Empirical Research in Vocational Education and Training, Vol. 4(2), 2012, 131–147

Training participation of a firm's aging workforce

Christian Pfeifer [a, b *], Simon Janssen [c], Philip Yang [d], Uschi Backes-Gellner [c]

[a] *Leuphana University Lüneburg, Germany,* [b] *IZA Bonn, Germany,*
[c] *University Zurich, Institute for Strategy and Business Economics, Switzerland,*
[d] *Leibniz University Hannover, Institute of Labor Economics, Germany*

Abstract

We use a long panel data set for four cohorts of male blue-collar workers entering into an internal labor market to analyze the effect of age on the probability of participating in different employer-financed training measures. We find that training participation probabilities are inverted u-shaped with age and that longer training measures are undertaken earlier in life and working career. These findings are consistent with predictions from a human capital model that incorporates amortization period and screening effects.

JEL Codes: J14, J24, M53

Keywords: age, human capital, internal labor markets, training

1. Introduction

Older workers have on average higher employment stability than younger workers but lower reemployment probabilities and often longer unemployment durations in most countries (e.g. Hutchens, 1988; Chan & Stevens, 2001; OECD, 2005; OECD, 2008; EU, 2009). This leads to hardships for unemployed older workers (e.g. loss in consumption standards, psychological burden due to loss of main activity and social networks) and to society, because tax payers have to finance unemployment benefits or early retirement schemes. Therefore, identifying the factors that might lead to employment barriers for older workers is of central importance in times of demographic change. One major economic explanation for employment barriers is that older workers might have a productivity that is lower than their wages. As productivity is largely determined by human capital investments, the relationship between training and aging can help us to understand disadvantages of older workers in the

* Corresponding author: Christian Pfeifer, Institute of Economics, Leuphana University, Scharnhorststrasse 1, 21335 Lüneburg, Germany; phone: +49-4131-6772301; e-mail: pfeifer@leuphana.de
This work was financially supported by the VolkswagenStiftung. We thank Jens Mohrenweiser, Patrick Puhani, participants of the Colloquium in Personnel Economics 2010 in Trier, and two referees of this journal for their comments.

labor market. If firms and workers invest less in human capital at later stages of workers' careers, and if firms cannot adjust individual wages (e.g. due to collective contracts or minimum wage legislations), it would be less profitable for firms to employ older workers. Productivity enhancing training might alter the incentives to expand employment contracts in current firms and to integrate older unemployed workers in new firms (Gruber & Wise, 1998). In times of rapid technological change, training becomes increasingly important because computer-based technologies demand a new range of abilities, which older workers need to acquire in order to avoid the depreciation of skills and competencies (Friedberg, 2003).

In this paper, we analyze the impact of aging on the participation probability in employer-financed training in an internal labor market to shed some light into the black box of training decisions in firms. For this purpose, we develop a simple model for the timing decision when to train a worker, which accounts for screening and amortization period effects and from which our econometric framework is generated. We further use a personnel data set that contains information on more than 10,000 yearly observations for 400 male blue-collar workers of a German company for four entry cohorts. The length of the panel is longer than 24 years. The data contain unique information about four different training measures: short training course, longer training course, longer vocational re-training, and academy of vocational training. Although results from personnel data are not necessarily representative, they have the advantage of overcoming unobserved firm and training course heterogeneity, which might bias results from survey data. To analyze the effect of age on training participation, we apply random effects Logit and multinomial Logit regressions. The main results of our econometric case study are that training participation is inverted u-shaped with age and that longer training is performed earlier in life. These findings are in line with predictions from our theoretical model.

The subsequent paper is structured as follows. The next section summarizes previous findings on age and training participation. In Section 3, we present a model for timing of training participation, from which we generate our main research hypotheses and estimation framework. Section 4 informs about the personnel data set and provides descriptive statistics. The regression analyses are presented in Section 5. The paper concludes with a short summary and a discussion of the results.

2. Previous research

Several empirical studies analyze worker characteristics to explain individual variation in training participation and find that education and age are the most important factors. Frazis et al. (2000) draw from a rich database of employer and employee surveys to analyze the educational effect on training participation in the U.S. They find significant positive effects of educational attainment on the incidence and intensity of formal training. Similar results are found in panel data of young U.S. workers (Veum, 1997) and in European data (Oosterbeek, 1998; Arulampalam et al., 2004; Arulampalam & Booth, 1997).

Theoretical models with respect to age and training emphasize two main arguments: the amortization (payback) period of training investments and the signaling function of training. The former explanation states that older workers are less likely to receive training due to lower total net returns associated with shorter time horizons until retirement (Becker, 1962; Becker, 1993). Therefore, the investments into older workers have to yield significantly larger gains to make their training profitable, especially when facing deferred payment schemes (Lazear, 1979). The signaling function of training refers to information asymmetries. After incurring hiring costs, firms still know little about the potential ability and productivity of the new employees. Training might reduce information asymmetries and is most effective early in workers' careers (Acemoglu & Pischke, 1998). Overall, both arguments (amortization period and signaling) predict a negative correlation between training and age.

Oosterbeek (1998) uses Dutch household panel data to estimate univariate and bivariate Probit models with linear age as explanatory variable. He finds small but significant negative age effects on training. Maximiano and Oosterbeek (2007) evaluate the impact of age on workers' willingness to receive training and employers' willingness to provide training. They also report a small but significant negative linear age effect. Studies with non-linear age specifications provide a more detailed view of the correlation between age and training incidence. Leuven and Oosterbeek (1999) include binary age categories as independent variables in Probit and linear probability models of training. The results are heterogeneous with respect to size, direction, and significance across different countries. Whereas Canada and the Netherlands suggest an inverted u-shaped relation between age and training probability, Switzerland and the U.S. reveal no significant association. O'Connell and Byrne (2009) extend the empirical investigations by controlling for binary age categories within a multinomial Probit regression. Training classification distinguishes between no training, general training, and specific training. The empirical results suggest an inverted u-shaped relationship between age and training, which exhibits weak robustness when including further control variables. An inverted u-shaped age curve for participation in training is also found by Sousa-Poza and Henneberger (2003), who use age, squared and cubed age as explanatory variables for training probability. The results provide small but robust age effects. Riphahn and Trübswetter (2006) also find an inverted u-shaped association between age and training in German microcensus data.

Whereas the downward-sloping part of the inverted u-shaped relationship, which has been found in several studies, can be explained by amortization period and signaling effects, the upward-sloping part cannot. We therefore develop a new simple model for the timing decision of training participation in the next section, from which we derive our econometric framework and research hypotheses.

3. A model for timing of training participation

The focus of our subsequent model about age and training participation is not on the question of whether a firm and a worker invest in human capital, which is the core of most models, but on the question of when the investment is undertaken. For simplicity, we do not distinguish between firm and worker decisions; instead we treat training as a joint decision.[1] As we discuss the effects on total rents, the rent-sharing aspect of human capital investments can be neglected and, hence, wages do not need to be incorporated into our model. Moreover, human capital investment is a binary choice variable, because our paper is about participation in training courses.

The basic mechanisms in our model are a «screening/learning effect» and an «amortization period effect» which have different directions. Younger workers are more engaged in job shopping and firms have to undertake more screening of younger workers, because uncertainty of their quality and willingness to stay in a specific job and firm is higher (e.g. Jovanovic, 1979; Topel & Ward, 1992; Farber & Gibbons, 1996; Lange, 2007). Consequently, firms and workers have less incentives to invest in (firm-specific) human capital at the start of an employment relationship when a worker is young. If the match between worker and firm proves to be of good quality, both parties have incentives to undertake human capital investments. The worker benefits from higher future wages due to higher future productivity and from signaling to the firm higher productivity and work attachment, which increases his promotion probability and long-term income prospects. The firm benefits from higher future productivity of workers. The firm furthermore might need some time to learn about workers' skills to determine training contents and to select participants. We therefore expect that the training participation probability is positively correlated with age for younger workers («screening/learning effect»). Investment incentives, however, decline with age because the amortization period decreases as a worker gets older and approaches retirement («amortization period effect»). While the total effect of age on training should be dominated by the «screening/learning effect» in the first years of workers' careers, the «amortization period effect» should dominate thereafter.

Let us now turn to the simple model. The decision to train a worker depends on total net rents of training R in equation (1).[2] The net rents are the total increase in the value of productivity due to training ΔP (compared to the situation in which a worker receives no training) over all years t until retirement is reached minus the total fixed costs C of the training course. The age at which training takes place is denoted by a and retirement age by r. The length of the amortization period in years is therefore $r–a$. In the subsequent discussion, we consider two cases. The first case assumes no depreciation of human capital acquired in the training course, which leads to a constant productivity increase over time ($\Delta P_t = \Delta P_0$), while the second case acknowledges human capital depreciation.

$$R[\mathrm{a}] = \Delta P - C = \sum_{t=1}^{r-a} \Delta P_t - C \qquad (1)$$

1 In principal, workers and firms face the same effects. Thus, we would obtain the same insights if the investment decisions were analyzed separately. An advantage of analyzing the joint decision is that we can neglect the rent-sharing aspect of human capital investments.

2 Table A.1 in the Appendix contains a list of the model's variables.

We begin by illustrating the «amortization period effect» for the first case. The total net rent is depicted in equation (2) and its first derivate with respect to age in equation (3). We see that one more year of age at training, which implies a reduction of the amortization period by one year, decreases the total net rent linear by the foregone higher value of productivity in that additional year.

$$R = \Delta P - C = \sum_{t=1}^{r-a} \Delta P_0 - C = (r-a)\Delta P_0 - C \tag{2}$$

$$\frac{\partial R}{\partial a} = -\Delta P_0 \tag{3}$$

In the next step, we introduce the «screening/learning effect». The «screening/learning effect» implies that the productivity increase is to some extent uncertain, which is represented through the expected total productivity increase as presented in equation (4). Firms and workers need to learn about the match quality and the willingness to engage in a long-term contract to benefit from returns of human capital investments. The firm further needs to learn about a worker's human capital stock to determine course contents. Both learning necessities can be introduced through a learning parameter γ, which is a non-linear function of worker's age at training. If training takes place later in life, the more has been learned about a worker, but with decreasing marginal returns to learning.[3] Because the learning parameter γ is restricted to values between zero and one, γ can be interpreted as the probability that a worker has the increased productivity after training and $(1-\gamma)$ as the probability that training does not increase productivity ($\Delta P = 0$).

$$E\,[\Delta P] = \gamma[a]\big((r-a)\,\Delta P_0\big) \quad \text{with} \quad \frac{\partial \gamma}{\partial a} > 0, \frac{\partial^2 \gamma}{\partial a^2} < 0 \tag{4}$$

Equation (5) presents the expected total net rent combining the «amortization period effect» and the «screening/learning effect». The first derivate in equation (6) shows that the expected total net rent increases with age as long as $\frac{\partial \gamma}{\partial a}(r-a)\,\Delta P_0 > \gamma[a]\ \Delta P_0$ and decreases with age if $\frac{\partial \gamma}{\partial a}(r-a)\ \Delta P_0 < \gamma[a]\ \Delta P_0$. It can be seen that the left hand side of the first order condition for the maximum expected total net rent in equation (7) decreases with age and that the right hand side increases with age. This is also reflected in the second derivate in equation (8). Overall, the age effect is non-linear with an inverted u-shaped relationship between the expected total net rent of training and the age at which training takes place.

$$E[R] = E[\Delta P] - C = \gamma[a]\big((r-a)\,\Delta P_0\big) - C \tag{5}$$

3 Note that learning in our model depends only on age. This can be reasoned by the fact that workers in our model are homogeneous with respect to entry age and tenure is age minus entry age. A rationale in a model with heterogeneous entry age would be that learning can also take place through previous work careers in other firms (e.g. experience, signals).

$$\frac{\partial E[R]}{\partial a} = \underbrace{\frac{\partial \gamma}{\partial a}(r-a)\Delta P_0}_{>0} - \underbrace{\gamma[a]\Delta P_0}_{>0} = 0 \tag{6}$$

$$\underbrace{\frac{\partial \gamma}{\partial a}(r-a)\Delta P_0}_{a\uparrow \Rightarrow \downarrow} = \underbrace{\gamma[a]\Delta P_0}_{a\uparrow \Rightarrow \uparrow} \tag{7}$$

$$\frac{\partial^2 E[R]}{\partial a^2} = \underbrace{\frac{\partial^2 \gamma}{\partial a^2}(r-a)\Delta P_0}_{<0} - \underbrace{2\frac{\partial \gamma}{\partial a}\Delta P_0}_{>0} < 0 \tag{8}$$

We now consider the second case with human capital depreciation, which leads qualitatively to same results as the first case. Human capital depreciation is introduced through the depreciation factor $(1+\beta)^t > 1$, i.e. the productivity increase due to training is lower in later periods than in earlier periods after training participation $\left(\Delta P_t = \Delta P_0(1+\beta)^{-t}\right)$. The new expected total net rents from training are presented in equation (9). From the first derivate in equation (10) and the second derivate in equation (11), we can again see that the relationship between expected total net rents and age at training is also inverted u-shaped if we account for human capital depreciation.

$$E[R] = \gamma[a]\sum_{t=1}^{r-a}\frac{\Delta P_0}{(1+\beta)^t} - C \tag{9}$$

$$\frac{\partial E[R]}{\partial a} = \underbrace{\frac{\partial \gamma}{\partial a}\sum_{t=1}^{r-a}\frac{\Delta P_0}{(1+\beta)^t}}_{>0} - \underbrace{\gamma[a]\frac{\Delta P_0 \ln(1+\beta)}{(1+\beta)^{(r-a)\beta}}}_{>0} \overset{!}{=} 0 \tag{10}$$

$$\frac{\partial^2 E[R]}{\partial a} = \underbrace{\frac{\partial^2 \gamma}{\partial a^2}\sum_{t=1}^{r-a}\frac{\Delta P_0}{(1+\beta)^t}}_{<0} - \underbrace{2\frac{\partial \gamma}{\partial a}\frac{\Delta P_0 \ln(1+\beta)}{(1+\beta)^{(r-a)\beta}}}_{>0} - \underbrace{\gamma[a]\frac{\Delta P_0 \ln(1+\beta)^2}{(1+\beta)^{(r-a)\beta}}}_{>0} < 0 \tag{11}$$

The probability of participating in training at a given age $(T_a=1)$ is depicted in equation (12) and depends on expected total net rents at that age. To be more precise, training takes place $(T_a=1)$ if the expected total net rents plus an idiosyncratic normally distributed error term ε with zero mean are larger than some threshold value z. Because we have shown that expected total net rents are inverted u-shaped with age, the training probability should also be inverted u-shaped with age.

$$\Pr\left[T_a = 1 | E[R[a]]\right] = \Pr\left[E[R[a]] + \varepsilon > z\right] \tag{12}$$

From equation (12), we can derive our econometric model applying a second order Taylor approximation to the expected total net rents ($E[R]$). Equation (13) states the basic Logit model we have to estimate, in which ρ_1 and ρ_2 denote the coefficients for age and squared age, λ are the coefficients for a vector of control variables X, and Λ is the cumulative density function of the logistic distribution.

$$\Pr\,[T_a = 1|a,X] = \Pr\,[\rho_1 a + \rho_2 a^2 + \lambda X + \varepsilon > z] = \Lambda[\rho_1 a + \rho_2 a^2 + \lambda X] \qquad (13)$$

To summarize, we can formulate our main research hypothesis on the timing of training, which is then tested using longitudinal personnel data and Logit models in the next sections.

> Hypothesis 1: The training participation probability is inverted u-shaped with age ($\rho_1 > 0$ and $\rho_2 < 0$).

Our model also allows us to generate an additional hypothesis. Longer and, consequently more expensive, training courses are likely to increase productivity (ΔP) by more than shorter training courses. Therefore, the «amortization period effect» ($-\Delta P_t$) is larger for longer training courses so that expected net rents are, ceteris paribus, maximized at earlier training age.

> Hypothesis 2: The training participation probability peaks at earlier age for longer training courses.

4. Data set and descriptive statistics

We use personnel data of a large German company from the energy sector located in West Germany. The company is subject to a collective contract and has a works council. Due to data protection reasons, we are neither allowed to name the company nor to provide any further information. The data contain a subsample of 438 blue-collar workers in the company's mining business. All of these workers entered the firm in four subsequent cohorts, from 1976 until 1979, and stayed in the company over the entire observation period up to the year 2002. The sample represents a share of about 25 percent of all employees in the company's operation unit and 3.5 percent of the company's entire workforce.

A disadvantage of our quasi-balanced panel design is that we have no information about workers who left the firm so that we cannot control for a potential selection bias. The data set is nevertheless adequate to study the long-term issues of an aging workforce and of career aspects in the context of human capital investments due to its large panel length. We include only German male blue-collar workers without missing values in the variables we use. This restriction reduces our sample by only 5 percent. The final sample contains 10,544 yearly observations of 415 different workers (1976: number of workers n = 105, panel length in years T = 27; 1977: n = 96, T = 26; 1978: n = 77, T = 25; 1979: n = 137, T = 24).

The data set allows us to use two kinds of training variables. The first variable is binary and takes the value one if a worker participated in training in a given year. Thus, we can apply a random effects Logit model. The second variable indicates what kind of training a worker received so that multinomial Logit models are appropriate. If a worker did not participate in training in a given year the value is zero. For training participation, we have information about four different training measures: (1) short training course («kurze Schulung») (one or two days), (2) longer training course («längere Schulung») (up to several weeks), (3) longer vocational re-training («längere Umschulung») (up to several weeks), and (4) longer academy of vocational training («Berufsakademie») (up to several weeks). Unfortunately, we do not have information about earnings of workers. We know however that workers are paid during the training measures and do not have to cover any direct costs. Table 1 presents summary statistics of the training measures. On average 6.3 percent of the workers in our sample participated in some kind of training in an average year, which results in 664 training cases in our observation period. About two thirds of all cases are short training courses, whereas the other training measures are nearly equally distributed.

Table 1: Descriptive statistics of training variables

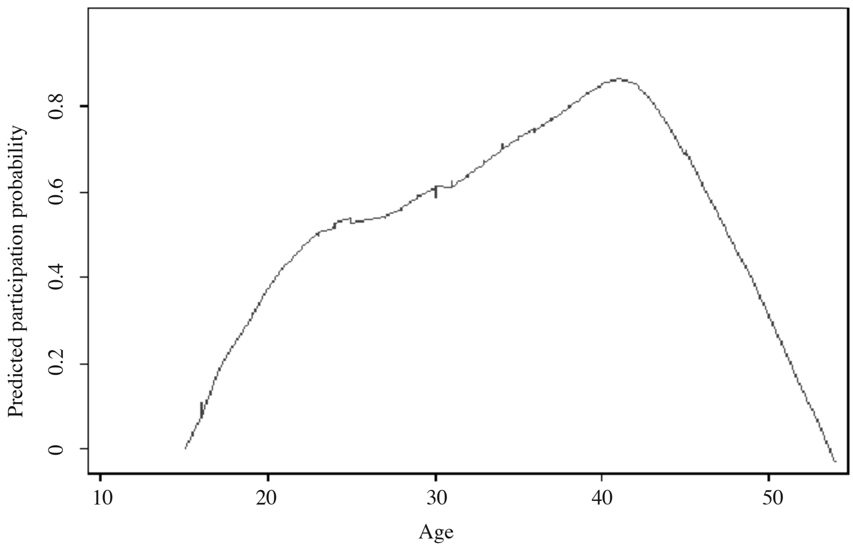
	Mean	Standard deviation	Training cases (total number)
Training (all) (dummy)	0.0630	0.2429	664
Training measures (reference (0) no training):			
(1) Short training	0.0405	0.1971	427
(2) Longer training	0.0073	0.0851	77
(3) Longer re-training	0.0068	0.0824	72
(4) Longer academy	0.0083	0.0910	88

Notes: Number of observations is 10,544 from 415 workers in a balanced panel design.

Our main interest lies in the impact of age on training participation. We specify age in two non-linear ways in the subsequent regression analyses. First, we use dummy variables for the age category. Second, we use age in years and its higher terms. Though most age variance stems from within as we observe workers for at least 24 years, between-age variance also exists as the workers were born between 1952 and 1963. We further consider dummy variables for schooling and apprenticeship degrees to account for skill differences of workers at the time they enter the firm. More information about the explanatory variables is given in Table A.2 in the Appendix.

First descriptive evidence for the impact of age on the overall training participation probability is depicted in Figure 1. The results are based on estimations using robust locally weighted regressions. This is a non parametric approach to smooth scatter plots based on multiple weighted linear regressions for every observation point (Cleveland, 1979). It can be seen that our expected inverted u-shape relationship is indeed confirmed by the data, which stresses the importance of non-linear specification of age when estimating the determinants of training participation.

Figure 1: Age and participation probability in training from locally weighted regressions



5. Regression analyses

At first, we estimate a random effects Logit model for the general participation probability in training. The likelihood ratio test rejects the null hypothesis that the variance of the random effects is zero. As our dependent variable is binary and has a rather low expected probability, linear regressions would yield a high share of outside predictions. We estimate two specifications, which reveal in principal the same results. The first specification includes dummy variables for age categories and the second specification includes polynomials of age in years (until the quartic term). The results of the binary random effects Logit regressions are presented in Table 2. Though we also present the coefficients, our main interest is on marginal effects at the means of all covariates as well as on predicted probabilities.

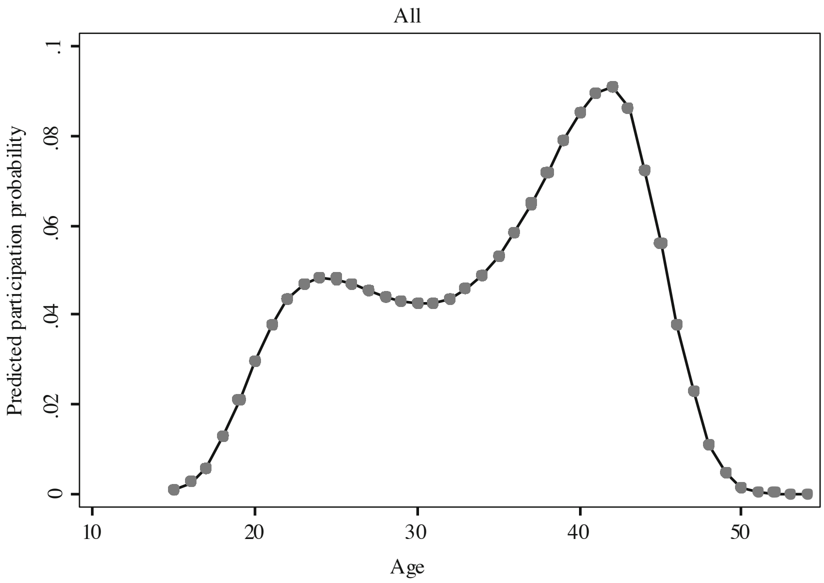
Table 2: Determinants of training participation

	(1-Coeff.)		(2-Coeff.)		(1-Mfx)		(2-Mfx)	
Age categories (reference:15–19)								
Age category 20–24 (dummy)	0.974 (0.343)	**			0.061 (0.028)	**		
Age category 25–29 (dummy)	1.149 (0.338)	***			0.073 (0.029)	***		
Age category 30–34 (dummy)	0.854 (0.342)	**			0.050 (0.025)	**		
Age category 35–39 (dummy)	1.699 (0.333)	***			0.129 (0.038)	***		
Age category 40–44 (dummy)	1.643 (0.337)	***			0.129 (0.041)	***		
Age category 45–54 (dummy)	0.759 (0.408)	*			0.047 (0.033)	*		
Age polynomials								
Age in years			9.153 (1.735)	***				
Age squared / 100			–44.873 (8.33)	***				
Age cubed / 1000			9.563 (1.737)	***				
Age quartic / 10000			–0.746 (0.133)	***				
Mfx: Age polynomials							0.001 (0.0005)	**
S*chooling* *(reference: no school degree)*								
Low school degree («Hauptschule») (dummy)	0.014 (0.153)		0.014 (0.152)		0.001 (0.007)		0.001 (0.007)	
Higher school degree (at least «Realschule») (dummy)	0.488 (0.227)	**	0.482 (0.226)	**	0.027 (0.015)	**	0.024 (0.011)	**
Apprenticeship *(reference: no apprenticeship)*								
Apprenticeship degree in firm	0.438 (0.166)	**	0.412 (0.165)	**	0.022 (0.009)	**	0.020 (0.008)	**
Apprenticeship degree outside firm	0.105 (0.144)		0.077 (0.144)		0.005 (0.007)		0.004 (0.007)	
Observations	10544		10544					
Wald test	111.83	***	99.2	***				
LR test of rho=0	65.82	***	65.67	***				

Note: Random effects Logit (coefficients and marginal effects). Standard errors in parentheses. Standard errors for marginal effects are calculated by using the delta method.*** p<0.01, **p<0.05, *p<0.10.

The first specification in Table 2 indicates that training participation is inverted u-shaped with age and peaks during the middle-age years, between 35 to 45, which is in line with our first hypothesis. We further use the results of our second specification to plot predicted probabilities in Figure 2. The participation probability is to some degree inverted u-shaped with age. As we have considered higher age polynomials, we do not smooth the age effect as we did in the robust locally weighted regressions in Figure 1 in the previous section. That we do not find a smoother u-shaped pattern is also reasoned by training course heterogeneity in the binary pooled training measure we employ. Therefore, a multinomial Logit model for different training measures is likely to identify age effects more accurately.

Figure 2: Age and predicted participation probability in training from random effects Logit



Our Logit estimates in Table 2 further show that workers with higher schooling (at least «Realschule») and workers with an apprenticeship degree earned in the firm have significantly higher training participation probabilities. Differences between low schooling and no school degree as well as between an outside apprenticeship and no apprenticeship degree are not significant. Higher schooling is likely to be associated with higher levels of general human capital, whereas an internal apprenticeship is associated with job specific human capital. Both kinds of human capital might have a self-productivity effect on further skill acquisition, which increases incentives to invest in training for the worker as well as the firm. The firm might also have better knowledge of qualifications and skills of their own former apprentices and can therefore determine training contents and predict outcomes (e.g. training success, productivity effects, willingness to stay in firm) more precisely.

In the next step, we use a multinomial Logit model to estimate participation probabilities in different training measures [(0) no training, (1) short training course, (2) longer training course, (3) longer vocational re-training, (4) longer academy of vocational training], which includes age polynomials and only dummies for higher schooling (at least «Realschule») and internal apprenticeship absolved within the firm. We leave out the other educational categories, because we would otherwise have the problem of perfect predictions in different outcome variables. As has been shown in the previous binary Logit estimates, the reduction of categories is reasonable because we have not found significant differences between workers without a school degree and workers with the lowest school degree («Hauptschule») and between workers without apprenticeship degrees and workers with apprenticeship degrees earned in other firms.

The multinomial Logit model is often criticized because of its reliance on the independence of irrelevant alternatives (IIA) assumption and some authors argue for using the Probit rather than the Logit approach (e.g. Alvarez & Nagler, 1995). Recent studies show however that the multinomial Logit model performs better in practice, even under serious violations of the IIA (Dow & Endersby, 2004; Kropko, 2008). We decided to use the Logit approach and carried out a test in order to check whether the IIA is violated in our special case. In detail we carried out the test proposed by Hausman and McFadden (1984). The null hypothesis that the odds of our different outcome categories are independent of other alternatives could not be rejected for any category. Table 3 informs about the multinomial Logit regression results.

Table 3: Determinants of participation in different training measures

	(1) Short training		(2) Longer training		(3) Longer re-training		(4) Longer academy	
Age polynomials								
Age in years	5.414 (1.951)	***	4.188 (6.479)		13.830 (6.901)	**	30.392 (12.897)	**
Age squared / 100	–29.896 (9.427)	***	–17.029 (33.584)		–69.635 (35.672)	*	–131.686 (60.284)	**
Age cubed / 1000	7.077 (1.972)	***	3.000 (7.590)		15.272 (8.037)	*	24.961 (12.362)	**
Age quartic / 10000	–0.600 (0.151)	***	–0.206 (0.631)		–1.238 (0.666)	*	–1.756 (0.938)	*
Schooling (Reference: no/low degree)								
Higher school degree (at least «Realschule») (dummy)	–0.005 (0.179)		0.662 (0.335)	**	0.108 (0.469)		1.428 (0.239)	***
Apprenticeship (Reference: no/external degree)								
Apprenticeship degree in firm (dummy)	0.227 (0.113)	*	0.430 (0.243)	*	–0.918 (0.360)	**	1.644 (0.234)	***

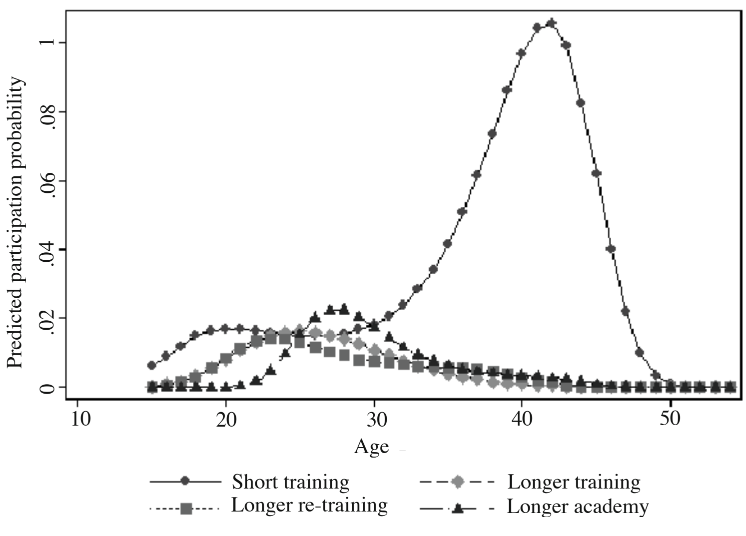
Table 3 continued

Constant	–39.833 (14.713)	**	–41.354 (45.986)		–104.887 (49.088)	**	–263.261 (102.144)	***
Observations	10544							
LR $Chi^2(24)$	517.200 ***							
Pseudo R^2	0.082							

Note: Multinomial Logit (coefficients). Standard errors in parentheses. *** p<0.01, **p<0.05, *p<0.10.

To make interpretation of the results in the multinomial Logit model easier, we plotted the predicted probabilities at different age levels for each training measure in Figure 3. Short training courses are the most frequently used measure, which peak in probability at age 42. Longer training and re-training courses have quite similar profiles with peaks between 23 and 25 years. Longer training in the academy is most likely to occur in the late 20s. For each training measure, we find an inverted u-shaped impact of age, which is more pronounced than in the previous binary Logit estimates for the pooled training probability. The results further indicate that longer and, hence, more costly training measures are more likely to be undertaken earlier in life, which supports our second hypothesis. Older workers seem to receive only short training to update their skills. Career enhancing training (academy) is mainly performed by middle-aged workers and training to close a qualification gap (longer training and re-training) is primarily performed by young workers.

Figure 3: Age and participation probabilities of different training measures from multinomial Logit



We can also see from the multinomial Logit results in Table 3 that workers with higher secondary schooling are more likely to receive longer training and to attend academy training. Workers with an internal apprenticeship are more likely to receive short and longer training as well as academy training but are less likely to get vocational re-training. The latter result is quite plausible as outside workers might have the wrong qualifications for the job and need re-training. Job-specific and firm-specific skills acquired during an internal apprenticeship might have a self-productivity effect on acquiring further specific skills, which might explain the enormous advantage of insiders in attending the academy for vocational training because in this training measure advanced skills are taught.

6. Conclusion

The main results of our econometric case study are that (1) training participation is inverted u-shaped with age, (2) longer training courses are mainly performed earlier in the career, and (3) old workers above the age of 50 years are unlikely to receive any training. Especially the low training probability of older workers, which is likely to be caused by shorter amortization periods, might explain disadvantages of older workers in the labor market (e.g. low re-employment probability). A possible policy intervention are training subsidies targeted at older workers that could counter the effect of decreasing amortization periods and, consequently, should increase the training participation probability, which would hopefully enhance productivity and employability of older workers. Because the amortization period decreases with age, the training subsidies should also increase with age to be effective.

References

Acemoglu, D. & Pischke, J.-S. (1998). Why do firms train? Theory and evidence. *Quarterly Journal of Economics,* 113, 79–119

Alvarez, M. R. & Nagler, J. (1995). Economics, issues and the Perot candidacy: voter choice in the 1992 presidential election. *American Journal of Political Sciences,* 39(3), 714–744

Arulampalam, W. & Booth, A. L. (1997). Who gets over the training hurdle? A study of the training experiences of young men and women in Britain. *Journal of Population Economics,* 10(2), 197–217

Arulampalam, W.; Bryan, M. L. & Booth, A. L. (2004). Training in Europe. *Journal of the European Economic Association,* 2(2–3), 346–360

Becker, G. S. (1962). Investments in human capital: a theoretical analysis. *Journal of Political Economy,* 7(5), 9–49

Becker, G. S. (1993). Human capital: a theoretical and empirical analysis with special reference to education. 3rd ed. Chicago: The University of Chicago Press

Booth, A. L. & Bryan, M. L. (2002). Who pays for general training? New evidence for British men and women. IZA Discussion Paper, 486. Bonn: IZA

Chan, S. & Stevens, A. H. (2001). Job loss and employment patterns of older workers. *Journal of Labor Economics,* 19(2), 484–521

Cleveland, W. S. (1979). Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association,* 74(368), 829–836

Dow, J. K. & Endersby, J. W. (2004). Multinomial probit and multinomial logit: a comparison of models for voting research. *Electoral Studies,* 23(1), 107–122

EU [European Union] (2009). Employment in Europe. Brussels: European Union

Farber, H. S. & Gibbons, R. (1996). Learning and wage dynamics. *Quarterly Journal of Economics,* 111, 1007–1047

Frazis, H.; Gittleman, M. & Joyce, M. (2000). Correlates of training: an analysis using both employer and employee characteristics. *Industrial and Labor Relations Review,* 53(3), 443–462

Friedberg, L. (2003). The impact of technological change on older workers: evidence from data on computer use. *Industrial and Labor Relations Review,* 56(3), 511–529

Gruber, J. & Wise, D. (1998). Social security and retirement: an international comparison. *American Economic Review,* 88(2), 158–163

Hausman, J. & McFadden, D. (1984). Specification tests for the multinomial logit model. *Econometrica,* 52(5), 1219–1240

Hutchens, R. M. (1988). Do job opportunities decline with age? *Industrial and Labor Relations Review,* 42(1), 89–99

Jovanovic, B. (1979). Job matching and the theory of turnover. *Journal of Political Economy,* 87, 972–990

Kropko, J. (2008). Choosing between multinomial logit and multinomial probit models for analysis of unordered choice data. Chapel Hill: University of North Carolina [Dissertation]

Lange, F. (2007). The speed of employer learning. *Journal of Labor Economics,* 25, 1–35

Lazear, E. P. (1979). Why is there mandatory retirement? *Journal of Political Economy,* 87(6), 1261–1284

Leuven, E. & Oosterbeek, H. (1999). The demand and supply of work-related training: evidence from four countries. *Research in Labor Economics,* 18, 303–330

Maximiano, S. & Oosterbeek, H. (2007). On the determinants of workers' and firms' willingness to train. Conference paper, European Association of Labour Economists (EALE), Oslo, 2007

O'Connell, P. J. & Byrne, D. (2009). The determinants and effects of training at work: bringing the workplace back in. ESRI Working Paper, 289. Dublin: ESRI

OECD (2005). Ageing and employment policies. Paris: OECD

OECD (2008). Employment Outlook. Paris: OECD

Oosterbeek, H. (1998). Unravelling supply and demand factors in work-related training. O*xford Economic Papers,* 50, 266–283

Riphahn, R. T. & Trübswetter, P. (2006). Population aging and continued education. IZA Discussion Paper, 2415. Bonn: IZA

Sousa-Poza, A. & Henneberger, F. (2003). The determinants and wage effects of course-related training of elderly workers in Switzerland. Research Institute for Labour Economics and Labour Law Discussion Paper, 94. St. Gallen: University of St. Gallen

Topel, R. H. & Ward, M. P. (1992). Job mobility and the careers of young men. *Quarterly Journal of Economics,* 107, 439–479

Veum, J. R. (1996). Training and job mobility among young workers in the United States. *Journal of Population Economics,* 10, 219–233

Appendix

Table A.1: Variable list for theoretical model

Variable name	*Variable description*
T	Binary training participation
R	Total net rents from training
ΔP	Total increase in productivity due to training
ΔP_t	Increase in productivity due to training in period t after training
C	Costs of training
a	Age when training takes place
r	Retirement age
t	Period after training
γ	Learning parameter
β	Depreciation rate

Table A.2: Descriptive statistics of explanatory variables

Age categories	Mean
Age category 15–19 (dummy)	0.0513
Age category 20–24 (dummy)	0.1582
Age category 25–29 (dummy)	0.1946
Age category 30–34 (dummy)	0.1968
Age category 35–39 (dummy)	0.1968
Age category 40–44 (dummy)	0.1545
Age category 45–54 (dummy)	0.0478
Age polynomials	
Age at end of year (years) [standard deviation: 7.85; min.: 15; max.: 54]	31.9303
Age squared / 100	10.8124
Age cubed / 1000	38.4732
Age quartic / 10000	142.5961
Schooling (reference: no degree)	
Low school degree («Hauptschule») (dummy)	0.7209
Higher school degree (at least «Realschule») (dummy)	0.0799

Table A2 continued

Apprenticeship (reference: no degree)	
Apprenticeship degree outside firm (dummy)	0.4803
Apprenticeship degree in firm (dummy)	0.2514

Note: Number of observations is 10,544 from 415 workers in a balanced panel design.