DISCUSSION PAPER SERIES

IZA

IZA DP No. 1078

Work-Related Training and Wages: An Empirical Analysis for Male Workers in Switzerland

Michael Gerfin

March 2004

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Work-Related Training and Wages: An Empirical Analysis for Male Workers in Switzerland

Michael Gerfin
University of Bern
and IZA Bonn

Discussion Paper No. 1078
March 2004

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
Email: iza@iza.org

IZA Discussion Paper No. 1078
March 2004

ABSTRACT

Work-Related Training and Wages: An Empirical Analysis for Male Workers in Switzerland*

Work-related training is considered to be very important for providing the workforce with the necessary skills for maintaining and enhancing the competitiveness of the firms and the economy. On the individual level, the primary effect of training should be an increased productivity of the trained workers. This paper provides estimates of the effects of training on wages which can be seen as a lower bound for the effects on productivity. Based on panel data from the Swiss Labour Force Survey (SLFS) I estimate these effects using nonparametric matching methods. Training is measured either as firm-sponsored training or as any work-related training. The data show that multiple participation in work-related training is not a rare event. This complicates the analysis considerably because the evaluation of dynamic treatments is not yet fully developed. As a solution to this problem a heuristic difference-in-differences approach to estimate the incremental effect of further training events is used. The results indicate that it is important to account for multiple training events. Taken together, there are significant effects of work-related training on wages of roughly 2% for each training event. There is some evidence that workers who already have high earnings profit more from continuous work-related training.

JEL Classification: I2, J31, C14

Keywords: Training, Wages, Nonparametric Matching

Michael Gerfin
Volkswirtschaftliches Institut
University of Bern
Gesellschaftsstr. 49
3012 Bern
Switzerland
Email: michael.gerfin@vwi.unibe.ch

* This paper is part of the National Research Project "Education and Employment" (NFP 43). I would like to thank the Swiss Science Foundation for financial support (Grant 610-062887), the Swiss Office of Statistics for data provision, and Michael Lechner, Robert Leu and Reto Nyffeler for helpful discussions. All errors are mine.

1. Introduction

Work-related training is considered to be very important for providing the workforce with the necessary skills for maintaining and enhancing the competitiveness of the firms and the economy (see e.g. OECD, 1995). On the individual level, the primary effect of training should be an increased productivity of the trained workers. However, it is difficult to measure individual productivity. The best proxy for productivity is usually the worker's wage which theoretically should be equal to the worker's marginal product. In the case of training this is more difficult to do, at least for general training. Becker (1964) has shown that the costs of general training will be paid by workers. The costs of firm-specific training will be shared by firm and worker. In both cases it is likely that workers pay for the costs with reduced wages. Hence, at least for some time, work-related training will lead to a wedge between productivity and wage. Hence, analysing the effects of training on wages will provide a lower bound for the effects on productivity. Empirical evidence based on data containing information on productivity indicates that the effects on productivity can be much larger than the effects on wages (Barron, Berger, and Black, 1999, and Goux and Maurin, 2000).

The major econometric problem in analysing the effects of work-related training on wages is that training participation is not a random event. In order to control for nonrandom selection into training I apply a difference-in-differences matching estimator. This estimator has been proposed by Heckman et al. (1997) and has been recently used by Eichler and Lechner (1999) and Bergemann et al. (2001). The difference-in-differences matching estimator combines the advantages of both difference-in-differences and matching. Matching removes all observable differences between the group of participants and the control group by appropriate econometric methods. Hence matching will yield unbiased estimates of the treatment effect when selection is only on observable factors. The major criticism against matching is that it may be hard to justify that there is no selection on unobservable factors like ability or motivation. As long as these unobservable factors are constant over time they can be eliminated by differencing over time. In this sense difference-in-differences matching corrects for both selection on observables and on unobservables.

Using data from the Swiss Labour Force Survey I estimate the effect of work-related training on individual earnings in the first and second year after training. Contrary to previous results for Switzerland I find only small and often insignificant effects. This finding suggests that if training increases productivity workers are not able to benefit from this increase, at least in the first two years after training.

2 A brief survey of the theory and empirical evidence on work-related training

2.1 Theory of work-related training

In his seminal work on human capital Becker (1964) made the crucial distinction between general and specific training. If he skills a worker aquires through on-the-job training are purely general, the wage on the external labour market will reflect the full marginal product from this training. Thus, the worker captures the entire return from their general human capital in a competitive labour market. On the other hand, training in perfectly specific skills has no effect on the worker's productivity in other firms, i.e. the wage he can get elsewhere will be independent of the amount of training he received. As a consequence, the return to specific human capital will be shared between employer and employee. Becker concluded that workers must bear all costs of their general training whereas the costs of specific training are shared between workers and firms.

This prediction, however, is at odds with empirical work on firm-sponsored formal training which is general in nature. Recent research has suggested several reasons why and under which circumstances firms may be willing to contribute to the costs of general training. One prominent explanation is based on informational asymmetries between training firm and potential future employers. If the outside market is not as well informed as the current employer about a worker's level of training or other relevant characteristics, the worker's general skills are no longer perfectly marketable and in essence become specific skills (Katz and Ziderman, 1990, Acemoglu and Pischke, 1998, 1999). An analoguous argument applies if there are labour market frictions created by search and hiring costs (Acemoglu, 1997). In both cases, workers receive less than their marginal product from general training which improves firms' investment incentives. Acemoglu and Pischke (1999) note further labour market imperfections where wages are below marginal product and rise less steeply than productivity so that the wedge between marginal product and (outside) wage is higher the more trained a worker is. They refer to this situation as a compressed wage structure. Kessler and Lülfesmann (2000) present a model based on the assumption that general and specific training are complements. They show that in this case employer and employee will share the costs and returns of general training even without market imperfections.

2.2 Empirical evidence

There is a large and growing literature on estimating the effect of work-related training on wages and job turnover. Methodologically, the papers vary between cross-section OLS regressions with and without selection correction, fixed effect estimators, and nonparametric matching approaches. Since it is rather unlikely that training is allocated randomly across workers estimates without taking account of selection are to be interpreted with caution. These studies often find returns to training that are larger than the returns to education (see Pfeiffer, 2000, for a recent survey). However, controlling for selectivity is difficult in the training context because it is hard to find variables that affect training decisions but do not affect earnings.[1] This may explain the very high estimates of over 20% for the Netherlands in Groot (1995) and for Germany in Pfeiffer and Reize (2000).[2] An alternative to control for selection is estimation by fixed effects, assuming that the unobserved variable determining training decisions and earnings can be eliminated by differencing over time. Examples for this approach are Pischke (2000) and Blundell et al (1999). Pischke uses data from the German Socioeconomic Panel and finds hardly any significant effect of training on wage levels or wage growth. Blundell et al. use data from the British National Child Development Survey, which is a unique panel data set following a birth cohort (born between March 3 and 9, 1958) over time. They analyse the effect of training between 1981 and 1991 on wage growth in this period. In addition to control for permanent unobserved heterogeneity by first differencing, they also control for transitory fluctuations between the determinants of training and wages by a selection term. They find significant effects of roughly 8% for employer-provided training on wage growth over 10 years, i.e. less than 1% per year. Lechner (1999) estimates the effect of enterprise-related training in East Germany in the early 1990s using matching methods. He finds significant effects in the second year after the training of about 350 DM (more than 10% of participants mean earnings prior to training).

Two interesting recent studies are Barron, Berger, and Black (1999) and Goux and Maurin (2000). Both studies are based on data for workers and firms. Barron et al. find only small effects of training on wages (based on fixed effect estimation), but large effects on productivity. Their results imply that firms bear most costs of training, but also get most of the returns to training. Goux and Maurin find a effect of about 5% for training when not controlling for selectivity.

[1] In principle such a variable is not necessary to estimate selection models which can be identified by functional form. In practice, however, identification by functional form only often yields very imprecise and volatile estimates.

However, when they control for selectivity using firm information the effect vanishes indicating that the returns are taken up by firms.

The only comparable study for Switzerland are Bänziger (1999) and Gerfin et al. (2002). Bänziger estimates the returns to training by uncorrected OLS using cross section data from the Swiss Labour Force Survey 1996 and finds effects between 4 and 6% for men. These numbers appear to be quite large, given that average labour productivity growth in Switzerland was 0.7% per year during the 1990s. Gerfin et al. employ fixed effects estimators using data from the 1998-2000 waves of the SLFS (which are also used in this paper). Their estimates for men are around 1.5%.

3. Econometrics

Estimating the effect of training is a classical *treatment effect* problem. To estimate a treatment effect we compare the value of some outcome variable (e.g. wages) for the treated individuals with the value the outcome variable would have taken in case of nontreatment. This hypothetical value is usually called *counterfactual*. It must be estimated using the group of the nontreated since we never observe anyone both as treated and nontreated. In order to get an unbiased estimate there must be no systematic differences between the treatment group and the control group selected from the nontreatment group, , i.e. selection into treatment must be random. However, in the case of work-related training workers are selected or select themselves based on observable and unobservable characteristics. If we do not control for this selection the estimates of the treatment effect are likely to be biased.

The framework for the empirical analysis in this paper is the potential-outcome approach to causality suggested by Roy (1951) and Rubin (1974). Let Y^p and Y^n denote the potential outcomes in case of participation in treatment, *p*, and nonparticipation, *n*.[3] Furthermore, let *X* denote variables that are unaffected by treatment. Finally, let *S* denote an indicator for participation (*S*=1). The observable outcome is thus $y_i = s_i y_i^p + (1-s_i) y_i^n$. It is obvious that the causal effect defined as the difference between the two potential outcomes can never be estimated because the counterfactual to the observable outcome y_i is not observable. However, what can be

[2] Both studies employ a switching regression framework using cross section data.
[3] In the following capital letters indicate quantities of the population and lower case letters denote the respective quantities in the sample. The units of the sample (*i=1,…,N)* are assumed to be the result of *N* independent draws from the population. The exposition closely follows Eichler and Lechner (1999).

estimated is the expected causal effect for the group of participants (the treatment effect on the treated), θ.

$$\theta = E(Y^p - Y^n \mid S=1) = E(Y^p \mid S=1) - E(Y^n \mid S=1) \tag{1}$$

$E(Y^p \mid S=1)$ can consistently be estimated by the sample mean of y_i in the subsample of participants. The problem is the term $E(Y^n \mid S=1)$. A central issue in the literature on causal models in statistics and selectivity models in econometrics is finding useful identifying assumptions to predict the unobserved expected non-treatment outcomes of the treated population using the observable non-treatment population. The most common approach is the standard selection model in which identification is achieved by parametric assumptions about the joint distribution of the error terms in the selection and in the outcome equation. It is well known that the selection model in most cases requires a variable that influences the selection, but not the outcome in order to be fully identified (the model is in principle identified by its nonlinearity, but in practice results often are volatile when no such variable exists). In the context of work-related training such a variable is hard to find, especially in typical labour force surveys. For this reason I use another approach outlined below.

One possible assumption to solve the identification problem is the conditional independence assumption (CIA) proposed by Rubin (1977). CIA can be stated as follows:

$$Y^n \coprod S \mid X = x, \forall x \in \chi \tag{2}$$

In words CIA means that participation is independent ($\coprod$) of the non-treatment outcome conditional on the values of the attributes x in the space χ. Thus $E(Y^n \mid S=1, X=x) = E(Y^n \mid S=0, X=x)$, and θ is identified. As opposed to model-based econometric approaches CIA allows to estimate treatment effects directly without imposing functional form or parametric assumptions necessary to estimate structural models.

A technical problem arises when X has a high dimension. A solution to this problem is the propensity score or the balancing score, respectively. Let $P(x) = P(S=1 \mid X=x)$ denote the propensity score, defined as the probability $P(x)$, $0 < P(x) < 1$, of participating in the treatment. If CIA holds Rosenbaum and Rubin (1983) show that $Y^n \coprod S \mid P(X) = P(x), \forall x \in \chi$ holds, so

$$E(Y^n \mid S=1) = E_X \left\{ E\left[Y^n \mid S=0, P(X)=P(x) \right] \mid S=1 \right\} \qquad (3)$$

In words, this implies that when the Y^n outcomes are independent of participation conditional on *X*, they are also independent of participation conditional on the propensity score. 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.

CIA and the propensity score property are the basis for the increasingly popular matching estimator of the treatment effect on the treated. A typical matching estimator takes the form

$$\theta^M = \frac{1}{n^p} \sum_{i \in I^p \cap S_P} \left[Y_i^p - \hat{E}\left(Y_i^n \mid S=1, P_i \right) \right] \qquad (4)$$

where

$$\hat{E}\left(Y_i^n \mid S=0, P_i \right) = \sum_{j \in I^n} W(i,j) Y_i^n \qquad (5)$$

and where $P_i = \Pr(S_i = 1 \mid X_i)$, I^p denotes the set of participants, I^n denotes the set of nonparticipants, S_P denotes the region of common support on *P*, and n^p is the number of persons in the set $I^p \cap S_P$. Common support is the subset of (0,1) for which values of *P* are present in both the participant and the nonparticipant sample. The match for each participant $i \in I^p \cap S_P$ is constructed as a weighted average over the outcomes of nonparticipants, where the weights *W(i,j)* depend on the distance between P_i and P_j. Matching estimators differ in the weights they attach to members of the comparison group. The most common matching estimator, the nearest neighbour (or one-to-one) matching estimator, sets *W* equal to one for the matched nearest neighbour and zero for all other members of the control group. Alternatives are kernel or local linear regression approaches for *W*.

In order to justify CIA it is necessary to identify and observe all variables that are mutually correlated with assignment and potential non-treatment outcomes. This implies that there is no important variable missing that influences non-treatment outcomes and assignment given a value of the relevant variable. It is unlikely that the SLFS data are sufficiently informative to justify CIA in the context of work-related training.

As a possible solution to this problem Heckman et al. (1997) proposed a generalisation of CIA. It is applicable when there is at least one observation of the outcome before the treatment and one after the treatment. The idea is that although CIA may not hold, it may be reasonable to assume that the resulting bias is the same for at least one date before training and for one date after training. If the true effect of the treatment is zero before the treatment takes place, the estimated treatment effect before treatment will be an estimate of the bias. This bias estimate can be used to correct the estimate of the treatment effect after treatment. This idea is of course very similar to a difference-in-differences estimator. For panel data the conditional difference-in-differences estimator is defined as

$$\theta^{DiDM} = \frac{1}{n^p} \sum_{i \in I^p \cap S_P} \left[\left(Y_{ti}^{\ p} - Y_{t'i}^{\ n} \right) - \hat{E}\left(Y_{ti}^{\ n} - Y_{t'i}^{\ n} \mid S_i = 1, P_i \right) \right] \qquad (6)$$

where

$$\hat{E}\left(Y_{ti}^{\ n} - Y_{t'i}^{\ n} \mid S_i = 1, P_i \right) = \sum_{j \in I^n} W(i, j) \left(Y_{tj}^{\ n} - Y_{t'j}^{\ n} \right) \qquad (7)$$

The empirical evidence presented below is based on estimating equation (7) using a balanced panel.

4. Data

I employ data from the Swiss Labour Force Survey (SLFS). The SLFS is conducted by the Swiss Federal Statistical Office on a yearly basis. Each year about 18’000 households are interviewed. The SLFS is designed as a rotating panel, i.e. individuals are interviewed at most in 5 consecutive years. In the years 1996 and 1999 there were special questionnaires relating to vocational training. The questions determine who had any training in the past twelve months, who had work-related training, whether this training was financed by the firm or took place during work time, whether training ended with a certificate, and duration of training. In addition, those not in training are asked whether they would have wanted to go into training but could not do so for some reason (no time, family reasons, etc). From these questions I constructed indicator variables for work-related training, work-related training (at least partially) sponsored by the firm, certified work-related training, and unfulfilled training intentions. In addition, all waves of the SLFS contain information on work-related training in the past twelve months. However, the information is much less detailed and does not allow a distinction according to who paid for the training. But using this reduced information it is possible to analyse the dynamics of training participation in Switzerland. This information proves to be very important for controlling for selection effects.

Unfortunately, there was a significant change in the questionnaire regarding income between 1995 and 1996.[4] Since the estimation method is based on the difference between the income before and after training it is impossible to use the 1995/1996 waves for the analysis. Hence I focus on the waves surrounding the 1999 wave. I constructed a unbalanced 3-years panel covering the years 1998-2000. Persons must be observed in the wave with the training questionnaire and in the previous wave. This yields wage data and individual characteristics prior to the training event we analyse. Because the training variables refer to he past twelve months it is necessary to use the 1998 characteristics in order to estimate training participation propensity scores. The data for the third period contain information on wages and job changes one year after training is completed.

I construct two potential control groups. Control group 1 consists of all workers who do not receive training. This is the control group used in most studies. For control group 2 I use additional information. It consists of those reporting that they want to participate in training but for some reason could not do so. If these reasons are random to the participation decision it would be possible to treat control group 2 as a experimental control group.

Only full-time working men are included in the sample. Work-related training is defined as training in the past 12 months that is either employer-financed or that takes place during work time. Training duration must be at least a week, and only completed training spells are considered. Table 1 displays descriptive statistics of some important variables for participants and nonparticipants in work-related training. It is obvious that there are significant differences between participants and control group 1 with respect to education, skill level, job position and firm size. Control group 2 appears to be more similar to the participant group but there are still some systematic differences. This will be reflected in the estimated propensity score in the next section. As a second treatment indicator I use participation in any work related training in the past twelve months. This is the training information available in each wave, whereas the more refined

4 Until 1995 respondents were asked to state their full labour income, including income from jobs other than their main job. Since 1996 the questionnaire differentiated between main and additional jobs.

training indicator discussed above is only available in the 1999 wave. Of course, both indicators are highly correlated, and the difference should be workers who finance their training themselves. This is the case for 20% of the workers reporting to have participated in work-related training (hence the overlap of the two indicators is 80%).

The final three rows display real monthly earnings by treatment status. It is obvious that the treatment group had much larger earnings in 1998, i.e. before the training that is being analysed had started. Using these numbers it is possible to compute simple difference-in-difference estimates without control variables. The effect of training using control group 1 is 63 CHF after one year and 113 CHF after two years. This amounts to an increase of roughly 2%. Using control group 2 the effects are 5 and 58 CHF, respectively. None of these estimates is significant (all t-values are smaller than one).

An interesting question concerns training dynamics. Using the training variable contained in all waves ("did you receive work-related training in the past twelve months?") Table 2 analyses the dynamics. 32% did not participate in any training between 1997 and 1999 (recall that training is measured retrospectively, so the 1998 wave contains training information for the year 1997). On the other hand, 22% received training in all three years. Of the remaining 46% 21% had least one training spell and 25% had two training spells. In other words almost 50% of the sample had at least 2 training spells in the three years 1997-1999. This causes additional problems regarding the endogeneity of training. If the focus is on the training spell in 1998 for which the more detailed information is available it is possible to treat training participation in the previous year as exogenous. But it is very important to control for this previous training spell in the matching since this spell clearly has an impact on both training participation in 1998 and on earnings in 1999 and 2000.

The more difficult problem is caused by training participation in 1999. From the point of view of training participation in 1998 the future participation is clearly endogenous. For this reason it is not possible to exclude observations with training participation in 1999. This implies that the control group will contain persons who participated in training in 1999 but not in 1998. In other words, the control group will be a mixture of workers without training at all and workers with a training event later than the training event that is evaluated. Hence the results for the effect on earnings two years after training have to interpreted with caution.

The evaluation of the effects of repeated training participation is not yet fully developed (see Miquel, 2003, for a recent analysis). In addition, the SLFS only provides detailed training information in the 1999 spell. In order to analyse the potential effects of repeated training I focus on the simple training indicator described used for Table 2. Methodologically, I follow a suggestion in Bergemann et al (2001). They propose a simple way to estimate the incremental effect of further training events. The idea is partition the sample into those with two training events and those with at most one training event. The propensity score for this sample separation is estimated as before, and matching is performed as in the standard case as well, only the respective control groups are different. The outcome variable is the earnings difference between 2000 and 1998. Hence I estimate the effects of two training events compared to at most one training spell. It should be noted, however, that while this approach has intuitive appeal it still lacks a formal proof. But it is very similar to the parametric difference-in-difference estimator for dynamic treatment effects outlined in Miquel (2003).

5. Results

Table 3 displays the estimation results of the training participation probit. Note that all control variables refer to the 1998 wave because the training variables from the 1999 wave refer to the past 12 months. Thus the situation in 1998 is relevant for training participation. Training participation is more likely for highly educated workers and workers with jobs requiring high skill levels. Training is more likely in large firms and in some sectors such as banking and insurance, and public administration. The most important determinant of training participation is previous training, indicating that training participation is highly correlated over time.

Table 4 shows the results of the nonparametric difference-in-difference estimation of equation (7). The results are based on nearest neighbour matching with replacement, imposing the common support restriction. Matching was performed using the Mahalanobis weighting matrix, with the estimated propensity score and real income in 1998 as matching variables. The latter variable was included because analysing the balancing properties of matching on the propensity score alone showed that earnings in 1998 were not balanced well at all. This is documented in Appendix Table A.1. Standard errors are computed as proposed by Adabie and Imbens (2002).

Inspection of Table 4 clearly shows that most estimated effects are insignificant. The only significant effects are for the wage increase in the first period (ΔIncome1) for both training indicators when control group 1 is used. With respect to firm-sponsored work-related training the effect on earnings is CHF 134 (1.8%), with respect to all work-related training the effect is 172 CHF (2.3%). These findings are similar to those in Gerfin et al (2003) and Gerfin (2003) based on parametric fixed effects estimation. The effects using control group 2 are larger in most cases, but the estimates are rather imprecise. The effects on earnings in the second year are all insignificant which might be explained by the problem discussed in the previous section concerning the dynamics of training.

In order to analyse the potential importance of repeated training events I estimate the incremental effect of a second training event along the lines sketched in the previous section. Due to the data limitations this is only possible for the training indicator "all work-related training". The first step consists of estimating the propensity score of two training events opposed to at most one training event. The results of this estimation are in the Appendix. The outcome variable is the earnings difference between 1998 and 2000 (ΔIncome2). The results are presented in the bottom line of Table 4. Both estimated incremental effects are relatively large and significant. In the case of control group 1 the estimated effect of two training events is somewhat larger than the effect of the training event in 1998, but the difference is not significant. The same is the case when control group 2 is used. While the estimated incremental effect of the second training event is very large it is not significantly different from the effect for the first training event. It is also not significantly different from the effect estimated using control group 1. Overall, these results indicate that it is important to take account of repeated training events. In other words, the estimated effects on ΔIncome2 using only the first training event appear to be quite misleading.

Not reported are estimates of the treatment effects by population subgroups. In all cases the remaining sample sizes were too small to estimate treatment effects with any precision. The considered subgroups were private sector, workers in large firms, and workers with higher education.[5] Separating the sample by earnings in 1998, however, provides one significant insight.[6] For workers with 1998 earnings above the median the incremental effect of a second training event is estimated to be CHF 415, which corresponds to an increase in earnings by 5% (see Table 5). For lower income workers this effect is much smaller and insignificant. This finding suggests that workers who already have high earnings profit more from continuous work-related training.

6. Conclusions

The aim of this paper was to estimate the effects of work-related training on earnings. Given the theoretical literature these estimated effects are only lower bounds for the effects of work-related training on productivity. International evidence suggests that these effects are much larger than the effects on wages. Based on panel data from the Swiss Labour Force Survey (SLFS) covering the years 1998-2000 I estimate these effects using nonparametric matching methods. Specifically, in order to control for permanent observable differences between training participants and non-participants I employ difference-in-differences matching. Training is measured either as firm-sponsored training or as any work-related training. Only the latter measure is available in each wave of the SLFS. Analysing the dynamics of this indicator clearly shows that multiple participation in work-related training is not a rare event. This complicates the analysis considerably because the evaluation of dynamic treatments is not yet fully developed. As a solution to this problem a heuristic difference-in-differences approach to estimate the incremental effect of further training events is used. The results clearly indicate that it is important to account for multiple training events. Taken together, the main results are that there are significant effects of work-related training on wages of roughly 2% for each training event. Focussing on firm-sponsored training the estimated effect is somewhat smaller but the difference is not significant. As argued above these estimates are a lower bound for the effects of training on productivity.

From a methodological point of view the results emphasise the importance to account for multiple treatment participation. The approach used in this paper is heuristic. More work is necessary to develop estimators that fully account for the dynamic nature of sequences of treatments.

[5] These results are available on request.

[6] Earnings in 1998 are reported before training takes place. Hence it is possible to treat these earnings as exogenous.

Literature

Acemoglu, D. und J-S Pischke (1999a), „The Structure of wages and Investment in General Training“, *Journal of Political Economy* 107,.539-572.

Acemoglu, D. und J-S Pischke (1999b), „Beyond Becker: Training in imperfect labor markets, *Economic Journal*, 109, F112-F142

Angrist, J. D., und A. B. Krueger (1999): "Empirical Strategies in Labor Economics", in O. Ashenfelter und D. Card (Hrsg.), *Handbook of Labor Economics*, Vol. III A, Kapitel 23, 1277-1366.

Bänziger, A, (1999), *Weiterbildung und Lohnunterschiede zwischen Frauen und Männern*, Chur/Zürich: Verlag Rüegger.

Barron, J.M., M.C. Berger and D.A. Black (1999), “Do Workers Pay for On-The-Job Training”, *Journal of Human Resources* 34, 235-252.

Becker, G. (1964), *Human Capital*, Chicago: University of Chicago Press

Bergemann, A, B. Fitzenberger, and S. Speckesser (2001), “Evaluating the Employment Effects of Public Sponsored Training in East Germany: Conditional Difference-in-Differences and Ashenfelters Dip”, mimeo, University of Mannheim.

Booth, A. and G. Zoega (2001), “Is Wage Compression a Necessary Condition for Firm-Financed General Training?”, CEPR Discussion Paper 2845.

Blundell, R., L. Dearden, and C. Meghir (1996), *The Determinants of Work-Related Training in Britain*, London: Institute for Fiscal Studies

Eichler M, and M. Lechner (1999), “An Evaluation of Public Employment Programmes in the East German State of Sachsen-Anhalt”, IZA Discussion Paper No 94, Bonn.

Gerfin, M. (2003), “Firm-sponsored work-related training in frictional labour markets – an empirical analysis for Switzerland, Discussion Paper, Department of Economics, University of Bern.

Gerfin, M, R. E. Leu and R. Nyffeler (2003), “Determinanten und Wirkung beruflicher Weiterbildung in der Schweiz”, Discussion Paper, Economics Department, University of Bern.

Goux, D. and E. Maurin (2000), “Returns to Firm-Provided Training: Evidence from French Worker-Firm Matched Data”, *Labour Economics* 7, 1-19.

Groot, W (1995), “The Wage Effects of Investments in Enterprise-Related Training”, *Empirical Economics* 20, 133-147.

Heckman, J.J., H. Ichimura, P. Todd (1997), “Matching as an Econometric Evaluation estimator: Evidence from Evaluating a Job Training Programme”, *Review of Economic Studies* 64, 605-654.

Heckman, J. J., R. J. LaLonde, and J. A. Smith (1999): "The Economics and Econometrics of Active Labor Market Programs", in O. Ashenfelter and D. Card (eds.): *Handbook of Labor Economics*, Vol. III A, 1865-2097, North-Holland, Amsterdam.

Lechner, M. (1999), “The Effects of Enterprise-Related Training in East Germany on Individual Employment and Earnings”, *Annales d’Économie et de Statistique* 55-56, 97-128.

Lynch, L (1994), *Training and the Private Sector – International Comparisons*, Chicago, University of Chicago Press.

Miquel, R. (2003), “Identification of Effects of Dynamic Treatments with a Difference-in.-Differences Approach, Discussion Paper, University of St.Gallen.

Pischke, J.S. (2001), “Continuous Training in Germany”, *Journal of Population Economics* 14, 523-548.

Stevens, M. (1994), „A theoretical model of on-the -job training with imperfect competition“, *Oxford Economic Papers*, 46, 537-562.

Table 1: Descriptive Statistics by Training Status, Balanced Panel. Training is firm-sponsored training

		Participants	Control Group 1	Control Group 2
	Age	38.56	37.05	36.86
	Experience	18.63	16.39	14.65
	Tenure	9.77	8.38	7.48
Educational Level:	Secondary	0.08	0.18	0.10
	Upper Secondary	0.51	0.53	0.52
	Tertiary	0.29	0.21	0.27
	Academic	0.12	0.08	0.12
ISCO Skill level:	Unskilled	0.02	0.05	0.02
	Skilled manual	0.21	0.37	0.25
	Skilled nonmanual	0.15	0.16	0.17
	Semi-Professional	0.25	0.21	0.29
	Professional	0.37	0.20	0.27
Job Position:	No Supervisionary Position	0.40	0.51	0.44
	Supervisionary Position	0.29	0.22	0.22
	Management	0.29	0.21	0.26
Firm Size:	< 10	0.12	0.21	0.24
	10<...<100	0.35	0.42	0.45
	> 100	0.53	0.37	0.35
	Temporary Work Contract	0.02	0.03	0.02
	Looking for new job	0.09	0.09	0.12
Sector:	Manufacture of Machinery	0.09	0.10	0.12
	Other Manufacturing	0.08	0.13	0.12
	Chemicals	0.06	0.05	0.03
	Energy	0.01	0.01	0.01
	Construction	0.06	0.12	0.09
	Trade	0.12	0.12	0.11
	Hotels and restaurants	0.01	0.01	0.03
	Transport	0.11	0.11	0.09
	Banking, Insurance	0.10	0.06	0.03
	Other services	0.12	0.12	0.18
	Public Administration	0.13	0.04	0.04
	Education	0.05	0.03	0.03
	Health and Social Work	0.04	0.05	0.07
Region of Residence:	Canton of Zurich	0.19	0.17	0.22
	North-West	0.19	0.15	0.14
	South-West (French and Italian part)	0.15	0.21	0.19
	East	0.15	0.15	0.18
	Central	0.23	0.23	0.20
	Foreigner	0.09	0.20	0.18
	Real Monthly Earnings 1998	7360	6234	6691
	Real Monthly Earnings 1999	7473	6283	6798
	Real Monthly Earnings 2000	7629	6390	6901
Number of observations		502	829	203

Source: Swiss Labour Force Survey 1998-2000, own calculations

Table 2: Training Dynamics

Sequence	Number of observations	Percentage of sample
000	464	0.320
010	84	0.058
001	127	0.088
011	121	0.084
100	101	0.070
110	133	0.092
101	99	0.068
111	319	0.220

SLFS 98-00, own calculations

Table3a: Participation Logit, Firm-Sponsored Work Related Training

	Control Group 1		Control Group 2	
	Coefficient	Standard Error	Coefficient	Standard Error
Age	0.021	0.329	-0.079	-0.801
Age squared	-0.026	-0.312	0.108	0.846
Tenure	0.001	0.065	0.023	1.040
Tenure Squared	0.032	0.690	-0.012	-0.264
Educational Level: Upper Secondary	**0.674**	2.951	0.038	0.106
Tertiary	**0.592**	2.352	0.012	0.032
Academic	0.391	1.338	-0.094	-0.219
ISCO Skill level: Unskilled	-0.576	-1.355	0.053	0.081
Skilled manual	-0.043	-0.196	0.464	1.387
Semi-Professional	0.096	0.450	-0.166	-0.543
Professional	**0.571**	2.608	0.448	1.423
Job Position: No Supervisory Position	-0.219	-1.396	-0.105	-0.463
Management	0.091	0.523	0.123	0.487
Temporary Work Contract	-0.112	-0.244	0.619	0.920
Looking for new job	-0.040	-0.174	-0.245	-0.806
Firm Size: < 10	*-0.379*	-1.916	-0.514	-1.941
> 100	0.525	3.552	**0.663**	2.922
Foreigner	-0.314	-1.541	*-0.496*	-1.805
Sector: Manufacture of Machinery	-0.249	-0.920	-0.280	-0.777
Other Manufacturing	-0.544	-1.996	-0.366	-0.983
Chemicals	0.149	0.459	0.584	1.161
Energy	-0.188	-0.316	-0.332	-0.416
Construction	*-0.353*	-1.187	-0.076	-0.183
Trade	0.042	0.170	0.467	1.357
Hotels and restaurants	-0.101	-0.153	-0.715	-0.964
Transport	-0.138	-0.525	-0.012	-0.031
Banking, Insurance	0.136	0.485	**1.416**	2.824
Public Administration	**0.978**	3.388	**1.301**	2.964
Education	0.324	0.897	0.605	1.102
Health and Social Work	-0.505	-1.488	-0.473	-1.087
Region of Residence: North-West	0.158	0.804	0.273	0.931
South-West (French and Italian part)	-0.276	-1.372	-0.235	-0.805
East	-0.109	-0.541	-0.193	-0.684
Central	-0.193	-1.066	-0.034	-0.127
Training Participation Previous Year	**1.126**	8.512	**0.919**	4.801
Number of observations	1331		705	

Source: Swiss Labour Force Survey, own calculations. All estimations included a constant term. Coefficients in *italic* are significant on the 10% level, coefficients in **bold** on the 5% level, and coefficients in ***bold italic*** on the 1% level.

Training is firm-sponsored training in 1998

Table3b: Participation Logit, Any Work Related Training

		Control Group 1		Control Group 2	
		Coefficient	t-value	Coefficient	t-value
	Age	0.005	0.071	-0.084	-0.771
	Age squared	-0.013	-0.157	0.124	0.865
	Tenure	0.001	0.058	0.023	0.963
	Tenure Squared	0.039	0.712	-0.014	-0.288
Educational Level:	Upper Secondary	**0.715**	3.150	-0.177	-0.438
	Tertiary	**0.625**	2.487	-0.023	-0.054
	Academic	**0.898**	2.983	0.339	0.669
ISCO Skill level:	Unskilled	-0.514	-1.246	-0.487	-0.745
	Skilled manual	-0.151	-0.701	0.488	1.376
	Semi-Professional	0.187	0.878	0.117	0.351
	Professional	**0.451**	2.030	0.455	1.320
Job Position:	No Supervisory Position	-0.160	-1.015	-0.063	-0.254
	Management	**0.410**	2.291	0.406	1.401
	Temporary Work Contract	*-0.868*	-1.756	-0.126	-0.185
	Looking for new job	-0.028	-0.121	-0.247	-0.756
Firm Size:	< 10	-0.122	-0.639	-0.276	-0.947
	> 100	**0.368**	2.435	**0.612**	2.435
Foreigner		*-0.359*	-1.788	*-0.555*	-1.867
Sector:	Manufacture of Machinery	-0.310	-1.127	-0.497	-1.280
	Other Manufacturing	**-0.532**	-1.966	-0.506	-1.265
	Chemicals	0.026	0.078	0.562	1.027
	Energy	-0.323	-0.540	-0.745	-0.882
	Construction	-0.042	-0.145	0.140	0.312
	Trade	0.088	0.357	0.453	1.196
	Hotels and restaurants	0.354	0.564	0.291	0.379
	Transport	-0.357	-1.332	0.036	0.088
	Banking, Insurance	-0.072	-0.251	**1.565**	2.611
	Public Administration	**0.630**	2.100	**1.094**	2.291
	Education	**0.855**	2.184	0.949	1.343
	Health and Social Work	0.024	0.068	0.429	0.752
Region of Residence:	North-West	0.196	0.979	0.359	1.114
	South-West (French and Italian part)	**-0.469**	-2.311	-0.259	-0.812
	East	0.001	0.004	0.062	0.194
	Central	-0.185	-1.009	0.074	0.248
	Training Participation Previous Year	**1.542**	11.678	**1.562**	7.276
	Number of observations	1331		705	

Source: Swiss Labour Force Survey, own calculations. All estimations included a constant term. Coefficients in *italic* are significant on the 10% level, coefficients in **bold** on the 5% level, and coefficients in ***bold italic*** on the 1% level.

Training is any work-related training in 1998 (firm-sponsored or privately financed)

Table 4: Estimates of Treatment Effect

	Employer-Sponsored Work Related Training		All Work Related Training	
	Control Group 1	Control Group 2	Control Group 1	Control Group 2
ΔIncome1, training in 1998	*134* (73.6)	75 (133.1)	172 (80.9)	298 (211.5)
ΔIncome2, training in 1998	117 (93.0)	162 (157.9)	62 (94.6)	210 (244.4)
ΔIncome2, training in 1998 and 1999	-	-	216 (107.2)	405 (127.8)

ΔIncome1 denotes the estimated income difference between 1999 and 1998 (in Swiss Francs), ΔIncome2 denotes the estimated income difference between 2000 and 1998. Standard errors computed according to Abadie and Imbens (2002) in parentheses

Table 5: Estimates of Treatment Effect, Subgroups by Pre-Training Earnings

	All Work Related Training	
	1998 Earnings < median	1998 Earnings > median
ΔIncome1, training in 1998	79 (93.5)	93 (120.8)
ΔIncome2, training in 1998	17 (108.1)	-35 (136.1)
ΔIncome2, training in 1998 and 1999	150 (119.8)	415 (156.6)

ΔIncome1 denotes the estimated income difference between 1999 and 1998 (in Swiss Francs), ΔIncome2 denotes the estimated income difference between 2000 and 1998. Standard errors computed according to Abadie and Imbens (2002) in parentheses. Results only for control group 1,

Appendix

Table A1: Match Quality

	Treatment Group	Control Group 1[a)]			Control Group 2[a)]		
		a	b	c	a	b	c
Educational Level: Upper Secondary	0.51	0.53	0.48	0.53	0.52	0.53	0.51
Tertiary	0.29	0.21	0.29	0.27	0.27	0.26	0.37
Academic	0.12	0.08	0.16	0.14	0.12	0.15	0.08
ISCO Skill level: Unskilled	0.18	0.05	0.01	0.02	0.02	0.02	0.01
Skilled manual	0.21	0.37	0.20	0.20	0.25	0.20	0.18
Semi-Professional	0.25	0.21	0.26	0.24	0.29	0.25	0.27
Professional	0.37	0.20	0.40	0.38	0.27	0.32	0.34
Firm Size: < 10	0.12	0.21	0.12	0.15	0.24	0.08	0.10
> 100	0.53	0.37	0.55	0.53	0.35	0.58	0.56
Public Administration	0.13	0.04	0.12	0.11	0.04	0.11	0.13
Training Participation Previous Year	0.67	0.33	0.67	0.68	0.42	0.68	0.69
Real Monthly Earnings 1998	7360	6234	7780	7335	6691	7746	7337

a: Unmatched sample; b: Sample matched only on propensity score; b: Sample matched on propensity score and real monthly earnings 1998.

TableA2: Participation Logit, Any Work Related Training, More Than One Event

		Control Group 1		Control Group 2	
		Coefficient	t-value	Coefficient	Standard Error
	Age	0.074	1.074	0.068	0.790
	Age squared	-0.098	-1.117	-0.094	-0.850
	Tenure	0.000	-0.011	0.001	0.070
	Tenure Squared	0.031	0.929	0.021	0.601
Educational Level:	Upper Secondary	0.412	1.682	-0.037	-0.108
	Tertiary	0.340	1.277	-0.203	-0.574
	Academic	*0.507*	1.668	0.092	0.232
ISCO Skill level:	Unskilled	-0.438	-0.957	-0.296	-0.465
	Skilled manual	-0.355	-1.503	-0.180	-0.606
	Semi-Professional	0.246	1.122	0.227	0.841
	Professional	**0.535**	2.372	*0.502*	1.825
Job Position:	No Supervisory Position	0.096	0.578	0.239	1.177
	Management	0.201	1.106	0.236	1.067
	Temporary Work Contract	-0.817	-1.462	-0.396	-0.623
	Looking for new job	0.124	0.519	0.090	0.317
Firm Size:	< 10	0.145	0.712	0.382	1.501
	> 100	**0.370**	2.349	*0.374*	1.905
Foreigner		**-0.607**	-2.622	**-0.632**	-2.290
Sector:	Manufacture of Machinery	-0.254	-0.864	-0.190	-0.556
	Other Manufacturing	-0.188	-0.646	0.160	0.449
	Chemicals	0.061	0.176	0.341	0.821
	Energy	-0.363	-0.552	-0.210	-0.285
	Construction	-0.224	-0.672	-0.053	-0.129
	Trade	0.116	0.451	0.447	1.424
	Hotels and restaurants	0.897	1.395	0.852	1.151
	Transport	-0.106	-0.377	0.249	0.737
	Banking, Insurance	-0.039	-0.134	*0.641*	1.756
	Public Administration	0.384	1.330	*0.579*	1.726
	Education	*0.669*	1.817	0.459	0.983
	Health and Social Work	0.282	0.822	0.695	1.588
Region of Residence:	North-West	-0.187	-0.917	-0.262	-1.043
	South-West (French and Italian part)	**-0.680**	-3.134	**-0.571**	-2.153
	East	-0.220	-1.046	-0.304	-1.173
	Central	**-0.539**	-2.817	**-0.493**	-2.073
	Training Participation in 1997	**1.347**	9.499	**0.935**	5.441
Number of observations		1331		705	

Source: Swiss Labour Force Survey, own calculations. All estimations included a constant term. Coefficients in *italic* are significant on the 10% level, coefficients in **bold** on the 5% level, and coefficients in ***bold italic*** on the 1% level.

Training is any work-related training in 1998 (firm-sponsored or privately financed)