The Impact of Labor Market Reform on the Effectiveness of Training Programs in Germany*

Preliminary version

Marton Csillag (University Toulouse)
Hilmar Schneider (IZA, DIW Berlin)
Arne Uhlendorff (DIW Berlin, IZA)
Zhong Zhao (IZA)

February, 2006

Abstract

In 2003, Germany reformed its active labor market policy. With respect to public sector sponsored training for unemployed individuals, the reform mainly consists of the introduction of a voucher system for program participants, a stricter selection of participants, and a matching process between program types and participants by the case workers. The aim of this reform is to increase the competition on the supply side of the "training market" and thereby to increase the quality of the training programs as well as to achieve a better match of participants and programs. This paper analyses the employment effects of public sector sponsored training in Germany for unemployed individuals before and after the reform. Using a rich administrative data set from 2000 to 2004, we apply propensity score matching methods and duration model to estimate the average treatment effects on the treated and lock-in effect of the trainings. We are able to distinguish among treatment effects of six different program types. The results show strong lock-in effects for all program types before and after the reform. For some of the programs we observe slightly but not sustainable positive effects on the employment probability. The treatment effects are heterogeneous and vary with the program type as well as with the previous unemployment duration. The reform seems to improve the effectiveness of the training programs.

JEL-Classification: J64, J68, H43

Keywords: Training programs, Voucher, Propensity Score Matching, Hartz Reform

* *Corresponding author*: Hilmar Schneider, Institute for the Study of Labor (IZA), P.O. Box 240, D53072Bonn, Germany, schneider@iza.org.

1 Introduction

In 2003, Germany reformed its active labor market policy (ALMP). With respect to public sector sponsored training for unemployed individuals, the reform mainly consists of the introduction of a voucher system for program participants, a stricter selection of participants, and a matching process between program types and participants by the case workers, based on the expected reemployment probability of the participants. The aim of this reform is to increase the competition on the supply side of the "training market" and thereby to increase the quality of the training programs as well as to achieve a better match of participants and programs. This paper analyses the employment effects of public sector sponsored training in Germany for unemployed individuals before and after the reform.

There exist several studies on the effectiveness of publicly financed training programs in Germany. Depending on the method, the investigation period and the underlying data set, either negative, non-significant or positive results have been found. Examples for insignificant or even negative effects are Lechner (1998, 1999, 2000), Hujer and Wellner (2000), Schneider et al. (2000) and Hujer, Thomsen and Zeiss (2004). Papers that find inconclusive results are Hübler (1997) and Kraus, Puhani and Steiner (1999). Papers for positive findings are Fitzenberger and Prey (1998, 2000) and Lechner et al. (2005a, 2005b). Summarizing the literature it can be stated that positive effects mainly occur, if at all, in the long run and that studies which find long-term positive effects are also reporting negative short-term effects.

Using a rich administrative data set from 2000 to 2004, we apply non-parametric estimation methods, i.e. apply propensity score matching methods, to estimate employment and unemployment probability differences between participants and non-participants over the time. Additionally we apply parametric duration methods to study lock-in effect. The proportional hazard rate models need several parametric assumptions and are therefore more restrictive than the non-parametric estimators, but the results from them give us a more detailed insight in the processes on the labor market, i.e. the dynamics of the job finding processes of the participants. This allows us to differentiate between lock-in effect and the treatment effect after the program, which is especially useful for long programs and relatively short observation periods.

In contrast to other studies we distinguish between six different program types and between five regional types. Moreover we compare the pre-reform period to the post-reform period before and after January 2003.

The results show strong lock-in effects for all program types before and after the reform. For some of the programs we observe slightly but not sustainable positive effects on the employment probability. The treatment effects are heterogeneous and vary with the program type as well as with the previous unemployment duration. The reform seems to have a positive impact on the treatment effects, and improve the effectiveness of the training program.

The rest of the paper is as follows. Section 2 of this paper gives a short description of the labor market reforms in Germany. Section 3 provides information on the data. Section 4 presents the econometric methods. Section 5 discusses the results and section 6 concludes.

2 Reforms of Further Vocational Training in Germany

A sweeping reform of further vocational training schemes was introduced in Germany in 2003 in order to encourage the (re-)integration of the unemployed into the labor market. Different instruments were set up as new devices to achieve this goal in a more effective way: (*i*) the introduction of a compulsory education voucher to match those who are entitled to participate and the training providers, also to enhance the quality of continuing training by introducing competition among the providers; (*ii*) a reinforced quality management by certifying the training institutions and their respective FbW-programs; (*iii*) a modification of accessing rules.

The changes conducted in the course of the reforms concern access to measures as well as benefit payments came into force in staggered stages starting in the beginning of 2003.[1] However, the largest cut was made in early 2003 in the wake of the reorganization of access procedures to further vocational training measures, brought about by the statutory introduction of education vouchers and the introduction of a new quality management in the course of the administration reform of the Federal Employment Agency (FEA).

(*i*) Education Vouchers

All eligible schemes of further vocational training were to be maintained after the reform. Nevertheless, former access-rules were abolished. According to the new regulations which have been in force since January 2003, job-seekers are issued education vouchers, and they choose an FbW-course provider themselves. The chosen provider will then directly

1 The so called "Job-AQTIVE" law came into force already on January 1, 2002 and concerned regulations of maintenance allowances of further vocational training. The practice of setting-off residual unemployment benefit claims against maintenance allowance continuity payments was introduced then.

charge the FEA for the training fees. Before staring the training, the selected educational institution is needed to present the education voucher to the FEA. What kind of program-scheme is selected depends on the decision of the FEA case-worker. However, the decision is a result of a profiling-process where the potential candidate has to participate. The maximum duration of the further vocational training measure, the educational target and its main focus are documented by the case-worker. Thus, the former twofold relationship where the FEA case-worker sent the unemployed individual directly to a specific provider without allowing for any active contribution of the unemployed – is now substituted by a triangular approach, which the preferences of the unemployed with regard to the FbW-provider play additional role.

However, in the first two months of 2003 a *de facto* double tracked admission practice for further vocational training measures had been established due to the seamless introduction of education vouchers. Further vocational training measures which were approved in 2002 according to the former allocation practice as well as further vocational training measures allocated by education voucher coexisted during this early period. Therefore, the exclusive allocation of further vocational training measures via education vouchers has started seriously from March 2003.

(*ii*) Quality Management

The goal of quality enhancement of the schemes should not solely be achieved by the unemployed who make use of vouchers. Thus, an additional quality management was set up. Since 2003 only those measures that have an overall continuance forecast of at least 70 percent for participants can be approved for providing training. In this case this is the time-related continuance rate of prior training measures (percentage of graduates whose unemployment spell ended within six months after the measure). Furthermore, the development of the regional labor market and expected labor demand are relevant for the rating of a measure. The measure-related continuance rate is also an indicator for the quality control of measures and educational institutions by the FEA, which is incorporated in the Agency's annual training target schedule.

The training target schedule of the Labor Agencies, which was first introduced for the year 2004 on October 31, 2003, will record the results of the quality check of measures and quality check of training providers conducted by the Labor Agencies once a year. The training target schedule is thus to facilitate the institutionalization of the above mentioned 70 percent regulations.

After the “Approval and Admission Ordinance of Continuing Training” went into effect on July, 1 2004, a third party certification agency has to certify that educational institutions and courses of further vocational training comply with legal requirements. Employment offices fulfilled this task themselves before. The ordinance postulates that the transition to the new certification and approval procedure will be finished by the end of 2005. This ensures that providers of continuing training and certifiers have reasonable time to prepare for the new procedure. The labor agencies pay maintenance allowances and training fees for further vocational training only if the third party certification agency has verified that the educational institution and its training courses are followed legal requirements. The independent, private third party certification agencies may award an educational institution a certificate after a successful assessment of the institution and its training courses, which may include an on site inspection. The ordinance governs details of the certification process and particularly substantiates the requirements in terms of quality of the educational institution and its training offers. Moreover, it facilitates and speeds up the approval for training courses of educational institutions which have quality management at their disposal. The qualification and independence of the certification agencies must be checked and confirmed by an approval department with the FEA. The approval department is supported by an advisory board, which may express recommendations on the approval and certification process.

(*iii*) Modification of Accessing Rules

The regulations regarding quality management also could be described as a kind of detailed approach of selecting the FbW-schemes and their providers.

The 70 percent regulations refer to both a subjective assessment of the re-employed probability of unemployed individual and a measure-related continuance rate. First, the assessment of the individual must be that it is “very likely” that a participant will be re-employed after the completion of a specific measure. Although a rate of 70 percent is often mentioned in this context, such a rule is usually not being observed in daily practice since no empirical experiences with the assessment of the individual case exist. Instead, the responsible official draws on the qualifications of the participants, the conditions of the local labor market, and the characteristics of further training, and gives a subjective assessment based on his own professional experience. The assessment of the re-employment probability is conducted in a compulsory counseling interview with an employment counselor or job placement officer in which information gathered during profiling may also be used. An examination by the medical or psychological service of the FEA may ensue after the interview.

3 Data

The data of our analysis are drawn randomly from the universe of administrative records, the so-called "Integrated Employment Biographies" (IEB). This rich data set is a combination of individual records from the program participants' master data set (MTG), the employees' history (BeH), the benefit recipients' history (LeH), and the job seekers' data base (ASU/BewA).

The MTG contains basic information about participation in ALPMs (including further vocational training) as well as important individual characteristics. Entries into ALMPs are identified from January 2000 on. The BeH comprises remuneration notifications of employers about employment subject to social contributions. This information is included in the IEB from 1990 on, but is incorporated in the IEB with a time lag of about 18 months. The LeH has information about phases of benefit receipt from 1990 and onwards. Those benefits mainly include unemployment benefits, unemployment assistance, and maintenance allowances.

The IEB easily allows identifying specific groups of unemployed individuals. For instance, one can construct subgroups of participants in different further vocational training measures. The random sample of the IEB has been drawn conditional both on quarterly stratification with regard to participation in six different types of further vocational training and on a stratification related to previous duration of unemployment, sex and information on the region. Hence, individuals had to be unemployed prior to participation in the further vocational training measure and information on the labor market region, on sex as well as on the type of further vocational training had to be available. Additionally, an age restriction was imposed (17 to 65 years of age). Moreover, every participant could only be drawn once per quarter. Thus, multiple FbW-program entries of a person within a given quarter are not accounted for. According to these rules, only a marginal number of cases were excluded. Therefore, this IEB-random sample can be considered as representative.

With an annual sample of 1,100 participants per FbW-program type, the size of the random sample of the IEB was selected to be appropriately large enough to account for heterogeneities with respect to program types, individual characteristics of participants, and labor market regions. Furthermore, as the observation window already starts in January 2000, the staggered stages of the reforms of further vocational training were adequately covered. In sum, our data comprises information about 29,700 entrants into six different types of further vocational training for 18 quarters from January 2000 until June 2004.

In order to apply the matching approach as described below, 80 non-participants were drawn per participant. In light of lessons learned from Heckman et al. (1997, 1998), those individuals had to be in the same labor market status as the corresponding participant prior to program entry, i.e. normally they had to be previously unemployed for the same duration. Note that unemployment is defined here as being *not* regularly employed.[2] Furthermore, non-participants are required to not having participated in the respective further vocational training measure within the given reference quarter.[3] In sum, our data includes information on 2,376,000 non-participants that serve as control group.

Individual records are observed up to June 2004 with respect to unemployment and up the December 2003 with respect to income-earned employment in our data. Therefore we face a substantial problem of right censoring for the post-reform period. Hence, the average observation window is substantially shorter in the post-reform period as compared to the pre-reform phase, especially if we want to evaluate the effects on employment probabilities.

4 Methodology

Our analysis is conducted by comparing employment and unemployment probabilities of participants and matched non-participants. For this purpose we follow the participants and the matched non-participants from the entrance of the programs onwards. Our evaluation strategy involves two steps. First we compare the employment and unemployment probabilities based on a matched sample non-parametrically. The main advantage of this approach is that we do not need any functional form assumptions.

However, in the context of training programs, there often exist strong lock-in effects. For the evaluation of long programs or for evaluation studies with a relatively short observation period this may lead to wrong conclusions if one relies on the static comparison of employment probabilities only. Therefore, in second step, we analyze the transition process of participants and non-participants from unemployment to employment applying parametric duration models. For this approach we use the matched sample of participants and non-participants, which ensures that the program participation can be assumed as exogenous.

2 Hence, according to this definition e.g. also individuals employed in job creation schemes (JCS) are considered as unemployed.
3 We do not rule out that non-participants also enter further vocational training measures later in time.

The duration model approach allows us to analyze whether the transition probabilities changed after the reform. We are able to disentangle the lock-in effect and the treatment effect of the program.

(*i*) Matching Approach

Evaluation generally has to deal with a serious problem if the effects of participating in a specific program should be quantified compared to that what would have been without doing so. This problem naturally arises because it is impossible to observe individuals in two different states of nature (participation and non-participation) at the same time and place. Therefore, it is the principle task of any evaluation study to find a credible estimate for the counterfactual state of nature.

Moreover, participants in active labor market programs are a selection sample of the unemployed in two aspects. They are self-selected and also have been selected by an agent of the FEA into the program under consideration. Therefore, a selection bias is supposed to arise because participants are *not* a random sample of unemployed persons, but may instead differ systematically in important individual characteristics that may be important determinants of labor market outcomes.

When the selection bias is only due to observables, matching is a useful tool to estimate treatment effects and to solve the evaluation problem. Compared to regression type estimators, the most attractive feature of matching is its non-parametric nature. Matching neither imposes functional form restrictions such as linearity nor assumes a homogeneous treatment effect in the population.

Using covariate matching to correct the bias due to observables is intuitive, since the source of the bias is the difference of observables between treatment and comparison group. Matching on covariates by definition removes this difference and hence the bias. Rosenbaum and Rubin (1983) have shown that while covariate is the finest balancing score, propensity score is the coarsest balancing score. Therefore, matching on covariates and matching on the propensity score will both make the distribution of the covariates in the treatment group the same as in the comparison group.

Formally, in the potential outcome framework, each individual has two potential outcomes (Y_{0i}, Y_{1i}) depending on the treatment status (e.g. participation vs. non-participation in further vocational training). Y_{1i} denotes the outcome if individual i is treated, and Y_{0i} is the outcome if individual i is not treated. Besides, let $T_i = 1$ indicate that individual i is treated, and $T_i = 0$ otherwise. With (Y_{0i}, Y_{1i}) we can define different treatment effects, such as those in Heckman and Vytlacil (1999), as follows:

$\Delta_i = Y_{1i} - Y_{0i}$	Treatment effect for individual i
$\Delta_{ATE} = E[\Delta_i]$	Average treatment effect for the population (ATE)
$\Delta_{TT} = E[\Delta_i \mid i : T = 1]$	Average treatment effect on the Treated (TT)

The average treatment effect at the population (or sub-population) level can be estimated without bias by observational data if the selection bias is only due to observables. That the selection bias is in fact only due to observables is formally characterized by the following two assumptions:

$[CIA]$	$(Y_0, Y_1) \perp T \mid X$	Conditional independence assumption
$[CSA]$	$0 < prob(T = 1 \mid X) < 1$	Common support assumption

where $\perp$ is the notation for statistical independence. The *CIA* is also commonly referred to as unconfoundedness assumption or exogeneity assumption.

Hence, it follows under *CIA* and *CSA*:

$$\begin{aligned}\Delta_{TT} &= E_{x|T=1}\{E[Y_1 \mid T = 1, X = x] - E[Y_0 \mid T = 1, X = x]\} \\ &= E_{x|T=1}\{E[Y_1 \mid T = 1, X = x] - E[Y_0 \mid T = 0, X = x]\} \qquad (1)\end{aligned}$$

Unbiased estimates of $E[Y_1 \mid T = 1, X = x]$ and $E[Y_0 \mid T = 0, X = x]$ can be obtained from the data and, thus, Δ_{TT} can also be estimated without bias. The same holds for Δ_{ATE} and Δ_S.

The so called balancing property formally looks as follows:

$$prop(X_i \mid T_i = 1, p(X_i) = p) = prob(X_i \mid T_i = 0, p(X_i) = p) = prob(X_i \mid p)$$

Using this property, Rosenbaum and Rubin (1983) have shown that *CIA* and *CSA* imply

$[CIA']$	$(Y_0, Y_1) \perp T \mid p(X)$
$[CSA']$	$0 < prob(T = 1 \mid p(X)) < 1$

It follows from *CIA'* and *CSA'*:

$$\begin{aligned}\Delta_{TT} &= E_{p|T=1}\{E[Y_1 \mid T = 1, p(X) = p] - E[Y_0 \mid T = 1, p(X) = p]\} \\ &= E_{p|T=1}\{E[Y_1 \mid T = 1, p(X) = p] - E[Y_0 \mid T = 0, p(X) = p]\} \qquad (2)\end{aligned}$$

Unbiased estimates of $E[Y_1 \mid T=1, p(X)=p]$ and $E[Y_0 \mid T=0, p(X)=p]$ can be obtained if $p(X)$ is known. The advantage of equation (2) over equation (1) is that instead of controlling for a high-dimensional vector X, formula (2) only needs to control for the scalar p.

The rich information from the administrative data ensures the *CIA* assumption is reasonable. The ratio between the number of treatment and control group observations is 1 to 80. The large pool of control observations ensures the quality of matched sample.

This approach combines exact matching on a number of covariates and propensity score matching.

First, propensity scores for each of the six types of further vocational training are estimated from the whole sample of participants and non-participants. The specification of propensity score is selected from different specifications using balancing test.[4] Therefore, the specifications of propensity score differ between the types of further vocational training.

After estimating the propensity score, we match one participant with non-participant by exact covariate matching plus propensity score matching. The variables used for exact matching are sex, previous duration of unemployment, region and quarter. Therefore, we stratify the sample by these four variables first, and then implement propensity score matching for each cell without replacement, i.e. after a participant is matched with his or her nearest non-participant, the matched non-participant is no longer part of the matching pool.

(*ii*) Duration Models

The process of leaving unemployment to participate labor market can appropriately be modeled by a transition rate approach. Two potential destination states *q* are considered reflecting transitions to employment (*q=1*) and alternative transitions like for example other labor market policy measure or retirement (*q=2*). The hazard rate is defined as the limit of the conditional probability for the ending of a spell in interval $[t;t+\Delta t]$ given that no transition occurred before the start of this interval:

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{P(t \le T < t+\Delta t \mid T \ge t)}{\Delta t} \qquad (3)$$

where *T* denotes the length of a spell. *T* is assumed to be a continuous, non-negative random variable. We assume proportional transition rates with covariates causing proportional shifts

4 We use psmatch2 ado file of Leuven and Sianesi (2003) to implementation of all propensity score estimations and balancing testing, and use nnmatch ado file of Abadie et al. (2004) to carry out the matching.

of a so-called baseline transition rate and interval constant covariates. The hazard rate $\lambda(t|x(t))$ corresponds to the sum of the two transition rates

$$\lambda(t|x(t)) = \sum_{q=1}^{2} \lambda_q(t|x(t))$$

with the transition probability to destination state q corresponding to

$$\lambda_q(t|x(t)) = \lambda_{0q}(t)\exp(x(t)\beta_q + \eta_q) \text{ with } (\eta_1, \eta_2) \sim N(0, 0, \sigma_1^2, \sigma_2^2, \rho). \qquad (4)$$

where $\lambda_{oq}(t)$ denotes the destination specific baseline transition rate, *x(t)* a time variant row vector of covariates, β_q a column vector of parameters and η_q a time invariant individual and destination specific error term, representing the joint influence of unobserved heterogeneity. We assume these error terms or random intercepts follow bivariate normal distribution with expected values 0, which allows for dependent competing risks.

Discrete-time measurement leads to the simplifying consequence that instead of continuous levels of $\gamma_{oq}(t)$ and $x(t)$ only their interval specific mean levels have to be taken into account. Assumed that the time axis is divided into intervals of unit length, a given spell consists of a number of k intervals, in the following referred to as subspells. The jth subspell covers a range from $t = j$-1 to t + 1, but excluding t + 1. The interval specific means of $\gamma_{oq}(t)$ and $x(t)$ are then denoted as $\gamma_{oq}(j)$ and x_j.

For the survivor function this implies:

$$S(j) = \exp\left(-\sum_{q=1}^{2}\sum_{k=1}^{j} \exp(x_{qk}\beta_k + \gamma_{qk} + \eta_q)\right) \text{ with } \gamma_{qk} = \ln\left(\int_{t_{k-1}}^{t_k} \lambda_{0q}(\tau)d\tau\right) \qquad (5)$$

The survivor function *S(j)* describes the probability that a spell lasts at least j intervals. The γ parameters capture the duration dependence of the baseline transition function. They may be interpreted as an interval specific mean of the baseline transition rate, which is equivalent to an interval specific constant baseline transition rate.

Following from this, the probability f of a transition to state r at a given interval j is given by the difference of two survivor functions multiplied by the share of the risk-specific transition rate at interval j related to the hazard rate at interval j.

$$f_r(j) = \frac{\exp(x_{rj}\beta_r + \gamma_{rj} + \eta_r)}{\sum_{q=1}^{2}\exp(x_{qj}\beta_q + \gamma_{qj} + \eta_q)}\left[S(j-1) - S(j)\right]. \qquad (6)$$

The likelihood contribution of a spell corresponds to[5]

---

[5] The corresponding likelihood is solved by applying Gauss-Hermite quadrature.

$$L(\beta,\gamma,\eta_1,\eta_2) = \frac{\exp(x_{1j}\beta_1 + \gamma_{1j} + \eta_1)^{c_1} \exp(x_{2j}\beta_2 + \gamma_{2j} + \eta_2)^{c_2}}{\sum_{q=1}^{2} \exp(x_{qj}\beta_q + \gamma_{qj} + \eta_q)} \left[cS(j-1) - (2c-1)S(j) \right] \tag{7}$$

whereby $c_1=1$ and $c_2=1$ indicate a transition to risk 1 and to risk 2 in interval j, respectively, and c corresponds to the maximum of c_1 and c_2. It implies that right-censored spells are assumed to be censored at the end of the related interval, but that transitions may occur somewhere between j-1 and j. The likelihood contribution is not separable into destination-specific components as suggested by Narendranathan and Stewart (1993b) because we do not assume that transitions can only occur at the interval boundaries (see Roed and Nordberg 2003 or Jenkins 2004 for similar approaches). Therefore we can not estimate destination specific models separately, even in a model without unobserved heterogeneity.

5 Results

(*i*) Selected Descriptive Statistics for Two Program Types

This part presents descriptive statistics on the whole sample as well as on the matched sample. We present statistics on one short program type, "occupation-related or general training program" (program type 1), and one long program type, "individual-scheme with occupation-related certificate" (program type 5).

Table 1 and Table 5 are descriptive statistics for selected variables of program type 1 and type 5.

The individuals in the program are generally more likely to be in marriage, and they are also more likely to have a kid under age 3 or age 14. For program type 1, 6.5% of participants have kid under age 3, and 24.3% have kid under age 14. But for non participants, these percentages are 4.9% and 19.1%, respectively.

The prior program earnings of participants are higher than the non-participants. The participants are more likely to be a native instead of migrants.

It is clear from Table 1 and Table 2 that there are significant difference of characteristics between participants and non participants. The differences of the variables in Tables are all significant at 5% level. But after propensity score matching, the differences of most variables between participants and non participants are no longer significant at usually level.

Table 1 Descriptive Statistics for Selected Variables of Program Type 1

			Mean			% reduct	Test
Variable	Label	Sample	Treated	Control	%bias	|bias|	p>|t|
female	Female	Unmatched	0.463	0.445	3.500		0.009
		Matched	0.463	0.449	2.700	22.800	0.162
kid03	Have kid under age 3	Unmatched	0.065	0.049	7.300		0.000
		Matched	0.065	0.066	-0.500	93.700	0.836
nkids	Number of kids	Unmatched	0.618	0.529	10.200		0.000
		Matched	0.618	0.617	0.100	99.000	0.943
kid414	Have kid under age 14	Unmatched	0.243	0.191	12.700		0.000
		Matched	0.243	0.240	0.800	93.800	0.703
marst1	Maried	Unmatched	0.411	0.403	1.600		0.239
		Matched	0.411	0.417	-1.200	23.900	0.516
school	Schooling	Unmatched	2.929	2.504	39.300		0.000
		Matched	2.929	2.914	1.400	96.500	0.463
inchalf1	Income in the 1st	Unmatched	23.014	18.772	26.200		0.000
	half-year prior training	Matched	23.014	23.063	-0.300	98.800	0.890
inchalf2	Income in the 1st	Unmatched	29.176	23.473	22.900		0.000
	half-year prior training	Matched	29.176	28.595	2.300	90.000	0.257
inchalf3	Income in the 1st	Unmatched	33.460	25.814	27.300		0.000
	half-year prior training	Matched	33.460	32.380	3.800	85.900	0.065
inchalf4	Income in the 1st	Unmatched	34.760	27.036	26.000		0.000
	half-year prior training	Matched	34.760	33.731	3.500	86.700	0.096
nat1	Native	Unmatched	0.926	0.878	16.000		0.000
		Matched	0.926	0.927	-0.300	98.100	0.843
noincome	No income	Unmatched	0.054	0.090	-13.700		0.000
		Matched	0.054	0.053	0.700	94.900	0.687

[to be completed: more descriptive statistics on other type of programs and more variables]

(*ii*) Selected Results for Two Program Types

In general, our analysis reveals considerable heterogeneous effectiveness of the different types of publicly financed vocational training. But the general dilemma faced by all program types is that the lock-in effect has to be compensated by improvement of employment chances after participation. The lock-in effect consists of two parts: the probability of leaving a program and the duration of the programs. The reform goes along with a shortening of program duration, which has weakened the overall lock-in effects. On the other hand the probability of leaving the program has decreased after the reform, indicated by an increase in the coefficients of the lock-in effects in the estimated hazard rate models for some program types. This indicates that the fit between participants and programs has been improved, i.e. it seems to be more attractive for the participants to finish the training program after the reform. This may be due to a higher quality of the programs or due to a better match between participants and programs.

Table 2 Descriptive Statistics for Selected Variables of Program Type 5

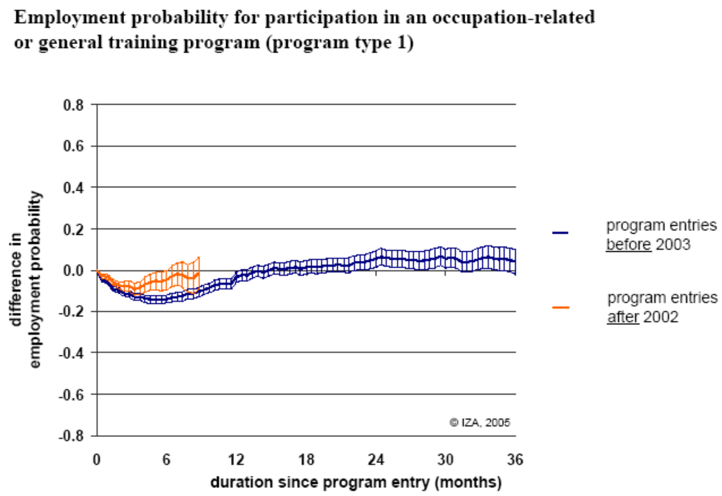
Variable	Label	Sample	Mean Treated	Mean Control	%bias	% reduct |bias|	Test p>|t|
female	Female	Unmatched	0.459	0.442	3.500		0.012
		Matched	0.459	0.458	0.200	94.500	0.913
kid03	Have kid under age 3	Unmatched	0.076	0.051	10.400		0.000
		Matched	0.076	0.071	2.100	80.000	0.328
nkids	Number of kids	Unmatched	0.608	0.522	9.800		0.000
		Matched	0.608	0.595	1.500	85.100	0.466
kid414	Have kid under age 14	Unmatched	0.246	0.191	13.300		0.000
		Matched	0.246	0.243	0.600	95.400	0.770
marst1	Maried	Unmatched	0.451	0.411	8.000		0.000
		Matched	0.451	0.460	-1.800	78.000	0.369
disable	Disability	Unmatched	3.973	3.894	16.300		0.000
		Matched	3.973	3.931	8.500	47.600	0.000
school	Schooling	Unmatched	2.676	2.478	18.700		0.000
		Matched	2.676	2.718	-4.100	78.100	0.039
inchalf1	Income in the 1st	Unmatched	24.499	19.076	32.100		0.000
	half-year prior training	Matched	24.504	24.567	-0.400	98.800	0.857
inchalf2	Income in the 1st	Unmatched	30.635	24.044	26.700		0.000
	half-year prior training	Matched	30.635	30.738	-0.400	98.500	0.849
inchalf3	Income in the 1st	Unmatched	33.907	26.495	26.700		0.000
	half-year prior training	Matched	33.907	33.800	0.400	98.500	0.859
inchalf4	Income in the 1st	Unmatched	34.171	27.761	21.900		0.000
	half-year prior training	Matched	34.171	34.146	0.100	99.500	0.968
nat1	Native	Unmatched	0.872	0.863	2.500		0.080
		Matched	0.872	0.876	-1.300	49.200	0.517
noincome	No income	Unmatched	0.058	0.097	-14.700		0.000
		Matched	0.058	0.058	0.000	100.000	0.998

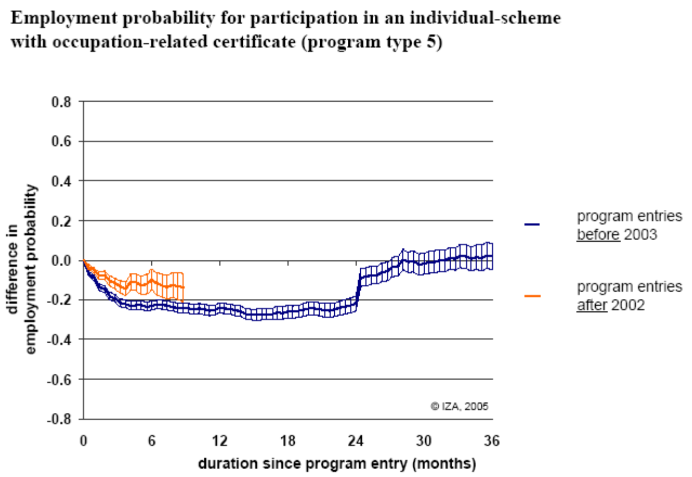
Some of the program types failed to improve employment probability, while others were moderately successful. However, the analogous observation of the probability of unemployment suggests that most measures had no significant positive effect before the reform. The phenomenon of increasing the probability of employment without decreasing the probability of unemployment is explained by the fact that participation in a training measure managed to cause persons, who otherwise would have withdrawn from the labor force, to re-enter into employment.

Due to the reform, the effects of measures concerning the probability of employment as well as of unemployment have noticeably improved for most of the six program types. The reform effect seems to be strongest, when program entries take place between the 4th and 12th month of unemployment.

Figure 1 presents the non-parametric estimation results of one short program type, “occupation-related or general training program” (program type 1), and one long program type, “individual-scheme with occupation-related certificate” (program type 5). The development of the differences between participants' and matched non-participants' probability of employment is being observed, starting from the point of the participant's entry into a measure. For the shorter program before the reform we observe a significantly positive effect after 24 months. The magnitude is about 5 percentage points and the effect seems to be not sustainable. For the long program type we do not observe any positive effect. However, this program type lasts for about two years.

Figure 1: Employment effects for selected program types



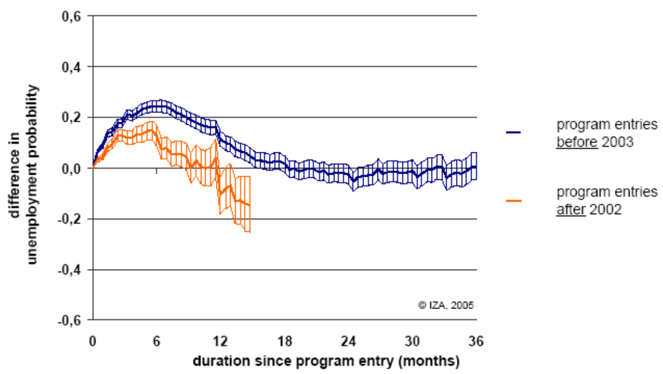


In Figure 2 the corresponding unemployment probabilities are presented. Participation in program type 1 does not decrease the probability of unemployment, although we observe an increased employment probability. Participation in program type 1 seems to increase the employment probability of the individuals who would otherwise leave the labor market, i.e. would not be employed nor unemployed.

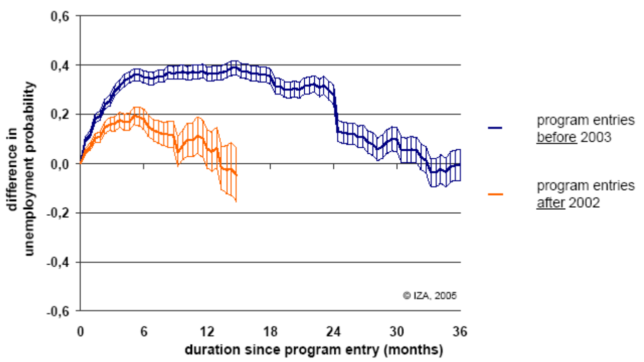
The reform increases the probability of employment for both program types during the first nine months. The unemployment probability is observed for a longer period (15 months) and a total reduction of the unemployment risk by around 10 percentage points is observed for program type 1 after the reform, 15 months after program entrance. For program type 5 we do not observe any positive effect but the reform leads to a clear decrease in the unemployment probability difference between participants and the control group.

Figure 2: Unemployment effects for selected program types

Unemployment probability for participation in an occupation-related or general training program (program type 1)



Unemployment probability for participation in an individual-scheme with occupation-related certificate (program type 5)



[to be completed: further discussion of treatment effects for several subgroups and program types]

6 Conclusion

In 2003, Germany reformed its active labor market policy. With respect to public sector sponsored training for unemployed individuals, the reform mainly consists of the introduction of a voucher system for program participants, a stricter selection of participants, and a matching process between program types and participants by the case workers, based on the expected reemployment probability of the participants. The aim of this reform is to increase the competition on the supply side of the "training market" and thereby to increase the quality of the training programs as well as to achieve a better match of participants and programs. This paper analyses the employment effects of public sector sponsored training in Germany for unemployed individuals before and after the reform.

Using a rich administrative data set for the years 2000 to 2004, we apply propensity score matching methods to estimate the average treatment effects on the treated. In contrast to other studies we distinguish between six different program types and between five regional types. Moreover we compare the pre-reform period to the post-reform period before and after January 2003.

We further analyze the reform by duration models. Our results show strong lock-in effects for all program types before and after the reform. For some of the programs we observe slightly but not sustainable positive effects on the employment probability. The treatment effects are heterogeneous and vary with the program type as well as with the previous unemployment duration. The reform seems to have a positive impact on effectiveness of training programs. In order to address the long-term effects of the reform as well as long-run effects of the different program-types, new data for 2004/2005 will be used in our future studies.

References

Abadie, Albert, David Drukker, Jane Leber Herr, and Guido W. Imbens (2004): Implementing Matching Estimators for Average Treatment Effect in Stata, *The Stata Journal* 4, 290–311.

Fitzenberger, B. and H. Prey (1998): Beschäftigungs- und Verdienstmöglichkeiten von Weiterbildungsteilnehmern im ostdeutschen Transformationsprozess: Eine Methodenkritik. In: Pfeiffer, F. und W. Pohlmeier (Hrsg.): Qualifikation, Weiterbildung und Arbeitsmarkterfolg. ZEW Wirtschaftsanalysen, Bd.31, Baden Baden: Nomos, S. 39-96.

Fitzenberger, B. and H. Prey (2000): Evaluating Public Sector Sponsored Training in East Germany, *Oxford Economic Papers*, 497-520.

Heckman, James J., Hidehiko Ichimura, Jeffrey A. Smith, and Petra E. Todd (1998): Characterizing Selection Bias Using Experimental Dada, *Econometrica* 66, 1017–1098.

Heckman, James J., Hidehiko Ichimura, and Petra E. Todd (1997): Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme, *Review of Economic Studies* 64, 605–654.

Heckman, James J., and Edward J. Vytlacil (1999): "Local Instrumental Variables and Latent Variables Models for Identifying and Bounding Treatment Effects", *Proceedings of the National Academy of Sciences*, 96, 2, pp. 4730 – 4734.

Hübler, O.(1997): Evaluation beschäftigungspolitischer Maßnahmen in Ostdeutschland, *Jahrbücher für Nationalökonomie und Statistik*, 216, 21-44.

Hujer, R. and M. Wellner (2000): Berufliche Weiterbildung und individuelle Arbeitslosigkeitsdauer in West- und Ostdeutschland: Eine mikroökonometrische Analyse, Mitteilungen aus der Arbeitsmart- und Berufsforschung, 33, S.405-419.

Hujer, R., S.L. Thomsen and C. Zeiss (2004): The Effects of Vocational Training Programmes on the Duration of Unemployment in Eastern Germany. mimeo.

Jenkins, Stephen P. (2004): Independent competing risks models for interval-censored survival time data. Unpublished Paper, ISER, University of Essex

Lechner, M. (1998): Training the East German Labour Force. Microeconometric Evaluations of Continuous Vocational Training after Unification. Heidelberg: Physica-Verlag.

Lechner, M. (1999): Earnings and Employment Effects of Continuous Off-the-Job Training in East-Germany after Unification, *Journal of Business Economic Statistics*, 17, 74-90.

Lechner, M. (2000):An Evaluation of Public-Sector-Sponsored Continuous Vocational Training Programs in East Germany, *Journal of Human Resources*, 35, 347-375.

Lechner, M., R. Miquel and C. Wunsch (2005a): The Curse and Blessing of Training the Unemploymed in a Changing Economy. The Case of East Germany After Unification, *IAB Discussion Paper* No.14/2005.

Lechner, M., R. Miquel and C. Wunsch (2005b): Long-Run Effects of Public Sector Sponsored Training in West Germany, *IAB Discussion Paper* No.3/2005.

Leuven, Edwin and Barbara Sianesi (2003), "PSMATCH2: Stata Module to Perform Full Mahalanobis and Propensity Score Matching, Common Support Graphing, and Covariate Imbalance Testing," unpublished manuscript.

Narendranathan, Wiji, and Mark .B. Stewart (1993): Modeling the Probability of Leaving Unemployment: Competing Risks Models with Flexible Base-line Hazards, *Applied Statistics*, 42, 63-83.

Roed, Knut and Morten Nordberg (2003): Temporary layoffs and the duration of unemployment, *Labour Economics,* 10, 381-398.

Rosenbaum, Paul R., and Donald B. Rubin (1983): "The Central Role of the Propensity Score in Observational Studies for Causal Effects", *Biometrika*, 70, pp. 41 – 55.