Training Participation of an Aging Workforce in an Internal Labor Market

Christian Pfeifer

Leuphana University Lueneburg, Institute of Economics; and IZA

Simon Janssen

University Zurich, Institute for Strategy and Business Economics

Philip Yang

Leibniz University Hannover, Institute of Labor Economics

Uschi Backes-Gellner

University Zurich, Institute for Strategy and Business Economics

Leibniz Universität Hannover

Discussion Paper No. 447

ISSN 0949-9962

May 2010

Corresponding author: Christian Pfeifer, Institute of Economics, Leuphana University Lüneburg, Scharnhorststr. 1, 21335 Lüneburg, Germany; phone: +49-4131-6772301; e-mail: pfeifer@leuphana.de.

Acknowledgements: This work was financially supported by the VolkswagenStiftung. We thank Jens Mohrenweiser, Patrick Puhani, and participants of the Colloquium in Personnel Economics 2010 in Trier for their comments.

Training Participation of an Aging Workforce in an Internal Labor Market

Abstract

We use a long panel data set for four entry cohorts into an internal labor market to analyze the effect of age on the probability to participate in different 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, respectively. These findings are consistent with predictions from a human capital model which incorporates amortization period and screening effects. Our results point to a market failure in the context of human capital investments to increase employability of older workers.

JEL Classification: 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 low reemployment probabilities and often long unemployment durations in most countries (e.g., Hutchens, 1988; Chan and Stevens, 2001; OECD, 2005; OECD, 2008; EU, 2009). This puts large burdens to 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. The growing concerns about the future of pension systems in the presence of demographic change have already moved policy away from the pull factors of early retirement (Gruber and Wiese, 1999). Most European countries face however still less than 50 percent labor market participation rates of workers aged 55 or above (Fourage and Schilis, 2009). Moreover, firms will face an increasing labor supply of older workers and consequently will have to employ a higher share of older workers due to the ongoing demographic change. The question what factors might lead to employment barriers for older workers is hence of central importance for today and tomorrow because the answers provide policy recommendations on how to mitigate them.

One major economic explanation for an employment barrier is that older workers have a productivity that is lower than their wages. As productivity is largely determined by human capital investments, the relationship between training and aging is of central importance. If firms as well as workers invest less in human capital at later stages of workers' careers and firms cannot adjust individual wages (e.g., due to collective contracts or minimum wage legislations), it would become 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 and Wiese, 1999). 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 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 model for the timing decision when to train a worker, which accounts for screening and amortization period effects. We further use a personnel data set which comprises information on more than 10,000 yearly observations of 400 male blue-collar workers of a German company for four entry cohorts. The length of the panel is more than 24 years. The data contain information about four different training measures: short training course, longer training course, longer vocational re-training, and longer academy of vocational training. To analyze the effect of age on training participation probability, 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. Our results therefore support the implication from human capital theory that the length of the amortization period strongly affects the decision to invest in human capital. From this might follow a market failure for training of older workers as firms as well as older workers themselves have fewer incentives to undertake human capital investments to maintain productivity. Policy could hence emphasize training subsidies which aim at the employability of older workers.

Our paper is also important in the context of internal labor market theories. First, our results for age can be interpreted from an internal career perspective because we use a very long balanced panel data set of entry cohorts in a large company. Second, we find that workers with an internal apprenticeship have in general higher training participation probabilities than outsiders. The kind of training courses largely differs between both types of workers. It seems that external candidates participate especially in longer training and re-training courses at early career stages in order to compensate the lack of job specific skills. Internal candidates are more likely to receive extensive academy training and to obtain training at later career stages than externals.

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

2. Human Capital Theory and Previous Research

Becker (1962) distinguishes in his seminal work between general and specific training within a setting of perfectly competitive labor markets. General training is assumed to be useful in all firms, whereas specific training enhances only the productivity of a worker in the training firm. The employer will hence only pay for general training if investments serve specific productivity enhancing purposes and if the firm is able to deduct some of the gains from training by paying the worker below his increased marginal product. Hashimoto (1981) supplements these thoughts by including long-term contracting without the possibility of renegotiation. Specific training costs and benefits from human capital investments will be shared by firm and employee if there are transaction costs of evaluation and bargaining over the worker's productivity.

Since Becker (1962), a branch of non-competitive theoretical models of training has emerged (Leuven, 2005).[1] Acemoglu and Pischke (1998) state that imperfect capital markets do not allow workers to pay for general training without smoothing consumption and consequently moving away from their desired level of consumption. Firms might invest in general training if they are able to compress wages due to asymmetric information and transaction costs of job search. Wage compression allows firms to pay the worker below or at his marginal productivity before and during training. If workers marginal productivity gains are larger than wage increases after training, firms deduct returns to training. In contrast to Becker (1962), Acemoglu and Pischke (1998) treat ability and training complementarily. Investments in high-skilled workers are thus most profitable for the employer as compressed wages are the largest for this group.

Empirical literature roots training participation to employer specific, institutional, and employee specific factors. Lynch and Black (1999) use a cross-sectional survey of U.S.

---

[1] This is mainly due to several empirical findings, disproving the results of competitive labor market predictions using panel surveys (Booth and Bryan, 2002), cross sectional samples (Loewenstein and Spetzler, 1998), matched employer-employee data (Loewenstein and Spetzler, 1999), and cross country comparisons (Pischke, 2001; Bassanini et. al., 2005). These studies find that training is often of general kind and employer-paid. Several authors explain these findings through limitations of the assumption of perfectly competitive markets (e.g., Eckaus, 1963; Katz and Ziderman, 1990; Chang and Wang, 1996; Acemoglu and Pischke, 1998).

establishments to investigate the impact of workplace practices and characteristics on the extent of employer provided training. They find that formal training programs are positively correlated with establishment size. Moreover, the share of trained workers is the highest in firm with high-performance work systems, high investments, and much physical capital (Booth and Zoega, 2000; Franzis et. al. 2000; Maximiano and Oosterbeek, 2007; Rinne, 2009).

Drawing from the predictions of Acemoglu and Pischke (1998), a number of studies examines the existence of market imperfections and training. Dustmann and Schönberg (2009) test wage compressing effects of unions and the related union effect on training. Using German linked employer-employee data, they find that unionization increases training. A positive effect of unionization is also observed in British data (Green et. al., 1999; Booth et. al., 2003). Empirical support for a negative training effect of minimum wages, as predicted by Hashimoto (1982), is miscellaneous. No significant effect of minimum wage variations and training probability is found for the U.S. by exploiting cross state variations of panel data for younger workers (Neumark and Washer, 2001), by using cross sectional industry data Grossberg and Sicilian (1999), or by focusing on a low wage sample (Acemoglu and Pischke, 2001). Metcalf (2004) shows that the first ever national minimum wage in the U.K., which covered about 5 percent of British workers, has even boosted the probability and intensity of training.

Studies, which focus on worker characteristics to explain variation in training participation, have found education to be one of the most important determinants. Franzis et. al. (2000) draw from a rich database of employer and employee surveys to analyze the educational effect on training in the U.S.. They find significantly positive effects of educational attainment on 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 and Booth, 1997).

Workers' age and training participation has gained increasing research attention in recent years. 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 states that older workers are less likely to receive training due to lower net returns associated with shorter time horizons until retirement. Therefore, the investments into older workers have to yield significantly larger gains to make their training profitable (Becker, 1993), especially when facing deferred payment schemes (Lazear, 1979). The signaling function of training refers to information asymmetries. After hiring costs have occurred, 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 and 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 report again a small but significant negative linear age effect. Studies with non-linear age specifications provide a more detailed view on 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. Cornell and Bryne (2009) extend the empirical investigations by controlling for binary age categories within a multinomial Probit regression. Training classification separates 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. The 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 robust age effects. Riphan 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 can be explained neither by amortization period nor signaling effects. We therefore develop a new simple model for the timing decision of training participation in the next section.

3. A Model for Timing of Training Participation

The focus of our subsequent model about age and training participation is not on the question if a firm and a worker invest in human capital, which is the core of most models, but on the question when the investment is undertaken. For simplicity, we do not distinguish between firm and worker decisions and treat training as a joint decision.[2] As we discuss the effects on total rents, the rent sharing aspect of human capital investments can be neglected and, hence, wages need not to be incorporated in our model. Moreover, human capital investments are 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. Consequently, firms as well as workers have lower incentives to invest in (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 higher productivity and work attachment to increase his promotion probability. 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 age ("amortization period effect"). While

[2] Workers and firms face in principal the same effects discussed in the following. Thus, we would obtain the same insights if the investment decisions are analyzed separately. An advantage of analyzing the joint decision is however that we can neglect the rent sharing aspect of human capital investments.

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 after the first years of employment.

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).[3] The net rents are the total increase in the value of productivity ΔP due to training (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 with a and retirement age with r. The length of the amortization period in years is therefore r-a.

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

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 (ΔP_t=ΔP_0), while the second case acknowledges human capital depreciation. At first, we illustrate 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.

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

[3] Table A.1 in the Appendix contains a list of the model's variables.

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

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 as well as 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 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.[4] 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-*γ*) as the probability that training does not increase productivity (ΔP=0).

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

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

[4] 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).

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 age at which training takes place.

$$(5) \qquad E\left[R\right] = E\left[\Delta P\right] - C = \gamma[a]\left((r-a)\Delta P_0\right) - C$$

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

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

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

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 ($\Delta P_t = \Delta P_0(1+\beta)^{-t}$). 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) can again be seen that the relationship between expected total net rents and age at training is also inverted u-shaped if we account for human capital depreciation.

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

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

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

The probability to participate 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 normal 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.

(12) $$\Pr\left[T_a=1\middle|E\left[R[a]\right]\right]=\Pr\left[E\left[R[a]\right]+\varepsilon>z\right]$$

From equation (12) we can derive our econometric model applying a second order Taylor approximation to the expected total net rents ($E\left[R\right]$). Equation (13) states the basic Logit model we have to estimate in the following, in which δ_1 and δ_2 denote the coefficients for age and squared age, λ the coefficients for a vector of control variables X, and Λ the cumulative density function of the logistic distribution.

(13) $$\Pr\left[T_a=1\middle|a,X\right]=\Pr\left[\delta_1 a+\delta_2 a^2+\lambda X+\varepsilon>z\right]=\Lambda\left[\delta_1 a+\delta_2 a^2+\lambda X\right]$$

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

Hypothesis 1: The training participation probability is inverted u-shaped with age ($\delta_1 > 0$ and $\delta_2 < 0$).

Our model allows to generate an additional hypothesis. Longer and consequently more expensive training courses are likely to increase productivity (ΔP) to a larger extent 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 give any further information. The data itself contains a subsample of 438 blue-collar workers in the company's mining business, who entered the firm in four subsequent cohorts from 1976 until 1979. All of these workers 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 balanced panel design is that we have no information about workers who left the firm. 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 used variables. 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 kind of training variables. The first variable is binary and takes the value one if a worker has 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 ('Schulung') (one or two days), (2) longer training course ('Schulung') (up to several weeks), (3) longer vocational re-training ('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 our training measures. On average 6.3 percent of the workers in our sample participate in some kind of training in an average year, which results into 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.

- insert Table 1 about here

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 since 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 when they enter the firm. The apprenticeship degree information is of special interest because we have three groups of workers. One group without any apprenticeship degrees (27 percent), one group with apprenticeship degrees earned in other firms (48 percent), and one group with apprenticeship degrees from the analyzed firm (25 percent). More information about the explanatory variables are 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 found in the data which stresses the importance of non-linear specification of age when estimating the determinants of training participation.

- insert Figure 1 about here

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 dummies 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 mean of all covariates as well as on predicted probabilities.

- insert Table 2 about here

The first specification in Table 2 indicates that training participation is inverted u-shaped with age and peaks at the middle age between 35 to 45 years. 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 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 used binary pooled training measure. Therefore, a multinomial Logit model for different training measures is likely to identify age effects more accurate.

- insert Figure 2 about here

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 significant 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 inside apprentices and can therefore more precisely determine training contents and predict outcomes (e.g., training success, productivity effects, willingness to stay in firm).

In the next step, we use a multinomial Logit model to estimate a reduced form for 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. The reduced form is necessary because otherwise we would have the problem of perfect predictions in different outcome variables. As has been shown in the previous binary Logit estimates, the reduced form 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 et al., 2000). Recent studies show however that the multinomial Logit model performs better in practice, even under serious violations of the IIA (Dow and Endersby, 2004; Kropko, 2008). We decided for 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.

- insert Table 3 about here

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 (see also Table A.3 in the Appendix). Short training courses are the most frequently used measure, which have their peak probability at an age of 42 years. 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 only to receive short training to update their skills.

- insert Figure 3 about here

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 wrong qualifications for the job and need re-training. Job- and firm-specific skills acquired during an internal apprenticeship might have a self-productivity effect on acquiring further specific skills, which might explain especially the enormous advantage of insiders in attending the academy for vocational training because in this training measure advanced skills are taught.

The timing of training participation is further analyzed by estimating the previous multinomial Logit model with additional interaction terms between the age variables and the dummy variable for an internal apprenticeship, which show significant differences (see Table A.4 in the Appendix). For an easier interpretation we plotted the predicted probabilities in Figure 4 separately for workers with an internal apprenticeship and other workers (no or outside apprenticeship degree) for each training measure. On the one hand, workers with an external or no apprenticeship are more likely to receive short and longer training as well as longer re-training early in life. This finding is consistent with our previous argument that outside workers might have wrong qualifications for the job and need (re-)training at the beginning of an employment relationship. On the other hand, workers with an internal apprenticeship are more likely to receive training later in the career, except for re-training. Workers with an internal apprenticeship are additionally more likely to attend academy training throughout the entire career. Overall, we find all profiles to some extent being inverted u-shaped and that older workers are likely to receive only short training.

- insert Figure 4 about here

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, (3) old workers are quite unlikely to receive any training, and (4) workers without an internal apprenticeship are more likely to participate in longer training and re-training at the beginning of an employment relationship. Especially the low training probability of older workers, which is likely to be enforced by shorter amortization periods, might explain disadvantages of older workers in the labor market (e.g., low re-employment probability). Because incentives to invest in human capital decline with age, workers as well as firms do not make appropriate investments in the employability of older workers. Thus, a market failure in this context is identified as private actors do not make the optimal investments from a welfare perspective. This market failure provides a rationale for policy interventions. Such an intervention could be subsidies for training targeted at older workers. Subsidies targeted at older workers can counter the effect of decreasing amortization periods and increase the training participation probability, which hopefully enhances 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., and Pischke, J.-S. (1998). Why do firms train? Theory and evidence. Quarterly Journal of Economics 113, 79-119.

Acemoglu, D., and Pischke, J.-S. (2001). Minimum wages and on-the-job training. IZA Discussion Paper 384.

Alvarez, M. R., and 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., and 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, 197-217.

Arulampalam, W., Bryan, M. L., and Booth, A. L. (2004). Training in Europe. Journal of the European Economic Association 2, 346-360.

Bassanini, A., Booth, A., Brunello, G., de Paola, M., and E. Leuven (2005). Workplace training in Europe. IZA Discussion Paper 1640.

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., and Zoega G. (2000). Why do firms invest in general training? “Good” firms and “bad” firms as a source of monopsony power. ILR Working Paper 058.

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

Booth, A. L., Francesconi, M., and Zoega, G. (2003). Unions, work-related training, and wages: evidence for British men. Industrial and Labor Relations Review 57(1), 68-91.

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

Chang, C., and Wang, Y. (1996). Human capital investment under asymmetric information: the Pigovian conjecture revisited. Journal of Labor Economics 14(3), 505-519.

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

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

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

Dustmann, C., and Schönberg, U. (2009). Training and union wages. Review of Economics and Statistics 91(2), 363-376.

Eckaus, R.S. (1963). Investment in human capital: a comment. Journal of Political Economy 71(5), 501-504.

EU (2009). Employment in Europe. Brussels, European Union.

Fourage, D., and Schilis, T. (2009). The effect of early retirement incentives on the training participation of older workers. LABOUR 23(5), 85-109.

Frazis, H., Gittleman, M., and 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.

Green, F., Machin, S., and Wilkinson, D. (1999). Trade unions and training practices in British workplaces. Industrial and Labor Relations Review 52(2), 179-195.

Grossberg, A. J., and Sicilian, P. (1999). Minimum wages, on-the-job training, and wage growth. Southern Economic Journal 65(3), 539-556.

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

Hashimoto, M. (1981). Firm-specific human capital as a shared investment. American Economic Review 71, 573-581.

Hashimoto, M. (1982). Minimum wage effects on training on the job. American Economic Review 72(5), 1070-1087.

Hausman, J., and 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.

Katz, E., and Ziderman, A. (1990). Investment in general training: the role of information and labour mobility. Economic Journal 100(403), 1147-1158.

Kropko, J. (2008). Choosing between multinomial logit and multinomial probit models for analysis of unordered choice data. Working Paper.

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

Leuven, E. (2005). The economics of private sector training: a survey of the literature. Journal of Economic Surveys 9(1), 91-111.

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

Loewenstein, M. A., and Spetzler, J. R. (1998). Dividing the costs and returns to general training. Journal of Labor Economics 16(1), 142-171.

Loewenstein, M. A., and Spetzler, J. R. (1999). General and specific training: evidence and implications. Journal of Human Resources 34(4), 710-733.

Lynch, L. M., and Black, S. E. (1998). Beyond the incidence of employer-provided training. Industrial and Labor Relations Review 52(1), 64-81.

Maximiano, S., and Oosterbeek, H. (2007). On the determinants of workers' and firms' willingness to train. Working Paper Presented at EALE conference 2007.

Metcalf, D. (2004). The impact of the national minimum wage on the pay distribution, employment and training. Economic Journal 114, c84-c86.

Neumark, D., and Washer, W. (2001). Minimum wages and training revised. Journal of Labor Economics 19(3), 563-595.

OECD (2008). Employment Outlook. Paris, OECD.

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

Oosterbeek, H. (1998). Unravelling supply and demand factors in work-related training. Oxford Economic Paper 50, 266-283.

Pischke, J.-S. (2001). Continuous training in Germany. Journal of Population Economics 14(3), 523-548.

Rinne, U. (2009). You live and learn: private-sector training in Germany. Paper presented at ESPE Conference 2009.

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

Sousa-Poza, A., and 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.

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

	Mean	Std. dev.	Min.	Max.
Age categories				
Age category 15-19 (dummy)	0.0513	0.2206	0	1
Age category 20-24 (dummy)	0.1582	0.3649	0	1
Age category 25-29 (dummy)	0.1946	0.3959	0	1
Age category 30-34 (dummy)	0.1968	0.3976	0	1
Age category 35-39 (dummy)	0.1968	0.3976	0	1
Age category 40-44 (dummy)	0.1545	0.3614	0	1
Age category 45-54 (dummy)	0.0478	0.2134	0	1
Age polynomials				
Age at end of year (years)	31.9303	7.8547	15.0000	54.0000
Age squared / 100	10.8124	5.0687	2.2500	29.1600
Age cubed / 1000	38.4732	25.9047	3.3750	157.4640
Age quartic / 10000	142.5961	123.3105	5.0625	850.3056
Schooling (reference: no degree)				
Low school degree ('Hauptschule') (dummy)	0.7209	0.4486	0	1
Higher school degree (at least 'Realschule') (dummy)	0.0799	0.2711	0	1
Apprenticeship (reference: no degree)				
Apprenticeship degree outside firm (dummy)	0.4803	0.4996	0	1
Apprenticeship degree in firm (dummy)	0.2514	0.4339	0	1

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

Table A.3: Predicted probabilities of different training measures at each age level

Age in years	(1) Short training	(2) Longer training	(3) Longer re-training	(4) Longer academy
15	0.0064	0.0004	0.0001	0.0000
16	0.0091	0.0009	0.0006	0.0000
17	0.0116	0.0017	0.0017	0.0000
18	0.0149	0.0036	0.0031	0.0000
19	0.0164	0.0056	0.0056	0.0000
20	0.0170	0.0081	0.0087	0.0001
21	0.0170	0.0109	0.0115	0.0005
22	0.0166	0.0131	0.0136	0.0018
23	0.0160	0.0149	**0.0143**	0.0049
24	0.0155	0.0159	0.0140	0.0098
25	0.0150	**0.0163**	0.0130	0.0154
26	0.0149	0.0161	0.0117	0.0201
27	0.0150	0.0153	0.0104	0.0223
28	0.0156	0.0141	0.0092	**0.0224**
29	0.0167	0.0126	0.0083	0.0204
30	0.0183	0.0110	0.0075	0.0174
31	0.0207	0.0094	0.0069	0.0143
32	0.0239	0.0078	0.0065	0.0117
33	0.0283	0.0064	0.0062	0.0095
34	0.0340	0.0052	0.0060	0.0077
35	0.0414	0.0041	0.0058	0.0064
36	0.0505	0.0032	0.0056	0.0054
37	0.0613	0.0024	0.0053	0.0047
38	0.0734	0.0018	0.0047	0.0042
39	0.0857	0.0014	0.0041	0.0038
40	0.0966	0.0010	0.0032	0.0034
41	0.1041	0.0007	0.0023	0.0032
42	**0.1054**	0.0005	0.0014	0.0030
43	0.0987	0.0004	0.0008	0.0025
44	0.0822	0.0003	0.0004	0.0016
45	0.0616	0.0002	0.0001	0.0010
46	0.0399	0.0001	0.0000	0.0005
47	0.0220	0.0001	0.0000	0.0002
48	0.0099	0.0000	0.0000	0.0001
49	0.0036	0.0000	0.0000	0.0000
50	0.0010	0.0000	0.0000	0.0000
51	0.0002	0.0000	0.0000	0.0000
52	0.0000	0.0000	0.0000	0.0000
53	0.0000	0.0000	0.0000	0.0000
54	0.0000	0.0000	0.0000	0.0000
Total (mean)	0.0405	0.0073	0.0068	0.0083

Notes: Predicted probabilities obtained from multinomial Logit model in Table 3. Maximum is marked in bold.

Table A.4: Determinants of participation in different training measures (internal apprenticeship)

	(1) Short training		(2) Longer training		(3) Longer re-training		(4) Longer academy	
Age polynomials:								
Age in years	5.414	**	115.651	**	12.403		10.830	
	(1.951)		(50.303)		(7.710)		(29.328)	
Age squared / 100	-0.299	***	-6.855	**	-0.622		-0.440	
	(0.094)		(2.987)		(0.397)		(1.505)	
Age cubed / 1000	0.007	***	0.179	**	0.014		0.008	
	(0.002)		(0.078)		(0.009)		(0.034)	
Age quartic / 10000	0.000	***	-0.002	**	0.000		0.000	
	(0.000)		(0.001)		(0.000)		(0.000)	
Schooling (Ref.: No/low degree):								
Higher school degree (at least 'Realschule') (dummy)	-0.005		0.641	*	0.098		1.421	***
	(0.179)		(0.337)		(0.470)		(0.240)	
Apprenticeship (Ref.: No/external degree):								
Apprenticeship degree in firm (dummy)	0.227	**	368.985		-4784.415	*	-339.300	
	(0.113)		(428.765)		(2600.853)		(284.814)	
Interaction effects:								
Internal apprenticeship * age	-39.833	**	-73.523		636.273	*	41.093	
	(14.713)		(62.282)		(344.097)		(37.715)	
Internal apprenticeship * age squared / 100	1.789	**	5.002		-31.529	*	-1.832	
	(0.661)		(3.447)		(16.975)		(1.862)	
Internal apprenticeship * age cubed / 1000	-0.037	**	-0.143	*	0.690	*	0.036	
	(0.014)		(0.086)		(0.370)		(0.041)	
Internal apprenticeship * age quartic / 10000	0.000	**	0.001	*	-0.006	*	0.000	
	(0.000)		(0.001)		(0.003)		(0.000)	
Constant	-76.259	**	-728.563	*	-94.192	*	-102.510	
	(25.167)		(314.861)		(54.978)		(211.940)	
Observations	9888							
LR Chi²(24)	631.860***							
Pseudo R²	0.103							

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

Figures and Tables Included in Text

Table 1: Descriptive statistics of training variables

	Mean	Std. dev.	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.

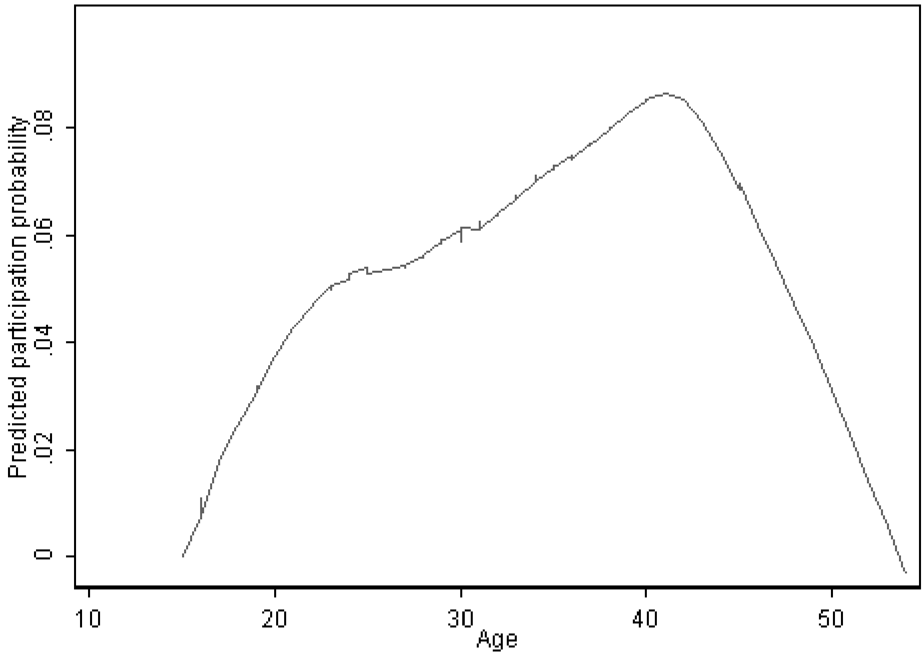


Figure 1: Age and participation probability in training from locally weighted regressions

Table 2: Determinants of training participation

	(1 - Coeff.)		(2 - Coeff.)		(1 - Mfx)		(2 -Mfx)	
Age categories (Ref.:15-19):								
Age category 20-24 (dummy)	0.974	**			0.061	**		
	(0.343)				(0.028)			
Age category 25-29 (dummy)	1.149	***			0.073	***		
	(0.338)				(0.029)			
Age category 30-34 (dummy)	0.854	**			0.050	**		
	(0.342)				(0.025)			
Age category 35-39 (dummy)	1.699	***			0.129	***		
	(0.333)				(0.038)			
Age category 40-44 (dummy)	1.643	***			0.129	***		
	(0.337)				(0.041)			
Age category 45-54 (dummy)	0.759	*			0.047	*		
	(0.408)				(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)	
Schooling (Ref.: No school degree):								
Low school degree ('Hauptschule') (dummy)	0.014		0.014		0.001		0.001	
	(0.153)		(0.152)		(0.007)		(0.007)	
Higher school degree (at least 'Realschule') (dummy)	0.488	**	0.482	**	0.027	**	0.024	**
	(0.227)		(0.226)		(0.015)		(0.011)	

Apprenticeship (Ref.: No apprenticeship)								
Apprenticeship degree in firm	0.438	**	0.412	**	0.022	**	0.020	**
	(0.166)		(0.165)		(0.009)		(0.008)	
Apprenticeship degree outside firm	0.105		0.077		0.005		0.004	
	(0.144)		(0.144)		(0.007)		(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. *** p<0.01, **p<0.05, *p<0.10.

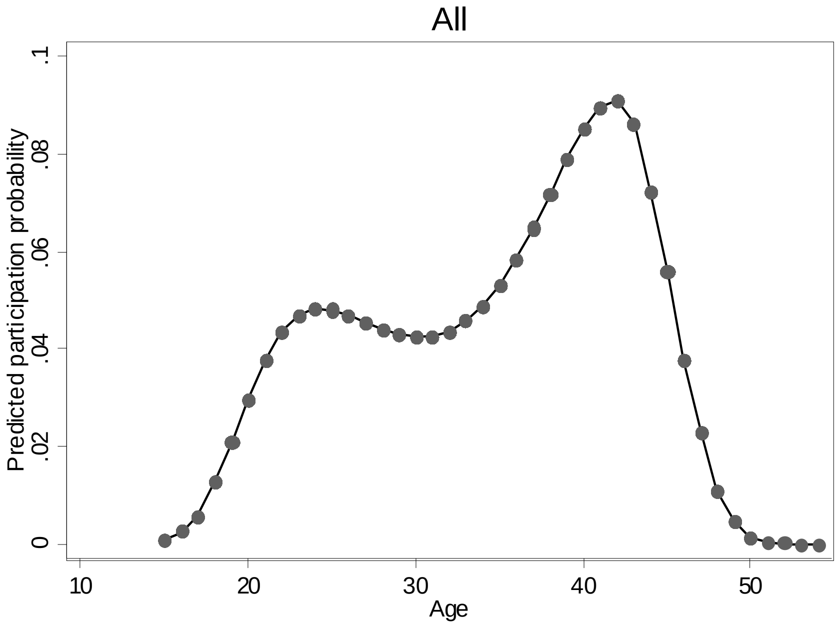


Figure 2: Age and predicted participation probability in training from random effects Logit

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	**	4.188		13.830	**	30.389	**
	(1.951)		(6.479)		(6.901)		(12.897)	
Age squared / 100	-0.299	**	-0.170		-0.696	*	-1.317	**
	(0.094)		(0.336)		(0.357)		(0.603)	
Age cubed / 1000	0.007		0.003		0.015	*	0.025	**
	(0.002)		(0.008)		(0.008)		(0.012)	
Age quartic / 10000	0.000		0.000		0.000	*	0.000	*
	(0.000)		(0.000)		(0.000)		(0.000)	
Schooling (Ref.: No/low degree):								
Higher school degree (at least 'Realschule') (dummy)	-0.005		0.662	**	0.108		1.428	***
	(0.179)		(0.335)		(0.469)		(0.239)	
Apprenticeship (Ref.: No/external degree):								
Apprenticeship degree in firm (dummy)	0.227	*	0.430	*	-0.918	**	1.644	***
	(0.113)		(0.243)		(0.360)		(0.234)	
Constant	-39.833	**	-41.354		-104.887	**	-263.261	***
	(14.713)		(45.986)		(49.088)		(102.144)	
Observations	10544							
LR Chi²(24)	517.200***							
Pseudo R²	0.082							

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

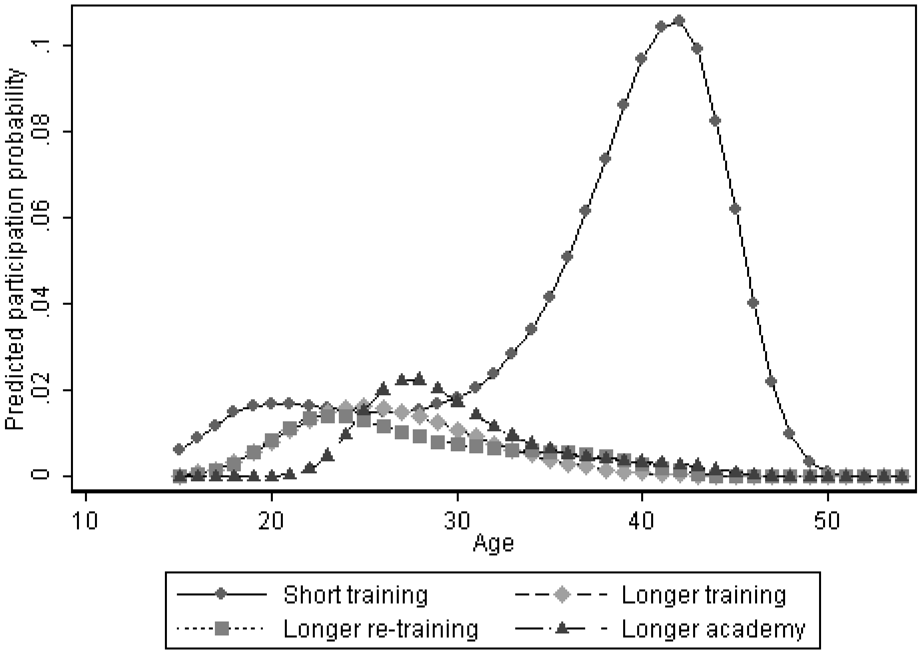
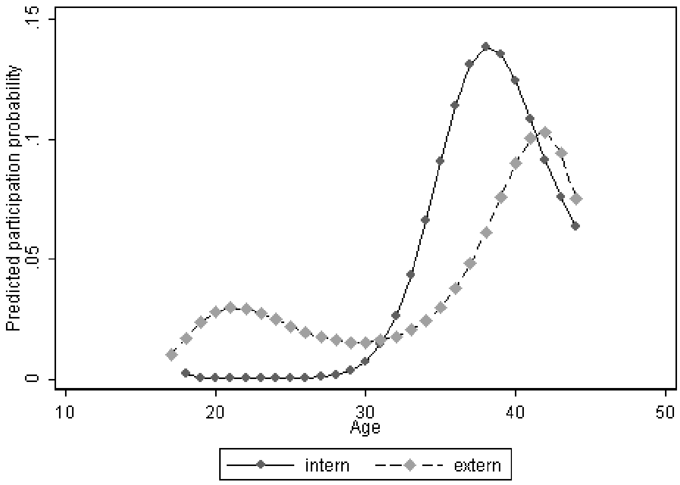
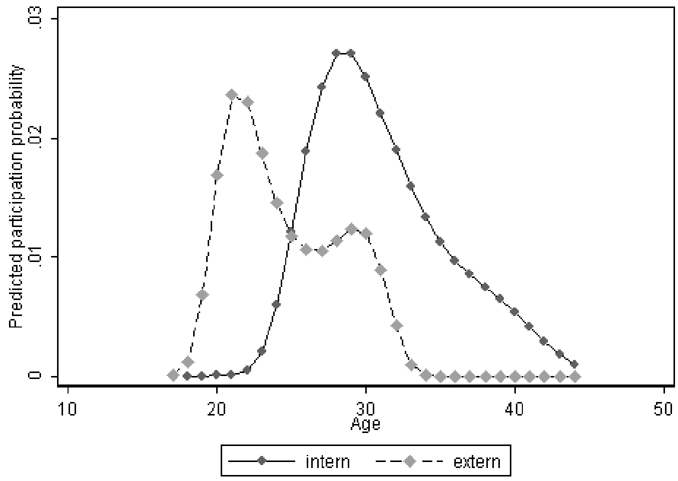


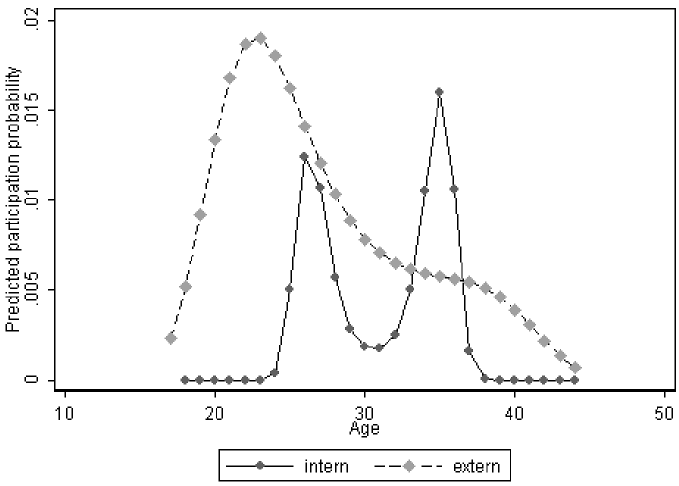
Figure 3: Age and participation probabilities of different training measures from multinomial Logit



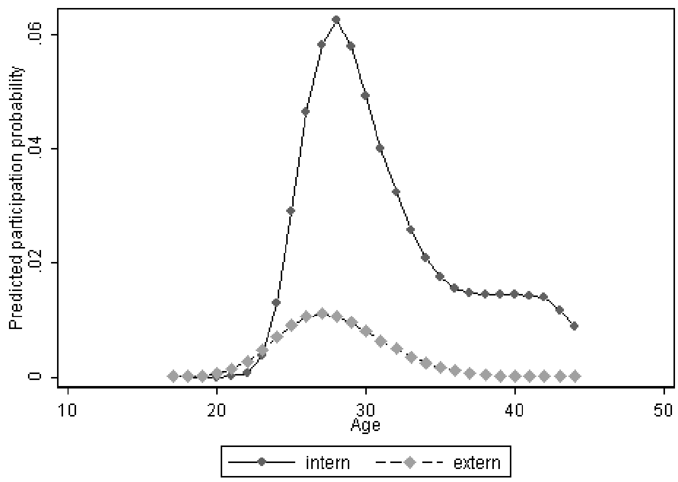
Short training



Longer training



Longer re-training



Longer academy

Figure 4: Age and participation probabilities of different training measures from multinomial Logit (internal apprenticeship)