The Impact of Training on Employment: A Survey of Microeconometric Studies

by

Reinhard Hujer, Kai-Oliver Maurer and Marc Wellner

Department of Economics
Johann Wolfgang Goethe-University, Frankfurt am Main
March 1996

Financial support from the Deutsche Forschungsgemeinschaft (#Hu 225/7-1; "Weiterbildung und Erwerbsverläufe") is gratefully acknowledged.

Reinhard Hujer, Kai-Oliver Maurer and Marc Wellner
Department of Economics
Johann Wolfgang Goethe-University
Mertonstraße 17
D-60054 Frankfurt am Main

Phone: +49 (0)69 798-28115
Fax: +49 (0)69 798-23673

E-Mail: hujer@wiwi.uni-frankfurt.de/K.Maurer@em.uni-frankfurt.de/Wellner@em.uni-frankfurt.de

WWW: http:/www.wiwi.uni-frankfurt.de/professoren/empwifo/

ABSTRACT

This paper provides a review of empirical evidence relating to the impact of training on employment performance. Since a central issue in estimating training effects is the sample selection problem a short theoretical discussion of different evaluation strategies is given. The empirical overview primarily focuses on non-experimental evidence for Germany. In addition selected studies for other countries and experimental investigations are discussed.

I. Introduction

Manpower training programs have always been an essential instrument of active labor market policy in different countries. In order to justify the costs caused by these programs it seems self-evident that an empirical evaluation of their effectiveness is necessary. In principle, this measurement is done by estimating the impact of training on different individual outcomes such as earnings, wages, or employment.

Since our research project focuses on the effect of training on individual employment histories in Germany, the overview of empirical studies that this paper gives, primarily concentrates on corresponding German studies with this particular outcome. Nevertheless also selected studies relating to other countries are included.

A central problem with evaluating training programs is the typically occurring selection bias among the participants. Hence it seems necessary to theoretically outline this issue as well as the evaluation strategies that are proposed to deal with it. This is provided in chapter II. In chapter III and IV we present training studies that use experimental and non-experimental or so-called observational data sets, respectively. However, the fact that these studies vary with respect to data they use, the types of training they evaluate and the evaluation methods they apply, makes a comparison difficult. Chapter V gives an outlook on our own evaluation strategy.

II. The Sample Selection Problem

When we aim to understand how various empirical studies attempt to determine the impact of training programs on various kinds of outcomes such as earnings, wages or our in case employment performance, it is useful to define clearly what is actually meant by impact of training. We try to do this by formulating the question of interest more precisely: *What is the average individual outcome gain of training participants compared to the situation in which they would not have participated in the training program?* Apparently, in order to answer this question one has to contrast the individuals' situation after training with the corresponding situation of the *same* individuals in the hypothetical case of not having participated in training. This points to the central problem of all training evaluations. They have to compare two situations which can never both occur. While one can observe the first situation, the latter is always hypothetical i.e. unobservable. Hence, it has to be estimated based on the information of *other* individuals who did not participate in the training program, members of a so-called control group. Using the formal notation of expectations, the average training effect of interest, denoted by α, is defined as:

$$\alpha := \mathrm{E}\left(\mathrm{Y}^{\mathrm{t}} - \mathrm{Y}^{\mathrm{c}} \middle| \mathrm{D}=1\right) = \mathrm{E}\left(\mathrm{Y}^{\mathrm{t}} \middle| \mathrm{D}=1\right) - \mathrm{E}\left(\mathrm{Y}^{\mathrm{c}} \middle| \mathrm{D}=1\right) \qquad (1),$$

where Y^{t} and Y^{c} describe the potential outcome of interest, namely the employment performance in the situation with and without training, respectively. D stands for a binary assignment indicator determining whether an individual participated in training ($\mathrm{D}=1$) or not ($\mathrm{D}=0$). The problem of not observing the hypothetical situation is now documented easily by the fact, that one only has information to estimate:

$$\alpha^{\mathrm{e}} = \mathrm{E}\left(\mathrm{Y}^{\mathrm{t}} \middle| \mathrm{D}=1\right) - \mathrm{E}\left(\mathrm{Y}^{\mathrm{c}} \middle| \mathrm{D}=0\right) \qquad (2).$$

It is obvious that α^{e} is a potentially biased estimator of the training impact of interest α, as $\mathrm{E}\left(\mathrm{Y}^{\mathrm{c}} \middle| \mathrm{D}=0\right)$ does not necessarily equal $\mathrm{E}\left(\mathrm{Y}^{\mathrm{c}} \middle| \mathrm{D}=1\right)$. This inequality evidently arises if trainees and non-trainees systematically differ in their observable or unobservable individual characteristics. A systematic difference typically appears because the selection or decision to participate in training is a result of different personal attitudes such as motivation, family

background or self confidence. Hence for the case of a negative (positive) selection a simple comparison of trainees and non-trainees will lead to an under-(over-) estimation of training impact since the negative (positive) characteristics not accounted for also influence the individual employment performance. Evaluations based on experimental and on non-experimental or so-called observational data sets, respectively, pursue two different strategies to overcome this *problem of selection bias.*

Experimental evaluations aim to solve the problem of selection bias through the instrument of randomization at the level of data collection. Individuals who are eligible training participants are *randomly* divided into a trainee group, which participates in the training program, and a control group, that does not. As all individuals are suitable for the training program but a control group was randomized out just prior to the beginning of the program, the two groups consist of individuals who should completely equal in all relevant observable and unobservable characteristics, except for the fact of having received training. Thus the difference in the labor market performance is only induced by the training program itself, i.e. the impact of training is isolated and a selection bias does not exist. In formal terms the role of randomization is precisely that the potential outcomes (Y^t and Y^c) are independent of the assignment to the training program (D). Under this circumstance, conditioning on participation in the training program is unnecessary. It follows:

$$E\left(Y^c \middle| D = 1\right) = E\left(Y^c \middle| D = 0\right) \qquad (3),$$

i.e. the hypothetical situation of not participation can be estimated consistently based upon information of the observed controls yielding α^e to be an unbiased estimator of α. Researchers, e.g. ASHENFELDER/CARD (1985), BJÖRKLUND (1989), HAM/LALONDE (1996) sometimes characterize a randomized experiment as the ideal approach and believe that it is the only possible strategy to obtain reliable estimates on the impact of training. Despite these arguments for an experimental approach there are some serious objections to it. First it has to be assured that the members of the control group do not try to participate in a similar program at another occasion. If this condition is not guaranteed a so-called *substitution bias* can occur. Second, especially for the case of estimating the long-term in contrast to the short-term impact of training on employment HAM/LALONDE (1996) demonstrate that experiments can involve another potential selection bias. It arises because *trainees' and controls' subsequent employment and*

unemployment duration (spells) are not generally random subsets of the initial randomly selected trainee and control group. Since the two initial groups have different sorting processes into the subsequent employment and unemployment spells a sample selection bias gets involved when comparing the spell lengths or transition rates of controls and trainees. A further issue in the context of estimating the training impact on unemployment and employment duration which is more difficult to handle when using experimental data compared to observational data is known as *initial condition problem*. It appears because in an experiment typically all observed left censored unemployment and employment spells belong to control group members. If the starting point of the program is also the origin of the experimental data set, the unemployment and employment spells of the controls are always in progress while the first observed unemployment and employment spell of a trainee always occurs after he exits the program and thus is measured from the beginning. Assuming negative duration dependence, i.e. the probability of leaving a spell declines with the time remaining in the spell, the expected length of a spell in progress (left censored) is greater than the expected length of a new spell. A comparison of transition rates from new spells of trainees to the transition rates from left censored spells of controls is thus likely to overstate the impact of training. An exclusion of these left censored spells from the evaluation is not a valuable solution to this problem since it can make matters worse due to introducing a further sample selection bias (HAM/LALONDE (1990)). Another more general objection to experiments which is often raised is the unpopularity due to *ethical arguments against the random selection process* involved in an experiment. It is easy to imagine that there would be considerable difficulties in gaining acceptance for such a randomization in the population and responsible political officials. A final practical problem of experiments are simply the high costs of realization. Especially the last two arguments provoke that there is only a very rare supply of experimental data sets.

In a non-experimental setting the group of trainees and non-trainees are not selected upon a randomized process as described above. Hence it is necessary to carefully consider the potentially occurring selection bias in the estimation method. In principle there are two different strategies pursued in non-experimental studies. We will differentiate between on the one hand the model based approaches which were developed in econometric literature (e.g. HECKMAN (1979) HECKMAN/ROBB (1985, 1986), HECKMAN/HOTZ (1989)) and on the other hand the matching approaches that show a close similarity to the experimental context and originated in statistic literature (RUBIN (1974), RUBIN (1977), ROSENBAUM/RUBIN (1983, 1985)).

In order to understand how econometric studies deal with the sample selection problem it is helpful to give a formal notation of the outcome equation of interest

$$Y_{it} = X_{it}\beta + d_i\alpha + U_{it} \qquad \text{where } t = 0,\ldots T. \tag{4}$$

The individual employment performance (Y_{it}) is assumed to depend on characteristics (X_{it}) as well as on a dummy variable (d_i) that equals 1 if the person participated in the training program and is 0 otherwise. β is the parameter vector of the X_{it}' s and α the coefficient that measures the impact of training. In this context, a selection bias due to training participation is present, when

$$E(U_{it}, d_i) \neq 0 \tag{5}$$

As HECKMAN/ROBB (1985) point out this stochastic dependence between d_i and U_{it} may occur for two not mutually exclusive reasons. If we describe the decision to participate in training in terms of an index function framework

$$IN_i = Z_i\gamma + V_i \tag{6}$$

with the i'th individual's training status given by

$$\begin{aligned} d_i &= 1 \qquad \text{if} \quad IN_i > 0 \\ d_i &= 0 \qquad \text{otherwise} \end{aligned} \tag{7}$$

the dependence may exist because of (a) a *selection on observables*, i.e. $E(U_{it}, Z_i) \neq 0$ or (b) a *selection on unobservables*, i.e. $E(U_{it}, V_i) \neq 0$. The typical econometric solution to this selection problem is to explain the decision (the selection procedure) of training participation endogenously in the estimated model. It can be demonstrated that in the presence of selection on observables controlling for the observed Z_i variables solves the selection problem. For the case where there is also selection on unobservables, estimation procedures have been proposed that incorporate the whole index function of training participation in the econometric model thereby invoking specific sets of assumptions about the distributions of Z_i, V_i and U_i and the form of dependence between V_i and U_i. In addition to the econometric estimation approaches HECKMAN/HOTZ (1989) have developed three types of specification tests, namely the pre-program test, the post-program test and the test of model restriction that focus to identify the

presence of any selection bias in non-experimental evaluations. Since especially the pre-program test was applied in non-experimental training studies we will only present the underlying idea of this test. It builds up upon a comparison of employment performance between trainees and non-trainees prior to the training program. Its implementation affords an additional dummy in the employment equation which takes the value of 1 for those individuals who will be participating in training program in the future but have not done before (future trainees) and 0 for the non-trainees. If the selection effects are properly controlled for by the econometric model the estimated value for the coefficient of the dummy should not be statistically significant different from zero since no observation has undertaken training.

The approaches that emerged from the statistical literature to evaluate the impact of training do not put so much emphasis on modeling output and training equations through various functional forms or distributions of error terms. Instead they were developed as a straight forward extension of the experimental approach, since they try to adopt this framework in a non-experimental setting. This means that the statistical approach also aims to overcome the sample selection problem at the level of data. The idea is to apply some kind of matching technique to the observational data set in order to select from a large number of non-trainees an adequate, i.e. comparable group of controls so that the group of trainees and the selected control equal in all relevant characteristics. If this can be done correctly, i.e. each member of the selected control group corresponds to a trainee as if he/she was his/her clone, a comparison between trainees and controls outcome will provide - analogous to an experimental evaluation - an unbiased estimation of the isolated impact of training. The central assumption under which such matching procedures provide a valid control group was formulated by RUBIN (1977) and is known as *ignorable assignment conditional on a covariate*. In such a case, i.e. if we correctly condition on the vector of all relevant covariates Z which determine the assignment to training, then -like in a randomized experiment- assignment to training is independent of the outcome of interest. If this assumption holds, then:

$$E\left(Y^c \middle| Z = z, D = 1\right) = E\left(Y^c \middle| Z = z, D = 0\right) \qquad (8).$$

Rewriting the average training effect (1), as:

$$\alpha = E\left[E\left(Y^t \middle| Z = z, D = 1\right) - E\left(Y^c \middle| Z = z, D = 1\right) \middle| D = 1\right] \qquad (1')$$

and insert equation (8), we obtain (DEHEJIA/WAHBA (1995)):

$$\alpha = E\Big[E\big(Y^t \big| Z = z, D = 1\big) - E\big(Y^c \big| Z = z, D = 0\big) \Big| D = 1\Big] \qquad (9).$$

An evaluation of the effect of training is now possible through first estimating the impact of training conditional on Z i.e. $E\big(Y^t \big| Z = z, D = 1\big) - E\big(Y^c \big| Z = z, D = 0\big)$ and then averaging this estimate over distribution of the set of covariates in the group of trainees $(D = 1)$. Whereas the final step is done quite easily by replacing the expectation through the corresponding estimates of sample means, the first one involves an implementation on how to condition on Z. In principle conditioning implies matching or grouping of the trainees with non-trainees according to exactly equal values of the covariates Z. Obviously if there are many relevant covariates this method of exact matching is limited due to the high number of possible matches. In order to reduce this problem of dimensionality ROSENBAUM/RUBIN (1983) suggest, to ground the matching procedure on the so-called propensity score function, the conditional probability of participating in training given a set of covariates. In this procedure trainees and controls with identical propensity scores have to be matched. The advantage of the score is that one only has to find matches on an univariate scale. The disadvantage however is that the probability to participate, i.e. propensity score is unknown and needs to estimated. In literature this is typically done by means of a simple logit or probit model.

The following two chapters review studies on the impact of training on employment. The focus will be primarily on non-experimental evidence in Germany. However we will first present selected literature relating to other countries and experimental investigation, since experimental data sets are not available in Germany.

III. Studies Using Experimental Data

HAM/LALONDE (1990) focus on the sensitivity of training impact estimations on employment histories due to different treatments of the initial condition problem in an experimental setting. The data they use is drawn from the National Supported Work (NSW) experiment which was conducted from 1975 to 1979. The program's goal was to ease the transition of women with welfare reception and poor work histories into the labor market, and thus to provide these women with the general skills and confidence necessary to be successful employees. A comparison of the mean unemployment duration of trainees and controls using all spells, i.e. including the left censored spell of the controls, indicates that training has a positive effect because it shortens the mean of the unemployment duration by approximately four months. Unfortunately however the authors obtain a quite opposite result when discarding left censored spells from their evaluation. Unemployment spells of trainees are now about two months longer compared to the controls. This sensitivity is confirmed by estimation results for the survivor functions. The advantage of using survivor functions is that they are unaffected by longer time frames for the controls and right censoring of spells. However, differences in time varying individual characteristics as well as differences in demand conditions associated with the spells, still pose a problem when interpreting these results. A comparison between the estimated unemployment survivor function for trainees and controls where all unemployment spells are incorporated indicates no training effect. Again a comparison where left censored spells of the controls are excluded leads to the result that controls tend to exit unemployment spell earlier than trainees. These empirical results confirm theoretical consideration about the initial condition problem. One is confronted with a tradeoff problem since both approaches to the initial condition problem lead to biased estimates, either by a specification error or otherwise by a sample selection problem.

In their paper of **1996 HAM/LALONDE** assess the separate effects of a training program on the duration of participants subsequent employment and unemployment spells using the same experimental data set as in their paper of 1990 which was drawn from the National Supported Work (NSW). Focusing on the sensitivity of the estimation to the initial condition and sample selection problem which were already discussed in HAM/LALONDE (1990) they now propose an estimator that addresses these problems. They apply a discrete time duration model with a hazard rate that follows a logistic distribution. They account for the transitions out of three states,

namely employment, unemployment and training. Due to the experimental data it is not necessary to address the transition into training. Their model captures the effects of duration dependence, conditions on personal characteristics, demand side variables and incorporates unobserved heterogeneity. Moreover it explicitly takes into account spells that are in progress at the baseline. The authors follow a suggestion of HECKMAN/SINGER (1984) according to which these interrupted spells have their own hazard and heterogeneity term. The model is estimated under five different assumptions about the specification of the unobserved heterogeneity. The first two specifications either do not account for unobserved heterogeneity or treat the unobserved heterogeneity terms across the different transitions as independent and thereby deal with the sample selection problem only inappropriate. The final three specifications attempt to address the sample selection problem more adequately by allowing possible correlation between the different transitions. To investigate in the initial condition problem the model is estimated once taking account only of spells that are observed from their beginning, i.e. fresh spells, and once including also the interrupted spells. The estimation results confirm the findings of HAM/LALONDE (1990) according to which the initial condition and sample selection problem are of more than theoretical interest. The two standard models which do not account for the sample selection in an appropriate way and only use fresh spells, produce results indicating that training raises the expected employment duration and surprisingly also the expected unemployment duration. In contrast the more sophisticated statistical frameworks that incorporate the problems of the initial condition and sample selection lead to more reasonable results, namely that training only lengthens the expected employment duration whereas it has not essential effect on the unemployment duration.

IV. Studies Using Observational Data

RIDDER (1986) evaluates the impact of a number of training, recruitment, and employment programs which were conducted in the Netherlands mainly in 1979 on subsequent employment and unemployment spells of the participants. The programs aimed to stimulate the re-employment probabilities, especially for unemployed who belonged to some disadvantaged group, namely individuals with a long unemployment spell, disabled individuals and workers of (ethnic) minorities. Due to data limitation the estimation does not differentiate between the different training, recruitment and employment programs. Instead RIDDER (1986) estimates the impact of three categories, namely training, recruitment and employment assuming that they consist of homogeneous programs. In other words the measured effects do not appraise for example a particular training program but rather the general average impact of a group of training programs. Since the study uses an observational data set which only includes information about program participants RIDDER (1986) employs a simple pre-treatment - post-treatment design without any kind of control or comparison group of non-trainees. The author decides in the presence of data which allows a measurement of spell duration by months to apply a duration model in continuous time. The model conditions on observed personal characteristics as well as demand side variables that are supposed to incorporate the effects of business cycles. The estimated treatment impacts on the average employment spell indicate, that employment programs are more effective than recruitment programs, and these are more effective than training programs. A distinction among demographic groups exhibits that females benefit most from the programs followed by minority workers and young workers in this order with the exemption that minority workers do not benefit from recruitment programs. Furthermore it seems that older workers do not benefit at all from programs. The estimation results of the treatment impacts on the average unemployment spell do not tend to support a positive assessment of the programs. It seems as if program participation lengthens the unemployment spells. RIDDER (1986) interprets these outcomes as a result of the non-random selection process by the program administrators which is not adequately captured by his model.

The study of **CARD/SULLIVAN (1988)** for the United States is concerned with the impact of subsidized training programs on movements in and out of employment. The authors utilize non-experimental data that contains pre- post-training information about a trainee group as well as

information about a group of non-trainees. The sample of trainees consists of adult males and is drawn from the Continuous Longitudinal Manpower Survey (CLMS) which provides facts on the participants of the Comprehensive Employment Act (CETA) in 1976. Trainees are differentiated between the distinct groups of classroom trainees (i.e. short-term vocational courses) and non-classroom trainees including participants of on-the-job training, of work experience programs and of public sector employment programs. The members of the group of non-trainees are taken from the 1976 Current Population Survey (CPS). The starting point of the empirical analysis is a simple comparison between the pre- and post-training employment probabilities for trainees and non-trainees. The findings are that CETA participation had a small but significantly positive impact on the post-training employment probabilities. Differentiation between classroom and non-classroom trainees exhibits that this positive impact is mainly due to classroom training, whereas non-classroom training apparently has no effect on employment. Since a comparison between trainees and non-trainees pre-training histories and characteristics provides that trainees are obviously a non-random sample of the population CARD/SULLIVAN (1988) proceed to compare trainees and controls with exactly the same pre-training history in order to avoid a possible sample selection bias. The results they obtain now shows a smaller effect for classroom participants and a larger effect for non-classroom participants, although the overall training effect remains the same. A further distinction between the retention rate, describing the probability of remaining employed and the accession rate, describing the probability of moving from unemployment to employment suggests that the training effect may have mainly worked through a positive accession rate for classroom trainees. The main problem of the above estimation procedure which is similar to the statistical matching approach described in chapter II is the strong assumption that the participation decision is independent of any unobservable determinants, in other words that trainees and non-trainees do not differ with respect to unobservable characteristics such as motivation. Therefore the final estimation builds up upon an econometric model based approach. Since the variables are only measured on a yearly basis the authors decide to apply a logit model of employment probability that includes a random-effects specification for the unobserved individual heterogeneity. Moreover the model incorporates an equation of training participation in order to permit interactions between the unobservable components of the employment probabilities and the individual-specific determinants of training status. The technique they use to specify the distribution for the unobserved heterogeneity component follows a nonparametric approach which was developed for transition data models by HECKMAN/SINGER (1984). The authors also incorporate state dependence in the employment

equation to account for the observed differences between the retention and accession probabilities. The results obtained from this more sophisticated model point to a somewhat larger increase in trainees employment rates than was obtained by the two previous approaches. However analogous to the results above most of the increase in employment rates is concentrated among classroom trainees. Furthermore the results from the model based approach exhibit that these positive effects are of transitory as well as permanent nature.

GRITZ (1993) presents a study for the United States that attempts to determine the impact of private and government sponsored training programs on the frequency and duration of employment. He makes use of the first four waves of the Youth Cohort of National Longitudinal Survey (YNLS) covering a time span from 1979 to 1982 and containing a nationally-representative sample of youths and a supplemental sample of Black, Hispanic, or economically disadvantaged non-Hispanic, non-Black youths. Since the data consists of individual youths who have participated in training programs as well as non-participants a comparison group is available. Moreover it is possible to determine the labor market state of an individual for every week during the sample period. This detailed information about the individual labor market histories induces GRITZ (1993) to apply a continuous time duration model. In order to evaluate the impact of training on the duration as well as on the frequency of employment it is necessary to introduce at two state duration model that distinguishes between employment and non-employment. The transition out of employment is relevant for the impact of training on employment duration. Modeling the transition out of non-employment is necessary to capture the impact of training on the frequency of employment. Finally the necessity to take into account the endogeneity of training, i.e. to address the problem of sample selection, efforts to consider the transitions into training. Hence the author has to utilize a three state duration model with six possible transitions among three mutually exclusive labor market states employment, non-employment and participation in training. Due to the nature of the data the model has to adjust for right censoring as well as initial conditions. Moreover GRITZ (1993) accounts for unobserved heterogeneity in the way that was proposed by HECKMAN/SINGER (1984), i.e. the distribution of the heterogeneity term is described by discrete distribution with a finite number of mass points. The estimated results for the impact of private training show that women benefit through an increased duration of employment as well as a decreased duration of non-employment (i.e. a higher frequency of employment). These benefits appear to exist over both, a short and long period of time. In contrast the implications are less clear for men since participation in private

training increases the length of employment and non-employment episodes. Finally for the case of government programs participation actually deteriorates the employment performance of women and men, however, this result is based upon a small number of government trainees and therefore has to be judged carefully.

HUJER/SCHNEIDER (1990) analyze for Germany the short- and long-run effects of training programs on employment probabilities of unemployed. The sample they use is drawn from the first four waves of the German Socio-economic Panel (GSOEP) covering the time period from 1983 to 1986 and consisting of men who were at least once unemployed during the observation time. In presence of data that allows information on monthly basis about unemployment and employment duration the authors decide to apply a transition model in continuos time. Thereby the baseline hazard is specified through a Weilbull distribution and unobserved heterogeneity is incorporated via a random term with a discrete distribution that is estimated as proposed by HECKMAN/SINGER (1984). In order to address the first question i.e. short term impact of training, the study assesses the effect of training on the transition rate out of unemployment into employment. The evaluation of the long run training effect is done by determining the influence of training on the transition rate out of employment back to unemployment. The estimated results indicate a significant positive short run effect of training whereas training seemed to have no positive long run effect i.e. training did not lower the transition rate out of employment back to unemployment. The authors conclude that due to the short observation time and small sample size the results have to be judged carefully.

For the case of Germany **BECKER (1991)** examines separately the determinants of participation in vocational training programs and the impact of this kind of training on changes on career ladder. According to the aim of our own study we are mainly interested in the second question. The sample on which the estimations are based on consists of West-German people who belong to the cohorts 1929-31, 1939-41 and 1949-51. It was drawn from data of the German Life History Study (MAYER/BRÜCKNER (1989)). The data provides information about the length of employment and training episodes by months. This leads BECKER (1991) to apply a duration model in continuos time. He devotes for the Cox-model and utilizes the partial likelihood approach for the estimation. The model is estimated independently for the effects of training on job changes that are combined with a decline in employment status and on job changes that lead to a raise in employment status. Furthermore these evaluations are done separately for women and men. The results indicate, that a participation in vocational training programs reduces the

probability of a decline in employment status significantly (this effect is more distinct for men than for women), while participation has no notable effect on a raise in employment status. In contrast if one looks at interaction variables that also incorporate the success of a vocational training program, i.e. if a certification was assigned, than the estimations show that successful training programs raises the promotion probabilities significantly for both women and men.

BÜCHEL/PANNENBERG (1994) are interested in the relation between on-the-job training, firm-internal career ladders and income changes in Germany. With respect to our intention mentioned above we will only present the results that describe the connection between on-the-job training and career ladders. The study is based upon data which was drawn from the first 8 waves of the GSOEP i.e. from 1984 to 1991. Since the study focuses on career ladders the relevant sample consists only of individuals who were employed in the same firm continuously during all 8 waves. BÜCHEL/PANNENBERG (1994) consider on-the-job training and movements on internal career ladders as simultaneous process and hence rely estimations on a bivariate probit model with one equation describing the entrance to on-the-job training and the other one describing the promotion probability. The results show that the probability of promotion rises for an individual significantly, if on-the-job training takes two to seven days and the individual has participated on various on-the-job training programs before.

HÜBLER (1994) examines for East Germany the interactions between participation in training and job search activities as well as possible impacts of these activities on working time. The estimation is based on the first four waves, i.e. from 1990 to 1991, of the 'Arbeitsmarktmonitor-Ost' , a representative labor market data set drawn from the population of the former GDR. It incorporates only individuals who were employed at the first wave and who had valid information for the following three waves. To address the first question, the relation between training and job search activities, HÜBLER (1994) estimates a bivariate probit model with one equation describing the decision to participate in training and the other equation describing the decision to search a job. With respect to the second question the introduction of an additional equation for hours of work is afforded whereby the upper bivariate probit model is extended to a selection model with doubled endogenous switch. The results obtained for the first model indicate that training activities provoke search activities but not vice versa. A distinction between training on the job and training in an institution outside the firm exhibits that the latter has a stronger impact on search activities than the former. A further differentiation between female and male employees shows that HÜBLER'S (1994) results are quite similar to the ones obtained by

GRITZ (1993) for the USA. It seems to be that only for women an impact of training on search activities can be measured whereas for men no significant impact is estimated. If one accepts the assumption that training within a firm is typically privately financed and training outside the firm is typically government induced, than further results of GRITZ (1993) are confirmed. Namely that on the one hand training within a firm (i.e. according to the upper convention private training) reduces the search activities of women, thereby lengthening the employment duration whereas it has no significant effect on search activities of men, while on the other hand training outside of the firm (i.e. according to the upper convention government sponsored training) has no significant effect for both, women and men. Finally the results of the second model, the selection model of working time, indicate that search activities lead to an extended working time, whereas training activities reduce the hours of work.

The work of **FITZENBERGER/PREY (1995)** explores the impact of training within a firm and training in an institution outside of the firm, on the future employment probability of trainees for the case of East Germany. The estimates are based on the first six waves of the "Arbeitsmarktmonitor Ost" covering a time span from 1990 to 1992. Since the evaluation uses an observational data set which is in addition plagued by serve attrition (panel mortality) the authors' central feature is to adequately account for possible sample selection problems. The estimated model is a simultaneous random effect probit model consisting of an employment equation, a qualification equation and an attrition equation. Moreover the evaluation includes a further probit equation that explicitly accounts for the initial condition problem and allows for state dependence in the employment equation. In order to identify the presence of any selection effects among training participants the authors apply the pre-program test procedure that was proposed by HECKMAN/HOTZ (1989). The estimation results indicate that when one neglects the existing selection bias, training outside of the firm shows a strong negative short run effect and nearly no long run effect on employment probabilities. For the participants of training in the firm one obtains a strong positive short run and a smaller but also positive long run effect. However a correction of these outcomes for the existing sample selection bias, as it is indicated by the coefficient for the pre-program test dummy leads to quite opposite results. For the case of training outside of the firm the pre-program test exhibits a significant negative selection bias, i.e. individuals who participate in this kind of training belong to a labor market group with bad prospects. Keeping this negative selection in mind, reverses the above results and indicates that training outside of the firm actually raises employment probabilities of this particular group. For

the case of training participants within a firm the pre-program test shows a positive selection, i.e. this kind of training seems to attract individuals with rather positive prospects. With respect to this positive selection only a small positive short run effect exists, whereas the long run effect of training actually is negative.

PFEIFFER/BRADE (1995) investigate the determinants of training participation, and the impact of training on earnings, thereby separating between the effect on employment, i.e. more precisely on working time and the effect on wages. According to the aim of this overview we will only present the results concerning the training effects on working time. The sample they use consists of employed West-German men with vocational experience between 0 and 30 years and is drawn from the 'Mikrozensus' of 1991, a representative non-experimental cross-sectional data set for Germany. Due to the cross-sectionality of the data the authors simply estimate OLS regression on working time. The obtained results indicate, that training approximately increases the monthly working time by 1.1%. A differentiation among the alternative training types exhibits that the highest positive training effect, an increase of almost 9% working time per month, results from a training participation of less than one month in a training program of the Chamber of Industry and Commerce (Industrie und Handelskammer). In contrast people who participate in training within a firm have an increased monthly working time of about 2.5%. As the authors state, the estimated results on working have to be judged carefully because of the small variation in monthly working time across individuals due to the restrictive working time rules in Germany.

STEINER/KRAUS (1995) are interested in the assessment of a special treatment of the active labor market policy in East-Germany after the unification, namely subsidized temporary jobs (Arbeitsbeschaffungsmaßnahmen, ABM). Therefore they evaluate whether participants in ABM had better reemployment chances than unemployed. The applied sample was drawn from the first six waves of the 'Arbeitsmarktmonitor-Ost' . The estimation is based on a discrete time hazard model where the transition rate is described by a multinomial logit model. The model allows for unobserved heterogeneity, which is specified along the lines of HECKMAN/SINGER (1984). The hazard model is separately employed for individuals with a transition out of ABM into two possible states, employment and nonemployment, and for individuals with a transition out of unemployment into three possible states, employment nonemployment and ABM. Due to the large sample of unemployed it is possible to estimate the transition out of unemployment separately for men and women. In contrast the small sample size of ABM participants forced the authors to model this distinction through an interaction term. To obtain the impact of ABM on

reemployment STEINER/KRAUSS (1995) construct two reference groups, one which has a high transition rate into ABM (ABM reference group) and another one which has an average transition rate into ABM (average reference group). For these two groups, they compare their transition rate into employment for being unemployed with the transition rate into employment for being an ABM participant. This comparison is done separately for women and men. For a man of the average reference group there seems to be only a small difference between his transition out of unemployment compared to that out of ABM. This result is also confirmed for a man of the ABM reference group. In contrast for women the authors find significantly different impacts of ABM, depending on the utilized reference group. If one looks at a woman of the average reference group transition rates out of unemployment and ABM differ rarely. In contrast, for a woman of the ABM reference group the transition rate out of ABM seems to be significantly lower compared to the transition rate out unemployment. These results exhibit the sensitivity of the estimation to the selected reference group and therefore emphasize the sample selection problem which is typically involved in all treatment estimations.

Whereas the studies presented so far either use an econometric approach or had access to experimental data the study of **LECHNER (1995)** applies the third strategy which we described in Chapter II, namely the statistical approach, to estimate the impact of training. The study aims to evaluate the effect of training programs outside of the firm for the case of East Germany. The balanced sample which was used for the estimation is drawn from the first five waves of GSOEP covering a time span from 1990 to 1994. The author argues that the key assumption of the statistical approach, namely the independence between the potential outcomes and the decision to participate in training conditional on observable covariates is justified due to the specific situation in East Germany after the unification. In order to find an adequate control group for the trainees, LECHNER (1995) uses a matching procedure based on the estimated propensity score as well as on some selected observable variables that could not be captured by the propensity score because they depend on the particular date of the beginning of training. The estimation of the propensity score function is done with a conventional binary probit model. Based on such a matched sample the author then simply compares the means of different post-training outcomes such as earnings and unemployment rates for trainees and matched controls. The results obtained by LECHNER (1995) indicate that there seems to be no robust positive effect of training outside of the firm on these outcomes. It has to be noted however, that these result have to be judged carefully, since they are based on a fairly small number of trainees and matched controls.

IV. Conclusion

This paper provided a selected overview of empirical studies on the impact of training on employment. Since a central problem with evaluating training effects is the intriguing selection problem we first theoretically outlined this issue and the different evaluation strategies that are typically proposed to deal with it. As our project intends to determine the impact of training on employment histories using an observational data set, namely the GSOEP, we have to carefully address this issue in our evaluations. We intend to pursue two different approaches. The first approach pursues the typical econometric strategy and employs a hazard model for grouped transition data (see HUJER/MAURER/WELLNER (1996) for a survey on these models) with instrument variables for the participation in training. In order to detect a possible remaining sample selection bias, that was not controlled for by the model we will use the pre-program test of HECKMAN/HOTZ (1989). The other strategy that we follow tries to combine the statistical and the econometric approach. In the first stage we intend to apply matching techniques to construct an adequate control group and thus overcome the sample selection problem. In the second stage we then proceed along the lines of the econometric approach and estimate a discrete hazard model based on the matched sample to obtain the effect of training on employment histories.

Literature

ASHENFELTER, ORLEY/ CARD, DAVID (1985): Using the Longitudinal Structure of Earnings to Estimate the Effect of Training Programs, in: *Review of Economics and Statistics*, Vol. 67, p. 648-660.

BECKER, ROLF (1991): Berufliche Weiterbildung und Berufsverlauf: Eine Längsschnittuntersuchung von drei Geburtskohorten, in: *Mitteilungen aus der Arbeitsmarkt- und Berufsforschung*, 2/91, p. 351-364.

BJÖRKLUND, ANDERS (1989): Evaluations of Training Programs: Experiences and Proposals for Future Research, Discussion Paper FS I 89 - 13, Wissenschaftszentrum Berlin für Sozialforschung.

BÜCHEL, FELIX/ PANNENBERG, MARKUS (1994): On-the-Job Training, innerbetriebliche Karrierepfade und Einkommensentwicklung, in: *Jahrbücher für Nationalökonomie und Statistik*, Vol. 213/3, p. 278-291.

CARD, DAVID/ SULLIVAN, DANIEL (1988): Measuring the Effect of Subsidized Training Programs on Movements in and out of Employment, in: *Econometrica*, Vol. 56 No. 3, p. 497-530.

DEHEJIA, RAJEEV, H./ WAHBA, SADEK (1995): Causal Effects in Non-Experimental Studies, Working Paper, Harvard University.

FITZENBERGER, BERND/ PREY, HEDWIG (1995): Assessing the Impact of Training on Employment: The Case of East Germany, Working Paper, Universität Konstanz.

GRITZ, R. MARK (1993): The Impact of Training on the Frequency and Duration of Employment, in: *Journal of Econometrics*, Vol. 57, p. 21-51.

HAM, JOHN C./ LALONDE, ROBERT J. (1990): Using Social Experiments to Estimate the Effects of Training on Transition Rates, in: Hartog, Joop/ Ridder, Geert u. Theeuwes, Jules (eds.): Panel Data and Labor Marketing Studies, p. 157-172.

HAM, JOHN C./ LALONDE, ROBERT J.(1996): The Effect of Sample Selection and Initial Conditions in Duration Models: Evidence from Experimental Data on Training, in: *Econometrica*, Vol. 64, No. 1, p. 175-205.

HECKMAN, JAMES J. (1979): Sample Selection Bias as a Specification Error, in: *Econometrica*, Vol. 47, No. 1, p. 153-161.

HECKMAN, JAMES J./ SINGER, BURTON (1984): A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data, in: *Econometrica*, Vol. 52, No. 2, p. 271-320.

HECKMAN, JAMES J./ HOTZ, V. JOSEPH (1989): Choosing Among Alternative Nonexperimental Methods for Estimating the Impact of Social Programs: The Case of Manpower Training, in: *Journal of the American Statistical Association*, Vol. 84, No. 408, p. 862-880.

HECKMAN, JAMES J./ ROBB, RICHARD (1985): Alternative Methods for Evaluating the Impact of Interventions, in: Heckman, James J./ Singer, Burton (eds.): Longitudinal Analysis of Labor Market Data, New York, u.a., p. 156-245.

HÜBLER, OLAF (1994): Weiterbildung, Arbeitsplatzsuche und individueller Beschäftigungsumfang - Eine ökonometrische Untersuchung für Ostdeutschland, in: *Zeitschrift für Wirtschafts- und Sozialwissenschaften*, Vol. 114, p. 419-447.

HUJER, REINHARD/ MAURER, KAI.-OLIVER/ WELLNER, MARC (1996): Models for Grouped Transition Data, Frankfurter Volkswirtschaftliche Diskussionsbeiträge No. 68, Depart niversität Frankfurt.

HUJER, REINHARD/ SCHNEIDER, HILMAR (1990): Kurz- und mittelfristige Auswirkungen von Umschulungs- und Fortbildungsmaßnahmen auf die Beschäftigungschancen von Arbeitslosen, Manuskript, Frankfurt/M.

LECHNER, MICHAEL (1995): Effects of Continuous Off-the-job Training in East Germany after Unification, Discussion Paper No. 95-27, Universität Mannheim.

MAYER, K. U./BRÜCKNER, E. (1989): Lebensverläufe und Wohlfahrtsentwicklung. Konzeption, Design und Methodik der Erhebung von Lebensverläufen der Geburtenjahrgänge 1929-1931, 1939-1941,1949-1951, Materialien aus der Bildungsforschung Nr. 35, Berlin: Max-Planck-Institut für Bildungsforschung.

PFEIFFER, FRIEDHELM/ BRADE, JOACHIM (1995): Weiterbildung, Arbeitszeit und Lohneinkommen, in: Steiner, V./ Bellmann, L. (Hrsg.): Mikroökonomik des Arbeitsmarktes, Beiträge zur Arbeitsmarkt- und Berufsforschung, Nürnberg, p. 289-326.

RIDDER, G. (1986): An Event History Approach to the Evaluation of Training, Recruitment and Employment Programs, in: *Journal of Applied Econometrics*, Vol. 1, p. 109-126.

ROSENBAUM, PAUL R./ RUBIN, DONALD B. (1983): The Central Role of the Propensity Score in Observational Studies for Causal Effects, in: *Biometrika*, Vol. 70, No. 1, p. 41-55.

ROSENBAUM, PAUL R./ RUBIN, DONALD B. (1985): Constructing a Control Group Using Multivariate Matched Sampling Methods that Incorporate the Propensity Score, in: *The American Statistical Association*, Vol. 39, No. 1, p. 33-38.

RUBIN, DONALD B. (1974): Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies, in: *Journal of Educational Psychology*, Vol. 66, No. 5, p. 688-701.

RUBIN, DONALD B. (1977): Assignment to Treatment Group on the Basis of a Covariate, in: *Journal of Educational Statistics*, Vol. 2, p. 1-26.

RUBIN, DONALD B. (1979): Using Multivariate Matched Sampling and Regression Adjustment to Control Bias in Observational Studies, in: *Journal of the American Statistical Association*, Vol. 74, No. 366, p. 318-328.

STEINER, VIKTOR/ KRAUS, FLORIAN (1995): Haben Teilnehmer an Arbeitsbeschaffungsmaßnahmen in Ostdeutschland bessere Wiederbeschäftigungschancen als Arbeitslose?, in: Steiner, V./ Bellmann, L. (Hrsg.): Mikroökonomik des Arbeitsmarktes, Beiträge zur Arbeitsmarkt