published in Journal of Business & Economic Statistics, 17, 74-90, 1999

Earnings and Employment Effects of Continuous Off-the-Job Training in East Germany After Unification

Michael Lechner

Universität Mannheim, Mannheim, Germany

First revision: October 1996

Second revision: September 1997

Address for correspondence:

Michael Lechner

Fakultät für Volkswirtschaftslehre

Universität Mannheim

D-68131 Mannheim, Germany

Email: Michael.Lechner@vwl.uni-mannheim.de

WWW: http://www.vwl.uni-mannheim.de/inst_vwl/lechner

Abstract

The effects of continuous off-the-job training (OFT) for East Germans after unification are analyzed in terms of their earnings and employment probabilities. Using the *potential outcome approach to causality* as general framework, different matching procedures are suggested for the estimation. They allow for permanent and transitory shocks that influence OFT participation and labor market outcomes. The matching procedures also take account of individually different starting dates. The data is taken from the German Socio-Economic Panel (GSOEP, 1990-1994). This dataset is very informative with respect to factors influencing the participation in training. The results show no positive effects of training.

Keywords

Evaluation of training programs, causal analysis, panel data, matching on the propensity score, panel data, East German labor market.

1 Introduction

Retraining the labor force to match the demands of a modern economy is an important task during the transition from central planning to a market oriented economic system. This need is particularly pressing in East Germany because the transition process is much faster there than in the rest of Eastern Europe. Therefore, substantial public and private resources are devoted to this purpose, and the need for an evaluation of the results is obvious.

This paper concentrates only on one particular aspect of training, namely off-the-job training (OFT). It attempts to identify average individual gains for the workers of the former GDR participating in OFT after unification compared to a hypothetical state of nonparticipation. The targets of the evaluations are labor market outcomes after the completion of training, such as current or expected earnings, labor market status, and career prospects.

Since experimental data is not available for Germany, potentially serious sample selection biases that influence the evaluation results because of the nonrandom selection of training participants are an issue (e.g. Heckmann and Robb, 1985). Various model-based procedures are suggested in the econometrics' literature to avoid such biases. Ashenfelter and Card (1985) and Lalonde (1986) come -among others- to the conclusion that the results are highly sensitive to the different stochastic assumptions made about the selection process. Both papers conclude that the econometric adjustment procedures are unreliable, and hence that social experiments are necessary to evaluate training programs. Recently, Dehejia and Wahba (1995a, 1995b) - using an approach very similar to the one chosen here - reevaluate the Lalonde (1986) data. By using nonparametric techniques, among them *matching* that will be discussed later on, they come to far more positive conclusions about the potential quality of inferences based on observational data than Lalonde himself. The recent interest in matching and other nonparametric evaluation methods is also documented by a series of papers, yet unpublished when this work was written, by James Heckman and coauthors (e.g. Heckmann, Ichimura, Smith, Todd, 1996, Heckman, Ichimura, Todd, 1997).

Project (or treatment) evaluation and the need of a definition of causality have a history in the statistics' literature as well. This literature stresses the need of nonparametric solutions to the identification problem, and the nonparametric estimation of the causal effects. Rubin (1974) seems to be the first to explicitly suggest a model of potential outcomes that clarifies that the individual causal effect of training - defined as the difference of the two potential outcomes, for example - is never identified. Therefore, the lack of identification has to be overcome by plausible, generally untestable assumptions that usually depend heavily on the problem analyzed and the data available.

The empirical results in this paper are obtained by using the potential outcome approach as a general framework to define causal effects of OFT on individual actual and expected post-training labor market outcomes. The paper argues that due to the specific situation in East Germany after unification and the rich data at hand, it is plausible to assume that the outcomes and the assignment mechanisms are independent conditional on observed attributes, including monthly pre-training labor force status. Hence, this assumption solves the identification problem that is inherent in causal analysis. This identification strategy as well as the estimation methods suggested try to avoid identification by using functional assumptions. Thus, the critique of evaluation results obtained by parametric econometric models does not apply.

Generally, the paper contributes to the discussion of the effectiveness of the training in East Germany by analyzing the participation decision as well as by identifying empirically important factors related to it, before obtaining evaluation results for several outcome measures related to the individual position in the labor market. Methodologically, matching procedures taken from the statistical literature (e.g. Rubin 1979, Rosenbaum and Rubin, 1983, 1985) are extended to allow an accommodation of the specific problems encountered in this study and to exploit monthly information on the labor force status that could be particularly valuable.

The results do not confirm previous positive findings of the effectiveness of work-force training in East Germany (e.g. Fitzenberger and Prey, 1996, Pannenberg and Helberger, 1994). Although only a few studies exist so far, they differ in many respects ranging from the database to the implementation of the evaluation. However, they share two common features that are absent here: they do not use an explicit causality framework, and they are based on modeling the distributions of the outcome variables given certain covariates.

The paper is organized as follows: The following section outlines a few features of the East German labor market after unification as the relevant economic environment. Section 3 describes the longitudinal data used in this study. All computational aspects of the evaluation are discussed in Section 4. Section 5 concludes. Appendix A contains additional information about the data. Finally, Appendix B consists of several more technical aspects of the econometric methods used.

2 East German labor markets and training

Unification came like a shock over the East German labor market. The transformation from a once centrally planned economic system to a West German type market economy led to considerable disequilibria in the labor market. For example, for the active working population under 50 of the late GDR (in 1990), the share of full-time employment declined from 100% in mid 1990 to about 70% in early 1991 and than stabilized at around 80%. The unemployment rate - below 2% before unification - increased steadily up to about 12 % by the end of 1993. Finally, the number of people taking part in some kind of training also increased steadily after unification. It reached a proportion of about 5% in 1992 and fell thereafter.

To smooth the transition to a market economy and to adjust the East German stock of human capital to the needs of the new economic system, the government conducted an active labor market policy. Among other things this policy provided significant funds for training and re-training opportunities (a total of about 26 bn DM from mid 1991 to 1993). For a detailed account of these policies, empirical facts, and institutional arrangements, the reader is referred, for instance, to Eichler and Lechner (1996). The focus here is on off-the-job training including subsidized continuous training intended to increase skills within the current occupation. Getting a subsidy usually depends on certain conditions related to the employment history, the approval of the course by the labor office, and the potential termination of unemployment or the avoidance of becoming unemployed soon. Payments cover in most cases the costs for the provision of the course as well as about two thirds of the previous net earnings. Using different data, the companion paper Lechner (1996) focuses exclusively on courses for which substantial subsidies have been obtained by the participants.

The main targets for the evaluation considered here are gross monthly earnings and labor force status. It is interesting to compare the development of these variables before and after training for the training participants. Prior to training unemployment rises up to a level of about 20%. After OFT it clearly falls. A similar picture arises when full-time employment or other measures incorporating reduced hours are considered. Comparing pre- and post OFT monthly earnings shows an even more optimistic picture. Mean earnings are about DM 1900 before OFT, and about DM 2600 in the years after OFT. Another 'naive' look at the evaluation problem is provided by Figure 1. It shows a comparison of the share of unemployed OFT participants before and after OFT with a randomly chosen group of nonparticipants.

[------------------------------------- Figure 1 about here ---

]

Considering the difference of unemployment rates of trainees after they finish training it appears that their unemployment rates are not below those of the rest of the population. It is also obvious that the group of training participants is fairly different from the rest of the population. In conclusion, more sophisticated methods seem to be appropriate to evaluate the effects of training in this case.

3 Data

The sample used for the empirical analysis is drawn from the German Socio-Economic Panel (GSOEP), which is very similar to the US Panel Study of Income Dynamics (PSID). About 5000 households are interviewed every year since 1984. About 2000 East German households were added in 1990. The GSOEP is rich in terms of socio-demographic information, in particular concerning current and past labor force status. For a more comprehensive English language description of the GSOEP see Wagner, Burkhauser, and Behringer (1993).

A very useful characteristic is the availability of monthly retrospective information for some variables, so that the current sample allows to follow the individual employment histories quite accurately from July 1989 to December 1993. The training information is taken from a special survey on continuous training included in the 1993 survey. It contains questions about the last three continuous training courses that were either completed during the last three years or are still going on at the time of the interview. The information provided includes the month of the beginning, the (approximate) duration, the number of weekly hours, the objective, whether training took place during working hours, and finally whether a certificate of participation has been obtained. See Appendix A for more details.

To be able to use the special survey as well as information concerning the labor force status in the GDR, a sample of individuals who responded in all of the first four of the yearly interviews (1990 to 1993) is selected. They were born between 1940 and 1970. The upper age limit is set to avoid the need of addressing early retirement issues. Since the population of interest is the labor force of the GDR, all selected individuals worked full-time just before unification. Furthermore, the self-employed in the former GDR (2%) are not observed taking part in OFT, so they are deleted from the sample. Finally, individuals reporting severe medical conditions are not considered.

Table A.1 in Appendix A gives a description of all variables used in the empirical analysis for those who received off-the-job training (OFT) and those who did not receive it. Individuals who did not complete OFT until Dec. 93 are deleted from the sample. The definition of OFT used is as follows: The purpose of OFT is qualification other than retraining for a different occupation. However, if retraining has a duration of less than three months it is also considered as OFT. The minimum duration is 16 hours, or one week (with fewer than 16 hours). Furthermore, training does not take place during regular working hours. The purpose of the definition is to obtain a not too heterogeneous group of trainees by excluding very short courses, on-the-job-training and (substantial) retraining for a different occupation. Those are all very different kinds of training with very heterogeneous objectives and very different selection rules. This definition does not exclude the possibility that OFT-participants receive some kind of other training before or after OFT-participation.

Table A.1 shows that OFT-trainees are not a random sample. For example, there are far more women in OFT than men. Individuals who accumulated more human capital and who reached a higher job position in the former GDR are more likely to obtain OFT. There are also regional differences: Individuals living in *East Berlin* are more likely to be observed in OFT than for instance people living in *Sachsen-Anhalt*.

[------------------------------------ Figure 2 about here --]

Figure 2 shows the sample distribution function of OFT duration in full-time equivalents (38 hours per week). About 50% of the courses have a duration of up to one month. Less than 5% last longer than 12 months. Therefore, the censoring problem due to the omission of courses not completed by the 1994 interview will have no serious consequences for the evaluations.

All information about individual costs and subsidies are only available for the particular course the individual believes is the most important one for the own career. For these courses 16% of the individuals obtained financial support by their employer, such as a continuation of their wage or salary. 44% obtained financial support from the labor office. 42% received nothing. Hence, there is a substantial part of courses not subsidized by the labor office. For more detailed information the reader is referred to Lechner (1995).

4 Econometric methodology and empirical implementation

4.1 Causality, potential outcomes, identification, and the propensity score

"What is the average gain of OFT participants compared to the hypothetical state of nonparticipation in OFT?" This question is in the center of the empirical analysis of this paper. It refers to potential outcomes or potential states of the world. The underlying notion of causality requires the researcher to determine whether participation in OFT effects the respective outcomes, such as earnings or labor force status. This is different from an empirical association, typically related to some kind of correlation between OFT and the outcome. See for example Holland (1986) and Sobel (1994) for an extensive discussion of concepts of causality in statistics, econometrics, and other fields. The framework that will serve as guideline for the empirical analysis is the potential-outcome approach to causality (i.e. Rubin, 1974). This idea of causality is inspired by the set-up of experiments in science. The main building blocks of the notation are *units* (here: individuals, *i*), here assumed to belong to the large population defined above, *treatment* (participating in OFT) and potential *outcomes*, that are also called *responses* (earnings, labor market status). Y_i^t and Y_i^c denote the outcomes (*t* denotes treatment, *c* denotes control, i.e. no treatment). As a notational convention, capital letters indicate quantities of the population or of members of the population, and small letters denote the respective quantities in the sample. The units of the sample (*n=1,...,N*) are supposed to stem from *N* independent draws in this population. Additionally, denote variables that are unaffected by treatments - called *attributes* by Holland (1986) - by X_i. It remains to define a binary *assignment* indicator S_i, that determines whether unit *i* gets the treatment ($S_i = 1$) or not ($S_i = 0$). If *i* participates in OFT the actual (observable) outcome (Y_i) is Y_i^t, and Y_i^c, otherwise. This notation points to the fundamental problem of causal analysis. The causal effect, for example defined as difference of the two potential outcomes, can never be estimated, because the *counterfactual* (Y_i^t or Y_i^c) to the observable outcome (Y_i) is never observed. However, under certain assumptions the average causal effect of OFT, denoted by θ^0 and defined in equation (1), can be identified:

$$\theta^0 := E(Y^t - Y^c | S = 1) = E(Y^t | S = 1) - E(Y^c | S = 1)\,. \tag{1}$$

The short hand notation $E(\cdot|S=1)$ denotes the mean in the population of all units *i* who participate in training ($S=1$). If the objective is to draw inference only in a subpopulation of $S=1$, defined by attributes contained in *X*, then this and the following expressions are changed in an obvious way.

Finally, to make the model's representation of outcomes adequate for causal analysis, the *stable-unit-treatment-value assumption* (SUTVA) has to be satisfied for all members of the population. Here, the most important implication of SUTVA is that the values of Y_i^t and Y_i^c do not depend on the treatment for units other than unit *i* (e.g. Rubin, 1991). In economic problems when Y_i^t and Y_i^c are outcomes of the market, the latter assumption cannot be precisely fulfilled. Due to the substantial amount of training going on in East Germany after unification such effects - also called general equilibrium effects - could be substantial (and more important than for typically small scale US programs). These effects have two dimensions: On the one hand there are effects on the outcomes due to the changed skill composition of the labor force after training. They effect both, trainees and nontrainees and the total effect appears to be ambiguous. On the other hand there are positive effects of OFT for the non-trainees during the time of training of OFT participants in so far as OFT participation reduces working time or job search intensity. However, satisfying identification of such effects is usually difficult even when the economy is in or close to an equilibrium. It is even more complicated for an economy during a transition process, which is characterized by many severe disequilibria in the labor and product markets. Therefore, the following analysis ignores general equilibrium effects.

The problem for the identification of θ^0 from a large random sample is the term $E(Y^c|S=1)$, because the pair $(y_n^c, s_n = 1)$ is not observable. Much of the literature on causal models in statistics and selectivity models in econometrics is devoted to find identifying assumptions to estimate $E(Y^c|S=1)$ by somehow using the observable pairs $(y_n^c, s_n = 0)$. If participation in OFT were random, the potential outcomes would be independent from the assignment mechanism and it would be true that $E(Y^c|S=1) =$ $E(Y^c|S=0)$. Hence, the untreated population could be used as a control group. Given a large enough sample, their corresponding sample moments are close to these population moments under standard regularity conditions.

However, a brief look at Table A.1 shows that the assumption of random assignment is hardly satisfied. There appear to be several variables that influence assignment as well as outcomes (gender, schooling, etc.).

A weaker condition, called random assignment conditional on a covariate (Rubin, 1977), is to assume that the assignment is independent (denoted by $\coprod$) of the potential nontraining outcome conditional on the value of a covariate or attribute (conditional independence assumption, CIA):

$$Y^c \coprod S \mid X = x \,. \qquad (2)$$

It is assumed that (2) is valid in all the support. If CIA holds, then $E(Y^c|S=1, X=x) = E(Y^c|S=0, X=x)$. Let P*(x)* denote the propensity score that is defined as the participation probability conditional on *x [P(S=1|X=x)]*. If $0 < P(x) < 1$ holds, then $E(Y^c \mid S=1)$ = $E[E(Y^c|S=0, X=x)|S=1]$ can be estimated in large samples using respective sample analogues. Note that the outer expectation operator is with respect to the distribution of *X* in the population of participants (*S=1*). The next section will argue that all variables that could be correlated with assignment and potential outcomes are at least approximately observable in the sample used, so that this restriction is reasonable in this context.

Rosenbaum and Rubin (1983) showed that if CIA is valid, then the estimation problem simplifies: If the potential nontreatment outcome is independent of the assignment mechanism conditional on *X*=*x*, then it is also independent of the assignment mechanism conditional on *P(X)*=*P(x)*, thus:

$$E[Y^c \mid S=1, P(X) = P(x)] = E[Y^c \mid S=0, P(X) = P(x)] \,. \qquad (3)$$

Hence, $E(Y^c \mid S=1) = E\{E[Y^c \mid S=0, P(X)=P(x)] \mid S=1\}$ can be used for estimation. The major advantage of this property is the reduction of the dimension of the estimation problem. The disadvantage is that the probability of assignment is unknown and has to be estimated. However, this estimation may lead to a better understanding of the assignment process itself. Having solved the identification problem, the next step would be nonparametric estimation of the causal effect. When the dimension of the balancing score is not too large and enough components are sufficiently smooth, nonparametric regression methods could be used. In other cases, using matching methods as proposed by e.g. Rubin (1979) and Rosenbaum and Rubin, (1983, 1985) is an attractive alternative. Such methods are used here.

4.2 Estimation of the propensity score

4.2.1 Variables potentially influencing training participation and outcomes

Reasons for participation in OFT are typically identified by supposing that individuals maximize the difference between the present values of future earnings streams for both states (thus ignoring leasure in the utility function). One would like to condition directly on these expected earnings streams, but since they are unobserved, they have to be decomposed into the cost of OFT and the returns of OFT.

Typical labor economists considerations about human capital accumulation, signalling, leisure-labor trade-offs, and costs of OFT that can be found in any labor economics textbook as well as considerations about the availability of subsidies from the labor office suggest that age, expected labor market prospects, actual labor force status, and other socioeconomic characteristics are major factors that could potentially influence the employment decision (for more details see Lechner, 1995). Before going in more details about the groups of variables used in the empirical analysis, I state two assumptions that are important for the particular situation in East Germany after unification, because they make CIA a justifiable assumption.

The first hypothesis is that the complete switch from a centrally planned economy to a market economy in mid 1990, accompanied by a completely new incentive system, invalidates any long term plans that connect past employment behavior to OFT participation. It was generally impossible for East German workers to forsee the impact and timing of the system change and adjust behavior accordingly. This realistic assumption allows me to use all pre-unification variables as attributes. The second assumption is related to the labor market in the rapidly contracting East German economy characterized by rapidly rising unemployment: No individual - having virtually no chances of getting rehired - gives up employment voluntarily to get easier access to training funds. This assumption allows to consider all pre-training information on labor force status as attributes.

Variables used in the empirical analysis to approximate and describe the above-mentioned categories of attributes are age, sex, marital status, educational degrees, and regional indicators. Features of the pre-unification position in the labor market are captured by many indicators including wages, occupation, job position, and employer characteristics. Individual future expectations are described by individual pre-unification predictions about what might happen during the next two years regarding job security, a change in job position or occupation, and a subjective conjecture whether it would be easy to find a new job. Details of the variables, as well as means and standard errors in the OFT and control group are given in Table A.1 (Appendix A). Furthermore, monthly labor force status information is available from mid 1989 to end 1993.

Are there important variables missing? One group could be described as motivation, ability and social contacts. They are approximated with several subjective variables together with the accomplishment of voluntary services in social organizations, memberships in unions and occupational associations before unification, as well as schooling degrees and professional achievements (see Appendix A for a complete list). Another issue is the discount rate implicitly used to calculate present values of future earnings streams. I assume that controlling for factors that have already been decided by using the individual discount rate, such as schooling and professional education, are sufficient.

Finally, empirical papers analyzing training programs in the US point to the importance of transitory shocks before training. Card and Sullivan (1988) for example find declining employment probabilities before training. Here, monthly labor force status data should take care of that problem. Ashenfelter (1978) observes declining mean earnings prior to training. There is no evidence of this phenomenon in the sample used here. This may seem surprising because of increased unemployment prior to OFT. However, during the transition process consumer price inflation adjusted earnings increased so dramatically that the increase in unemployment is easily compensated. Hence, instead of a dip, there is only a flattening (compared to nonparticipants) of the mean earnings profile of OFT participants prior to training.

4.2.2 Econometric considerations

The estimation of the propensity score is not straightforward, because there are potentially important variables, labor force status in a particular month prior to OFT for example, that are related to the months or years previous to the beginning of OFT. Since these dates differ across OFT participants (see Figure 3), they are not clearly defined for the control group. An approximation that might be appealing at first sight is to choose an arbitrary date for the controls and compute the value of these variables regarding this date. However, to have the same date for all controls and different dates for the OFT participants leads to a dependence of this variable on OFT participation, the dependent variable. This dependence is aggravated by the rapidly changing labor market conditions. Therefore, such a variable cannot be considered an attribute or an exogenous variable, so that for example a probit estimation would lead to inconsistent estimates of the propensity score. Consequently, three different ways are suggested below as possible solutions for that problem.

[-- Figure 3 about here -----------------------------------]

The first approach is based on the idea of partioning the propensity score *P(x)* in two components: One part containing the time constant variables (*V*), and another part containing the (problematic) time variant pre-training variables (*M*). Furthermore, it will be more convenient not to partition the propensity score but a function *b(x)* that fulfills *P[S=1|b(x)] = P(x),* called a balancing score by Rosenbaum and Rubin (1983). The underlying reason for this partioning is that *M* is well defined for all controls given a particular start (ie. given a particular group of OFT participants). Hence, *M* cannot be included in a straightforward way in an estimation of a conditional participation probability, but it can nevertheless under some assumptions be used to estimate $EY^c|X$ for a given date of entry into OFT.

Suppose now that the propensity score can be formulated in the following way: $P(x) = P[V\beta^0 + f(M,U) > 0| \quad V = v, M = m]$. U denotes attributes not included in X that are independent of the potential outcomes, but influence OFT participation. β^0 is a fixed parameter vector. If the potential outcomes are independent of *S* conditional on $P(x)$, then it is also true that they are independent from *S* conditional on *b(x)* = ($V\beta^0 = v\beta^0$, $M = m$), because $(v\beta^0, m)$ is a balancing score. Note that the use of the linear index $v\beta^0$ instead of *v* can still lead to a dramatic reduction in the dimension of the conditioning set. The problem then is the consistent estimation of $v\beta^0$ up to scale (and a constant that does not vary in the population). In the following a binary probit model is estimated by maximum likelihood. The basic condition for the consistent estimation of the linear index up to scale is that the conditional expectation of the dependent variable is correctly specified:

$$P(S=1|V\beta^0 = v_n\beta^0) = \Phi(v_n\beta^0), \qquad n = 1,\ldots,N. \tag{4}$$

$\Phi(\mathrm{v}_\mathrm{n}\beta^0)$ denotes the cumulative distribution function of the standard normal distribution evaluated at $v_n\beta^0$. The first of two sufficient conditions for equation (4) to hold is that the propensity score has the additive form $P(x)= P[V\beta^0 + f(M,U) > 0|V = v, M = m]$. This assumption is not so restrictive, because *V* may contain flexible functional forms for the attributes, such as polynomials or interaction terms. The crucial assumption is that:

$$[f(M,U)|[V\beta^0 = v_n\beta^0] \sim N(0,1). \tag{5}$$

$N(0,1)$ denotes the normal distribution with mean 0 and variance 1. Neither the assumption of mean zero nor of unit variance is a problem, because required identification is only up to scale and location. The crucial assumptions are normality and independence with respect to $V\beta^0$. Conditional homoscedasticity (implied by independence) and normality is tested using conventional specification tests (similar to Bera, Jarque, and Lee, 1984, Davidson and MacKinnon, 1984, and Orme, 1988, 1990, see Lechner 1995 for details). The use of semi-parametric methods does not appear to be necessary, because the specification tests indicate no violation of the distributional assumptions required for consistency of the probit model. Furthermore, the consistency property of the specification tests, in particular of such omnibus tests like the information matrix test, might eventually detect any other dependence of $V\beta^0$ and *f(M,U)*.

The drawback of this approach is that it depends on to some extend arbitrary additional assumption beyond CIA and that to some extent it looses the simplicity of the original approach. The results of this approach are compared to two other ways to define propensity scores. Both generate well-defined start dates for all controls. Therefore, the time varying variables can be included in the estimation of the propensity score, so that particularly the above independence assumption is no longer necessary. The first approach assigns each control unit a starting date by drawing in the discrete distribution of start dates as estimated from the OFT participants. The second approach artificially inflates the control pool by treating each control observation in each month from July 1990 to December 1992 (30 months) as a separate observation with the respective specific start date. In the following, I will abbreviate the three approaches as *partial*, *random*, and *inflated*. Note however, that along the lines of the latter two methods, many other ways to 'find' starting dates could be defined, and that although they are intuitively plausible, it appears difficult to derive the properties of these refinements analytically.

4.2.3 Results

For all three definitions of propensity scores a probit model is estimated by maximum likelihood. For *inflated*, the implicit choice based sampling is taken into account by reweighting the observations appropriately. Note that standard test procedures could only be used for *partial*, because they neither take account of the randomness of the start variable (*random*), nor of the correlation of blocks of the size of 30 observations as in *inflated*.

Table 1 presents the results of the probit estimation and the specification tests against heteroscedasticity for *partial*. The estimated coefficients for *random* and *inflated* are similar (see note of Table 3 for the exact *M*-variables included in *random and inflated*). These results as well as the tests against missing variables are available on request. All variables that are not contained in Table 1, but described in Table A.1, as well as different functional forms for the (approximately) continuous variables, and interaction terms between *gender* and variables related to job position and education, are subjected to score tests against omitted variables. None of them appears to be significantly missing at conventional significance levels.

[------------------------------------ Table 1 about here --]

Let us sketch these results briefly. They show regional differences with respect to the specific situation in Berlin (positive) and the northern state of Mecklenburg-Vorpommern (negative). The differences with respect to gender and education manifest themselves cp. basically through a large and significantly higher conditional participation probability for the relatively small group of women with university education (5% of the sample). Taken together, the results in the first part of Table 1 suggest that having a low educational and professional level in the former GDR reduces the probability of OFT participation. This finding is confirmed by the significantly positive coefficient of a high job position.

The estimated effect of gross earnings in 1990 (in 1993 DM) is nonlinear. It attains its maximum at about 1500, which implies a positive earnings effect for the first third of the earnings distribution and a negative one for the remainder part. Individuals who obtained some kind of training while being full-time employed in 1990 have a significantly higher OFT probability. The results show also marked differences regarding occupation and sector: production workers and people working in trade and most service sectors are cp. significantly less likely to be found in OFT. With one exception none of the subjective expectation variables (in 1990) play any role in the (partial) propensity score.

A comparison of Table 1 and Table A.1 reveals that many variables related to marital status, the federal states, motivations and general attitudes, memberships in job related organizations, finer groupings of job positions, occupations and professional degrees, remaining differences between federal states and the sizes of the cities and villages are all superfluous in the estimation of the partial propensity score.

[------------------------------------ Table 2 about here --]

It remains to check some of the stochastic assumptions implied by the mutual independence of the error term $f(M,U)$ and $V\beta^0$, and the normality of $f(M,U)$. First of all, note that the last two columns of Table 1 largely do not contradict the assumption of conditional homoscedasticity. When rejections occur, statistics based on different estimates of the covariance matrix of the test indicators suggest different decisions regarding whether to reject the null of no misspecification or not. This could suggest that in these cases the asymptotic distribution ($\chi(1)$) is a poor approximation in small samples. Resolving this puzzle is left to future work. The normality test as well as the information matrix test given in Table 2 do not reject. In conclusion, the results of the various tests do not provide enough evidence to reject the probit model of the *partial* propensity score. Nevertheless, *partial* could be misspecified in a direction for which the tests have no power.

4.3 Nonparametric estimation of causal effects and matching

The considerations in the previous sections suggest to estimate the causal effects by nonparametric methods to avoid potentially incorrect functional form restrictions. The balancing score property could be used to reduce the dimension of the estimation problem. However at least for *partial,* the balancing score is still high dimensional and has several components that are binary indicators or discrete variables with limited support (see note of Table 3) measuring pre-training labor force status, hence nonparametric regressions or similar methods are unattractive.

For these reasons the matching approach is used. The basic idea is to find for every treated observation a control observation that is as close to it as possible in terms of a balancing score. When an identical control observation is found, the estimation of the causal effects is unbiased. In case of 'mismatches', it is often plausible to assume that using local regressions on these differences will remove the bias (see below). Note that compared to nonparametric regressions, there is typically an efficiency loss, because OFT observation n_t and its closest neighbor in the control population only - instead of possibly larger number of close neighbors - are used. In the language of nonparametric regression, the typical variance-bias tradeoff is ignored in order to minimize bias only.

A basic requirement for the successful implementation of a matching algorithm is a sufficiently large overlap between the distributions of the conditioning variables in both subsamples. In this respect, the results for all approaches used are similar. Although the mass of the distribution of the controls is to the left of the treated, there is overlap for almost all of the treated distribution (details available on request).

The first part of Appendix B contains an exact description of the matching algorithms used for the three different ways of handling time-varying variables described above. Here, it is sufficient to recall that *partial* uses a high dimensional balancing score that includes the partial propensity score; *random* and *inflated* however, use propensity scores that already incorporate the *M*-variables. Therefore, the matching algorithm used for the latter are fairly simply. They draw a treated observation randomly and select the control observation that is closest to it in terms of the propensity score. This is repeated until no OFT observation is left. No individual is used more than once.

The matching algorithm for *partial* is more elaborate, because (i) the balancing score consists of several components, and (ii) the necessity of the additional conditional independence assumption for consistent estimation of $v\beta^0$ might be problematic, despite the nonrejection of the specification tests. Using some components of *v* separately is an additional safeguard against any impact due to inconsistent estimation of $v\beta^0$. The details of the matching algorithm used are described in Appendix B.1. It follows Rosenbaum and Rubin's (1985) suggestion of "matching within calipers of the propensity score" with the exception that window sizes (caliper widths) depend explicitly on the precision of the estimate $v_{n_t}\hat{\beta}$ for the OFT observation n_t. The more precise $v_{n_t}\hat{\beta}$ is estimated, the smaller is the width. The idea is that actual conditioning is on $v_{n_t}\hat{\beta}$ instead of $v_{n_t}\beta^0$. The asymptotic standard error (computed using the delta method) of $v_{n_t}\hat{\beta}$ resulting from the estimation of $\hat{\beta}$ can be considerable and ranges from 0.19 to 0.96 (median: 0.29) in the OFT sample, and from 0.17 to 1.55 (median: 0.28) in the control sample. This is similar to downweighting imprecisely estimated components of the partial propensity score. The exact *V* and *M*-variables are given in the note to Table 3. It should be noted that all three estimation procedures as they are stated above implicitely make the assumption that calendar time has only an effect by changing the value of other time varying variables included in the balancing score.

[------------------------------------- Table 3 about here ---]

Table 3 shows descriptive statistics of a selection of interesting attributes in the treated and in different control samples as well as two measures summarizing the match quality of the different approaches. Column (2) gives the marginal means for the entire unmatched control sample. For the matched control sample, these numbers are given in col. (3) for the matched sample using the *partial* approach, in col. (4) for the matched sample using the *random* approach, and in col. (5) using the *inflated* approach. Finally, col. (6) contains the corresponding numbers for OFT participants. With respect to $v\hat{\beta}$ and $x\hat{\beta}$, it is not surprising that *inflated* is clearly superior, because by considering every single month as a single observation, the control pool is much larger than for *random* or *partial*. With respect to gender, region, and earnings the means are very similar across the matched control and the treated group, although the standard deviation of earnings is too high. However, considerable differences appear with respect to education, and in particular to job position and monthly pretraining labor force status. The worst here is clearly *inflated*. By 'optimizing' the match on $x\hat{\beta}$, it substitutes high education that has a positive correlation with participation by high unemployment (having also a positive correlation). This leads to a very high share of unemployed people (for example) and a very low share of people with high education and job position. Although this is formally what matching on a summary statistics like the propensity score is all about, it clearly casts considerable doubts about the evaluation results. If a positive effect of OFT on unemployment was found, there would always be the suspicion that this is just because there are too many unemployed individuals in the control group. It is very likely that the mismatch for *inflated* could be reduced by including these mismatched variables explicitly as additional variables in the matching. However, since the major attractiveness of *inflated* (and *partial*) is that the propensity score is the only matching variable, this is not done here. The *random* approach is also plagued to some extend by this problem in particular with respect to the variable job position, whereas *partial* appears to be the least affected one. The final two rows of Table 3 contain two summary statistics of the match for the variables mentioned in that table. MSB denotes the median of the absolute biases of the means (ie. the differences in means) normalized by the average standard deviation (e.g. Rosenbaum and Rubin, 1985) for the variables mentioned in the table. JW denotes a quadratic distance measure for the mean biases weighted by the inverses of their covariance matrix (see note on Table 3 for details). Both statistics rank the matching methods in the same order. They suggest that *random* is best and *inflated* is worst. However, for MSB *partial* and *random* are very close. The order of *random* and *partial* is somewhat surprising given the marginal means. The technical reason for the large value of the quadratic measure JW is the mismatch of the partial propensity score with a standardized bias of 16%. One might want to argue that the difference between *partial* / *random* and *inflated* points to a more substantial problem affecting all three methods. This may be true, but one should keep in mind that they treat the time dimension fairly differently. Section 4.4 comes briefly back to the issue of match quality.

$\hat{\theta}_{N^t}$,

4.4 *Evaluation*

4.4.1 Outcomes

Unemployment and full-time employment are measured on a monthly basis. Earnings are measured once a year. For those being employed, earnings is defined as the gross monthly earnings in the month previous to the interview. For those not being employed, either zeros or imputed benefits are used instead (see Appendix A for details). Using zeros makes the outcome measure a proxy for productivity, whereas if benefits are included it is interpreted as the gross earnings result of the training for the individual. In addition to these variables, labor market prospects are measured once a year as individual expectations. Those results are available on request from the author.

4.4.2 Econometric issues

To simplify notation, assume that observations in the sample are arranged such that the first N^t observations receive OFT, and the remaining (N-N^t) observations do not. Define the differences in matched pairs in the sample as $\Delta y_{n_t} = y^t_{n_t} - y^c_{n_t}$, $\Delta x_{n_t} = x^t_{n_t} - x^c_{n_t}$, $n_t = 1,\dots,N^t$, where $y^c_{n_t}$ and $x^c_{n_t}$ denote values of an observation from the pool of individuals not participating in OFT (controls) that is matched to the treated (OFT) observation n_t. If the outcomes are continuous variables, e.g. earnings, then Δy_{n_t} is continuous. Otherwise, the outcomes are measured with indicators (0, 1) and $\Delta y_{n_t} \in \{-1,0,1\}$. The estimate of the average causal effect and the respective standard error are computed as:

$$\hat{\theta}_{N^t} = \frac{1}{N^t}\sum_{n=1}^{N^t} \Delta y_n, \qquad Var(\hat{\theta}_{N^t}) = \frac{1}{N^t}(S^2_{y^t} + S^2_{y^c}). \tag{6}$$

$S^2_{y^t}$ and $S^2_{y^c}$ denote the square of the empirical standard deviation of Y in the OFT sample and in the sample matched to the OFT-sample, respectively. The variance estimate exploits the fact that the matching algorithm given in Appendix B.1 never chooses an individual twice. As mentioned in the previous section, when a perfect match is achieved, implying that $\Delta b(x_n) = b(x_n) - b(x_j) = 0$, $n = 1,\ldots,N^t$, these estimates are unbiased (cf. Rosenbaum and Rubin, 1983). When the sample is large enough, the normal distribution can be used to perform tests and compute asymptotic confidence intervals.

Equation (6) gives the basic nonparametric estimate of the causal effect to be refined in the following to take account of the time before and after OFT. Denote by N_τ^t, $\tau \in \{\ldots,-3,-2,-1,1,2,3,\ldots\}$ the number of pairs observed at any distance to OFT. Let $\iota_\tau(n)$ = 1 if observation n is observed at distance τ, so that:

$$\hat{\theta}_{N_\tau^t} = \frac{1}{N_\tau^t} \sum_{n=1}^{N^t} \iota_\tau(n) \Delta y_{n,\tau}\,, \qquad \tau \in \{\ldots,-3,-2,-1,1,2,3,\ldots\}\,; \tag{7}$$

The variances are computed appropriately. When τ is negative, then $\hat{\theta}_{N_\tau^t}$ denotes the mismatch in period τ before OFT, otherwise it denotes the effect of training in period τ after OFT. $\hat{\theta}_{N_T^t}^T$ indicates the accumulated effect T periods after OFT. These effects are also computed for subpopulations defined by attributes or training characteristics.

Let us now consider the case when the match is not perfect. In general, equation (8) holds:

$$E[\Delta y_n | \Delta b(x_n) = 0] = E\{E[\Delta Y | b(X) = b(x_n)] | S = 1\} = \theta^0\,. \tag{8}$$

However, $\Delta b(x_n)$ may not be exactly zero. The exact type of the suggested correction depends on whether the outcome variables are continuous or discrete. In the case of continuous variables it is reasonable to assume that the conditional expectation of the dependent variable is linear in $\Delta b(x_n)$, because matching has already removed almost all differences in the balancing scores, so that the $\Delta b(x_n)$ are local deviations:

$$E[\Delta y_n | \Delta b(x_n) = \eta_n] = \theta^0 + \eta_n \lambda^0, \qquad n \le N^t. \tag{9}$$

λ^0 denotes an unknown coefficient vector. Local smoothing using a linear conditional expectation is not very restrictive; standard linear regression methods can be used to estimate the average treatment effect θ^0 by regressing the differences in outcome on the balancing score and a constant (cf. Rubin, 1979).

With binary outcomes the treatment effect can be written as:

$$\theta^0 = E(\Delta Y | S=1) = P(\Delta Y = 1 | S=1) - P(\Delta Y = -1 | S=1) \tag{10}$$

A consistent estimate of the average treatment effect can be obtained by substituting sample analogues for the population probabilities:

$$\hat{\theta}_{N^t} = \frac{1}{N^t} \sum_{n=1}^{N^t} \{P[\Delta y_n = 1 | \Delta b(x_n) = 0] - P[\Delta y_n = -1 | \Delta b(x_n) = 0]\} \tag{11}$$

Using a linear approximation for these differences of probabilities is not so attractive as before, except $\Delta b(x_n)$ is very small. A more parsimonious specification is the following: In a first step, a three-group-ordered probit model is estimated with Δy_n as dependent variable and $\Delta b(x_n)$ plus a constant as independent variables (one bound and the variance of the underlying latent linear model is normalized). As a second step, the above probabilities are directly derived from this model and computed for the individual observations using the estimated coefficients of the ordered probit model. Finally, the variance of $\hat{\theta}_{N^t}$ is approximated from the variance of the estimated coefficients of the ordered probit model by the delta method. Note that the functional form assumption for the conditional mean of ΔY is asymptotically unimportant as long as the differences in $\Delta b(x_n)$ disappear.

A similar approach as for the mismatch corrections is chosen to check whether the treatment effects vary either with characteristics of the courses, such as its duration, or with characteristics of the individuals participating in OFT by including the levels of such variables as additional regressors. This procedure is not nested in the previous one, because now the assumption that either the treatment effect is stable or varies in a particularly specified way is indispensable. Therefore, splitting the samples in subpopulations and performing estimations in these subpopulations that do not require such an assumption is an attractive alternative for discrete attributes. However, when the attributes have too many different values some modeling is required given the size of the sample. For more details see Lechner (1995).

4.4.3 Results

A selection of the results is given in Figure 4 and Tables 4 and 5. They show the differences between the control and the OFT group for specific time spans before and after the training for the unemployment and earnings variable. Figure 4 and Table 4 cover up to 18 months, Table 5 up to 3 'years' (the first year is only the difference from the start of training to the last interview) before the training, and up to 24 months or 3 years after OFT. They show the mean effect (solid line; + for mismatch corrected estimate) and its 95% confidence interval based on the normal approximation (dashed line; ∇, Δ for the mismatch corrected estimates), respectively its standard deviation.

The number of observations available to compute the respective statistics decrease the longer the distance to the incidence of OFT is (see Tables 4 and 5). The implications of this are threefold: First of all, the variance increases. Although this is reflected in the widening of the confidence interval, the accuracy of the estimated interval itself may deteriorate, because the normal distribution may not be a very good approximation of the sample distribution when the sample gets too small. Finally, a mismatch correction may be impossible or very imprecise, because there may be too few observations to identify and estimate the parameters of the ordered probit model. Note here is also an important difference between *partial* and the other two approaches. For the latter two it is allowed and it appears many times in the application to compare treated and controls at different points of time. Hence, the overall time window that could be used for estimation (both treated and controls must be observed) is smaller than for *partial*, and hence fewer observations are available to compute the training effects some time after OFT.

[-------------------------------------- Figure 4 about here ---------------------------------]

[-------------------------------------- Table 4 about here ---------------------------------]

Figure 4 presents the monthly unemployment status using the *partial* approach. The part left to the 0 vertical mark allows a judgment about the quality of the matches concerning the particular variable. Testing whether these lines deviate significantly from zero is in the same spirit as the tests suggested by Rosenbaum (1984) to use overidentifying restrictions to try to invalidate CIA. The pre-OFT outcomes here are denoted as unaffected outcomes in his terminology. Table 4 compares the result of partial with the other two matching approaches for some selected time intervals. With resepect to pre-training unemployment neither *partial* nor *random* exhibit any problems. However, as expected after the discussion of Table 3, there are significantly too many unemployed individuals in the control group produced by *inflated*.

Figure 4 shows that the effect of training appears to be higher unemployment in the months directly following its end. This is plausible if one takes into account that for those unable to keep their previous occupation job search is required. Since this is time consuming, it may not be performed with full intensity until OFT ends. Meanwhile, more members of the control group already found a new employment. These effects disappear entirely after about 6 to 12 months. This view is confirmed when considering only a sample of individuals not employed during OFT. However, for the remaining sample, which is employed during OFT, the considerations about a negative initial effect are not important. Here, OFT does not seem to have any impact at all. These conclusions are confirmed by considering full-time employment instead of unemployment. The corrected estimates do not deviate substantially from the uncorrected. This might be an indication that matching with *partial* could be considered successful.

For *random*, about the same picture emerges for the unadjusted estimates, however the initial negative effects are absent from the corrected estimates (not shown in Table 5). Probably due to the excess amount of unemployed individuals in the *inflated* control group, no initial negative effect appears and the unadjusted estimates indicate even a significantly positive effect of OFT several months later. However, the adjusted estimates do not confirm any positive effect. These finding cast considerable doubt on the results based on *inflated*, but they also seem to indicate that the adjustment mechanism works in the right direction. These results are qualitatively confirmed when for example employment is used as outcome measure instead of unemployment.

[------------------------------------- Table 5 about here ---]

The second group of outcome variables is only measured once a year, such as gross monthly earnings (Table 5), being very worried about keeping one's job, and expected improvement or decline in the professional career during the next two years. On the one hand, there are no significant differences for the pre-training outcomes. On the other hand, the same is true for the post-treatment period. This general result is valid for all yearly variables. It is also robust concerning other functional forms (such as logs) or the coding of earnings for nonemployed persons.. Similarly, no qualitative differences appear with respect to the use of the three different control groups.

To check whether there might be differences of the average treatment effects in specific sub-groups, the sample is divided according to gender, job position, occupational degree, and as already mentioned, whether the individual was employed during OFT (see Table A.1 for the value of N^t corresponding to the particular partition). No significant differences appear. Finally, to check the results for sensitivity with respect to the definition of OFT, the courses used in the estimation are split in several subsamples according to whether (i) they began not earlier than January 1991 (N^t = 108), (ii) they have a minimum duration of one week (N^t = 95), (iii) the objective is qualification for promotion (N^t = 45) or (iv) the adjustment of skills (N^t = 84), and whether (v) a certificate has been obtained by the participant that could be helpful for future job applications (N^t = 101). As a final sensitivity check I also considered a control and treatment group that did not participate in any other form of training (N^t = 108). None of the subsamples reveals a substantial difference compared to the results presented above. The conclusions drawn above regarding match quality and nonexisting OFT effects are adequate for the second perspective of time, i.e. specific dates, as well. As a further check against the influence of possible mismatches on outcomes the *partial* matching method is also changed to allow the use of a single control observation more than once. However, this does not change any of the qualitative results, which is not surprising because mismatch is not an important problem for *partial* (90% of the observations are only used once).

Summarizing the results presented in this subsection, it should be stressed that no robust positive effects of OFT can be found, and even some temporary negative effects surfaced. There is (at least) one caveat however. Figure 2 shows that the median duration is only one month (in full-time equivalents), hence there are many short courses. With this sample size it would be very difficult to detect their effects, and this would make it more difficult to detect overall positive effects. However, applying the same methodology, Lechner (1996) studies public-sector-sponsored training in East Germany. Those courses, some of them are also included here, have a median duration of about 9 months (no intensity information available), but still the result is the same: no positive effects.

These results are in contrast to more positive findings obtained in a recent study by Fitzenberger and Prey (1996, FP). FP use the first eight waves of the *Arbeitsmarktmonitor* that is not as informative as the GSOEP (there is no monthly labor force status information and much less information about training, e.g. no exact information about duration, etc.), but contains considerably more observations. To correct for observed and unobserved selectivity they model the outcome and the participation process using joint normality and a particular random effects specification for the joint error covariance matrix. However, it is my opinion that FP shares the problems of all model based evaluation procedures by identifying the training effect with a combination of (latent) linearity of conditional expectations and distributional assumptions (joint normality and covariance restrictions) of the error terms. It is difficult to discuss the validity of this kind of identifying assumptions (that are partly necessary because of the not so informative data) in terms of the economic problem at hand.

5 Conclusion

The major empirical result of this paper is that no robust positive effects of OFT were found. There are three possible reasons for this: First of all, the true effects can be so small that they are impossible to determine with the available sample size. Secondly, there could be positive effects in the longer run that could not be seen yet. Finally, it could be that there are no positive effects at all. However, although the study raises serious doubts, one should be cautious to conclude that the training part of the active labor market policy in East Germany has no positive impact even in the shorter run. The definition of off-the-job training used in this paper includes many courses that are not subsidized.

The results are obtained by using the potential outcome approach to causality as a general framework to define causal effects of off-the-job training on individual actual and future post-training labor market outcomes. The paper argues that due to the specific situation in East Germany after unification and due to the rich data available, the assumption that the outcomes and the assignment mechanisms are independent conditional on observed attributes, including monthly pre-training labor force status, is very plausible. Hence, the identification problem inherent in causal analysis is solved that way. Estimation is performed using three variants of a suitably adapted nonparametric matching approach. In conclusion, this nonparametric approach appears to be well suited for such an analysis.

Future research should jointly investigate the effects of different types of training, such as on-the-job training versus off-the-job training versus no training at all. Furthermore, other ways of handling the problem of different starting dates might be an interesting topic. This paper is just a starting point showing that matching could fruitfully employed even when there are different entry dates for the trainees. More systematic research exploring the exact properties and underlying assumptions of different ways to tackle this problem would be surely useful.

Acknowledgements

Part of this paper was written while I was visiting the Center for European Studies (CES), Harvard University, Cambridge, during the academic year 1994 / 1995. Financial support from the CES and the Deutsche Forschungsgemeinschaft (DFG) is gratefully acknowledged. The paper benefited much from ideas disseminated in lectures by Guido Imbens and Donald Rubin. I thank Klaus Kornmesser for competent help with the data and Irene Bertschek, Richard Blundell, Bruno Crepon, Martin Eichler, Bernd Fitzenberger, Francois Laisney, Markus

Pannenberg, Friedhelm Pfeiffer, Hedwig Prey, Viktor Steiner, and participants of workshops at the University of Freiberg, the IFS, London, and of the Winter-meeting of the Econometric Society in Copenhagen, 1996, for helpful comments and suggestions on a previous draft of this paper. The comments of two referees and an Associate Editor of this journal proved to be particularly helpful in improving the paper considerable. Furthermore, thanks also to Viktor Steiner for supplying a valuable piece of data, and to Angelika Klein and Anisa Boumrifak for carefully reading the manuscript. All remaining errors are my own.

Appendix A: Data

This appendix briefly explains the coding of the start, duration, and end date of OFT courses. Furthermore, the exact definition of earnings variables used in the evaluations are given. Finally, Table A.1 shows descriptive statistics of all variables used in the estimation.

The first month of the course is directly indicated by the individual. If there are several courses classified as OFT, the start date is coded as the earliest one. A problem is the imprecise measurement of the duration and, therefore, also of the ending date of OFT, because there is only categorical information available (Categories: 1 day, up to 1 week, up to 1 month, up to 3 months, up to 1 year, up to 2 years, more than 2 years). In the empirical analysis, the monthly durations are computed by using the mid-point of the duration intervals multiplied by the appropriately rescaled hours per week. However, this problem is reduced by combining the information in the calendar variables with the special-survey variables to adjust the duration and ending dates. In cases of several OFT courses, the single durations are added. The last month of each course is computed using the endpoint of the duration intervals to make sure that post-training outcomes are really *post*-training. Note that this is only important for courses of a duration of more than one month. The resulting measurement error for these courses is reduced by using additionally monthly calendar information on training. In cases of several courses, the end date is coded as the end date of the last course.

There is a special problem related to the training survey: about 19% of training participants attended more than 3 courses. No information is available on these additional courses. However, the 'lost courses' have been rather short and/or began very early (that is before unification) to fit into the three year time span used by the special survey. Hence, they are unimportant for this study.

[---------------------------------- Table A.1 about here ---]

Gross monthly earnings is only measured for those employed. Due to the selection criteria that creates a sample of full-time employees in mid 1990 it is not a problem for 1990, but for the following years. For those unemployed, unemployment benefits are computed using 67% of the last *gross* earnings. After performing these imputations, it is ensured that earning levels are not below average social assistance levels (Bundesministerium für Arbeit und Sozialordnung, 1994, Table 8.16A). Finally, all earning variables are converted to 1993 DM by using the private consumption price index for East Germany (Bundesministerium für Arbeit und Sozialordnung, 1994, Table 6.9, and Institut der Deutschen Wirtschaft, 1994, Table 8).

Appendix B: Econometrics

B.1 Matching protocol

This section gives the details of the matching protocol used for the *partial* approach.

Step 0: Estimate a probit model to compute $v\hat{\beta}$ and its conditional variance $Var(V\hat{\beta}|V=v)$ for each observation. The latter is derived from $Var(\hat{\beta})$ by the delta method.

Step 1: Split observations in two exclusive pools according to whether they participated in OFT (T-pool) or not (C-pool).

Step 2: Draw randomly an observation in T-pool (denoted by n_t) and remove from T-pool.

Step 3: Define caliper of partial propensity score for observation n_t in terms of the predicted index $v_{n_t}\hat{\beta}$ and its conditional variance $Var(V\hat{\beta}|V = v_{n_t})$.

Step 4: Find observations in C-pool (denoted by j) obeying $v_j\hat{\beta} \in [v_{n_t}\hat{\beta} \pm c\sqrt{Var(v_{n_t}\hat{\beta})}]$. The constant c is chosen so that the interval is identical to a 90% confidence interval around $v_{n_t}\hat{\beta}$.

Step 5: (a) If there is only one or no observation in this interval: find observation j in C-pool that is closest to observation n_t, so that it minimizes $(v_j\hat{\beta} - v_{n_t}\hat{\beta})^2$.

(b) If there are two or more observations in this set generated by Step 4: take these controls and compute the variables m in relation to the start date of observation n_t. Denote these and perhaps other variables already included in V as $\tilde{m}_j$ and $\tilde{m}_{n_t}$, respectively. Define a distance between each control j and i as $d(j, n_t) = (v_j\hat{\beta}, \tilde{m}_j)' - (v_{n_t}\hat{\beta}, \tilde{m}_{n_t})'$. Choose control j so that it has the smallest Mahalanobis distance $d(j, n_t)'Wd(j, n_t)$ within the caliper. W denotes the inverse of the estimated variance of $(v\hat{\beta}, \tilde{m})'$ in the C-pool (computed for December 1991).

Step 6: Remove j from C-pool.

Step 7: If there are any observations in the T-pool left, start again with step 2.

This matching protocol is close to the one proposed by Rosenbaum and Rubin (1985) and Rubin (1991). They find that this kind of protocol produces the best results in terms of 'match quality' (reduction of bias). The difference is that instead of using a fixed caliper-width (based on considerations about the true propensity score) for all observations, I allow the widths to vary individually with the precision of a monotone function of the partial propensity score (Step 4). The (unbounded) linear index $v_{n_t}\hat{\beta}$ is used instead of the (bounded) partial propensity score $\Phi(v_{n_t}\hat{\beta})$, because matching on the latter with this kind of symmetric metric leads to an undesirable asymmetry when $\Phi(v_{n_t}\hat{\beta})$ is close to 0 and 1, depending on which side of the control *j* is. Furthermore, defining the balancing score in terms of $(v_j\hat{\beta}, \tilde{m}_j)$ has also the advantage for the *partial* approach to make it easier to state exactly under what conditions this type of condition has similar properties as conditioning on the -for *partial* unknown and not estimable- propensity score itself.

Note that the random and inflated approach have a different initial step:

Step 0 (*random*): Estimate a probit model to compute $x\hat{\beta}$. The values of the time-varying part of *x* for the nonparticipants is computed with respect to a date that is obtained for each individual by an independent draw in the distribution of start dates (see Figure 3).

Step 0 (*inflated*): Estimate a probit model to compute $x\hat{\beta}$. Use participants only as a single observation each. Use every single month between July 1990 and December 1992 of the nonparticipants as a separate observation and compute the values of the time-varying part of *x* with respect to that date. Take account of this kind of choice based sampling by reweighting the likelihood function.

The algorithm used for matching with *random* and *inflated* is simpler, because the propensity score is the only matching variable. Therefore, it uses only Steps 1, 2, 5 (a), 6, and 7 ($v\hat{\beta}$ should be substituted by $x\hat{\beta}$). For *inflated*, a modification of Step 6 is used, because 30 monthly observations from the same individual are in the C-pool.

Step 6 (*inflated*): Remove *j* from C-pool. Remove also all observations from C-pool that belong to the same individual as observation j.

Not using the same individual twice greatly simplifies the variance computations for the unadjusted and adjusted estimates.

B.2 Correction for mismatches: homogenous effects

The question here is whether the price to pay for the use of the suggested regression methods to adjust for differences in attributes and course characteristics is the assumption of a homogenous treatment effect. This can be seen by considering whether such a regression can identify the mean causal effect θ^0, even if the individual causal effect is not constant in the population. Assume that the following linearity condition holds (given matching has already be performed in an unspecified way):

$$E(\Delta Y|S=1, \Delta X = \Delta X_i; \theta_i^0, \lambda^0) = \theta_i^0 + \Delta X_i \lambda^0 \,. \tag{B.1}$$

λ^0 is an unknown coefficient vector. For illustration assume that the matches remain imperfect, so that ΔX_i may be different from 0. Note that, for simplicity, this is again an argument about identification in the population only. Define the following population means: $\overline{\Delta Y} = E\Delta Y|S=1$, $\theta^0 = \bar{\theta} = E\theta|S=1$, $\overline{\Delta X} = E\Delta X|S=1$, $\overline{\Delta XX} = E(\Delta X'\Delta X)|S=1$, and $\overline{\Delta X\theta} = E[\Delta X'(\theta - \bar{\theta})]|S=1$. $\hat{\theta}_\infty$ denotes the population (probability) limit for the constant term of an OLS regression of a constant and the difference in attributes on the difference of an outcome. It can be computed by using the Frisch-Waugh-Lovell theorem (cf. Davidson and MacKinnon, 1993):

$$\hat{\theta}_\infty = \theta^0 + [1 - \overline{\Delta X}(\overline{\Delta XX})^{-1}\overline{\Delta X}'\,]^{-1}\overline{\Delta X}(\overline{\Delta XX})^{-1}\overline{\Delta X\theta}' \,. \tag{B.2}$$

Generally, the estimated OLS coefficient of the constant will not converge towards the population mean, unless $\overline{\Delta X\theta'}$ is zero. This is true if the difference regressors and the causal effects are uncorrelated. A very important case is if θ is the same for all members of the population, another important case is the case of perfect matches (in this case the notation has to be changed to allow for noninvertible matrices). However, note that the bias is reduced when the match quality increases and when effects are becoming more homogenous in the (sub-) population. Similar arguments apply to the nonlinear case.

References

Ashenfelter, O. (1978): "Estimating the Effect of Training Programs on Earnings ", *The Review of Economics and Statistics*, 58, 47-57.

Ashenfelter, O. and D. Card (1985): "Using the Longitudinal Structure of Earnings to Estimate the Effect of Training Programs", *The Review of Economics and Statistics*, 67, 648-660.

Bera, A., C. Jarques, and C. F. Lee (1984): "Testing the Normality Assumption in Limited Dependent Variable Models", *International Economic Review*, 24, 21-35.

Bundesministerium für Arbeit und Sozialordnung (1994), *Statistisches Taschenbuch 1994: Arbeits- und Sozialstatistik*, Bonn.

Card, D. and D. Sullivan (1988): "Measuring the Effect of Subsidized Training Programs on Movements in and out of Employment", *Econometrica*, 56, 497-530.

Davidson, R. and J.G. MacKinnon (1984): "Convenient Specification Tests for Logit and Probit Models", *Journal of Econometrics*, 25, 241-262.

---------------------- (1993): *Estimation and Inference in Econometrics*, Oxford: Oxford University Press.

Dehejia, R. and S. Wahba (1995a): "A Matching Approach for Estimating Causal Effects in

Non-Experimental Studies", Harvard University, *mimeo*.

---------------------- (1995b): "Causal Effects in Non-Experimental Studies", Harvard University, *mimeo*.

Eichler, M. and M. Lechner (1996): "Public Sector Sponsored Continuous Vocational Training in East Germany: Institutional Arrangements, Participants, and Results of Empirical Evaluations", forthcoming in R.T. Riphan, D.J. Snower, K.F. Zimmermann (eds.): *Employment Policy in the Transition: Lessons from German Intergration*.

Fitzenberger, B. and H. Prey (1996): "Training in East Germany: An Evaluation of the Effects on Employment and Earnings", *Unpublished manuscript*, University of Konstanz.

Heckman, J.J. and V.J. Hotz (1989): "Choosing Among Alternative Nonexperimental Methods for Estimating the Impact of Social Programs: The Case of Manpower Training", *Journal of the American Statistical Association*, 84, 862-880 (includes comments by Holland and Moffitt and a rejoinder by Heckman and Hotz).

Heckman, J.J. and R. Robb (1985): "Alternative Methods of Evaluating the Impact of Interventions", in: J.J: Heckman and B. Singer (eds.), *Longitudinal Analysis of Labor Market Data*, New York: Cambridge University Press.

Heckman, J.J., H. Ichimura, J. Smith, and P. Todd (1996): "Characterizing Selection Bias Using Experimental Data", *mimeo*.

Heckman, J.J., H. Ichimura, and P. Todd (1997): "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Program", *mimeo*.

Holland, P.W. (1986): "Statistics and Causal Inference", *Journal of the American Statistical Association*, 81, 945-970 (includes comments by Cox, Granger, Glymour, Rubin and a rejoinder by Holland).

Institut der Deutschen Wirtschaft (1994): *Zahlen zur wirtschaftlichen Entwicklung der Bun-*

desrepublik Deutschland 1994, Köln.

Lalonde, R.J. (1986): "Evaluating the Econometric Evaluations of Training Programs with Experimental Data", *American Economic Review*, 76, 604-620.

Lechner, M. (1995): "Continuous Off-the-job Training in East Germany after Unification", *ZEW-discussion paper*, 95-27.

Lechner, M. (1996): " The Effects of Enterprise-related Continuous Vocational Training in East Germany on Individual Employment and Earnings", *Beiträge zur angewandten Wirtschaftsforschung*, # 542-96, University of Mannheim Discussion Papers.

Maddala, G.S. (1983): *Limited-dependent and Qualitative Variables in Econometrics*, Cambridge: Cambridge University Press.

Orme, C. (1988): "The Calculation of the Information Matrix Test for Binary Data Models", *The Manchester School*, 56, 370-376.

Orme, C. (1990): "The Small Sample Performance of the Information Matrix Test for Binary Data Models", *Journal of Econometrics*, 4, 529-559.

Pannenberg, M. and C. Helberger (1994): "Kurzfristige Auswirkungen staatlicher Qualifizierungsmaßnahmen in Ostdeutschland: Das Beispiel Fortildung und Umschulung", to appear in: *Schriftenreihe des Vereins für Sozialpolitik*.

Rosenbaum, P.R. (1984): "From Association to Causation on Observational Studies: The Role of Tests of Strongly Ignorable Treatment Assignement", *Journal of the American Statistical Association*, 79, 41-48.

Rosenbaum, P.R. and D.B. Rubin (1983): "The Central Role of the Propensity Score in Observational Studies for Causal Effects", *Biometrica*, 70, 41-50.

------------------------ (1985): "Constructing a Control Group Using Multivariate Matched Sampling Methods That Incorporate the Propensity Score", *The American Statistician*, 39, 33-38.

Rubin, D.B. (1974): "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies", *Journal of Educational Psychology*, 66, 688-701.

--------------- (1977): "Assignment of Treatment Group on the Basis of a Covariate", *Journal of Educational Statistics*, 2, 1-26.

--------------- (1979): "Using Multivariate Matched Sampling and Regression Adjustment to Control Bias in Observational Studies", *Journal of the American Statistical Association*, 74, 318-328.

--------------- (1991): "Practical Implications of Models of Statistical Inference for Causal Effects and the Critical Role of the Assignement Mechanism", *Biometrics*, 47, 1213-1234.

Sobel, M.E. (1994): "Causal Inference in the Social and Behavioral Sciences", in: G. Arminger, C.C. Clogg and M.E. Sobel (eds.): *Handbook of Statistical Modeling for the Social and Behavioral Sciences*, New York: Plenum Press.

Wagner, G.G., R.V. Burkhauser, and F. Behringer (1993): "The English Language Public Use File of the German Socio Economic Panel", *Journal of Human Ressources*, 28, 429-433.

Titles and Notes for the Figures

Figure 1: Difference of unemployment rates in OFT group and randomly chosen control group in %-points

Note: *95%-interval* denotes the 95% confidence interval for the mean based on the normal distribution.

Figure 2: Empirical distribution functions for duration of training

Note: The remaining part (12 to 18 months) of the cdf is omitted. Censored refers to the sample with completed spells by December 1993. The dashed line denotes the median. Duration is measured in full-time equivalents, assuming 38 hours per week.

Figure 3: Distribution of OFT start dates

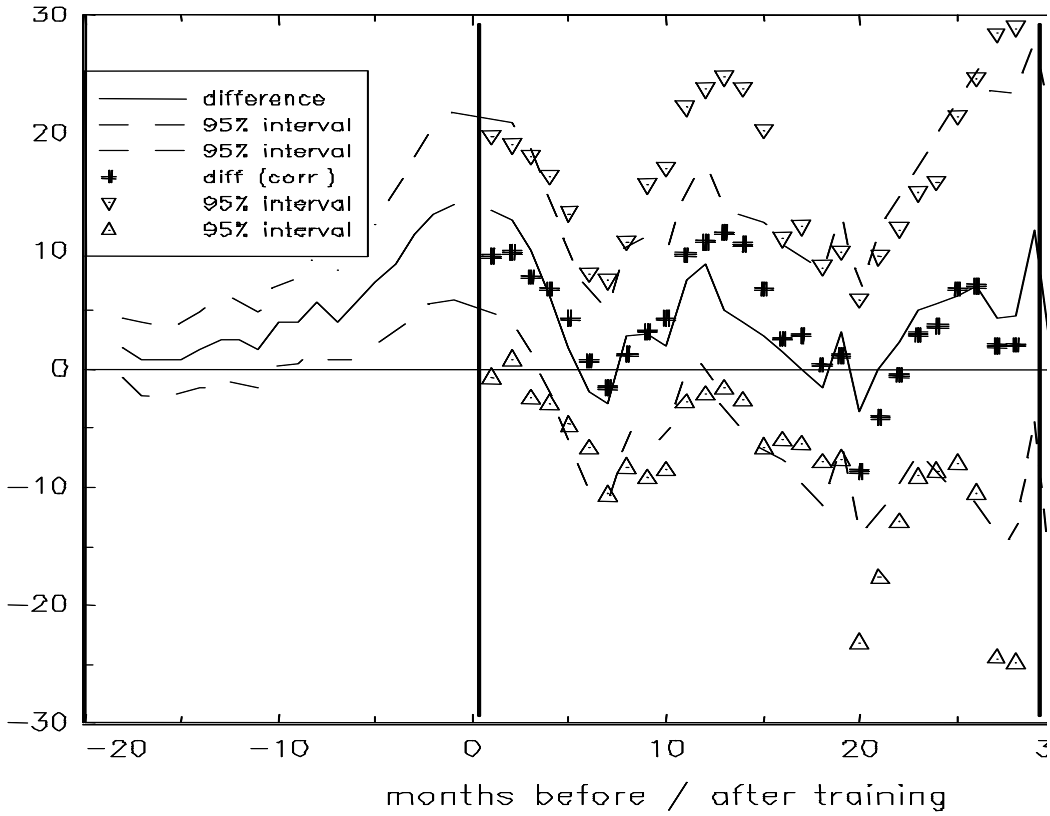
Note: Monthly information. Censored refers to the sample with completed spells by December 1993.

Figure 4: Registered unemployment (partial)

Note: N_t = 131. *Difference* denotes the difference of the means in the treated and the matched control group. *95% interval* denotes the 95% confidence intervals of the respective differences based on the normal distribution. *(corr.)* indicates that the estimates are adjusted for mismatches.

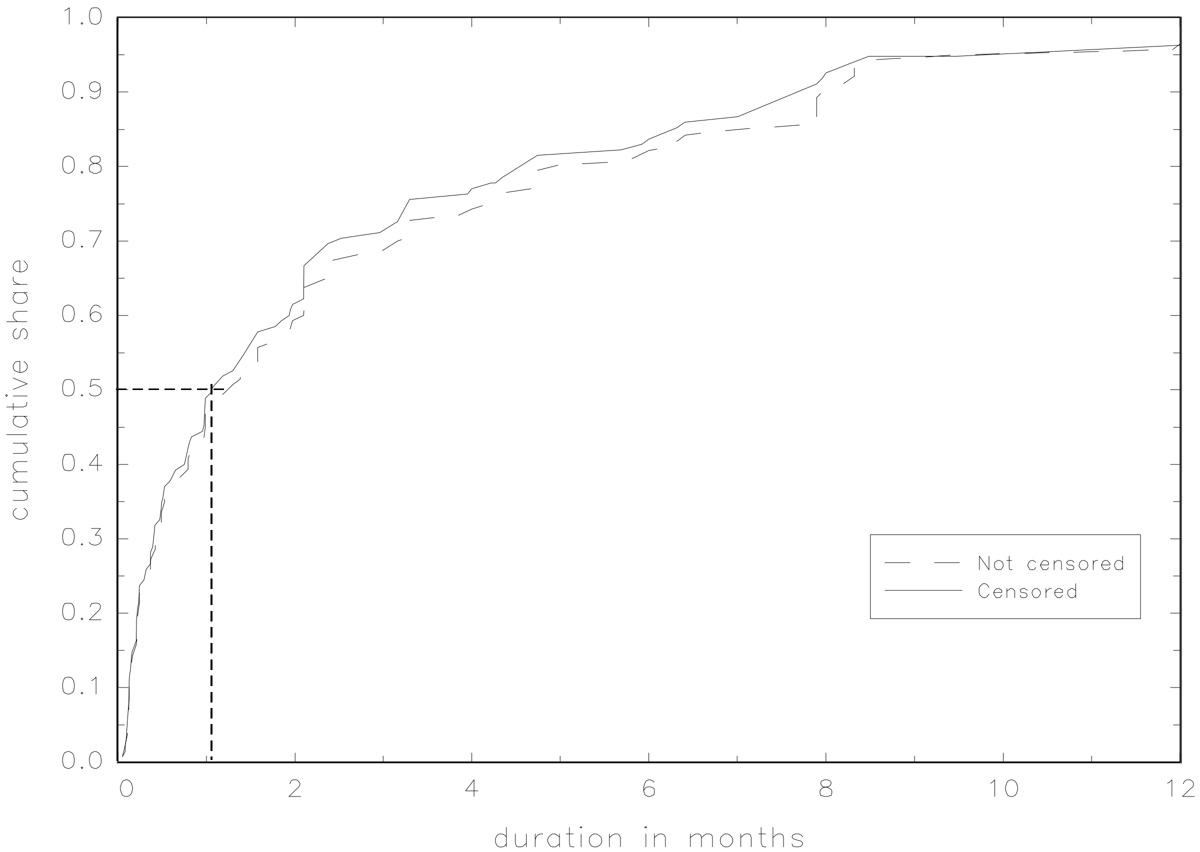
Figures

Figure 1: Difference of unemployment rates in OFT group and randomly chosen control group in %-points



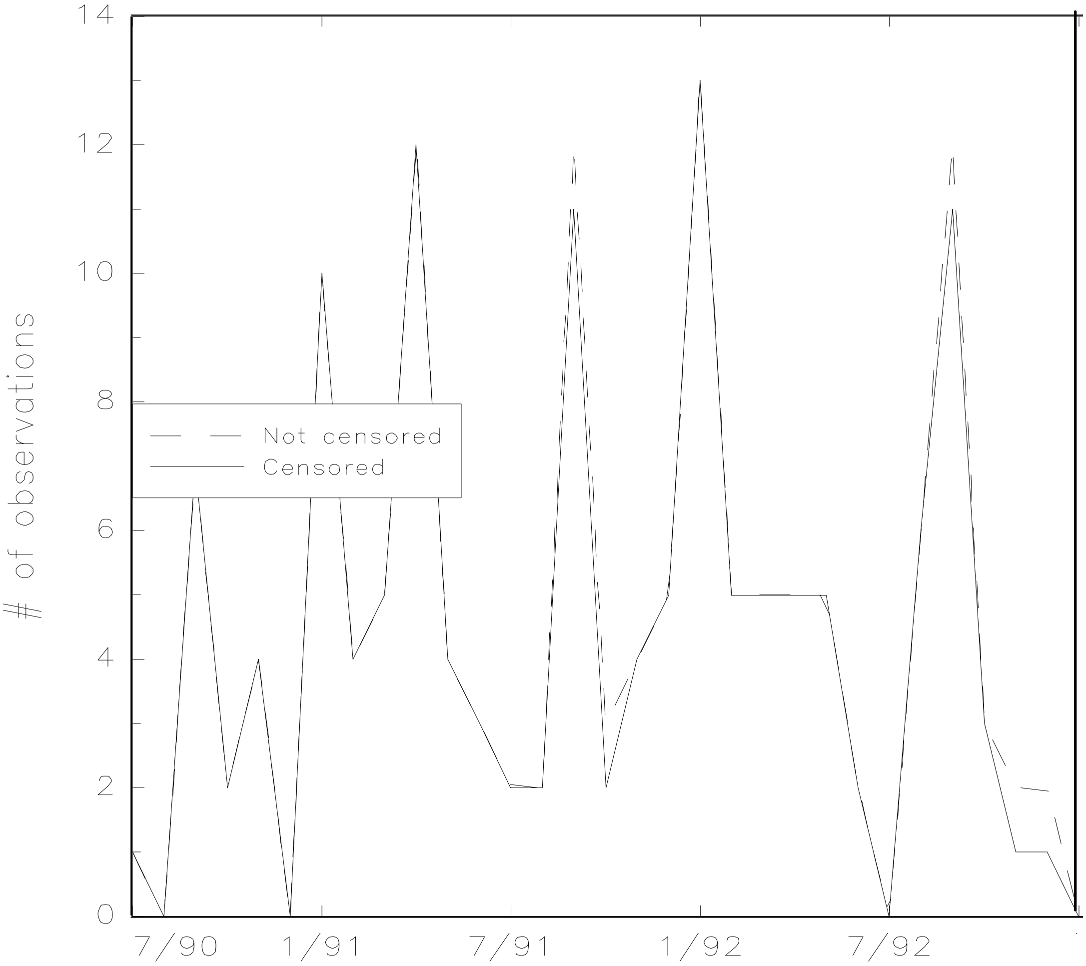
Note: *95%-interval* denotes the 95% confidence interval for the mean based on the normal distribution.

Figure 2: Empirical distribution functions for duration of training



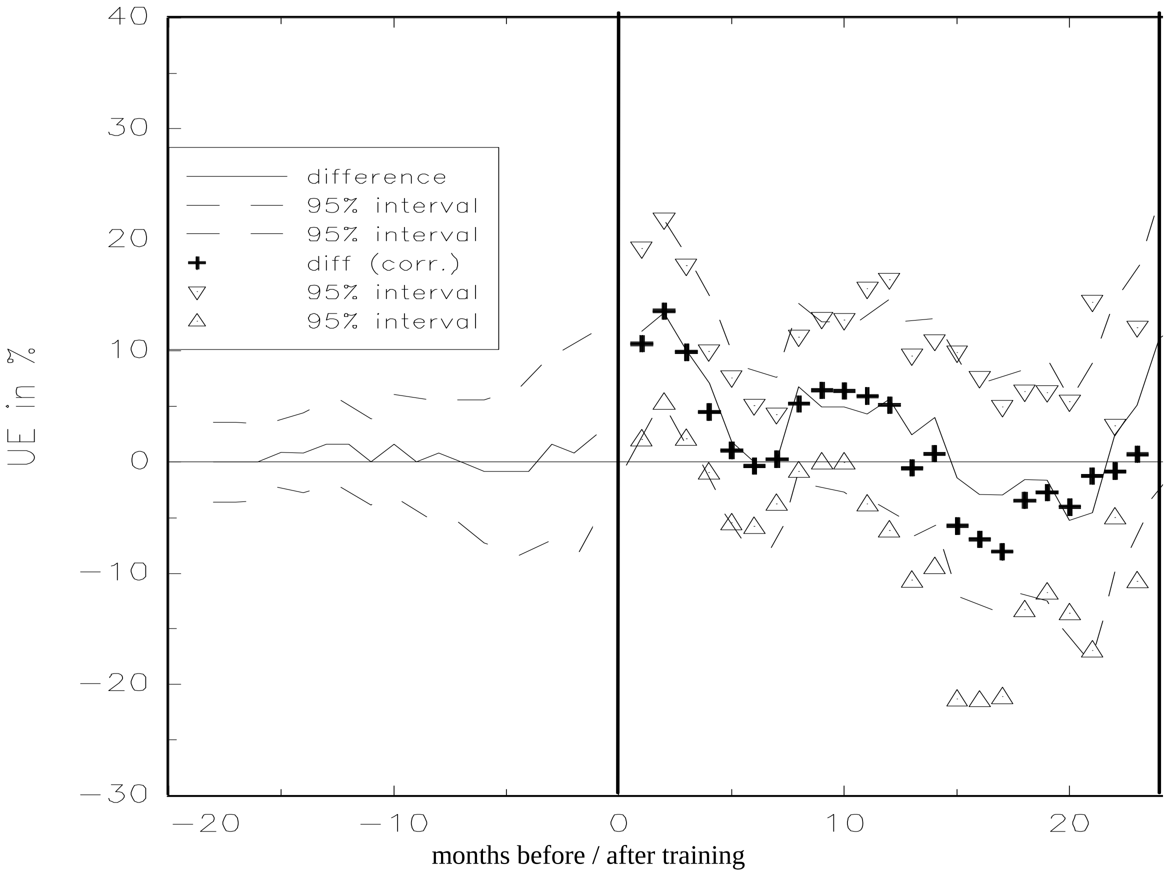
Note: The remaining part (12 to 18 months) of the cdf is omitted. Censored refers to the sample with spells not completed by December 1993. The dashed line denotes the median. Duration is measured in full-time equivalents, assuming 38 hours per week.

Figure 3: Distribution of OFT start dates



Note: Monthly information. Censored refers to the sample with completed spells by December 1993.

Figure 4: Registered unemployment (partial)



Note: N_t = 131. *Difference* denotes the difference of the means in the treated and the matched control group. *95% interval* denotes the 95% confidence intervals of the respective differences based on the normal distribution. *(corr.)* indicates that the estimates are adjusted for mismatches.

Table 1: Results of the estimation and the specification tests against heteroscedasticity for the participation probit (partial)

	estimation			heteroscedasticity test	
Variable	coef.	std.err.	p-val. in %	$\chi^2(1)$	p.-val. in %
Gender: female	0.14	0.14	33	0.3	62
Federal states (Länder) in 1990					
Berlin	0.34	0.19	7.1	2.2	14
Mecklenburg-Vorpommern	**-0.39**	0.18	3.2	1.6	20
Years of schooling (highest degree) in 1990					
12	0.27	0.27	33	4.9	2.8
10	**0.43**	0.17	1.1	3.9	4.7
Highest occupational degree in 1990					
university	0.07	0.35	84	0.8	38
university and female	**0.74**	0.29	1.2	0.6	44
engineering, technical college	0.31	0.18	8.7	2.0	15
master of a trade / craft	**0.47**	0.20	2.0	0.0	97
Job position in 1990: highly qualified, management	0.24	0.20	22	0.1	71
Job characteristics in 1990					
real wage or salary per month / 1000	**-1.95**	0.76	1.1	0.9	33
ln (real wage or salary per month)	**2.95**	1.37	3.1	1.2	27
temporary job contract	-0.10	0.29	73	1.2	27
training (unspecified) while full-time employed	**0.40**	0.16	1.4	0.2	67
Occupation in 1990 (ISCO)					
scientific, technical, medical	-0.25	0.17	14	1.2	27
production	**-0.74**	0.17	0.001	0.0	99
services, incl. trade, office	-0.26	0.16	10	0.3	60
Employer characteristics in 1990: industrial sector					
agriculture	-0.52	0.29	7.6	0.5	50
mining	-0.71	0.44	10	0.5	48
heavy industry	-0.56	0.31	6.6	0.5	49
light industry, consumer goods, electronics, printing	-0.19	0.26	47	0.0	86
machine building and vehicle construction	0.01	0.28	96	8.5*)	0.4*)
construction	-0.30	0.31	33	0.2	65
trade	**-0.66**	0.31	3.6	0.2	68
communication, transport	**-1.05**	0.39	0.7	0.2	65
other services	-0.42	0.27	12	0.4	51
education, science	-0.43	0.28	12	0.1	71
health	**-0.63**	0.29	3.2	3.0	8.4
Optimistic about the future in general in 1990	**0.29**	0.13	3.0	0.1	82
Expectations for the next 2 years in 1990					
redundancies in firm: certainly not	-0.37	0.28	18	0.7	39

Note: **Bold** letters: t-value larger than 1.96. N = 1339. (1199 controls). This table contains results for the estimation of the *partial* propensity score only. *random* and *inflated* lead to very similar results. All specifications contain a constant term. Asymptotic standard errors and score tests are computed using the GMM (or PML) formula given in White (1982). When other estimates of the covariance matrices of the tests leads to different inference using conventional significance levels, they are marked by an asterisk. For details of the computation see Lechner (1995).

Table 2: Other specification tests for the participation probit (partial)

	$\chi^2(df)$	df	p.-val.
Score test against nonnormality	0.3	2	88
Information matrix test			
All indicators	440	394	5.3
Only main diagonal indicators	43	30	6.1

Note: The information matrix tests are computed using the second version suggested in Orme (1988) that has good small sample properties. *Only main-diagonal indicators* refers to a statistic using as test indicators only the main diagonal of the difference between OGP and expected hessian. See also note on Table 1.

Table 3: Descriptive statistics of selected variables of OFT and control sample: different matching algorithms

	Controls				OFT (131)
	all (1105)	matched samples (131)			
		partial	*random*	*inflated*	
(1)	(2)	(3)	(4)	(5)	(6)
Variable	mean (std), share in %	mean (std), share in %	mean (std), share in %	mean (std), share in %	mean (std), share in %
$v\hat{\beta}$ *- partial*	-1.61 (.63)	-0.98 (.53)	-	-	-0.89 (.51)
$x\hat{\beta}$ *- random*	-1.76 (.77)	-	-0.75 (.61)	-	-0.71 (.66)
$x\hat{\beta}$ *- inflated*	-1.75 (.76)	-	-	-0.72 (.65)	-0.72 (.65)
Gender: female	42	62	68	62	64
Federal states (Länder) in 1990: Berlin	7	10	12	9	13
Years of schooling (high. deg.) 1990					
12	17	27	26	24	31
10	60	64	60	67	63
Highest occupat. degree in 1990: university	11	23	21	21	25
Job position 1990: highly qualified, manag.	19	40	30	22	43
Job characteristics in 1990: monthly wage / salary (in 1993 DM)	1714 (526)	1714 (380)	1726 (536)	1725 (779)	1736 (398)
Unemployment in month before OFT	-	18	21	34	20
Full time employment in month before OFT	-	69	69	40	67
Median absolute standardized bias (MSB)	-	5.7	4.5	14.3	-
Joint Wald test for paired mean diff. (JW)	-	26.6	7.2	44.4	-

Note: (2) no matching;

(3) matched on $v\hat{\beta}$, selected v-variables and m (monthly, yearly)-variables (*partial*). Ratio of variance of $v\hat{\beta}$ in OFT sample over variance in control sample is 0.71. Average width of a caliper is 0.98. *v*-variables used for the additional conditioning are: *gender*, *Berlin*, *university*, *10 years of schooling*, *expectation of no redundancies in firm for the next two years* (1990), *highly qualified or management job position* (1990), *monthly wage / salary* (1990), *training (unspecified) while full-time employed* (1990). *m*-variables are *unemployment / short time work / full-time work* 1 month before OFT, average of the last 4 months before OFT, and average of all months before OFT; unemployment / full-time employment / self-employment in the year before OFT.

(4) matched on $x\hat{\beta}$ (*random*); (5) matched on $x\hat{\beta}$ (*inflated*); (6) OFT sample;

$$MSB = \underset{k}{median}(b^k / \sqrt{(s^2(x_i^k) + s^2(x_{(i)}^k))/2}), \quad JW = N^t b'[s^2(x_i^k - x_{(i)}^k)]^{-1} b \cdot \; b^k = 1/N^t \sum_{i=1}^{N^t} (x_i^k - x_{(i)}^k) \cdot \; x_{(i)}$$

denotes the value of for the control observation matched to the treated observation *i*. S^2(a) denotes the empirical variance of a. The median is taken with respect to the K (=11)- variables presented in the table.

$b = (b^1,...,b^k,...,b^K)'$. Asymptotically, $\chi^2(K)$ should be good approximation for the distribution of *JW* when there are no systematic differences of the K (=11) attributes given in the table for the matched pairs. The corresponding p-values for JW are 0.3, 44, 0.

See also note of Table A.1.

Table 4: A Comparison of three matching methods for unemployment rate differences (in %-points) at selected months before and after OFT

months before / after OFT	*partial* mean	*partial* std.	*partial* N^t	*random* mean	*random* std.	*inflated* mean	*inflated* std.
-12 months	0.7	2.0	131	0.9	2.2	-6.0	3.2
-6 months	0.0	3.6	131	-0.9	3.4	-11.1	4.3
-1 month	3.1	4.9	131	-1.7	5.3	-0.9	5.5
1 month	8.8	4.6	125	11.7	4.6	-0.9	5.5
6 months	-0.9	4.1	112	0.0	4.2	-9.9	5.2
12 months	1.1	5.1	88	2.7	5.1	1.4	6.0
24 months	10.8	6.3	37	-21.4	14.4	-6.2	13.2

Note: A mean of *X* denotes an unemployment rate that is *X*%-points higher for the OFT participants. Mismatch adjusted results available on request from the author. See also note on Figure 4.

Table 5: Gross monthly income (in 1993 DM)

years before / after OFT	*Nonemployment earnings coded as 0*					*Nonemployment earnings coded as available benefits*			
			corrected					corrected	
	mean	std.	mean	std.	N^t	mean	std.	mean	std.
-2 years	44	137	44	137	80	53	69	53	69
-1 year	-88	110	-88	110	131	-34	90	-34	90
1 year	-140	183	61	234	120	-62	152	61	173
2 years	49	236	21	340	92	-59	196	-36	271
3 years	195	317	105	560	39	186	251	189	456

Note: The first year is only the difference from the start / end of training to the last / next interview. For the corrected estimates standard errors are computed using a heteroscedasticity robust estimator. The particular variant used is labeled as HC_2 by Davidson and MacKinnon (1993, p.554). See also note on Figure 4 and on Table 5.

Table A.1: Descriptive statistics

Variable	No OFT mean/share in %	No OFT std	OFT mean/share in %	OFT std
Age in 1990	35.2	8.1	35.4	7.5
Gender: female	42		64	
Marital status in 1990				
married	78		78	
single	16		13	
divorced, separated	7		9	
Very desirable behavior / attitudes in society in 1990				
performing own duties	72		63	
achievements at work	72		72	
increasing own wealth	29		20	
Voluntary services in social organizations in 1990:	38		47	
Federal states (Länder) in 1990				
Berlin	7		13	
Brandenburg	15		18	
Mecklenburg-Vorpommern	10		6	
Sachsen	31		32	
Sachsen-Anhalt	20		15	
Thüringen	17		16	
Size of city / village				
< 2000	25		21	
2000 - 20000	28		34	
20000 - 10000	25		24	
> 100000	22		20	
Years of schooling (highest degree) in 1990				
12	17		31	
10	60		63	
8 or no degree	22		6	
Highest occupational degree in 1990				
university [1)]	11		25	
engineering, technical college[2)]	16		33	
master of a trade / craft	6		6	
skilled worker[3)]	64		34	
no degree	2		2	
Job position in 1990				
highly qualified, management	19		43	
master of a trade / craft[4)]	8		7	
skilled blue and white collar[5)]	57		40	
Job characteristics in 1990				
wage / salary per month	1240	381	1256	288
tenure in years	10.5		9.6	
temporary job contract	4		4	
professional degree in other than current profess.	36		31	
already fired	4		7	
training (unspecified) while full-time employed	7		16	

Table A.1 to be continued ...

Table A.1: Descriptive statistics: continued

Variable	No OFT mean/share in %	OFT mean/share in %
Occupation in 1990 (ISCO)		
scientific, technical, medical	19	39
production	43	13
managerial	3	5
administrative	9	11
trade	5	2
agriculture	3	2
services	8	5
services, incl. trade, administrative	23	21
Memberships in 1990		
union	75	80
professional association	7	8
cooperative (LPG / PGH)	8	4
Employer characteristics in 1990		
firm size (number of employees)		
0-19	10	10
20-199	27	25
200-1999	37	39
2000 and more	26	26
industrial sector		
agriculture	11	7
energy and water	3	4
mining	3	2
heavy industry	10	4
light ind., consumer goods, electronics, printing.	16	18
machine building and vehicle construction	5	10
construction	7	4
trade	7	5
communication, transport	8	1
other services	11	13
education, science	10	20
health	7	9
redundencies announced	46	52
Finding a similar new job is (in 1990)		
impossible	11	16
difficult	69	70
easy	20	13
Very worried about job security in 1990	37	39
Optimistic about the future in general in 1990	17	18
Not at all optimistic about the future in general in 1990	7	9
Not enjoying work	5	6
Very confused by new circumstances	5	4
Income very important for subjective well-being	65	54

Table A.1 to be continued ...

Table A.1: Descriptive statistics: continued

Variable		No OFT mean/share in %	No OFT std	OFT mean/share in %	OFT std
Expectations for the next 2 years in 1990					
redundancies in firm: certainly		32		40	
redundancies in firm: certainly not		7		3	
losing the job: certainly		5		7	
losing the job: possibly		35		38	
losing the job: certainly not		12		7	
improvements in career: certainly		1		1	
improvements in career: certainly not		43		38	
decline in career: certainly		3		3	
decline in career: certainly not		49		42	
new occupation: certainly		4		7	
new occupation: certainly not		48		40	
Employment status (monthly)					
Unemployment	July 1989	0		0	
	December 1990	4		6	
	December 1991	6		10	
	December 1992	10		13	
	December 1993	11		15	
Short time work	July 1989	0		0	
	December 1990	12		15	
	December 1991	6		12	
	December 1992	2		0	
	December 1993	2		1	
Full time work	July 1989	95		97	
	December 1990	82		78	
	December 1991	84		68	
	December 1992	80		73	
	December 1993	78		78	
Earnings (yearly in 1993 DM)	1990	1724	541	1736	398
	1991	1893	852	1897	727
	1992	2259	1171	2184	963
	1993	2489	1561	2439	1180
	1994	2634	1560	2714	1278

Note: 1) University and 'Fachhochschule'; 2) 'Ingenier- und Fachschule', not 1); 3) 'Berufsausbildung', 'Facharbeiter', 'sonstige Ausbildung', not 1), 2) or master of a trade / craft; 4) Includes 'Brigadier', 'Meister im Angestelltenverhältnis'; 5) 'Facharbeiter', 'Angestellte mit qualifizierter Tätigkeit'.
1990 relates to the date of interview which for almost is earlier than July 1990 (EMSU).