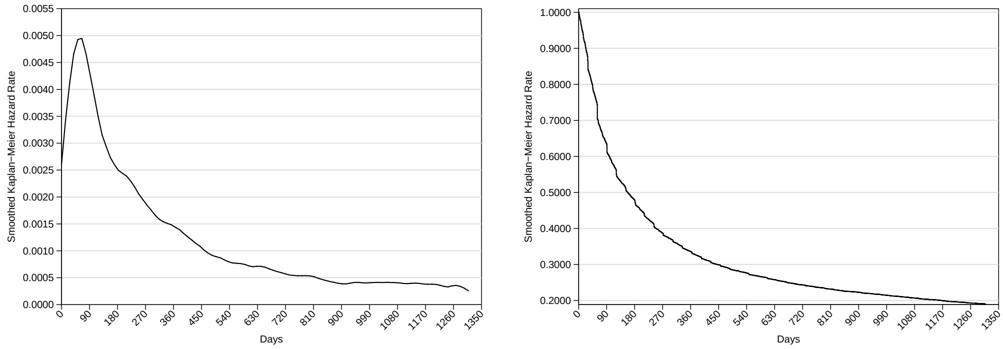
Table 2 presents means and frequencies of the observable covariates used in the analysis to highlight equalities and differences. As mentioned above, Kurtz (2003) points out that important determinants for the decisions of caseworkers to promote job-seekers by TM are the placement chances after participation, 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

[12] We take account of 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

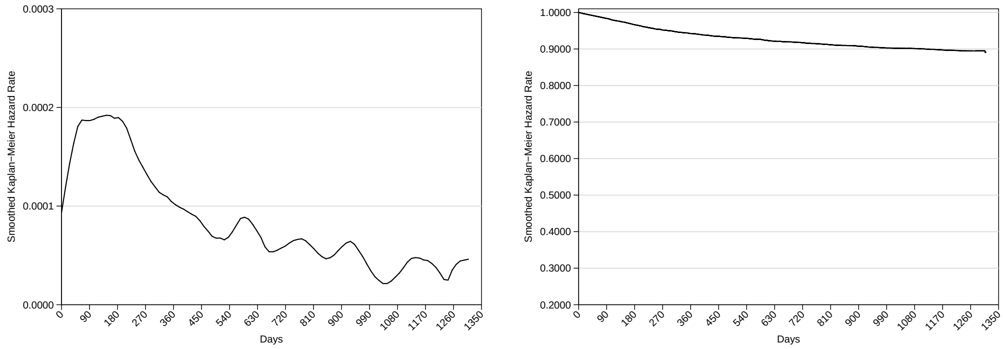
Transition into Employment

Smoothed Hazard Rate[a] | Survivor Function



Transition into Programme

Smoothed Hazard Rate[a] | Survivor Function



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

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.

In addition, labour market performance depends on the lifecycle-position of the individuals. To characterise its influence, we consider *age*, but also *gender* (women), *marital status* and the *number of children* of the job-seeker. Moreover, we incorporate the labour market attachment and occupational group of the individual by using information on *application for full time job* 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. For

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 non-employment 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.

the variables covering lifecycle-position, labour market attachment and occupational group of the individuals as well, the figures in table 2 show that participants and non-participants are very similar. 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

Let us start the discussion of the effects of TM with the results of the basic model (table 3). In this model, the treatment effect is specified as a constant and permanent shift of the hazard rate. Our main interest is in

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

Variable	Transition Rate into Employment Coeff.	t-Value	Transition Rate into Training-Programme Coeff.	t-Value
Baseline Hazard				
$\lambda_{90\geqslant Y<180}$; $\lambda_{180\geqslant S<360}$	0.3292	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	-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*).

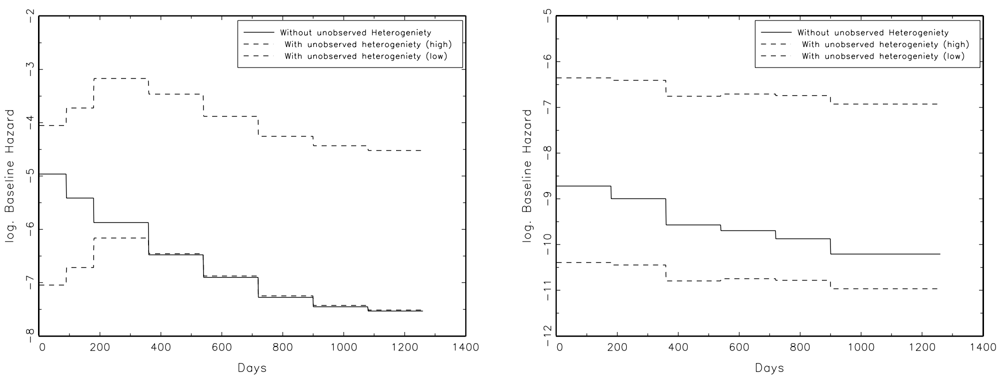
parameter μ, i.e., the causal impact of participation in a TM on the hazard rate into employment. We find a clear positive treatment effect of $\exp(0.3915) = 1.48$ which could be interpreted in the following way: At the point an individual enters a TM, the hazard rate into employment is shifted by 1.48. 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 *number of children*, and corresponding to the reference group with *single parent* and *married*. These variables describe responsibility for closely related persons, who are apparently be more willing to seek actively for a job. Moreover, women do find jobs faster than men and *skilled employee's* as well as persons owning a *CSE* leave unemployment more quickly than unskilled persons. Persons who seek for jobs in *manufacturing industry, technical occupations* or in *service professions* do also have a higher transition rate into employment than the reference groups. In contrast to that, other characteristics reduce the search efficiency of the individuals, like *age* or being an *applicant for a full time job*. From the dummies for the entry dates into unemployment, seasonal differences become obvious. Persons who became unemployed in July 2000, i.e., before and at the start of summer, have a higher hazard rate than persons who became unemployed in August. Persons becoming unemployed in October have the lowest hazard rate.

For the transition rate into programmes, the influence of the observable covariates is not as clear. The majority of the estimates do not provide a reasonable guidance for the selection process due to statistical insignificance. However, it could be established that persons with *children* are considered for participation in a TM earlier. Moreover, persons with an *off-the-job training* are favoured compared to persons without vocational education. The same result could be established for persons owning an *O-level* degree (compared to persons without graduation). Another interesting finding is that persons who apply for a job in a *technical occupation* are privileged. One reason may be the provision of skills and techniques within the TM. Finally, the dummies for the unemployment entry show that persons who became unemployed in August 2000 have increased participation chances.

FIG. 2: ESTIMATED BASELINE HAZARDS FOR WEST GERMANY

Transition into Employment | Transition into Programme



To test the sensitivity of our results with respect to the unobserved heterogeneity distribution, we have estimated a model that only accounts for selection on observable's (see table A.1 in the appendix for estimation results). This model only imposes one point of support for the constant term. The estimated treatment effect is smaller with $\mu = \exp(0.1881) = 1.21$. 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, we find that the hazard rates into employment as well as into programmes 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).[13] 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 programmes. In the model regarding 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 consider the effect of the treatment on the survivor function and the expected unemployment duration. These effects are comparable to average 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

$$\Delta(t)_{sk} = \bar{F}_e(t|s, x, v_e) - \bar{F}_e(t|k, x, v_e). \tag{14}$$

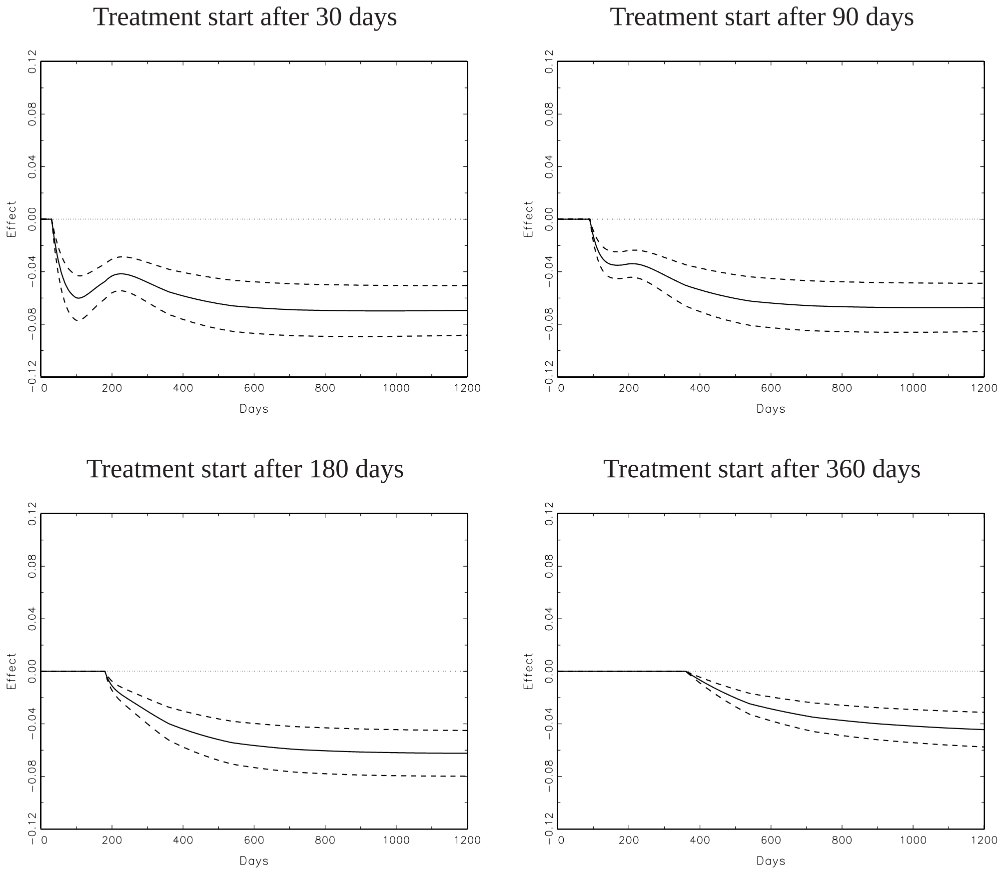
Note, that in this set-up treatments are characterised by the time when they occur. The effect in terms of the survivor function implies a time path of the treatment effect which is determined by the effect of a treatment on the hazard rate. As Fredriksson and Johansson (2004) note, this estimand is 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.,

$$\int_0^\infty \Delta(t)_{sk} dt = E[T_e|s] - E[T_e|k]. \tag{15}$$

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

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

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 the 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 programme 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 programmes are started earlier compared to when started later. 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 programme. However, it is reasonable to expect treatment effects to vary over time. On the one hand, effects may need some time to become effective. This may be the case if participation in a TM is associated with a certificate (e.g., computer course) that would be handed out after the end of the course. Programme effects may also be delayed if participants have to practice their newly received job application advices for some time. On the other hand, programme effects may vanish after a certain amount of time. To give an example: participants of the second module of TM are informed about available jobs they could apply for. However, the information becomes obsolete after some time and the effect ('being informed') decreases.

TAB. 4: TIME VARYING TREATMENT EFFECT

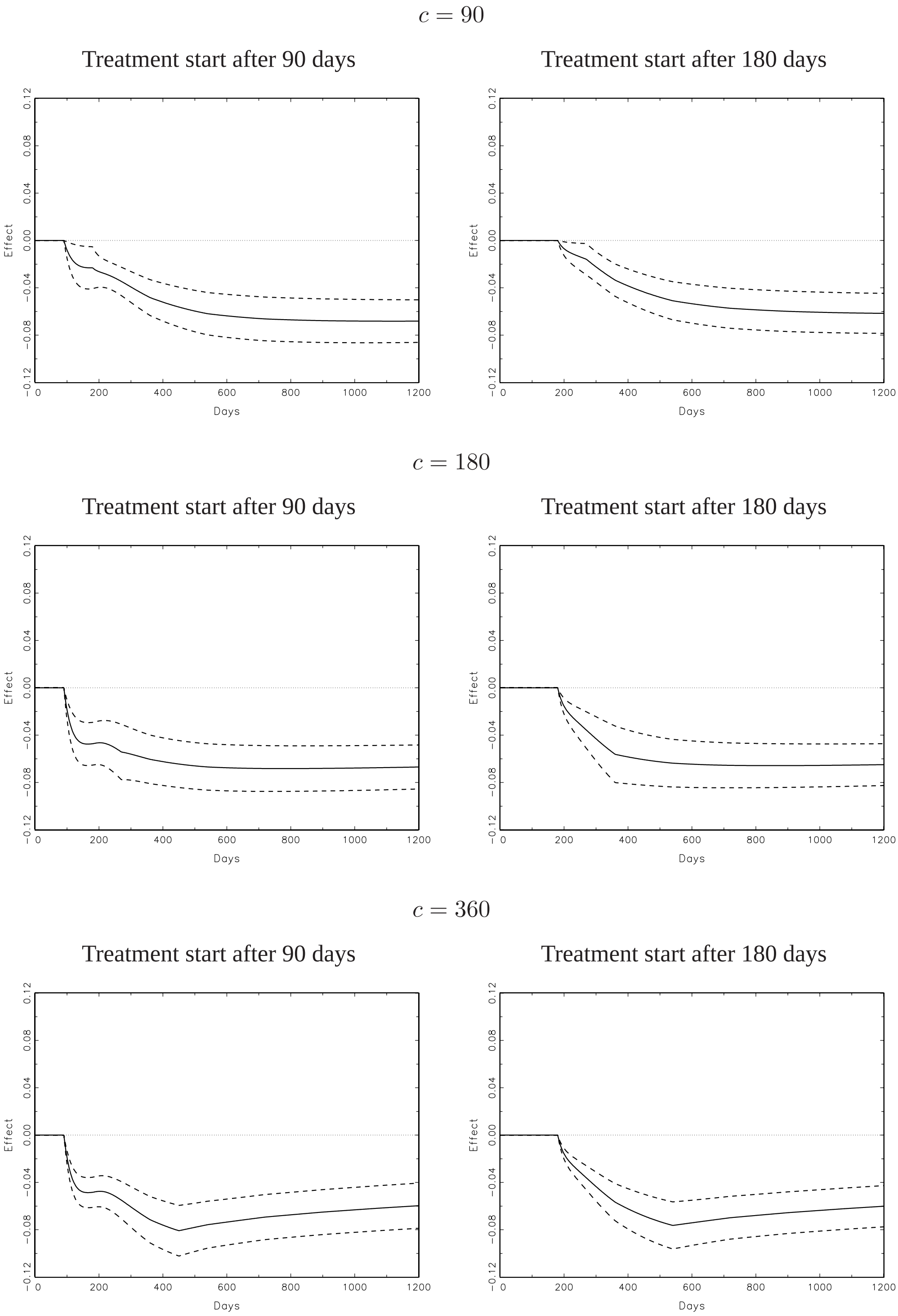
	$c = 90$		$c = 180$		$c = 360$	
Effect	Coeff.	t-Value	Coeff.	t-Value	Coeff.	t-Value
μ_1	0.2578	2.50	0.5297	5.37	0.5381	8.03
μ_2	0.4104	7.23	0.3412	5.26	0.1152	1.15
Log-Likelihood	-186,601.20		-186,600.92		-186,595.62	

In order to analyse 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, where c is 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.[14]

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 programm-start and after 6 months the effect starts slightly to decrease. Obviously 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., the first year after programme start with about 71 percent. Hence, as there is no effect of TM afterwards, programme effects have completely vanished one year after start of programmes. 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 programme. The set-up of TM provides necessary skills, techniques but also incentives for job-seekerks to apply for jobs. Apparently, after a certain amount of time negative effects of unemployment, like discouraged worker effects, stigmatisation etc., overlay the positive treatment effects.

[14] 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 the Delta-Method.

In analogy to the basic model we also 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 (figure 4). The pictures show some interesting features of the treatment effects when allowed to vary over time. Under the assumption that the treatment effect shifts after 90 days ($c = 90$), the effects on the survivor function are almost similar to those from the basic model. In contrast, if we assume the treatment effect to change after

180 days post programme start, the picture is clearly different compared to the basic model. In particular during the first 180 days after programme start we find a more pronounced positive effect of TM than in the basic model. Again, we could establish stronger effects if programmes 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 the 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, afterwards it decreases and is almost identical to that of the basic model three years after. These results support the finding from above. The effect on the predicted survivor functions points towards a treatment effect during the first year after programme start only.

The last aspect we want to analyse is whether treatment effects are heterogeneous due to individual characteristics. In particular, we analyse to what extent low qualified men with some work experience are affected by TM. In addition, we compare the effects to groups that differ in single characteristics. Namely, we estimate the effects for low qualified men who lack any work experience, but also for high qualified men with work experience (university or advanced technical college level). At last, we compare the results of men to that of 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, we employ the specifications of baseline hazards, systematic part and unobserved heterogeneity of the basic model and do not report the estimates.[15] 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 above that of the basic model. Unfortunately, for the higher educated and for persons without occupational experience no difference could be found which may be due to the small number of individuals in those groups. 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, as the hazard rate into employment for low qualified, but experienced men (women) shifts by about 62 (36) percent due to participation, TM are clearly successful in improving the search efficiency.

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

This paper provides the first empirical evidence on the impacts of TM on the search process for employment of the participating individuals in Germany. Based on unique data of the FEA we estimate the effects for persons who became unemployed in June, August and October 2000 until December 2003. Recent empirical

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

literature has emphasised the necessity to consider the timing of treatments in the unemployment spell as well as observable and unobservable characteristics for evaluation of programme effects. We take account of these issues by applying a multivariate mixed proportional hazards model. In addition, we extend the basic model for analysis of heterogeneity in the effects. First, instead of assuming the treatment effect to be a constant and permanent shift of the hazard rate into employment we allow programme impacts to vary over time, i.e., we explicitly regard the possibility of programme effects to develop or degenerate over time. Second, we consider differences in the effects due to individual characteristics.

Since TM are the most important intervention of German ALMP in terms of the number of individuals promoted, our results are of significant political importance. The estimates show that TM clearly reduce the time individuals search for employment in West Germany, i.e., they 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 programmes and are particulary successful in the short- to mid-run.

In addition, the results of the extended model for treatment effects that vary over time indicate that programme effects are strongest during months 3 to 6 after the begin of TM and decrease afterwards. More than 12 months after participation, programme effects have vanished completely. The third step of the analysis (heterogeneity due to individual characteristics) provides gender differences in the effectiveness. Although low qualified persons with some work experience benefit from programmes, the impacts are larger for men than for women. In summary, our results show that TM are successful in reducing the unemployment duration of participating individuals and improve the employment chances of job-seekers clearly.

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.

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.

CALMFORS, L., A. FORSLUND, AND M. HEMSTRÖM (2001): “Does Active Labour Market Policy Work? Lessons From The Swedish Experience,” *Swedish Economic Policy Review*, 85, 61–124.

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

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.

HUJER, R., S. L. THOMSEN, AND C. ZEISS (2006): “The Effects of Vocational Training Programmes on the Duration of Unemployment in Eastern Germany,” *Allgemeines Statistisches Archiv*, forthcoming.

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.

LAYARD, R., S. NICKELL, AND R. JACKMAN (1991): *Unemployment - Macroeconomic Performance and the Labour Market*. Oxford University Press, New York.

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 onthe individual transition rate from unemployment to work,” *Swedish Economic Policy Review*, 8(2), 175–213.

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.

WUNSCH, C. (2005): “Labour Market Policy in Germany: Institutions, Instruments and Reforms since Unification,” Working Paper, SIAW.

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	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*).