

Training and low-pay mobility. The case of the UK, the Netherlands and Germany

Dimitris Pavlopoulos* Ruud Muffels Jeroen K. Vermunt

Abstract

This paper investigates the effect of training on low-pay mobility in the UK, the Netherlands and Germany. To estimate the 'true' effect of training we correct for measurement error and transitory fluctuations of earnings. This is accomplished by using a random-effects multinomial logit model with a latent structure to correct for measurement error. Our results indicate that in the UK and in the Netherlands, training increases the likelihood for moving from low to higher pay. In Germany, we find a similar but not significant effect. In all three countries, we find that training reduces the likelihood for a transition from higher pay to low pay. However, this result is only significant in the UK. No complementarities are found in the effect of training with the effect of education and labour market experience.

Keywords: Low pay, training, measurement error, Markov models.

JEL-code: J31, C23.

*Corresponding author: Dimitris Pavlopoulos, CEPS/INSTEAD, 44 rue Emile Mark, BP48, L-4501 Differdange, Luxembourg, tel. +352 585855 556, email: dimitris.pavlopoulos@ceps.lu.

1 Introduction

The issue of low-pay mobility is receiving increasing interest in economic and political debate (OECD, 1996, 1997, 2003; Acemoglu, 2003b, 2003a). Low-pay mobility may have an equalizing effect on the earnings of workers at the bottom of the wage distribution. The higher the level of upward low-pay mobility in a country, the greater the chances low-paid workers have of improving their earnings level. Enhancing the participation in training programs has been suggested as one important means of increasing the (upward) mobility opportunities of workers. Through training low-paid workers may improve their skills and their productivity and therewith increase their wage in the same or in a different job. Previous research suggests that training has a positive effect on wages especially when the worker stays in the same job. However, the effect of training on low-pay mobility has not been explicitly investigated.

Furthermore, studies on low-pay mobility typically do not control either for measurement error or for the fact that some of the true observed mobility is completely transitory and therefore is not explained by any economic process. The presence of measurement error in data from household surveys results into a severe overestimation of mobility (Hagenaars, 1994; Pischke, 1995; Gottschalk, 2005). Rendtel et al. (1998) find that approximately half of the observed poverty transitions from the German Socio-Economic Panel (GSOEP) are spurious. Except for the overestimation of average mobility, measurement error underestimates the effect of the usual covariates of earnings (Bound et al., 2001). When failing to control for classification error, the dependent variable in an earnings transition model is full of noise. Therefore, the effect of covariates in such a model will, most probably, be underestimated.¹ The effect of the presence of 'randomness' in low-pay mobility is similar to measurement error. If the wage of individuals is very close under the low-pay threshold then even a light 'churning' in the wage distribution - unrelated to any individual factors, such as experience accumulation or job change - may lead him above the threshold. In this way, the overall low-pay mobility will increase and the effect of the covariates on earnings will be attenuated.

The aim of this paper is to investigate the effect of training on low-pay mobility while accounting for measurement error and 'randomness' in mobility. For this purpose, we

¹This is not always the case. It rather depends on whether there is error in the measurement of the covariates and on whether this error is correlated with the error in earnings.

develop a panel multinomial logit model for low-pay transitions with a latent structure in order to correct for measurement error. This model is a Mixed Latent Markov model that is advancing the model of Rendtel et al. (1998). While Rendtel et al control for measurement error in aggregate transition probabilities, we also correct for observed and unobserved heterogeneity and moreover work with a much longer time series. In this way, we relax the unattractive property of population homogeneity that is assumed in most of the studies using Markov models on labour market transitions. In our analysis, we distinguish between two earnings states, low-paid and higher-paid, as well as the state of non-employment. For low pay, we apply the most common definition: the threshold is set to the two-thirds of the median wage (OECD, 1996). The analysis is performed in three countries with very different labour markets: the UK with a liberal-unregulated labour market, Germany with a strongly-regulated labour market and the Netherlands with a semi-regulated labour market.

The rest of the paper is organized as follows: Section 2 reviews the literature on the relationship between training and earnings. Section 3 elaborates on the model we apply. The three datasets we use are presented in section 4. In section 5, we discuss the results of our data analysis. Finally, section 6 contains the conclusions of our study as well as some issues for further research.

2 The relationship of training with earnings

The relationship of human capital with earnings is well documented in economics (Becker, 1962; Mincer, 1986). Standard economic theory suggests that two types of human capital affect earnings formation. General human capital that concern skills that a worker accumulates from education and from labour market experience and firm-specific human capital that refers to skills that a worker acquires on the job and are usually not transferable across employers. General education and formal vocational training, such as apprenticeship, provide skills that increase the productivity of the worker throughout his working career. However, the effect of short-term training programs is more ambiguous. Some of them also provide skills and qualifications that the worker can transfer from job to job, but others - especially on-the-job training programs - provide skills that are job-specific. Human capital theory predicts that training has a negative effect on earnings during the

period of training provision, as the worker and the firm share the costs, and a positive effect thereafter when the worker can increase his productivity using the new skills he acquired during the training period.

Empirical evidence is rather in accordance with the predictions of theory. Mincer (1988) finds that American workers receiving training have a 4-6% higher wage growth than the rest of their colleagues. He also finds that training creates steeper wage profiles regardless of whether the worker changes a firm or not. Parent (1999) suggests that wage gains from training exist and are transferable across employers for young American workers. Booth (1991) suggests that wages of British male workers are 11.2% higher when receiving training. For the female workers the training premium is even higher, namely 18.1%. Lynch (1992) distinguishes between off-the-job and on-the-job training of young workers in the US. She finds that previous off-the-job training, previous apprenticeship and current on-the-job training increases wages. Moreover, she suggests that there is quite some heterogeneity in the returns of training. These returns are higher for the medium and highly educated as well as for the unionized workers. Duncan and Hoffman (1979) find that in the US, the returns to training are rather uniform between men and women as well as between native and immigrant workers. Nevertheless, they suggest that differences in the amount of training account for as far as 20% of the earnings gap between black and white workers and 10% of the earnings gap between male and female workers. Evidence for Germany is rather mixed. While Pischke (2001) finds no significant association between work-related training and earnings, Kuckulenz and Zwick (2003) find a positive effect of training on earnings even after controlling for selectivity into training. For the Netherlands, Leuven and Oosterbeek (2002) find no effect of training on earnings and suggest that the effect that is typically found in the relevant studies is actually the return of unobserved characteristics.

Studies that investigate low-pay mobility use training as a covariate although their focus is never training per se. Sloane and Theodossiou (1996) find that, in the UK, recent training increases the probability for a low-to-higher pay transition. Similarly, Stewart and Swaffield (1999) find that, in the UK, training reduces by 5-10% the probability of remaining in low pay. Blázquez Cuesta and Salverda (2007) reach the same conclusion for the Netherlands. They also find that the incidence of training is lower for the low paid than for their higher paid colleagues. However, no study has ever focused explicitly on modeling the effect of training on low-pay mobility. Moreover, almost all of the papers studying the

effect of training on wages consider only training programs that the individual received while being employed.

In this paper, we focus on the effect of training on low-pay mobility. We consider all training programs that were followed in the year prior to the survey, regardless of whether the individual was employed or not during the training period.

3 A Mixed Latent Markov model

Specification of the model

Our aim is to investigate the effect of training on the year-to-year transitions from and to low pay. Therefore, we start from random-effects multinomial logit model. By applying a random-effects instead of a cross-sectional multinomial logit regression we also control for the endogeneity of training. This problem of endogeneity is caused by the fact that sometimes individuals with more intrinsic motivation self-select themselves to training programs. Thus, what appears as the effect of training on earnings may be just capturing the effect of unobserved heterogeneity.

Furthermore, we want to study the effect of training on 'real' low-pay mobility, i.e. low-pay mobility net of measurement error and transitory fluctuations. For this purpose, we impose a latent structure to the multinomial logit model in the framework of the Latent Markov models (van de Pol & Langeheine, 1990; Langeheine & van de Pol, 1990; Vermunt et al., 1999; Bassi et al., 2000). The simplest form of this model is depicted in Figure 1. According to this model, the true state X_{it} of an individual i at a time point t cannot be observed; it is a latent state. We rather observe state Y_{it} , which might differ from the true (latent) state X_{it} . Y_{it} and X_{it} are probabilistically related.² The observed states at different time points are mutually independent, conditional on the true latent states. In

²To understand how this model estimates measurement error, let us assume a fictitious transition matrix for a discrete variable X with two categories and between two time points. We further assume that there is error in the observation of the variable X . Instead of X_1 and X_2 , we rather observe the states Y_1 and Y_2 . The model for the joint distribution of Y_1 and Y_2 has the form of a Latent Class model for two time points. More specifically, the joint distribution of the observed states Y_1 and Y_2 can be expressed as follows:

$$P(Y_1 = y_1, Y_2 = y_2) = \sum_{\mathbf{Y}_1, \mathbf{Y}_2} [P(X_1 = x_1)P(X_2 = x_2|X_1 = x_1) P(Y_1 = y_1|X_1 = x_1)P(Y_2 = y_2|X_2 = x_2)]. \quad (1)$$

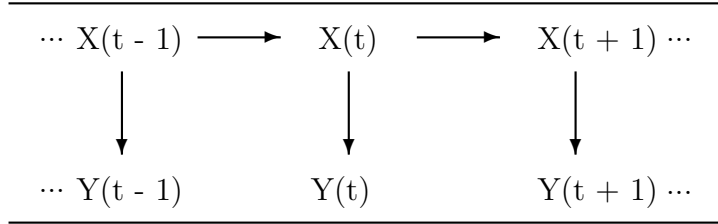


Figure 1: Path diagram for the Latent Markov model

other words, we assume that measurement error is not serially correlated in any way. This means that the independent classification error (ICE) assumption is made (Bassi et al., 2000).

The true state X_{it} follows a Markov process. Thus, the state of an individual i at time point t , X_{it} , is independent of the state at time point t' , $X_{it'}$, where $t' < t - 1$, conditionally on the state at $t - 1$, X_{it-1} . An arrow indicates a direct effect, for example of the state at one time point on the state at the next time point. In our study, X_{it} and Y_{it} are the true and observed earnings state, respectively that are assumed to take on three values: low-paid, higher-paid and non-employed ('other').³

It is important to stress the fact that our model corrects for observed and unobserved heterogeneity and therefore tackles the issue raised by Shorrocks (1976) against the use of Markov models in income mobility. Observed heterogeneity is controlled with the approach

In the above probability expression $P(X_1 = x_1)$ denotes the probability of being in the latent (true) state x_1 at the first time point and $P(X_2 = x_2|X_1 = x_1)$ the probability of being in the latent state x_2 at the second time point, conditional on being in the latent state x_1 at the first time point. The other two terms refer to the relationship between the latent and observed states, and represent the measurement error component. $P(Y_1 = y_1|X_1 = x_1)$ denotes the probability of observing the state y_1 conditional on being in the latent (true) state x_1 . The expected observed transition probability is:

$$P(Y_2 = y_2|Y_1 = y_1) = \frac{P(Y_1 = y_1, Y_2 = y_2)}{P(Y_1 = y_1)} = \frac{P(Y_1 = y_1, Y_2 = y_2)}{\sum_{Y_2} P(Y_1 = y_1, Y_2 = y_2)} . \quad (2)$$

To illustrate the impact of measurement error, assume that $P(X_2 = x_2|X_1 = x_1) = .05$ for $x_1 \neq x_2$ and that $P(Y_1 = y_1|X_1 = x_1) = P(Y_2 = y_2|X_2 = x_2) = .05$ for $y_1 \neq x_1$ and $y_2 \neq x_2$. Using equations (1) and (2) one can easily verify that the probability $P(Y_2 = y_2|Y_1 = y_1)$ for $y_1 \neq y_2$ equals .136 . In other words, even a small amount of classification error (5%) results in a large increase in the number of the observed transitions, here by a factor of 2.72 (13.6% observed versus 5% real transitions).

³It is obvious that our definition of earnings states includes a state where the individual has no income from paid employment, the 'other' or non-employment state. For reasons of simplicity, however, we will refer to these states as 'earnings states'.

suggested by Vermunt et al. (1999). Specifically, we allow the covariates \mathbf{Z}_{it} to affect the latent transition probabilities between latent states X_{it-1} and X_{it} . These covariates are assumed to be uncorrelated to the error. As mentioned earlier, to control for unobserved heterogeneity we use the standard random-effects approach.

The problem of 'initial conditions' is also tackled by our model. This problem refers to the fact that previous experiences of low pay may affect the low-pay transition probability and therefore the group of individuals being in low pay at a certain time point t may not be random (Stewart & Swaffield, 1999; Cappellari & Jenkins, 2004b). Our full model assumes that the joint distribution of the individual effects affecting the transition probability F_{1i} and the individual effects affecting the initial state F_{2i} follow a bivariate normal distribution that is characterized by two variances and one correlation to be estimated:

$$\sigma_1^2 = \text{var}(F_{1i})$$

$$\sigma_2^2 = \text{var}(F_{2i})$$

$$\rho = \text{corr}(F_{1i}, F_{2i})$$

The joint probability of having a particular state path conditional on covariate values can be expressed as:

$$\begin{aligned}
 P(\mathbf{Y}_i = \mathbf{y}_i | \mathbf{Z}_i) &= \int \int \sum_{x_0=1}^3 \sum_{x_1=1}^3 \dots \sum_{x_T=1}^3 P(X_{i0} = x_0 | \mathbf{Z}_{i1}, F_{1i}) \\
 &\quad \prod_{t=1}^T [P(X_{it} = x_t | X_{it-1} = x_{t-1}, \mathbf{Z}_{it}, F_{2i})] \\
 &\quad \prod_{t=0}^T P(Y_{it} = y_{it} | X_{it} = x_t) f(F_{1i}, F_{2i}) dF_{1i} dF_{2i}, \quad (3)
 \end{aligned}$$

where $i = 1, \dots, I$ is the index for the individual, $t = 0, \dots, T$ represents the time points and $f(F_{1i}, F_{2i})$ is the joint density function for the individual effects F_{1i} and F_{2i} .

Because this full model is computationally very intensity and, furthermore, it adds little to the model fit, the model that we finally use assumes a perfect correlation between the two individual effects F_{1i} and F_{2i} .

The probability $P(Y_{it} = y_{it} | X_{it} = x_t)$ is supposed to represent the measurement or classification error. This model, however, does not only filter out measurement error. What

the model actually does is derive from the longitudinal information for all individuals a pattern of ‘regular transition behaviour’ for individuals belonging to state x (Vermunt, 2004). A spurious transition results in a violation of the first-order Markov process. However, a true but ‘unexpected’ transition may also be classified as spurious. This may be the case if the position of the worker in the wage distribution in $t - 1$ was so close to the low-pay threshold that a small overall change in the distribution moves him above this threshold in t . Thus, the ‘true’ transitions we estimate are the transitions between the states x_j and x_k when accompanied by a change in transition ‘behavior’; from the transition ‘behaviour’ corresponding to individuals in state x_j to the transition ‘behaviour’ of individuals in state x_k . A further discussion on the validity of this model can be found in Pavlopoulos (2007).

For identification reasons, we restrict the probability of observing a state Y_{it} conditional on the true state X_{it} to be constant over time, so $P(Y_{t-1} = s | X_{t-1} = r) = P(Y_t = s | X_t = r)$ for every t . With these restrictions, the model is identified with at least three time points (Vermunt et al., 1999).

Parameter estimation

The estimates for the parameters of our model are obtained by means of maximum likelihood. Specifically, we use a variant of the well-known Expected Maximization (EM) algorithm (Dempster et al., 1977), which switches between an E step and a M step until it achieves convergence. The E-step of the EM algorithm involves computing the expected value of the complete data log-likelihood or, more intuitively, filling in the missing data (here the unobserved class memberships and the unobserved random effects) with their expected values given the current parameter values and the observed data. In the M step, standard estimation methods are used to update the model parameter such that the expected complete data log-likelihood is maximized. In our case the M step involves using the filled-in expected values as though these were observed data in logistic regression analysis. The E and M steps cycle until a certain converge criterion is reached.

The relevant variant of EM, which is called the forward-backward or Baum-Welch algorithm, is implemented in the recent syntax version of the statistical software LatentGOLD (Vermunt & Magidson, 2008). The standard EM algorithm cannot be applied for Latent Markov models for many time points T , as the time and storage needed for computation

increases exponentially with T (Vermunt et al., 1999). The extended version of the forward-backward algorithm we applied supports multivariate analysis and control for unobserved heterogeneity, features that are required for our analysis. Details on this algorithm can be found in Vermunt et al. (2008).

4 Data and main concepts

The study uses data for the period 1991-2004 from three national panel datasets. For the UK, we use waves 1 to 14 of the British Household Panel Survey (BHPS) (Taylor et al., 2006), covering the years 1991-2004. For Germany, we make use of 14 waves of the German Socio-Economic Panel (GSOEP) (Wagner et al., 1993), which cover the period 1991-2004. For the Netherlands, our data come from the Socio-Economic Panel (SEP) (CBS, 1991). We make use of the last 9 waves of the panel, covering the years 1994-2002.⁴

The labour markets of the three countries under scrutiny present important differences. In the liberal-unregulated labour market of the UK, there is a much higher level of job and wage mobility than in the regulated labour market of Germany and the semi-regulated labour market of the Netherlands. The UK and Germany are countries with a high incidence of training while the Netherlands, training is less common than in the other two countries (Pischke, 2001; Arulampalam et al., 2004). Investigating low-pay transitions that are corrected for measurement error and random fluctuations will be more informative concerning the real extent of cross-country differences in the effect of training on low-pay mobility than by simply looking at observed transitions.

Since we focus on earnings transitions of employed individuals, our sample consists of prime age males (aged 25-55). We restrict our sample to males in order to avoid the problem of endogeneity of female labour supply. Our main economic variable is the earnings state of the individual, defined as the level of the hourly wage. Since there is no direct information available on an individual's hourly wage, this is computed by dividing the earnings of last month from paid employment by the total amount of the monthly hours worked. Unusual overtime pay and bonuses are not included in the earnings of last month. As in the SEP

⁴The BHPS data were made available by the Data Archive at Essex University. The GSOEP was provided by the German Institute for Economic Research. The SEP was made accessible by Statistics Netherlands.

and the GSOEP, only retrospective wage information is available, the wage in t is derived from wave $t + 1$. We define two real earnings states, low paid and higher paid, as well as an ‘other’ (non-employment) state.

Only individuals reporting paid employment as their main employment status are classified in one of the two earnings states. The self-employed are clustered in the ‘other’ state (non-employment state). Individuals who are in education or in apprenticeship - especially relevant for Germany - are also classified as non-employed. This ‘other’ state is very heterogeneous implying that transitions to and from ‘other’ cannot be expected to have a clear interpretation. However, the inclusion of such a state in our dependent variable is important from both a substantial and methodological point of view. Several studies, such as Cappellari and Jenkins (2004a) and Stewart (2007) show that transitions to non-employment are common for low-paid workers. Moreover, ignoring the non-employment state would make it impossible to define a Latent Markov model as the latent states should not only be mutually exclusive but also exhaustive.

The three datasets that are used include detailed information on training incidents. From the BHPS, for the first 7 waves, we use the question ‘Since September 1st last year, have you taken part in any education or training, other than training that was part of any job you may have?’. For the rest of the waves, we use the question, ‘Have you taken part in any other training schemes or courses at all since September 1st (of the previous year) or completed a course of training which led to a qualification?’. From the GSOEP, we use the question ‘Are you currently receiving training or education? Are you in school, college, career training, or are you attending a further education course?’. As this question refers to training that the individual is currently receiving, we include the lag of the relevant variable as a covariate in the model. We were not able to retrieve the questionnaire of the Dutch Socioeconomic Panel. Although the questions look similar, they measure different incidents of training in the three countries. In the UK and in the Netherlands, the question refers to all types of courses (part-time and full-time) while in Germany only in full-time courses.⁵ Therefore, we expect training incidence in the UK and in the Netherlands to be much higher than in Germany.

Each individual is included in the analysis from the time point he first enters the

⁵For Germany, there is also a question about all kinds of training courses followed. The variable that is derived by this question is used by Pischke (2001). However, this question is not asked every year and the retrospective information that it collects is unreliable.

survey. Using maximum likelihood estimation with missing data, we deal with the fact that at some occasions information for the earnings state of the individual may be missing, due to non-response or temporary attrition. This approach does not cause any bias as long as non-response is random conditionally on observed values, that is, as long as the missing data is missing at random (MAR). Missing values in covariates were imputed by interpolation when possible.⁶ The remaining missing values were imputed by the mean of the relevant variable.

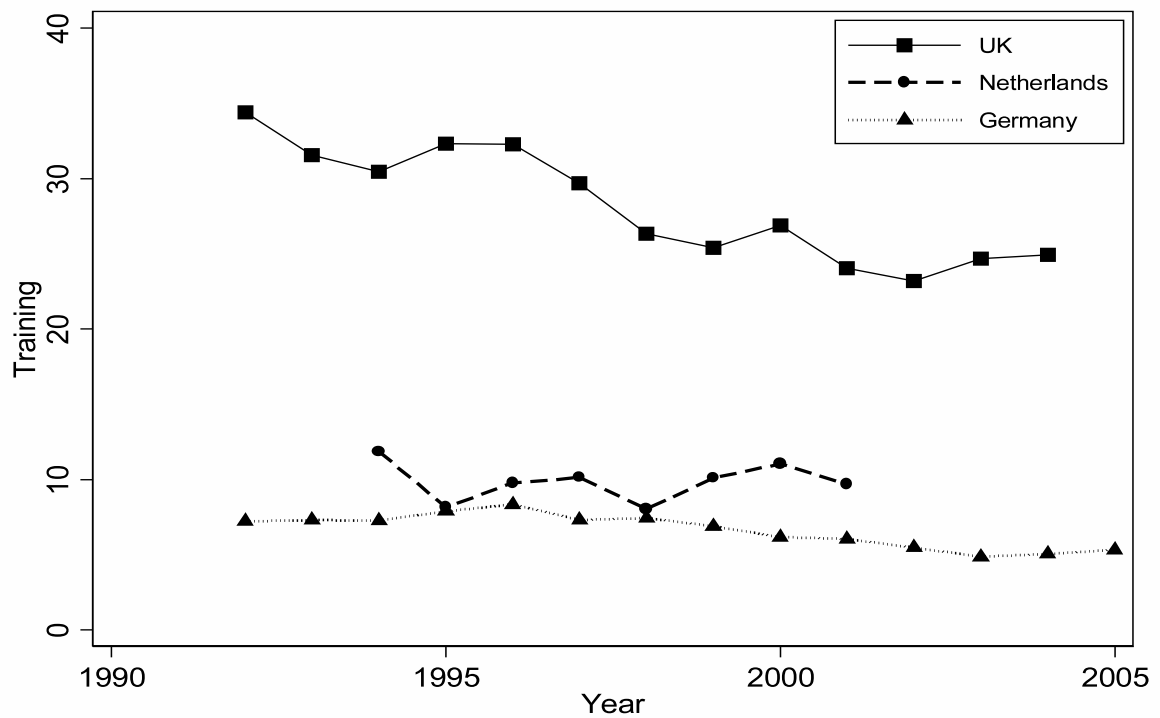


Figure 1: Training incidence across countries in percentages.

⁶For example if the individual reported ‘higher education’ in $t - 1$ and $t + 1$, and the value for education was missing for t we imputed the value for education in t as being ‘higher education’.

5 Results

Some descriptives

The incidence of training over time is depicted in Figure 1. As expected, work-related training is much more common in the UK than in Germany. In the UK 23.5-34.4% of male employees go through some type of training every year. However, the percentage of trainees is decreasing. In the Netherlands, the incidence of training varies between 8% and 11.9%, while in Germany between 4.9% and 8.3%. Considering the definition of the training variable, it appears that training incidence in the Netherlands is low. This finding is in accordance with Arulampalam et al. (2004).

Table 1: Low-pay mobility conditional on training incidence (in percentages)

State	UK			Netherlands			Germany		
	no training	training	Total	no training	training	Total	no training	training	Total
Low pay	53.3	45.6	51.7	48.5	39.8	47.5	48.9	37.7	47.9
Higher pay	36.9	48.2	39.4	42.9	54.8	44.3	37	39.8	37.3
Other	9.7	6.2	8.9	8.6	5.4	8.2	14.1	22.5	14.8
Total	100	100	100	100	100	100	100	100	100
Pearson (prob)	30.07 (0.000)			4.97 (0.083)			13.16 (0.001)		
Cases	2,316	702	3,018	815	101	916	1,942	191	2,133

Table 1 presents some descriptives on the effect of training on low-pay mobility. In all three countries, the percentage of low-paid workers that cross the low-pay threshold in a one-year period is higher for those that have followed a training course. In the UK, 48.2% of the low paid that received training moved to higher pay, while only 36.9% of their colleagues that didn't receive training moved to higher pay. In the Netherlands, 54.8% of the low-paid workers that finished a training course crossed the threshold, while only 42.9% of those not having followed a training course. In Germany, differences are less pronounced. 39.8% of the low-paid workers that followed a full-time training course moved to higher pay while 37% of those not having followed a full-time training course. The Pearson chi-square statistic shows that the association between the earnings state and training incidence is significant (in the Netherlands only at the 10% level). In Germany, the Pearson chi-square statistic for the cells (not presented here) is insignificant for the transitions to higher pay,

which indicates that the overall association that is found is not driven by the differences in transitions to higher pay. However, the association of training with low-pay mobility may just be capturing the effect of observed and unobserved heterogeneity. The multivariate analysis will be more informative on this issue.

Table 2: Model comparison

	UK		Netherlands		Germany	
	LL	BIC (LL)	LL	BIC (LL)	LL	BIC (LL)
1. Markov	-21,015.9	42,695.8	-9,260.6	18,894.3	-22,809.8	47,109.1
2. Latent Markov	-20,205.3	40,966.9	-8,948.7	18,321.4	-22,017.1	44,815.3
3. Mixed Markov	-20,612.4	41,763.2	-8,994.8	18,396.6	-22,452.9	45,505.1
4. Mixed Latent Markov	-20,118.5	40,829.1	-8,925.3	18,308.6	-21,962.9	44,579.6

Note: LL refers to the Log Likelihood and BIC (LL) refers to the Bayesian Information Criterion that is based on the Log Likelihood value.

Results of the multivariate analysis

In total, we applied four versions of the model described by equation (3); namely, a standard Markov transition multinomial logit model, a model with a latent structure (Latent Markov model), a model with random-effects (Mixed Markov), and a random-effects model with a latent structure (Mixed Latent Markov model) correcting for observed and unobserved heterogeneity. Models 2 and 4 correct for measurement error. In models 3 and 4, we assume perfect correlation between the unobserved effects that affect the initial state and the transition probability. The Log-Likelihood values and the BIC values for these models are reported in Table 2. This Table shows that Model 2 fits the data considerable better than Model 1 and Model 4 better than Model 3. This indicates that correcting for measurement error is important, regardless of whether we control for observed and unobserved heterogeneity. Also controlling for unobserved heterogeneity - and thus also for selectivity into training - improves the fit of the model, as can be seen by comparing the fit of either Models 1 and 3 or Models 2 and 4.

Table 3: Results from the Mixed Markov and the Mixed Latent Markov model

Transition	UK		Netherlands		Germany		
	Without error correction	With error correction	Without error correction	With error correction	Without error correction	With error correction	
Training	low pay to higher pay	0.238**	0.287**	0.595**	0.826**	0.276	0.298
	low pay to other	-0.157	-0.380	-0.424	-1.526**	0.514*	0.928***
	higher pay to low pay	-0.328***	-0.304***	-0.306	-0.692	-0.862***	-0.185
	higher pay to other	-0.172**	-0.243***	-1.805***	-1.959***	0.505***	0.029
Apprenticeship	low pay to higher pay				0.572*		-0.155
	low pay to other				1.026***		0.891**
	higher pay to low pay				-0.327		-1.758
	higher pay to other				0.711***		0.776***
Education (ref. lower than high school)							
High school	low pay to higher pay	0.062	0.079	0.213	0.261	0.168	0.157
	low pay to other	0.119	-0.038	0.263	0.042	-0.296**	-0.544***
	higher pay to low pay	-0.072	-0.113	-0.704***	-1.164***	-0.522***	-0.654***
	higher pay to other	-0.044	-0.043	-0.356***	-0.342	-0.420***	-0.509***
Higher education	low pay to higher pay	0.466***	0.527***	1.041***	1.241***	0.725***	0.639**
	low pay to other	0.158*	0.245*	1.800***	1.527***	-0.190	-0.672**
	higher pay to low pay	-0.746***	-0.803***	-1.557***	-2.124***	-2.060***	-2.804
	higher pay to other	-0.160***	-0.087	-0.364**	-0.340**	-0.982***	-1.018***
Experience (Age for the Netherlands)	low pay to higher pay	-0.070	0.026	0.011	0.037	0.008	0.088
	low pay to other	-0.713***	-0.894***	0.005	0.031*	-0.048***	-0.054***
	higher pay to low pay	-0.515***	-0.567***	-0.043***	-0.088***	-0.112***	-0.161***
	higher pay to other	-0.703***	-0.871***	-0.004	-0.003	-0.085***	-0.097**

* significant at 10%; ** significant at 5%; *** significant at 1%

The dependent variable is the earnings state. It takes three values: low pay, higher pay and other. Transitions between all states are modelled. However, here we only present the estimates on the transitions from low to higher pay, from low to the 'other' state and from higher to low pay. The control variables are calendar time, marital status, age, age squared, experience squared (only in the UK and in Germany). Other job characteristics are not included as covariates as these are not observed for the individuals being in the 'other' state and therefore their inclusion would make estimation of the model infeasible.

The estimates of training and other covariates from models 3 and 4 are presented in Table 3. One commonly-used specification is to allow covariates to affect the probability of being in a certain state at a time point t . We use another more flexible specification in which covariates have an effect on making a particular type of transition. For example, the model estimates the effect of training on making a transition from low pay to higher pay rather than just estimating the effect of training on being in low pay. Our model estimates the effect of training and other covariates on all six possible transitions between the three states (low pay, higher pay, other). In Table 3, we only present the estimates for the transitions from low pay and from higher pay as these are mostly important with respect to our research questions.

Our results indicate that training has a positive effect on the likelihood for a low-to-higher pay transition. In all three countries the relevant coefficient is positive. However, in Germany it is not significant, which indicates that a full-time training course does not have a clear effect on low-pay mobility. Results for the likelihood of the opposite event, i.e. a transition from higher pay to low pay, are consistent as all coefficients are negative although not always statistically significant. More specifically, in the UK, this effect is clear as the relevant coefficient is statistically significant, while in the Netherlands and in Germany coefficients are negative but not significant. Results are mixed with respect to transitions from employment to non-employment. In Germany, following a full-time course increases the likelihood of moving out of employment regardless of the initial pay level. In the UK and in the Netherlands, following any training course reduces the likelihood for moving out of employment. In the UK, however, only the coefficient for the high-to-other state transition is significant. It seems therefore that training strengthens the employment prospects - or the work-oriented attitude as we do not distinguish between quits and lay offs - of the worker in the UK and in the Netherlands but not in Germany. In Germany, workers following full-time training courses are inclined to search for different career opportunities than employment.

The correction for measurement error and transitory earnings fluctuations strengthens most of the effects. By comparing the estimates of model 3 and model 4, it is seen that most of the estimated coefficients of model 4 are larger than those of model 3. This difference in the coefficients was tested with a Hausman test and was found significant. This finding holds for the estimates of all covariates included in Table 3.

The results with respect to education are in accordance to expectations. In all three countries, having a higher education increases the likelihood for a low-to-higher pay transition and decrease the likelihood for a higher-to-low pay transition compared to having low education. Having high-school education has a similar effect although the coefficients for transitions from low pay to higher pay are not significant for Germany and the Netherlands. Having apprenticeship qualifications has an effect similar to training in Germany; it has no significant effect on transitions from low to higher pay and vice versa and it has a positive effect on transitions to non-employment. For the low-to-higher pay transitions, this results may sound puzzling as apprenticeships are supposed to ensure workers a 'smooth' transition transition to regular employment. However, it has been documented that wage careers of ex-apprentices are flatter than their colleagues that did not follow an apprenticeship (Winkelmann, 1996). No significant positive effect for labour market experience is found concerning transitions from low pay to higher pay. On the contrary, we find a negative effect of experience on the likelihood of moving from higher pay to low pay.⁷

Previous studies suggest that there are complementarities in the effect of training, education, experience and job change on earnings (Lynch, 1992). The presence of such complementarities in the effect of these variables on low-pay mobility was tested with the inclusion of several interaction effects between training and these variables.⁸ None of these interaction effects was significant, which suggests that such complementarities do not exist for low pay mobility.

6 Conclusions and discussion

The aim of this paper was to investigate the effect of training on low-pay mobility in the UK, the Netherlands and in Germany. This topic has hardly ever been the focus of research in labour economics. We employed a modelling approach that enabled us to study the effect of training on low-pay mobility net of measurement error and transitory moves out and into low pay. We distinguished between three states, low pay, higher pay and non-employment ('other') and we used the most common definition of low pay threshold , i.e. the two

⁷For the Netherlands, we report the coefficients for age as a proxy for labour market experience. The Dutch Socio-Economic Panel does not allow to retrieve information on labour market experience.

⁸The estimates for these interaction effects are not presented here but are available on request.

thirds of the median hourly wage. Our approach combined the virtues of a random-effects multinomial logit model and latent class modelling. The use of a random-effects model allowed us to control for self-selection into training.

The results of our analysis indicate that training improves the chances for upward wage mobility for the workers of the lower segment of the wage distribution reduces the risk of higher paid worker to 'fall' into low pay. In this respect, our results complement the findings of studies suggesting that training improves earnings (Duncan & Hoffman, 1979; Booth, 1991; Lynch, 1992; Sloane & Theodossiou, 1996; Blázquez Cuesta & Salverda, 2007). In the UK and in the Netherlands we find a strong positive effect of training on the likelihood of moving from low to higher pay. In Germany, no such significant effect was found. In the UK, we also found that training reduces the likelihood for a higher-to-low pay transition. However, due to data limitations our training variable for Germany is different than in the other two countries. In Germany, it refers to full-time training courses, while in the other two countries in both full-time and part-time courses. Our study also verifies that besides training, general human capital in the form of formal education increases the chances of low-paid workers to improve their wage as well as the chances of higher paid workers to avoid 'falling' to low pay.

Previous studies suggest that the effect of training is not homogeneous to population subgroups. The access to training and the pay-off of training differ according to education, gender, age and immigrant status. Such differences were not found in our analysis. However, this may be due to the fact that our sample was more homogeneous than in other studies. Further research could provide new insight on whether also the pay-off of training with respect to low-pay mobility depends on human capital and demographic characteristics.

Appendix: Description of the variables

Education: This is the highest educational level completed by the individual. It can take three values, lower than high school, high school and higher education.

Training: It takes the value 1 when the individual received formal training during the year prior to the survey and 0 in all other cases.

Labour market experience: Measured in months. This is available only for the UK and

for Germany. It is constructed by combining data from the yearly files and the employment history files of BHPS and GSOEP.

Age: Measured in years.

Job change: It takes the value 1 when the individual changed an employer during the year prior to the survey and 0 in all other cases. It also takes the value 0 when the individual moves from or to non-employment as well as when he remains in non-employment. This variable was not included for the Netherlands.

References

- Acemoglu, D. (2003a). Cross-country inequality trends. *Economic Journal*, 113, 121-149.
- Acemoglu, D. (2003b). Patterns of skill premia. *Review of Economic Studies*, 70(2), 199-230.
- Arulampalam, W., Bryan, M. L., & Booth, A. L. (2004). Training in Europe. *Journal of the European Economic Association*, 2(2-3), 346-360.
- Bassi, F., Hagenaars, J. A., Croon, M. A., & Vermunt, J. K. (2000). Estimating true changes when categorical panel data are affected by uncorrelated and correlated classification errors. *Sociological Methods and Research*, 29(2), 230-268.
- Becker, G. S. (1962). Investment in human capital: A theoretical analysis. *The Journal of Political Economy*, 70(5, part 2), 9-49.
- Blázquez Cuesta, M., & Salverda, W. (2007, January). *Low pay incidence and mobility in the Netherlands exploring the role of personal, job and employer characteristics*. Amsterdam Institute of Advanced Labour Studies, University of Amsterdam, Working Paper Number 06/46.
- Booth, A. L. (1991). Job-related formal training: Who receives it and what is it worth? *Oxford Bulletin of Economics and Statistics*, 53(3), 281-294.
- Bound, J., Brown, C., & Mathiowetz, N. (2001). Measurement error in survey data. In J. J. Heckman & E. Leamer (Eds.), *Handbook of econometrics* (Vol. 5, p. 3705-3843). Amsterdam: Elsevier.
- Cappellari, L., & Jenkins, S. P. (2004a). Modelling low income transitions. *Journal of Applied Econometrics*, 19, 593-610.
- Cappellari, L., & Jenkins, S. P. (2004b). *Modelling low pay transition probabilities, accounting for panel attrition, non-response and initial conditions*. CES Working Paper no. 1232.
- CBS. (1991). *Sociaal-economisch panelonderzoek. inhoud, opzet en organisatie*. Voorburg/Heerlen: Centraal Bureau voor de Statistiek.
- Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B*, 39(1), 1-38.
- Duncan, G. J., & Hoffman, S. (1979). On-the-job training and earnings differences by race and sex. *Review of Economics and Statistics*, 61(4), 594-603.
- Gottschalk, P. (2005). Downward nominal-wage flexibility: Real or measurement error. *Review of Economics and Statistics*, 87(3), 556-568.
- Hagenaars, J. A. (1994). Latent variables in log-linear models of repeated observations. In A. von Eye & C. C. Clogg (Eds.), *Latent variable analysis: Applications for developmental research* (p. 329-352). Thousand Oaks, CA: Sage Publications.
- Kuckulenz, A., & Zwick, T. (2003). *The impact of training on earnings differences between participant groups and training forms*. ZEW - Centre for European Economic Research, Discussion Paper No. 03-57.
- Langeheine, R., & van de Pol, F. (1990). A unifying framework for markov modelling in

- discrete space and discrete time. *Sociological Methods and Research*, 18, 416-441.
- Leuven, E., & Oosterbeek, H. (2002, January). *A new approach to estimate the wage returns to work-related training*. Institute for the Study of Labor (IZA), IZA DP No. 526.
- Lynch, L. M. (1992). Private-sector training and the earnings of young workers. *American Economic Review*, 82(1), 299-312.
- Mincer, J. (1986). Wage changes in job changes. In R. G. Ehrenberg (Ed.), *Research in labor economics* (Vol. 8, p. 171-197). Greenwich: JAI Press.
- Mincer, J. (1988). *Job training, wage growth and labor turnover*. NBER Working Paper.
- OECD. (1996). *Employment outlook 1996*. OECD.
- OECD. (1997). Earnings mobility: Taking a longer-run view. In *Employment outlook 1997* (p. 27-61). OECD.
- OECD. (2003). *Employment outlook 2003*. OECD.
- Parent, D. (1999). Wages and mobility: The impact of employer-provided training. *Journal of Labor Economics*, 17(2), 298-317.
- Pavlopoulos, D. (2007). *Wage mobility patterns in Europe*. Doctoral Dissertation, Tilburg University. Tilburg.
- Pischke, J.-S. (1995). Measurement error and earnings dynamics: Some estimates from the PSID validation study. *Journal of Business and Economic Statistics*, 13(3), 305-314.
- Pischke, J.-S. (2001). Continuous training in Germany. *Journal of Population Economics*, 14(3), 523-548.
- Rendtel, U., Langeheine, R., & Berntsen, R. (1998). The estimation of poverty dynamics using different measurements of household income. *Review of Income and Wealth*, 44(1), 81-98.
- Shorrocks, A. F. (1976). Income mobility and the markov assumption. *Economic Journal*, 86, 566-578.
- Sloane, P. J., & Theodossiou, I. (1996). Earnings mobility, family income and low pay. *Economic Journal*, 106, 657-666.
- Stewart, M. B. (2007). The inter-related dynamics of unemployment and low-wage employment. *Journal of Applied Econometrics*, 22(3), 511-531.
- Stewart, M. B., & Swaffield, J. K. (1999). Low pay dynamics and transition probabilities. *Economica*, 66(127), 23-42.
- Taylor, M. F. e., Brice, J., Buck, N., & Prentice-Lane, E. (2006). *British Household Panel Survey User Manual volume A: Introduction, technical report and appendices*. Colchester: University of Essex.
- van de Pol, F., & Langeheine, R. (1990). Mixed markov latent class models. *Sociological Methodology*, 20, 213-247.
- Vermunt, J. K. (2004). Comment on: An analysis of classification error for the revised current population survey employment questions. *Survey Methodology*, 30, 141-144.
- Vermunt, J. K., Langeheine, R., & Böckenholt, U. (1999). Discrete-time discrete-state latent markov models with time-constant and time-varying covariates. *Journal of Educational and Behavioral Statistics*, 24, 178-205.

- Vermunt, J. K., & Magidson, J. (2008). *LG - syntax user's guide: Manual for Latent GOLD 4.5 syntax module*. Belmont Massachusetts: Statistical Innovations Inc.
- Vermunt, J. K., Tran, B., & Magidson, J. (2008). Latent class models in longitudinal research. In S. Menard (Ed.), *Handbook of longitudinal research: Design, measurement, and analysis* (p. 373-385). Burlington, MA: Elsevier.
- Wagner, G. G., Burkhauser, R. V., & Behringer, F. (1993). The English language public use file of the German Socio-Economic Panel study. *Journal of Human Resources*, 28(2), 429-433.
- Winkelmann, R. (1996). Employment prospects and skill acquisition of apprenticeship-trained workers in Germany. *Industrial and Labour Relations Review*, 49(4), 658-672.