Does Work-related Training Reduce the Discrepancy between Function Requirements and Competencies?*

Eelco R. Kappe

Econometric Institute
Erasmus University Rotterdam

Govert E. Bijwaard†

Econometric Institute
Erasmus University Rotterdam

Econometric Institute Report EI 2005-42

*This paper is based on the Bachelor thesis of the first author at the Econometric Institute of Erasmus University Rotterdam. A substantial part of the research was carried out while Kappe was at SEOR. We are indebted to the comments from Jaap de Koning, director of SEOR.

†Bijwaard's research is financially supported by the Netherlands Organisation for Scientific Research (NWO) nr. 451-04-011. Corresponding author: Erasmus University Rotterdam, Econometric Institute, P.O. Box 1738, NL 3000 DR Rotterdam, The Netherlands; Phone: (+31) 10 40 81424; Fax: (+31) 10 40 89162; E-mail: bijwaard@few.eur.nl

Does Work-related Training Reduce the Discrepancy between Function Requirements and Competencies?

Abstract

The issue of lifelong learning is high on the political agenda. However, despite this political interest and the large economic literature on human capital, the impact of work-related training on the discrepancy between function requirements and the skills of the employee has been ignored. In this paper we use an ordered probit model to analyze the perceived change in discrepancy. Based on the bi-annual OSA panel from 1998 till 2002 for The Netherlands, we show that taking a work-related course decreases the discrepancy significantly. We correct for the endogeneity between the decision to take a course and the change in discrepancy and we argue that ignoring the selective decision to take a course leads to misleading conclusions about the effect of these courses on the change in discrepancy.

Some respondents of the OSA-panel drop out between two waves. To correct for the possibility of selective attrition we develop an Inverse Probability Weight (IPW) estimation method for the ordered probit with an endogenous binary regressor. From the implied marginal effects of the IPW estimation we conclude that taking a course increases the probability to change the fit between skills and function requirements from Bad to Good with 16 percent-point.

Key words: work-related training; ordered probit; endogenous regressor; inverse probability weighting.

1 Introduction

It is increasingly acknowledged that education and training of the adult workforce is a key policy issue to meet the challenge posed by technological changes and improving career prospects. In Europe a shift towards lifelong learning has been observed. A clear confirmation of this trend is the increasing number of part-time students. The issue of continuous education and training is crucial in the context of transitional labour markets, see Schmid (1998), Schmid (2002) and De Koning (2002). It refers to the employability of adult workers during their whole career. The knowledge-based economy demands a higher level of knowledge and skills and the employees have to adapt to that. The OECD and the European Commission also recognize the importance of lifelong learning, see OECD (2003).

Despite the policy interest and the large economics literature on human capital, there are, to our knowledge, no papers that examine the impact of work-related training on the discrepancy between function requirements and the skills of the employee. The primary contribution of our paper is to provide evidence for the impact of work-related training on the change of this discrepancy for The Netherlands. Analysis of the impact of training has focused almost entirely on estimating the impact on wages, see Lillard and Tan (1992), Lynch (1992) and Bartel and Sicherman (1999) for studies using US panel data. See Booth (1991, 1993) for evidence for the UK. Winkelmann (1994) uses German data. Bassanini et al. (2005) provide a recent comparison of work-related training in Europe.

For the Netherlands, time series indicate that the fit between function requirements and education/experience has not changed since the beginning of the seventies, SER-report (2001)[1]. Because job-certainty has decreased recently in the Netherlands, employability has become more important. This would increase the attendance of work-related training and create a highly mobile and skilled workforce that is well-adapted to the requirements of the modern knowledge-based economy. However, the Dutch workforce is not highly mobile. In The Netherlands labour mobility appears to be very cyclic, with a high labour mobility during the economic boom around the millennium and a low labour mobility during the recession of the early 90s, see SER-report (2001).

Ideally, if we observe the exact function requirements and the skills and knowledge of each employee we could derive the exact relation between them. However, function requirements and skills are hard to quantify, and difficult to compare between different functions. The OSA supply panel survey data contain subjective ordinal questions on the perceived fit between skills and function requirements for the waves 1998, 2000 and 2002. This longitudinal survey of a random sample of Dutch households, about 4500 individuals, is conducted every two years. It also provides many other characteristics of the sampled employees. For our analysis we use these last three waves.

The difference between the perceived fit is again an ordinal variable and, therefore, we choose an ordered probit model for the analysis. However, it is very likely that the decision to take a course depends on the perceived discrepancy between skills and function

[1] CBS Netherlands provide the data online at http://statline.cbs.nl/

requirements of the employee. Correcting for such a endogeneity is possible by introducing a selection equation with a error term that is jointly distributed with the error of the ordered probit model. If we assume a binary probit model for the decision to take a course the errors terms have a bivariate normal distribution. Thus, another reason to use the ordered probit model is the relative ease it can accommodate such an endogenous regressor.

However, endogeneity of the regressor is not the only issue we face. Like in many other panel surveys, the OSA panel suffers from respondents that drop out. If this attrition is selective on either the dependent variable, change in discrepancy, or an important explanatory variable, the decision to take a course, then the ordered probit model (with a selection equation for taking a course) will give biased results. The common distinction is between attrition that is selective on variables unobserved in the data,*unobservables*, and attrition that is selective on variables observed in the data, *observables*. The available methods to correct for attrition on observables are relatively easy to implement and rely on relatively weak assumptions, this in contrast to the methods that are required to correct for attrition on unobservables.

Attrition on observables can further be distinguished between attrition that leads to ignorable selection and non-ignorable selection on observables. Selection is ignorable if the attrition probability only depends on variables already in the (probit) model. Then, there is no need to adjust the model. If the selection is non-ignorable an alternative estimation procedure is required. Under the assumption of selection on observables it is possible to compute the complete population density of the dependent variable given the explanatory variables by using normalizing weights. These weights are based on the ratio of the attrition probability conditional on explanatory variables only and the attrition probability conditional on explanatory variables and auxiliary variables (see Fitzgerald et al. (1998)). This implies that people who have a higher attrition probability get more weight in the model because they are under-represented.

Our secondary contribution is that we are the first to develop an Inverse Probability Weighted method for an ordered probit model with an endogenous regressor. The Inverse Probability Weighted (IPW) estimation procedure exploits the weighting of those who have a higher attrition probability by maximizing the weighted likelihood. Wooldridge (2002b) provides an extensive discussion on the Inverse Probability Weight estimation method, including efficiency considerations.

The structure of the paper is as follows. In section 2 the OSA panel survey data is discussed. Section 3 introduces an ordered probit model for analyzing the perceived changes in the fit between skills and function requirements. We extend this model with a selection equation for the decision to take a course to allow for endogeneity between this decision and the change in discrepancy. Section 4 introduces the problem of selective attrition and describes different forms of attrition. We develop an Inverse Probability Weight estimation procedure to correct for selective attrition on observables. Section 5 gives the estimation results. Finally, the conclusion and comments are provided in section 6.

2 Data Used for Analysis

In this paper longitudinal survey data from the Organisatie voor Strategisch Arbeidsmarktonderzoek (OSA) is used. For this panel survey a random sample of households in The Netherlands has been followed over time. The study concentrates on individuals who are between 15 and 65 years of age, and are not full-time students. The survey is conducted every two years, except for the second wave (1986), since 1985. Data till 2002 is available. In the OSA-panel an effort is made to survey a fixed panel of around 4500 individuals. However, after every wave people drop out of the panel. These people are replaced with care by a refreshment panel. Another issue is that some questions in the survey are new or rephrased in a next wave. Because only from 1998 on it is explicitly asked how the fit between work and knowlegde/skills is, we restrict our analysis to the last three waves of the panel.

We provide some descriptive statistics of the relevant questions for our topic of research. For the waves of 1998, 2000 and 2002 the survey contained questions on the fit between knowledge/skills and work requirements, as it is perceived by the employee. Table 1 presents some basic statistics of these questions. First note that no clear trend is observed in the evolution of the discrepancy between knowledge/skills and work-requirements. A possible reason for this is the booming economic situation in The Netherlands around 2000. Another reason for the absence of a clear trend is the short time period considered, induced by the limited number of relevant waves of the panel. Around 80% of the employees report a good fit between knowledge/skills and work, while only around 5% (2% in 2000) report a bad fit. Males experience slightly more often a good fit than females. A change in the labour market situation for the individual seems to improve the fit at work. Employees with a long lasting relation with the same employer perceive a better fit between their skills and work-requirements. The fit improves with age. Finally, both the over-education of employees (knowledge is more) and the under-education (knowledge is less) seem to have declined after 1998.

– put table 1 about here–

Table 2 presents some basic statistics on the work-related courses taken by the individuals in the sample. Taking work-related courses for employees has become more common after 1998. The percentage of respondents reporting that they have taken any course in the two preceding years increased from 16% in 1998 to 46% in 2000 (and decreased to 37% in 2002). The number of taken courses has also dramatically increased in 2000. Before drawing hard conclusions from this it should be kept in mind that the Dutch economy was booming between 1998 and 2000. In a booming economy companies have more money left to spend on educating their workers. The increase in the percentage of people taking a course and the average number of taken courses may, therefore, be just a conjunctural effect. When we look at the differences between employees we see some interesting facts. Males, for example, attend more often and more frequent work-related courses than females. Individuals for whom the labour market situation has changed take less courses than individuals who remained in the same labour market situation. The more educated somebody is the higher the probability (s)he takes a work-related course. The regional difference in taking work-related courses is small: in the big cities and in the north of The Netherlands slightly less courses are taken. People who have been hired recently hardly take courses. Finally, it seems that companies invest less in old (45+) people than in young, as the latter take more courses.

– put table 2 about here–

3 Modelling the Change in Discrepancy Between Function Requirements and Competencies

We are interested in unravelling the impact of work-related courses, taken by employees, on the discrepancy between the requirements of the employee's work and their skills. Ideally, if we observe the exact function requirements and the skills and knowledge of each employee we could derive the exact relation between the two. However, function requirements and skills are hard to quantify, and difficult to compare between different functions. The OSA panel provides, as discussed in the previous section, the subjective discrepancy between function requirements and skills as perceived by the individual employee. These subjective discrepancies are measured as ordered responses, running from Good, Reasonable, Moderate to Bad. The most commonly used model for an ordered response is the ordered probit model. It assumes normally distributed errors and it can relatively easily accommodate endogenous regressors and attrition. Below follows a description of the probit model and how we extend it to accommodate possible selective choice to take courses.

3.1 An Ordered Probit model

The ordered probit model is used when the dependent variable takes on ordinal values. The intuitive idea behind the model is that the observed ordinal variable represents ranges of a unobserved latent variable with unknown cutoffs. The latent variable y^*_{1i}, representing in our application the exact change in discrepancy between function requirements and skills, is determined by the following linear regression:

$$y^*_{1i} = X_i\beta + \epsilon_{1i} \qquad , i = 1, \dots, N. \tag{1}$$

This latent variable is mapped onto an ordered categorical variable:

$$Y_{1i} = j \quad \text{if } \alpha_{j-1} < y^*_{1i} \leq \alpha_j \qquad , j = 1, \dots, J. \tag{2}$$

Here j represents the category in which y_{1i} is classified. In our application we have seven different categories. The latent variable can take on any value, so the thresholds $\alpha_0 = -\infty$ and $\alpha_J = \infty$. To preserve the ordering the thresholds α_i must satisfy $\alpha_0 < \alpha_1 < \dots < \alpha_J$.

Because Y_{1i} is only observed as an indicator, the coefficients are only identified up to scale and the intercept is not identified. In the probit model it is assumed that the error term has a standard normal distribution. The contribution to the likelihood for observation i is given by the following expression:

$$\begin{aligned} Pr(Y_{1i} = j|X_i) &= Pr(\alpha_{j-1} < y_{1i}^* \leq \alpha_j) \\ &= Pr(\alpha_{j-1} - X_i\beta < \epsilon_{1i} \leq \alpha_j - X_i\beta) \\ &= \Phi(\alpha_j - X_i\beta) - \Phi(\alpha_{j-1} - X_i\beta), \end{aligned}$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution. The parameters of the ordered probit model are estimated by maximizing the log-likelihood function

$$l(\beta, \alpha_1, \alpha_2, \ldots, \alpha_{J-1}) = \sum_{i=1}^{N} \sum_{j=1}^{J} I[y_{1i} = j] \cdot \log\big(\Phi(\alpha_j - X_i\beta) - \Phi(\alpha_{j-1} - X_i\beta)\big),$$

where $I[y_{1i} = j]$ is an indicator function which is one if y_{1i} is classified in category j.

It very plausible that the decision to take work-related courses depends on the perceived discrepancy of the function requirements and skills of the employee. This implies that taking a course is an endogenous variable in the model. If some of the explanatory variables are endogenous standard estimation of the model is biased, because the covariates and the error-term are correlated. A way to control for such an endogeneity bias is to introduce a second, selection, equation with a error term that is jointly distributed with the error of (1). It is convenient we have chosen an ordered probit model, because if we also assume a binary probit model for the selection equation the error terms have a bivariate normal distribution. This is accomplished by adding the following equation to the model:

$$\begin{aligned} y_{2i}^* &= Z_i\beta + \epsilon_{2i} \qquad , i = 1, \ldots, N \\ Y_{2i} &= \begin{cases} 0 & \text{if } y_{2i}^* \leq 0 \\ 1 & \text{if } y_{2i}^* > 0. \end{cases} \end{aligned} \tag{3}$$

The error terms $(\epsilon_{1i}, \epsilon_{2i})$ thus have a bivariate normal distribution with mean zero, variances one and covariance ρ. The contribution to the likelihood of individual i now depends on the selection and is given by

$$\begin{aligned} Pr(y_{1i} = j, y_{2i} = 0|X_i, Z_i) =& Pr(\alpha_{j-1} < y_{1i}^* \leq \alpha_j, y_{2i}^* \leq 0) \\ =& \Phi_2(\alpha_{j-1} - X_i\beta, -Z_i\delta, \rho) - \Phi_2(\alpha_j - X_i\beta, -Z_i\delta, \rho), \\ Pr(y_{1i} = j, y_{2i} = 1|X_i, Z_i) =& Pr(\alpha_{j-1} < y_{1i}^* \leq \alpha_j, y_{2i}^* > 0) \\ =& \Phi_2(\alpha_{j-1} - X_i\beta, Z_i\delta, \rho) - \Phi_2(\alpha_j - X_i\beta, Z_i\delta, \rho), \end{aligned}$$

where Φ_2 represents the cumulative standard bivariate standard normal distribution with ρ the correlation evaluated at the two cutoff points. The model is estimated using the log-likelihood:

$$l(\beta, \delta, \alpha_1, \alpha_2, \ldots, \alpha_{J-1}, \rho) =$$
$$\sum_{i=1}^{N} \sum_{j=1}^{J} I[y_{1i} = j, y_{2i} = 0] \cdot \log\Big(\Phi_2(\alpha_j - X_i\beta, -Z_i\delta, \rho) - \Phi_2(\alpha_{j-1} - X_i\beta, -Z_i\delta, \rho)\Big)$$
$$+ I[y_{1i} = j, y_{2i} = 1] \cdot \log\Big(\Phi_2(\alpha_j - X_i\beta, Z_i\delta, \rho) - \Phi_2(\alpha_{j-1} - X_i\beta, Z_i\delta, \rho)\Big). \quad (4)$$

4 Correcting for Attrition

Attrition is a problem in many panel surveys. In the OSA-panel each wave about 30% of the sampled individuals dropped out. Each wave these attritors were replaced by a refreshment panel. However, the major problem is not simply missing data, but the possibility that they are missing for a variety of self-selection reasons. For instance, interview institutions often lose the address of people who get a new job for which they have moved. In Van den Berg and Lindeboom (1998) more possible reasons for attrition are mentioned, where it should be noticed that sometimes groups of individuals have reasons that decrease their probability of dropping out and simultaneously reasons that increase their probability of dropping out.

– put table 3 about here–

Table 3 presents a comparison between the attritors and non-attritors in 2000 and 2002. There are apparent differences between the two groups. The people who drop out of the panel, are less often female, have less children, are less often married and, are more often single (never married and not cohabiting). The labour market situation also differs between the two groups. Attritors are less often working for a company (people in employment) and more often out of the labour market. The education level is slightly less for attritors.

If the probability of attrition is correlated with the response of interest, in our application the change in discrepancy between skills and function requirements, or with important explanatory variables, in our application work-related courses, then traditional econometrical techniques will lead to biased and inconsistent estimates of the effect of these variables on the responses. Table 3 reveals that attritors perceive a slightly better fit between their knowledge and their function requirement and if the fit is bad it is less often a problem at work. Finally, attritors take, on average, less work-related courses than non-attritors.

The common distinction is between attrition that is completely random, attrition that is selective on variables *unobserved* in the data, and attrition that is selective on variables *observed* in the data. The latter can further be distinguished between attrition that leads to ignorable selection or non-ignorable selection on observables. Attrition does not necessary introduce bias in the estimates of interest. Some studies report that the bias caused by attrition is small. While other studies exist with a significant bias caused by attrition.

It is easier to test for selective attrition on observables than on selective attrition for unobservables. In particular, this leads to a sequence of tests on attrition bias. First, given that there is sample attrition, test whether or not there is selection on observables. Second, if there is selection on observables, test whether this attrition is ignorable, and thus does not bias the estimates of interest. If attrition is non-ignorable, the analyses need to correct for attrition since otherwise selection leads to biased inferences about relevant parameters. The available methods to correct for attrition on observables are relatively easy to implement and rely on relatively weak assumptions, this in contrast to the methods that are required to correct for attrition on unobservables. While selective attrition on unobservables potentially remains a problem even after the analyses account for selection on observables, using as much information as possible about selection on observables in the panel assists to reduce the amount of unexplained variation in the data due to attrition. Therefore, controlling for selection on observables will likely reduce the biases causes by selection on unobservables.

More formally, consider the survey wave at time t and the object of interest is the conditional population density $f(y_t|x_t)$, where y_t is the dependent variable and x_t are independent variables. We introduce a dummy A_t that indicates whether an individual is still in the panel at time t. It is equal to 1 if no attrition occurs, and it is equal to zero if the individual drops out of the panel. For identification, we assume that x_t is observed for both attritors and non-attritors, as would be the case if it consists of only time-invariant or lagged variables. In the sample only the individuals with $A_t = 1$ are present and only the density $g(y_t|x_t, A_t = 1)$ is observed. Additional information or restrictions are necessary in order to infer the density of primary interest, $f(y_t|x_t)$.

The additional information can come from the attrition probability, $Pr(A_t = 1|y_t, x_t, z_t)$, where z_t are auxiliary variables that are assumed to be observed for all individuals but are different from x_t. In particular z_t can include lagged values of the dependent variable (which are observed up to $t-1$ for respondents leaving the panel), as well as characteristics of the respondents, fixed or time-varying. For example, Fitzgerald et al. (1998) suggest using lagged dependent variables as a candidate for z_t. We assume a linear attrition probability, i.e.,

$$A^*_{it} = \delta_0 + z_{it}\delta_1 + x_{it}\delta_2 + v_{it} \tag{5}$$

$$A_{it} = \begin{cases} 0 & \text{if } A^*_{it} \leq 0 \\ 1 & \text{if } A^*_{it} > 0. \end{cases} \tag{6}$$

where A^*_{it} is a latent index and attrition of individual i at time t occurs if this index is greater than zero and v_{it} is a mean-zero random influence on the attrition probability. Fitzgerald et al. (1998) classify attrition into selection on observables and selection on unobservables. Attrition exhibits *selection on observables* if

$$\Pr(A_t = 1|y_t, x_t, z_t) = Pr(A_t = 1|x_t, z_t) \tag{7}$$

that is, if, conditional on x_t and z_t, the attrition probability is independent of the dependent variable y_t and therefore the on the error-term, ϵ_{1i}, in (1). Attrition exhibits *selection* *on unobservables* if

$$\Pr(A_t = 1|y_t, x_t, z_t) \neq Pr(A_t = 1|x_t, z_t), \tag{8}$$

This implies that, if selection on unobservables occurs, v_{it} is not independent of the distribution of the errors in (2) conditional on the observed characteristics, x_{it}. Note, that (8) can also occur for other reasons, like measurement errors or misspecification of the model.

Selection on observables is *ignorable* if the ignorability conditions:

1. $Pr(A_t = 1|x_t, z_t) = Pr(A_t = 1|x_t)$,
2. y_t and z_t are independent conditional on x_t and $A_t = 1$

hold. The first condition implies that the attrition probability is independent of the auxiliary variables z_t. Ignorable selection on observables implies that estimation of the model (1) and (3) on the basis of the observed data on non-attritors leads to unbiased estimates of the effect of taking work-related courses on the function requirement discrepancy. In the sequel we call this the *unweighted* estimator. If, however, selection on observables is *non-ignorable*, that is if the ignorability conditions are violated, estimation of the model based only on non-attritors data will lead to biased inference. Then an alternative estimation technique is required. This estimation technique is discussed in more detail in section 4.2. Note that selection on observables is not equivalent to exogenous selection, because selection can still be based on endogenous variables.

Under the assumption of selection on observables it is possible to compute the complete population density $f(y_t|x_t)$ from the conditional density $h(y_t|z_t, x_t, A_t = 1)$:

$$f(y_t|x_t) = \int h(y_t|z_t, x_t, A_t = 1) w(z_t, x_t) dz_t \tag{9}$$

where

$$w(z_t, x_t) = \frac{Pr(A_t = 1|x_t)}{Pr(A_t = 1|x_t, z_t)} \tag{10}$$

are the normalized weights (see Fitzgerald et al. (1998)). The numerator of (10) is the probability of remaining in the panel, the non-attrition probability, conditional on x_t and the denominator is the probability of remaining in the panel conditional on x_t and z_t. The weights $w(z_t, x_t)$ can be estimated from the data when both z_t and x_t are observed, as when z_t contains time-invariant or lagged time-varying characteristics of the respondents.

The estimation technique based on this weighting is Inverse Probability Weighting that is discussed in section 4.2. The intuition behind this technique is that we weight every observation in the panel with the inverse of the probability that an observation is included. Because both the weights and the conditional density h are identifiable and estimable from the data, the complete population density $f(y_t|x_t)$ is estimable as well as its moments. This implies that the effect of taking work-related courses (and other employee characteristics) can be estimated without bias, with a method based on these inverse probability weights.

Inspection of (9) and (10) also reveals the cases attrition can be ignored. If z_t does not determine attrition the weights in (10) are equal to one and no attrition bias is present. If y_t and z_t are independent conditional on x_t and $A_t = 1$ the density h in (9) factors out and the unconditional density $f(y_t|x_t)$ equals the conditional density and there is no attrition bias.

4.1 Testing for Attrition Bias

Testing for attrition bias due to selection on unobservables is possible if we include an equation for the attrition index in the model. The identification of such models with panel data is, however, complicated because valid instruments that allow identification are hard to find. Examples of such instruments are interview payments, characteristics of the interviewer etc. Such variables are, however, absent in the OSA panel data. We, therefore, have to rely on other ways to detect selection on unobservables. One indirect test for selection on unobservables suggested by Fitzgerald et al. (1998) is to compare the results with the results from other available data with less (or no) attrition. Unfortunately, no other valid data set is available for our application. Finally, a way to get an impression of the severity of selection on unobservables is by first testing whether significant selection on observables exists. Then, if no selection on observables is detected, estimate the attrition equation and check for correlation with the main equation. However, misspecification of the model may also lead to a correlation between the two equations.

Due to this limited ability to detect selective attrition on unobservables with the data used in our application, we do not discuss this approach further nor do we perform the corresponding tests. However, the more lagged information of the panel is put in the model, the lower the unexplained attrition variation becomes. As a consequence this reduces the need for selection on unobservables, but with a danger of overfitting the data.

Testing for selection bias due to selective attrition on observables is, on the other hand, not difficult to perform. Two sufficient conditions that render the selection on observables through attrition *ignorable* are either (i) z_t does not affect the attrition probability or (ii) z_t is independent of y_t conditional on x_t for those that remain in the sample, $A_t = 1$. One test is simply to determine whether candidate variables for z_t significantly affect A_t.

If there is no evidence of attrition bias from this specification test, it suggests that attrition on observables is ignorable. Then, under the assumption that there is no attrition bias due to unobservable factors, the model can be estimated without adjustment for attrition. If the test detects non-ignorable selection on observables, the resulting bias in the parameter estimates can be corrected by a Inverse Probability Weight estimation method, see the next section.

However, it is impossible to test the ignorability conditions for any possible variables z_t. This is a frequently used argument against selection on observables. A way out might be to test whether there is selection on observables after estimating the model. Based on the principle of Hausman (1978) a simple test on endogeneity can be constructed. The weighted estimator is consistent if one can consistently estimate the sampling probabilities, Wooldridge (2002b). It is, however, not an efficient estimator when the selection is ignorable. The unweighted estimator is only consistent when the selection on observables is non-ignorable, that is, when it only depends on x_t. It is then a more efficient estimator than the weighted estimator. The Hausman-test exploits this.

Thus, the choice for using an Inverse Probability Weight estimation method depends on whether the attrition is selective on non-ignorable observables. Wooldridge (2002b) provides an extensive discussion about this issue. Wooldridge concludes that in general there is no clear-cut answer. In the next section we discuss the Inverse Probability Weight estimation method and show how it can be implemented for our application.

4.2 Inverse Probability Weighting

The idea behind the Inverse Probability Weight estimation method is that people who have higher attrition probability get more weight in the model because they are under-represented.[2] Assume that $q(w, \theta)$ is an objective function, like the likelihood function. Here, w is a random vector and the distribution of w depends on $\theta \in \Theta$. To maximize the objective function assume that θ_0 is the unique solution to

$$\max_{\theta \in \Theta} \mathrm{E}[q(w, \theta)].$$

If one has a random sample w_i for $i = 1 \ldots N$, from the population one can use the sample equivalent and solve $\max_{\theta \in \Theta} \sum_{i=1}^{N} [q(w_i, \theta)]$ to obtain θ_0.

For the ordered probit model with endogenous work-related course choice it applies that θ_0 is the unique solution to the score function

$$\max_{\theta \in \Theta} E\left(\sum_{i=1}^{N} \sum_{j=1}^{J} l_{ij} \right)$$

where $l_{ij}(\theta)$ is the contribution of the i^{th} individual in category j to the log-likelihood function in (4) and $\theta = (\beta, \delta, \alpha_1, \alpha_2, \ldots, \alpha_{J-1}, \rho)$. Because, in the sample with attrition only the individuals with $A = 1$ are present only the density $g(y|x, A = 1)$ is observed. Often A is a function of some of the elements of w, but as mentioned before A can also depend on unobservables. When attrition occurs, one has to review the identification of θ_0. The estimator is now defined by

$$\max_{\theta \in \Theta} \sum_{i=1}^{N} A_i \sum_{j=1}^{J} l_{ij}.$$

This is called the unweighted estimator, $\hat{\theta}_u$. One can now estimate the parameters of the density $f(y_t|x_t)$ by putting weights in the equation based on the probability of dropping

[2] The discussion in the section is based on Wooldridge (2002b).

out of a panel

$$\max_{\theta \in \Theta} \sum_{i=1}^{N} \frac{A_i}{\Pr(A_i = 1|v)} \sum_{j=1}^{J} l_{ij},$$

where w is a subset of v that is observed for all the people. The parameter vector that solves this, $\hat{\theta}_w$, is called the Inverse Probability Weighted estimator. Thus, v contains both the auxiliary variables z_t, influencing only the attrition probability, and the individual characteristics x_t that also explain the dependent variable. Note that both v and z_t are observed for all individuals that do not drop out. Wooldridge (2002b) proves consistency of this estimation method and shows that this is independent of the way the weights are obtained. If the ignorability conditions, mentioned above, are not satisfied, the weighted model is still consistent but not efficient, see Wooldridge (1999).

In Başer et al. (2004) a test is developed for comparing weighted and unweighted least squares estimators. Traditionally, the Hausman test for model adequacy is used under the assumption of homoscedasticity, see Hausman (1978). The weighted estimator is consistent but not efficient if attrition is selective on non-ignorable observables. The unweighted estimator is only consistent when the selection on observables is non-ignorable. The Hausman-test is based on the following statistic:

$$H = (\hat{\theta}_w - \hat{\theta}_u)'(V_w - V_u)^{-1}(\hat{\theta}_w - \hat{\theta}_u),$$

where V_w and V_u are the (asymptotic) covariance matrices of the weighted and unweighted estimators. The test statistic is asymptotically $\chi^2(\nu)$ distributed, where ν is the number of parameters.

5 Estimation Results

We now estimate the influence of taking a work-related course on the reduction of the discrepancy between work requirements and the skills a employee possesses using the OSA panel data for The Netherlands. The decision to take a course may be induced by the perceived skill deficiency and therefore lead to an endogeneity problem. Another problem is that the panel may have selective attrition. To account for these problems we apply the methods mentioned in the previous section. We discuss whether a correction for a selective decision to take a course and a correction for attrition are needed. We will employ three different models. First, a standard ordered probit model for the change in discrepancy is estimated that assumes an exogenous decision to take a course, see (1) and (2). In the second model we extend the model by introducing an endogenous selection equation for the decision to take a course. The change in discrepancy and the decision to take a course are then assumed to follow a bivariate (ordered) probit as is defined in (3). For comparison we also estimated the decision to take a course for the unrelated model. Finally, the model with an endogenous selection is corrected for possible attrition on observables through an IPW estimator.

5.1 Ordered Probit With and Without Correction for Endogeneity

First an standard ordered probit model is estimated.[3] The dependent variable, change in discrepancy, can take seven different values. The maximum observed value is +3 which occurs if an employee reports that her discrepancy has improved from Bad to Good. The minimum observed value is −3 which occurs if an employee reports that her discrepancy has deteriorated from Good to Bad. An improvement from Reasonable to Good (or v.v.) is considered equivalent to an improvement from Bad to Moderate (and Moderate to Reasonable). We think this is a reasonable assumption. Thus, we need to estimate 6 threshold variables, see (2). Of course, a change in discrepancy between function requirements and skills can occur for a number of reasons, not just due to increased knowledge of a work-related course, like because somebody changed its labour market situation. Personal (gender) and work-related (tenure) characteristics may also influence the change in discrepancy. Finally, the change in discrepancy may also be induced by a conjunctural effect. The estimated effects of all these variables can be found in the upper part of the first column of table 4.

The lower part of the first column of table 4 contains the estimated parameters of the binary probit model for the decision to take a course. In the first model we assume that this decision is independent of the change in discrepancy. Thus, the correlation parameter in the bivariate probit model is set to zero. The second column of table 4 presents the estimation results when the change in discrepancy and the decision to take a course are correlated. Note that the estimated parameters in the second model are nearly the same as the estimated parameters of the first model. However, the coefficient of the influence of work-related courses on the change in discrepancy alters sign and the correlation between the main equation and the selection equation is positive. This indicates that ignoring the positive correlation between the decision to take a course and a change in discrepancy leads to misleading impact of work-related courses on this change.

The most important parameter in the model is the effect of the decision to take a course. If we do not correct for possible endogeneity between the change in discrepancy and this decision, it does not have a significant effect on the change in discrepancy. When we allow for endogeneity we find a significant effect. A change in the labour market situation and the years of tenure are also significant in explaining the change in discrepancy.

In discussing the estimation results we should bare in mind that in a probit model, like in many nonlinear models, the coefficients cannot directly be interpreted (even the sign is not ambiguous). The effect of a variable on the probability to get one of the seven ordered discrepancy change is better measured by the marginal effects. Only the marginal effects for the main equation are given, because the selection equation is only used to estimate the parameters of the main equation correctly. The marginal effect represents the change in the probability of 'choosing' a category, due to a change in the explanatory variable. For a dummy this can be interpreted as the difference between 0 and 1 on the

[3] All estimations methods are programmed and estimated in Gauss version 3.2.

chosen category and for a continuous variable it can be interpreted as the change in the probability to choose a category due to a one percent rise of that variable. These marginal effects are shown in table 5, see for a discussion on the computation Wooldridge (2002a).

Somebody with a change in the labour market situation has a 19 percent-point higher chance of no change in discrepancy, category a4, and almost 9 percent-point lower chance to remove the bad fit between skills and function requirements completely. The influence of taking a course clearly increases the probability to decrease the discrepancy between work requirements and skills (category a1 till a3). The probability the discrepancy decreases from Bad to Good is 15 percent-point higher for somebody who takes a course. Taking a course also decreases the probability to remain at the same discrepancy level (with 29 percent-point) and to the probability to move one discrepancy level down, category a5, (with 2.8 percent-point).

In the binary probit model explaining the decision to take a course, most parameters are significant, except for the dummy that indicates that someone thinks (s)he has a good salary. The decision to take a course is significantly influenced by low-education level, gender, change in labour market situation, daily education (not defined as a work-related course) and, residence in one of the three big cities of The Netherlands (Amsterdam, Rotterdam or The Hague).

5.2 Results after Correction for Attrition

As mentioned before, with panel data we face the problem of attrition. If the attrition in the OSA-panel is correlated with the change in discrepancy or if attrition is correlated with the decision to take a course the analysis in the previous section is biased. In section 4 we discussed how to test for selective attrition and how to correct for it. We shall only correct for (possible) selection on observables. We apply an IPW-method as introduced in section 4.2 to correct for the selective attrition in our ordered probit model with a selective decision to take courses (model 2 in the previous section).

First, we need to estimate the IPW weights. We use a binary probit model with two possible endogenous variables that are not in the main equation to estimate the weights. The first variable is a dummy that indicates whether the bad fit between skills and function requirements is a problem at work. The second variable is the number of courses somebody has taken in the two years before the first wave. The values of these variables are observed for both the attritors and the non-attritors. The other variables in the attrition equation are age, squared age and a dummy for a change in labour market situation in the two years before the first wave.

The estimation results for this probit model are given in table 6. The results are clear, the more courses are taken the smaller the chance of attrition and if a bad fit causes problems at work the chance of attrition is also smaller to drop out of the panel. This can be explained by the phenomenon that unhappy people are more likely to respond to surveys, because that is a possibility to spout their comments. Further, it is important to note that all the used variables have significant coefficients.

Now we can calculate the weights for the IPW-method and estimate the weighted version of the ordered probit model. The estimation results are given in the third column of table 4. The obtained parameter values differ from the unweighted parameter values. The parameters have the same signs and roughly the same significance, except that gender is now significant. To have make a better comparison between the weighted and the unweighted model the marginal effects of the covariates are given in table 7.

The marginal effects of the IPW estimation of the ordered probit model differ in size but show the same pattern of sign change over the different categories. The large impact of a change in the labour market situation on the probability to improve the discrepancy between skills and function requirements increases if we correct for selective attrition. Somebody who has experienced a change in the labour market situation has a 21 percent-point lower probability to remove the bad fit between her/his skills and function requirements completely. Such an individual now has a 34 percent-point higher probability to maintain the same discrepancy. The impact of correcting for attrition on the marginal effect is not so large for the indicator of taking a course. Now the taking a course increases the probability to remove a bad fit completely with 16 percent-point. The probability to remain at the same discrepancy level is now decreased by 31 percent-point. Again, the other covariates hardly influence the change in discrepancy.

The significance of all estimated parameters improved with the IPW-method compared to the unweighted estimation because the standard errors decreased. This is an indication that the attrition selection is non-ignorable. To check formally whether the IPW method leads to a more efficient estimation a Hausman test is carried out, as described in section 4.1. All 19 parameters, both from the main equation and from the course selection equation, are included in the test, because all these parameters are weighted in the IPW method. The value resulting from the Hausman-test is 30.08, which is significant at a 10 percent level and almost significant at a 5 percent level. Given that the Hausman test may has low power (see Nicoletti (2002)), we conclude that there is significant selection on observables and that the model estimated with the IPW method is more efficient.

6 Conclusion and Comments

During the past decades changes in labour markets across Europe have accelerated. In the context of transional labour markets (see Schmid (1998)) employability has become more important. Employability implies to empower people by raising their productive capacities. The productive capacities of workers are optimal if their skills and competencies match their function requirements. Work-related training may lead to a better match.

In this paper we use an ordered probit model to analyze the perceived change in discrepancy between function requirements and skills of employees. Based on the bi-annual OSA panel from 1998 till 2002 for The Netherlands, we show that taking a work-related course decreases the discrepancy significantly. We argue that the decision to take a course is endogenous to the change in discrepancy. To correct for this endogeneity problem we estimate a selection equation that is correlated with the change in discrepancy. It appears that ignoring the selective decision to take a course leads to misleading conclusions about the effect of these courses on the change in discrepancy.

Like with any other survey panel we face the problem of panel attrition. In the OSA-panel about 30% of the respondents drop out each wave. Only if the attrition is correlated to the change in discrepancy or with the decision to take courses it will bias the analysis. Such selective attrition can be attributed to variables observed in the data, to variables unobserved in the data or to both. Due to limited ability to check to detect selective attrition on unobservables with the data at our disposal we only adjust for selective attrition on observables.

To correct for the possibility of selection bias we develop an Inverse Probability Weight (IPW) estimation method for the ordered probit with an endogenous binary regressor. The results indicate that selective attrition on observables is a problem in our application. Therefore, the IPW estimator is more efficient than an unweighted estimator. From the implied marginal effects of the IPW estimation we can conclude that taking courses increases the probability to change the fit between skills and function requirements from Bad to Good with 16 percent-point.

One possible route for further research is to incorporate data concerning the general labour market condition, like the unemployment rate or some other economic indicator, into the model. We have the feeling that a substantial part of the increase in the percentage of employees taking courses after 1998 in The Netherlands is due to the economic boom at that time. In this paper we assumed that the behavior of individuals are uncorrelated over time. Another possible extension of the model is, therefore, to allow for individual specific effect, or adding state dependent variables. Finally, testing to distinguish between weighting or not, or equivalently, testing whether selection on observables exists or not, is still at its infancy and asks for elaboration.

References

Başer, O., C. Bradley, J. Gardiner, and C. Given (2004). Estimation from censored medical cost data. *Biometric journal 46*, 351–363.

Bartel, A. and N. Sicherman (1999). Technological change and wages: An interindustry analysis. *Journal of Political Economy 107(2)*, 285–325.

Bassanini, A., A. Booth, G. Brunello, M. D. Paola, and E. Leuven (2005). Workplace training in Europe. Report prepared for Seventh European Conference of the Fondazione Rodolfo Debenedetti.

Black, S. E. and L. M. Lynch (1996). Human-capital investments and technology. *American Economic Review 86*, 263–267.

Black, S. E. and L. M. Lynch (2001). How to compete: The impact of workplace practices and information technology on productivity. *Review of Economics and Statistics 83*, 434–445.

Booth, A. (1991). Job-related formal training: Who receives it and what is it worth? *Oxford Bullitin of Economics and Statistics 53*, 281–294.

Booth, A. (1993). Private sector training and graduate earnings. *Review of Economics and Statistics 75*, 164–170.

Cragg, J. G. and R. Uhler (1970). The demand for automobiles. *Canadian Journal of Economics 3*, 386–406.

De Koning, J. (2002). Transitional labour markets: An economists view. In H. Mosley, J. O'Reilly, and K. Schömann (Eds.), *Labour Markets, Gender and Institutional Change. Essays in Honour of Günther Schmid*, pp. 290–315. Cheltenham, UK and Northhampton, MA, USA: Edward Elgar.

Ester, P. and H. Vinken (2000). *Van later zorg: verwachtingen van Nederlanders over arbeid, zorg en vrije tijd in de* 21^e *eeuw: Het OSA Toekomst van de arbeid survey.* Tilburg: OSA. (in Dutch).

Fitzgerald, J., P. Gottschalk, and R. Moffitt (1998). An analysis of sample attrition on the second generation of respondents in the Michigan Panel Study of Income Dynamics. *The journal of Human Resources 33*(2), 300–344.

Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica 46*, 1251–1271.

Lillard, L. A. and H. W. Tan (1992). Private sector training: Who gets it and what are its effects? *Research in Labour Economics 13*, 1–62.

Lynch, L. M. (1992). Private sector training and the earnings of young workers. *American Economic Review 82*, 299–312.

Nicoletti, C. (2002). Non-response in dynamic panel data models. Working paper 31, Institute for Social and Economic Research.

OECD (2003). *Employment Outlook*. Paris.

Schmid, G. (1998). Transitional labour markets: A new European employment strategy. Discussion paper FS I 98-2006, Wissenschaftszentrum Berlin für Sozialforschung.

Schmid, G. (2002). Towards a theory of transitional labour markets. In G. Schmid and B. Gazier (Eds.), *The dynamics of full employment: Social integration through transitional labour markets*, pp. 151–195. Cheltenham, UK and Northhampton, MA, USA: Edward Elgar.

SER-report (2001). Levensloopbanen: Gevolgen van veranderende arbeidspatronen. Technical report, SER. Rapport van de commissie sociaal-economische deskundigen (in Dutch).

Van den Berg, G. J. and M. Lindeboom (1998). Attrition in panel survey data and the estimation of multi-state labor market models. *The journal of Human Resources 33*(2), 458–478.

Winkelmann, R. (1994). Training, earnings and mobility in Germany. Discussion paper no. 982, Centre for Economic Policy Research.

Wooldridge, J. M. (1999). Asymptotic properties of M-estimators for variable probability sampling. *Econometrica 67*, 1385–1406.

Wooldridge, J. M. (2002a). *Econometric analysis of cross section and panel data*. Cambridge: MIT press.

Wooldridge, J. M. (2002b). Inverse probability weigted M-estimators for sample selection, attrition and stratification. *Portugese Economic Journal 1*, 117–139.

A Tables

Table 1: Fit between knowledge and function requirements.

	Good fit			Bad fit		
	1998	2000	2002	1998	2000	2002
Total	79%	85%	77%	6%	2%	5%
Male	81%	87%	78%	5%	2%	5%
Female	77%	82%	76%	7%	3%	6%
Labour situation changed	83%	87%	82%	4%	1%	4%
No change labour situation	72%	81%	77%	8%	3%	5%
Tenure						
< 1 year	90%	100%	70%	5%	0%	9%
1-2 year	67%	76%	73%	11%	5%	5%
3-5 year	75%	84%	76%	7%	1%	5%
6-10 year	81%	82%	77%	5%	4%	5%
11-20 year	82%	84%	79%	4%	3%	5%
>20 year	83%	89%	83%	4%	1%	2%
Age						
16-29	71%	77%	67%	9%	4%	8%
30-44	79%	84%	76%	5%	2%	5%
45+	85%	88%	82%	4%	2%	4%
	Knowledge is more			Knowledge is less		
	1998	2000	2002	1998	2000	2002
Total	33%	25%	26%	10%	3%	4%

[a] Residence in big city: The Hague, Rotterdam or Amsterdam.
[b] Residence in the north of The Netherlands.

Table 2: Descriptive statistics work-related courses.

	courses taken			Average # of courses		
	1998	2000	2002	1998	2000	2002
Total	16%	46%	37%	0.22	0.79	0.69
Male	20%	52%	42%	0.29	0.92	0.84
Female	11%	40%	32%	0.15	0.67	0.56
Labour situation changed	14%	35%	32%	0.21	0.53	0.58
No change labour situation	18%	42%	41%	0.25	0.63	0.76
Primary school	7%	28%	20%	0.09	0.50	0.31
High-school	17%	50%	40%	0.25	0.87	0.75
BA-degree	25%	57%	46%	0.35	1.03	0.94
MA-degree	20%	51%	45%	0.34	0.89	0.96
Big city[a]	15%	36%	32%	0.23	0.61	0.63
No Big city	16%	47%	37%	0.22	0.81	0.70
North[b]	13%	46%	34%	0.18	0.80	0.60
No North	16%	46%	37%	0.22	0.79	0.71
Tenure						
< 1 year	3%	27%	18%	0.03	0.43	0.29
1-2 year	15%	47%	46%	0.17	0.82	0.88
3-5 year	17%	55%	46%	0.25	0.95	0.90
6-10 year	21%	52%	44%	0.31	0.93	0.84
11-20 year	21%	56%	43%	0.31	0.99	0.88
>20 year	23%	53%	44%	0.34	0.92	0.84
Age						
16-29	16%	44%	41%	0.22	0.74	0.82
30-44	18%	50%	43%	0.27	0.89	0.83
45+	13%	42%	29%	0.17	0.71	0.53

[a] Residence in big city: The Hague, Rotterdam or Amsterdam.
[b] Residence in the north of The Netherlands.

Table 3: Comparison of attritors and non-attritors.

	2000		2002	
	Non-attritors	attritors	Non-attritors	attritors
Average age	42	42	44	44
Female	51%	49%	52%	50%
Big city[a]	8%	11%	10%	11%
North[b]	12%	13%	12%	13%
No children	22%	31%	21%	29%
Av. # of children	1.80	1.60	1.83	1.68
Married	77%	67%	72%	65%
Cohabiting	8%	11%	8%	9%
Divorced	4%	4%	5%	6%
Never married and not cohabiting	10%	16%	13%	18%
People in employment	71%	65%	70%	63%
Self-employed	5%	5%	5%	5%
Unemployed	3%	3%	3%	3%
Non-participants	21%	25%	20%	27%
Change in labour situation since first wave	70%	67%	66%	69%
Primary school	6%	8%	5%	7%
High-school	34%	33%	32%	31%
BA-degree	16%	13%	22%	18%
MA-degree	5%	4%	6%	6%
Av. fit Knowledge/skills	1.35	1.38	1.22	1.23
Fit knowledge-skills is problem at work	19%	17%	7%	5%
Number of courses preceding two years	0.25	0.19	0.80	0.78

[a] Big cities are The Hague, Rotterdam, Amsterdam.
[b] North are the provinces in the north of The Netherlands.

Table 4: Parameters and standard errors of the models.

	Model 1	Model 2	IPW
	Equation for change in discrepancy.		
Threshold 1	$-0.99(0.71)$	$-1.44^{*}(0.12)$	-1.84^{*} (0.077)
Threshold 2	$-0.58(0.71)$	$-1.07^{*}(0.12)$	$-1.48^{*}(0.068)$
Threshold 3	$-0.037(0.72)$	$-0.57^{*}(0.13)$	$-0.98^{*}(0.065)$
Threshold 4	$2.53^{*}(0.71)$	$1.73^{*}(0.20)$	$1.28^{*}(0.11)$
Threshold 5	$3.42^{*}(0.72)$	$2.53^{*}(0.23)$	$2.02^{*}(0.13)$
Threshold 6	$4.06^{*}(0.73)$	$3.12^{*}(0.25)$	$2.53^{*}(0.15)$
Female	$-0.0046(0.32)$	$-0.062(0.046)$	$-0.11^{*}(0.021)$
Survey 1998 2000	$0.029(0.19)$	$0.042(0.050)$	$-0.016(0.031)$
Labour situation changed	$0.98^{*}(0.046)$	$0.83^{*}(0.040)$	$0.42^{*}(0.021)$
Years at current employer	$-0.14^{*}(0.014)$	$-0.11^{*}(0.011)$	$-0.065^{*}(0.0063)$
Course taken	$0.047(0.044)$	$-0.84^{*}(0.17)$	$-0.52^{*}(0.063)$
	Selection equation for taking a course (=1).		
Primary school	$-0.34^{*}(0.063)$	$-0.30^{*}(0.061)$	$-0.18^{*}(0.036)$
High salary	$-0.043(0.085)$	$-0.015(0.079)$	$0.0094(0.049)$
Female	$-0.16^{**}(0.065)$	$-0.16^{*}(0.046)$	$-0.13^{**}(0.052)$
Labour situation changed	$-0.13^{*}(0.044)$	$-0.14^{*}(0.034)$	$-0.091^{**}(0.037)$
Daily education	$-0.049^{***}(0.030)$	$-0.061^{**}(0.028)$	$-0.028(0.018)$
Big city[a]	$-0.21^{**}(0.085)$	$-0.20^{*}(0.074)$	$-0.10^{**}(0.043)$
Intercept	$0.48^{*}(0.16)$	$0.50^{*}(0.095)$	$0.34^{**}(0.14)$
Correlation	0	$0.55^{*}(0.096)$	$0.57^{*}(0.064)$
Log-likelihood	-5149.57	-5148.15	-5203.22
Number of observations	3383	3383	3383

Model 1 is an Ordered probit for the change in discrepancy and a binary Probit for selection equation, Model 2 is model 1 with correlation between the main equation and the selection equation. (*) significant at 1 percent, (**) significant at 5 percent and (***) significant at 10 percent.

[a] Big cities are The Hague, Rotterdam, Amsterdam.

Table 5: Marginal effects on the change in discrepancy for the ordered probit model with endogenous decision to take a course.

	a1[1]	a2	a3	a4	a5	a6	a7
Female	0.0062	0.0052	0.0089	-0.015	-0.0046	-0.00056	-7.9E-05
Survey 1998 2000	-0.0044	-0.0036	-0.0061	0.011	0.0030	0.00036	5.1E-05
Labour situation changed	-0.087	-0.068	-0.11	0.19	0.067	0.0097	0.0016
Years at current employer	0.00014	0.0013	0.0025	-0.0035	-0.0015	-0.00019	-2.8E-05
Course taken	0.15	0.099	0.071	-0.29	-0.028	-0.0038	-0.00021

[1] a1...a7 are the seven different categories for the change in discrepancy, with a1 indicates an improvement from Bad to Good, a4 no change and, a7 a deterioration from Good to Bad.

Table 6: Probit estimations for attrition equation.

	Probit model
Bad fit is problem at work	0.469(0.083)
Number of courses	0.091(0.025)
Age	0.279(0.014)
$(\text{Age}/10)^2$	$-0.328(0.016)$
Labour market situation changed	0.115(0.044)
R^2	0.081[1]

All variables are significant at the 5 per cent level.
[1] R^2 of Cragg and Uhler (1970) is used.

Table 7: Marginal effects on the change in discrepancy for the IPW estimations of the ordered probit model.

	a1[1]	a2	a3	a4	a5	a6	a7
Female	0.0031	0.0036	0.0079	-0.0055	-0.0077	-0.0013	-0.00023
Survey 1998 2000	-0.0027	-0.0028	-0.0058	0.0061	0.0044	0.00065	0.00011
Labour situation changed	-0.21	-0.085	-0.067	0.34	0.017	0.0012	0.00013
Years at current employer	0.00022	0.00024	0.00050	-0.00048	-0.00041	-6.3E-05	-1.1E-05
Course taken	0.16	0.086	0.099	-0.31	-0.030	-0.0028	-0.00033

[1] a1...a7 are the seven different categories for the change in discrepancy, with a1 indicates an improvement from Bad to Good, a4 no change and, a7 a deterioration from Good to Bad.