Discussion Paper Series – CRC TR 224

Discussion Paper No. 382
Project C 01

Training in Late Careers
A Structural Approach

Teresa Backhaus [1]

December 2022

[1] Institute of Applied Microeconomics, University of Bonn, Email: tbackhaus@uni-bonn.de

Support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through CRC TR 224 is gratefully acknowledged.

Collaborative Research Center Transregio 224 - www.crctr224.de
Rheinische Friedrich-Wilhelms-Universität Bonn - Universität Mannheim

Training in Late Careers
A Structural Approach

Teresa Backhaus*

December 19, 2022

Abstract

This study investigates the role of on-the-job training in the employment outcomes of less educated men in their late careers. Using survey data from the German National Education Panel Study adult cohort, I estimate a structural dynamic discrete-choice model reflecting the trade-offs of the employees' training participation decision. The data set enables me to distinguish whether non-participation is due to lack of availability of training or due to individual cost-benefit considerations. As a consequence, I can investigate whether future policy interventions should target the provision of training or the individual participation incentives. I find that on-the-job training has a positive impact on the employees' employment prospects. Counterfactual simulations show that a reduction of the individual training costs would increase training participation and positively affect the employment rate near retirement. In contrast, an increase in the general availability of training would not be effective.

JEL codes: E24, J14, J22, J24, M53
Keywords: on-the-job training; late career; less educated; structural model

*I want to thank Philipp Albert, Richard Blundell, Uta Bolt, Yves Breitmoser, Monica Costa-Dias, Eric French, David Goll, Peter Haan, Anna Hammerschmid, Steffen Huck, Daniel Kemptner, Johannes Leutgeb, Fabien Postel-Vinay, and Roel Van Veldhuizen for their valuable comments and inputs. This paper uses data from the National Educational Panel Study (NEPS): Starting Cohort Adults, doi:10.5157/NEPS:SC6:11.0.0. From 2008 to 2019, NEPS data was collected as part of the Framework Program for the Promotion of Empirical Educational Research funded by the German Federal Ministry of Education and Research (BMBF). As of 2014, NEPS is carried out by the Leibniz Institute for Educational Trajectories (LIfBi) at the University of Bamberg in cooperation with a nationwide network. Financial support by the Deutsche Forschungsgemeinschaft (CRC TRR 190 and CRC TR 224) is gratefully acknowledged. Author address: Institute of Applied Microeconomics, University of Bonn, Adenaueralle 24-42, 53113 Bonn, Germany. E-mail: tbackhaus@uni-bonn.de. Phone: +49 228 73-9219.

1 Introduction

In view of aging populations, pay-as-you-go public pension systems face severe challenges: Decreasing numbers in younger generations and increasing life expectancy are threatening the system. In response many OECD countries have reformed their retirement policies, for example, they have raised the normal retirement age to encourage longer working lives (e.g. Blundell et al., 2016b; Staubli and Zweimüller, 2013).

Yet, older employees' labor market chances are worse than those of younger employees (Daniel and Heywood, 2007; Göbel and Zwick, 2012), which is one of the main reasons why increases in the statutory retirement age do not translate into a one-to-one extension of the working life, even if financial incentives for staying employed are high. Lack of employment is especially prevalent and problematic among the less-educated (see, for example, Blundell et al., 2016b; Börsch-Supan and Ferrari, 2017).

Therefore, it is crucial to understand the labor market outcomes of less-educated employees in their late careers and to investigate which instruments can foster employment. One instrument often discussed in this context is training (see, for example, Sanders et al., 2011). It is meant to keep employees' skills up to date so they meet the demands of today's tasks on the labor market, such that their productivity improves. Consequently, it increases the firms' incentives to keep them employed and preserve or even raise their wages (e.g. Picchio and Van Ours, 2013; Zwick, 2011; Bellmann et al., 2013). Many countries use training in their active labor market policy portfolio (see, for example, Kluve, 2010). A study by Gohl et al. (2020) supports the relevance of this instrument in the context of aging populations as it finds positive effects of an increase in statutory retirement age on the prevalence of training.

However, to date, it has not yet been resolved whether policy-makers should increase training supply or incentivize individual training take-up to foster overall training participation. Moreover, it is not clear how this increase in training participation would affect the employment outcomes of less-educated employees in their late careers.

In this paper, I use a structural dynamic discrete-choice model to answer this question: I investigate the role of on-the-job training for the employment outcomes of less-educated employees in their late careers and evaluate potential channels for policy interventions. First, I explicitly model the cost-benefit trade-offs that these individuals face when deciding on whether to invest in their human capital, by participating in training, or not. Second, in contrast to other studies, I use a data set that enables me to identify different channels of frictions related to training participation: The data of the German National Education Panel Study (NEPS) provides information about the availability of firm-sponsored training. Thus, I can separate non-participation due to the lack of availability of training from non-participation due to the individual cost-benefit trade-offs. This allows me to quantify first, the benefits of training for the employee, and, second, to show in counterfactual simulations how different types of policy interventions affect the employee's employment prospects. Should policy interventions target the training supply at the firm side or the training take-up incentives of individuals?

My model focuses on less-educated male employees aged above 50. Each period they decide whether they want to continue working and whether they want to participate in training, conditional on their available choice options: That is, the employer may not offer training, such that the employee cannot choose training, or the employee may lose his job in which case he automatically becomes unemployed. Employees decide on whether to invest in training by trading off the benefits of training with respect to future employment prospects, wages, and retirement benefits against instantaneous utility costs of training. I estimate the model parameters reflecting the utility costs and the benefits of training and employment with the maximum-likelihood method.

The estimated parameters determining the training choices and labor market outcomes of the men in my sample are in line with the literature: I find very small and insignificant effects of training on wages, and a positive effect of training on employment prospects. Training further decreases the estimated disutility of working. On the other hand, training participation creates significant utility costs.

In my counterfactual simulations I show that a policy intervention that fosters the availability of training in firms would not be effective to increase the employment rates of less-educated employees near retirement: If training was available for everyone, the training participation would increase by 30% but the employment rate near retirement would only increase by 0.5%.

In contrast, a policy intervention that seeks to reduce the individual utility costs of training has the potential to positively impact employment near retirement. Under a full compensation of the individual utility costs of training, the training participation would quintuple to 50% and the employment rate in the year before retirement could be increased by almost 5%. However, training in its current form is not able to fully counterbalance the decreasing employment rates of less-educated employees near retirement.

This paper is related to the literature in various ways: There is a number of reduced-form studies discussing the effects of training on productivity, wages, and employment, indicating the importance of this topic. See Leuven (2005) for a review of the theoretical literature. The overall evidence of empirical studies on further training is mixed.[1] Papers investigating effects of training on *wages* mostly find insignificant or very small effects: Pischke (2001) investigates the link between training and subsequent wage growth using German SOEP data and only finds insignificant positive estimates. Conti (2005) does not find positive wage effects using Italian panel data. Jürges and Schneider (2004) use GSOEP data to investigate effects of on-the-job training on wages with different approaches and find insignificant effects. Bassanini (2006) only finds positive effects on wages for high-educated and young employees using European Community Householdpanel (ECHP) data. Fouarge et al. (2013) argue that wage returns to on-the-job training are positive and do not significantly differ by education level using Dutch data but admit selection problems. Finally, Görlitz (2011) finds insignificant short-term impact of on-the-job training on wages in Germany, and Ehlert (2017) only finds significantly positive short-run effects on wages for employer financed mandatory training using NEPS data. Other papers have looked at firm level productivity (Göbel and Zwick, 2013; Zwick, 2002) and found mixed results.[2]

Papers investigating the impact on *employment* find mostly positive effects: E.g. Cairo and Cajner (2018) conclude that on-the-job training is the reason for different volatility levels in employment (via job separations) between high- and low-educated employees in the US. Picchio and Van Ours (2013) find that firm provided training significantly increases future employment prospects, even for older workers. Likewise, Bassanini (2006) finds positive results on employment security. Further, a study by Dauth (2020) finds positive effects of subsidized training on employment duration of low-skilled workers in Germany.

In contrast to these reduced-form studies a structural set-up allows to explicitly model endogeneities and trade-offs of individual decisions. Further, it allows for counterfactual simulations to find out which policy interventions are effective in increasing employment rates of less-educated employees near retirement: Interventions that tar-

[1] The largest part of the empirical training literature looks at vocational training, training in early careers, or (public) training programs for unemployed.

[2] Göbel and Zwick (2013) find no effects, Zwick (2002) finds positive association of training intensity with productivity.

get the provision of firm sponsored training or interventions that target the individual participation incentives?

Existing structural papers on this matter either do not use training data to identify their effects (Kuruscu, 2006; Fan et al., 2017), or focus on middle-aged women (Blundell et al., 2019), who arguably face different trade-offs than male employees in their late careers. None of these studies uses data that allows for adapting individual choice options in the model to the availability training. Thus, this paper closes a gap in the literature by providing a structural model that is explicitly designed to understand the training decisions of less-educated male employees in their late careers and by using a data-set that provides information to distinguish whether non-participation in training is due to a lack of the provision of firm-sponsored training or due to individual cost-benefit trade offs.

The remainder of this paper is structured as follows: Section 2 introduces the data set and provides first descriptive evidence. Section 3 contains all details of the structural model. Section 4 presents estimation results and the model fit. Section 5 shows the results of the counterfactual simulations, and section 6 concludes.

2 Data and descriptive evidence

2.1 Data

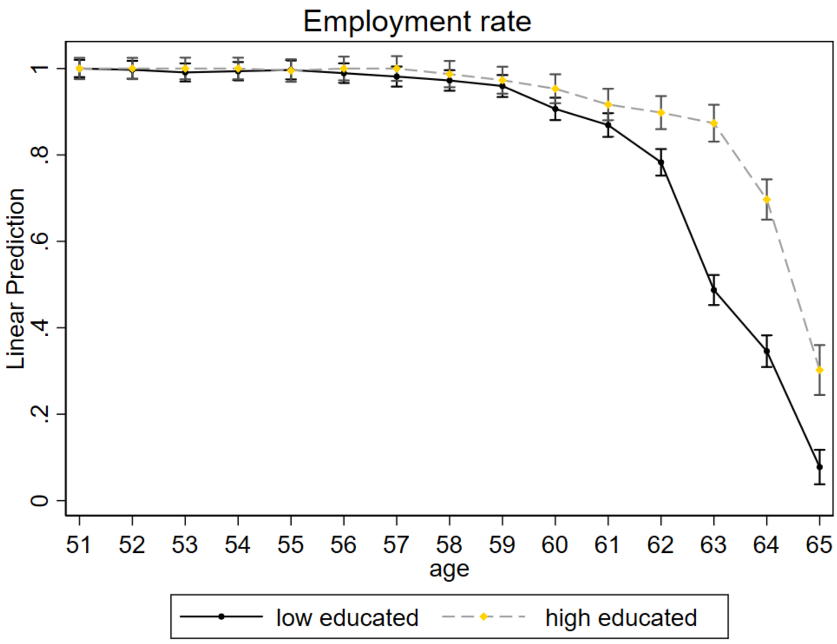
For the analysis, I use adult-cohort data from the National Education Panel Study (NEPS) – see Blossfeld et al. (2011). This is a comprehensive survey data set focusing on adult education and lifelong learning. The earliest observations in the data set were collected in 2007[3], while the NEPS itself started in 2009 and has been repeated every year since. The main advantage of this data set is that it is specifically designed for collecting information about further education and training among adults and therefore contains very detailed information about it. It contains information about training participation, type of training, attitudes with respect to work and training, and, importantly, the availability of training support by the company. The latter is the key feature that allows me to separate individual training costs from the availability of training in the firm in my model. For my analysis, I transformed the data into an annual panel

[3]The initial survey was called ALWA (Arbeiten und Lernen im Wandel (Working and Learning in a Changing World) run by the Institute for Employment Research (IAB))

format.[4]

In my analysis I will focus on male employees in their late careers with up-to-intermediate education, i.e. those who are aged above 50 and who do not have completed high-school but may have completed vocational training. I call them “low” or “less” educated as abbreviation.[5] In line with the findings of the previous literature (see, for example, Cairo and Cajner, 2018) the less-educated employees in my sample have lower employment rates than high-educated employees.[6] Figure 1 shows the employment rate of college educated and less-educated (no high-school diploma) male employees. It shows significantly lower employment rates after age 60 for less educated men.

Figure 1: Employment rate by age and education



Source: NEPS; own calculations based on estimation sample. For male employees only. Low educated denotes people with no high-school diploma, high-educated denotes people with college degree.

[4]The NEPS consists of several data sets with different formats: Some as spell data, some as panel data, which can be merged in several ways depending on data requirements. Details on the data processing are available upon request.

[5]This classification is based on the education classification by Blundell et al. (2019) who use the three groups – up-to-intermediate education, high-school degree, and college degree.

[6]Considering only men aged 50-64 who are in dependent employment when entering the sample, the group with up-to-intermediate education is largest with 51% of the observations, college educated are 34%, and the group of men with high-school degree but no college education is smallest with only 15%.

Furthermore, Figure 1 shows that, while most college-educated men in my sample leave the labor market at age 65, a large proportion of the less educated group already leaves the labor market at age 63. From this age, the German public pension system allows very long-term insured (those who have contributed for more than 45 years) to retire early without deductions.[7] People whose health status does not allow them to continue working may retire before age 63. To avoid confounding from this type of retirement and differing attitudes with respect to work of these people, I drop all those employees who state bad health status before age 63.[8] I further drop all remaining observations indicating retirement before age 63.[9]

Training data The most prevalent form of training among older employees in Germany is non-formal on-the-job training (Ehlert, 2017; Kruppe and Trepesch, 2017), i.e. training conducted while the individual is being employed and receives a regular salary without the awarding of any official certificate. The NEPS provides records of participation in such non-formal training for all employed individuals in every survey wave (for details see Kruppe and Trepesch, 2017). It also contains information about the availability of financing for such training.

In summary, my **sample** includes only men aged 50-63 without a high-school diploma; who are in dependent full-time employment when they enter the sample; who have at least two observations, no missing data for wages and training participation; and who state at least intermediate health status before age 63.

Note that this study investigates the role of on-the-job training, i.e. training of employed people, in future employment prospects. Less-educated men who are still in employment at age 50 are not representative of all less-educated people, who often suffer from unemployment at multiple points of their career. The reintegration of less-educated unemployed into the labor force, e.g. with public sector sponsored training programs, is not the subject of this study.

[7] For long-term insured (at least 35 years) it is allowed to retire *with deductions* from that age. From the year of birth 1953 onwards, the age limit for this deduction-free pension will gradually increase. For all those born in 1964 or later, the age limit is 65 years (Deutsche Rentenversicherung, 2020). Also for people born before 1953 it was possible to retire early after unemployment (with deductions).

[8] Possible answers are "very good", "good", "intermediate", "bad", "very bad". I drop all individuals who stated at least once "bad" or "very bad".

[9] After removing people with bad health status, there is only few (20) observations left who still retire before age 63. For those people early retirement could be due to partial retirement plans or retirement after unemployment.

2.2 Descriptives

2.2.1 Training offers and training participation

One of the requirements for an employee to participate in on-the-job training is the employer's support for such activities. A key feature of the NEPS is that it provides information about the availability of training support in the employee's firm as stated by the survey participant: Does the firm provide company agreements, further education planning, financing for training, or a responsible person? Throughout my analysis I use the availability of firm-sided funding for training to proxy whether the employee has the possibility to participate in on-the-job training or not – denoted as "**training offers**" (TO) in the following.

I use this indicator for three reasons: First, it shows to be a necessary condition for training participation.[10] Second, stated training offers based on this indicator are unlikely to be determined by the employees' demand for training: Many who state that funding is available still do not train. Third, there is no difference between employees with and employees without training offers in terms of beliefs about the usefulness of training and in terms of self stated laziness.[11] Further, there is no indication that people with better employment prospects select into firms which offer training: The employment rate is not higher for individuals who have (or used to have) training offers, see Figure 20. If anything it is lower for people with training offers near retirement.

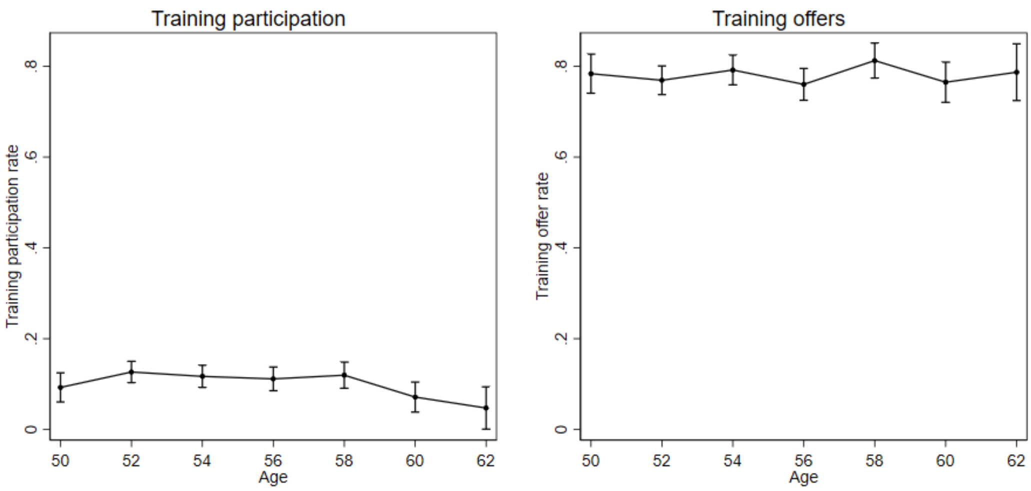
To indicate **training participation** I will use a binary variable following Blundell et al. (2019). In my case an employee is denoted as having participated in training if he did at least 20 hours of training in the past 12 months. Figure 2 shows the training participation and training-offer rate by age.

The training-offer rate is, at close to 80%, much higher than the training-partic-ipation rate, at around 10%. While the training-offer rate remains constant with age, the training-participation rate decreases. The decreasing training-participation rate reflects the lower returns compared to the costs of training for higher ages. Yet, it does not decrease to zero.

[10] If funding is not available 96% don't do any training.

[11] See appendix Figure 19.

Figure 2: Training participation and training offer rate



Notes: Whiskers depict 95% confidence intervals. *Source:* NEPS; own calculations based on estimation sample.

2.2.2 Training and individual characteristics

In this subsection I examine differences in characteristics between training participants and non-participants as well as employees with and without training offers to check whether people with specific employment-related characteristics select into either group.

The NEPS includes questions about people's career ambitions and attitudes in some waves. Hence, I can check whether the responses differ between training participants and non-participants in my sample.[12] Table 1 shows the average response by training participation. The career ambitions of training participants are very similar to those of non-participants. The ambitions for status maintenance, for career advancement, to perform tasks better, and the general importance of the career are slightly higher for training participants. However, the importance of job security, for keeping up with colleagues, and self stated laziness are the same between the two groups. Note that both, training participation and ambitions may evolve with age. Therefore, I provide figures with ambitions and attitudes broken down by age in the appendix (Figure 21 and 23). They do not show distinct patterns in training participants' ambitions.[13] It behaves

[12] Possible answers for ambitions range from 1 "very important" to 5 "very unimportant" and for self stated laziness from 1 "not lazy at all" to 5 "very lazy".

[13] When looking at the ambitions at a single age 55 there are no significant differences for most ambitions.

Table 1: Individual characteristics by training participation

	No training	Training
Ambitions		
Importance of status maintenance	1.784	1.709
	(1.123)	(1.077)
Importance career advancement	3.420	3.247
	(1.054)	(1.109)
Importance perform tasks better	1.944	1.814
	(0.917)	(0.733)
Importance job security	2.054	2.258
	(1.301)	(1.422)
Importance of keeping up	2.097	2.064
with colleges	(1.089)	(1.040)
Attitudes		
Lazy	2.212	2.212
	(1.102)	(1.147)
Importance of career	2.765	2.680
	(1.046)	(1.049)
Wages		
Monthly gross-wage	3568.4	4040.5
	(1473.7)	(1522.0)
Monthly net-wage	2391.1	2657.9
	(991.1)	(950.5)

Mean values, standard deviations in parentheses. Ambitions: 1 = very important, 5= very unimportant. Laziness: 1= not lazy at all, 5= very lazy. For a brake down by age see Figures 21 and 23 . Differences in wages are not significant when controlling for wage-level endowments, as I will do in my model. *Source*: NEPS data, low educated male employees in full time employment only.

similarly with training *offers*: the groups of people with and without training offers are very similar in terms of their ambitions (Table 4 and Figures 22 and 23). Hence, selection into training participation or firms with training offers based on ambitions and attitudes is not a problem in my sample of less-educated male employees in their late careers.

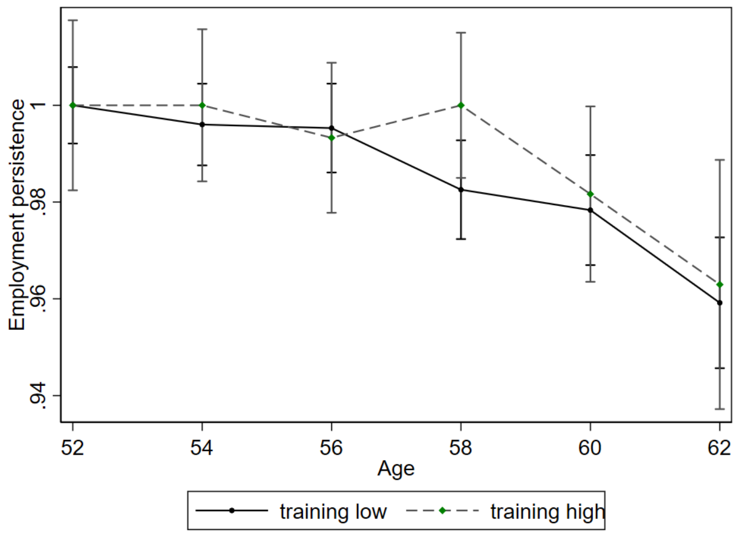
Stated gross- and net-wages of training participants are significantly higher compared to non-participants.[14] Thus, it will be important to control for wage-levels later in the

[14]In the NEPS all participants are asked about their gross- and net-wages last month. If they do not

analysis.

2.2.3 Training and employment

Figure 3: Persistence in employment by human capital gained from training



Source: NEPS; own calculations bases on estimation sample. Whiskers represent 95% confidence intervals. High training is defined as having at least 0.9 units of discounted human capital of training, i.e having participated at least once in 20 hours of training within the last two years or having participated more than once more than two years ago would suffice. For a formal definition of human capital of training see section 3.3.

In order to get first evidence on whether on-the-job training participation correlates with employment in my data, I look at the difference in employment persistence between training participants and non-participants. For this I define each person's *human capital of training*, that is, the human capital gained from training, as the discounted sum of past training participation.[15] Figure 3 shows the employment rate conditional on being employed in the previous year for the group of employees who have a hu-

know the exact number they are asked to classify themselves into an income category. First with a rough grid with 3 categories and then with a finer grid with 9 categories, i.e. 3 finer categories depending on the previous response. The income variable I use takes the most precise available value. In case of categories the midpoint of the range is used.

[15] See section 3.3 for a formal definition of the human capital of training.

man capital of training of at least 0.9, that is, for example, who participated in at least 20 hours of training per year within the last 2 years since age 50, and those who did not. For most ages, the persistence in employment is slightly higher for people with training, indicating that training might affect employment security. But the mean differences are not significant.

Table 2 shows simple linear-probability model regressions of employment status (columns 1 and 2) and wage growth (columns 3 and 4) on the human capital of training, conditional on previous period employment, while controlling for wage levels (and training offers). I find significantly positive coefficients for the human capital of training on employment persistence. This remains true when controlling for training offers (columns 1 and 2). Regressing wage-growth on the human capital of training yields insignificant coefficients in both specifications with and without training offer controls (columns 3 and 4). These outcomes are in line with the findings of the literature.

Table 2: Simple regression

	(1) Employment	(2) Employment	(3) Wage growth	(4) Wage growth
Human capital of training	0.00568*	0.00585*	0.000941	0.000494
	(0.00241)	(0.00245)	(0.00328)	(0.00334)
Age	-0.00366***	-0.00366***	-0.00147	-0.00146
	(0.000557)	(0.000557)	(0.000767)	(0.000767)
Wage level	-0.00280	-0.00268	-0.00489	-0.00521
	(0.00194)	(0.00196)	(0.00266)	(0.00270)
Training offer		-0.00166		0.00438
		(0.00437)		(0.00599)
Constant	1.197***	1.199***	0.119**	0.116**
	(0.0315)	(0.0317)	(0.0433)	(0.0435)
R^2	0.0145	0.0145	0.00219	0.00237
N	3050	3050	2975	2975

Notes: Wage level is defined as initial wage divided by 1000. Age relative to age 50. Standard errors in parentheses. Significance codes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

This is first evidence to indicate that the human capital of training may indeed play a role in less-educated men's employment prospects in their late careers. However, this descriptive evidence is unable to reproduce the dynamic trade-offs that individuals face when considering training participation. The employee's decision to work or participate in training is based on a dynamic cost-benefit trade-off. Therefore, a reduced form model can not identify the impacts of training on the share of job-separations and employment prospects of employees. Further, it could not identify the individual costs of training, which are necessary for counterfactual analyses to evaluate potential policy interventions. Only then I can investigate whether policy interventions should target the supply of training or the individual participation incentives.

In the next section I will turn to the design of the structural model, which allows me to evaluate these interventions later on.

3 Model

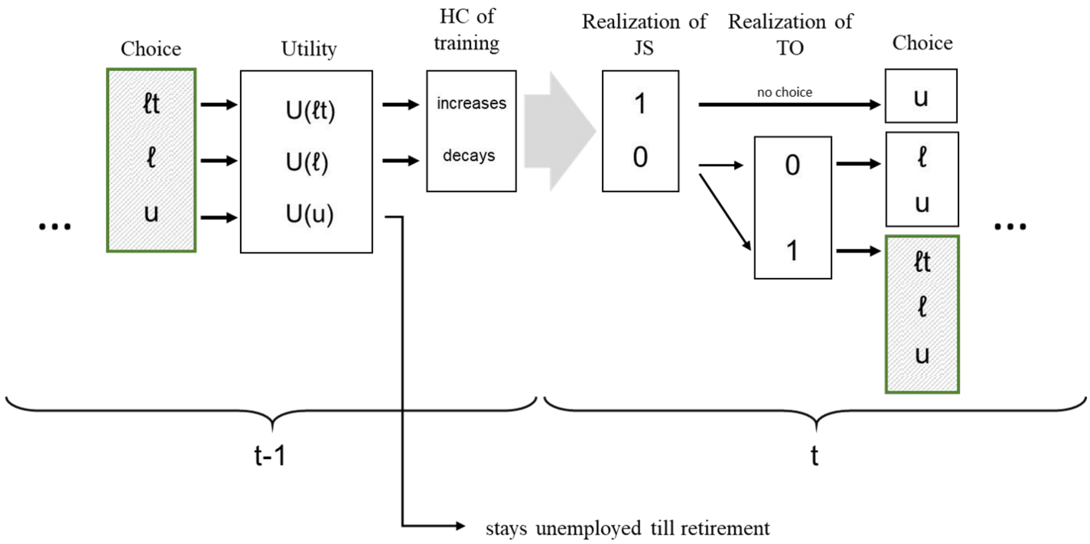
The previous section has shown that less-educated men in my sample indeed have lower employment rates close to retirement and that on-the-job training exhibits to be positively correlated with employment. I will now turn to the structural model, which explicitly models the trade-offs that individuals face when making choices about labor market participation and human capital investments.

3.1 Outline of the model

Employed men enter the model at age 50. In each period, they make a decision about whether to continue working and whether to invest in training, depending on their choice set. The choice set is determined by their employer's training offers (TO) and by the job separation rate (JS). Both, training and working are associated with a disutility. On the other hand, training and working can have positive effects on wages and future employment prospects. Therefore, they represent investments in monetary returns from the labor market.

During working life the individual has up to three different choice options available (unemployment u, working ℓ, working and training ℓt). Figure 4 shows a stylized sketch of the timing of the events in my model for the case where an individual has all three choices available in $t-1$: After the individual has made his decision ($d_t \in \{\ell t, \ell, u\}$), he receives his reward and the human capital (HC) of training is adjusted according to the decision. If the person chooses unemployment, he will be unemployed for the rest of his working life until retirement. If the person has chosen to work, realizations of the job separations (JS) and training offers (TO) occur in the next period. Depending on these realizations the individual faces one of the three possible choice sets, represented by the boxes, and again makes a decision. This choice process

Figure 4: Timeline from choice to choice



Notes: This is a stylized sketch of the model to illustrate the timing of events during working life of the individuals. It does not represent all interdependences between variables. For details on functional forms and dependent variables see section 3.2 ff. Choice sets depend on realizations of job separations (JS) and training offers (TO). Abbreviations: Choices: work and training (ℓt); work (ℓ); unemployment (u).

continues until a person becomes unemployed or retires at age 63. The utility and the job-separation rate in the model also depend on human capital investment decisions that the individual has made previously (details follow below).

3.2 The individual's optimization problem

At every age t the individual maximizes the following optimization problem. I drop individual subscripts for convenience.

$$max_{d \in D} \quad E_t \sum_{s=t}^{\bar{T}} \delta^{s-t} U(R_s, d_s), \tag{1}$$

with choice set $D_t \subseteq \{u, \ell, \ell t\}$ (u unemployment, ℓ work, ℓt work and training) during working life and $D_t = \{r\}$ (retirement) from age $t \geq 63$[16], and last period $\bar{T} = 85$.[17] Following Low et al. (2010) and Haan and Prowse (2014) I use a **utility function** that allows me to relate costs of work and training directly to the utility of consumption, which is set equal to rewards R_s in my model[18] (the rewards are defined in section 3.4).

$$U(R_s, d_s) = \begin{cases} \frac{\alpha}{1-\eta}\left[R_s(\ell t)\left(1-\zeta-\zeta_{age}*age-\tau*train-\nu\right)\right]^{1-\eta}+\varepsilon_t & \text{if } d_s = \ell t \\ \frac{\alpha}{1-\eta}\left[R_s(\ell)\left(1-\zeta-\zeta_{age}*age-\tau*train\right)\right]^{1-\eta}+\varepsilon_t & \text{if } d_s = \ell \\ \frac{\alpha}{1-\eta}\left[R_s(d_s)\right]^{1-\eta}+\varepsilon_t & \text{if } d_s \in \{r,u\} \end{cases} \quad (2)$$

The parameter ν is the disutility of training. The disutility of labor ζ is allowed to differ by age ζ_{age} and can depend on the human capital of training τ.[19] The disutility of labor can vary by age as leisure might become more attractive with deteriorating physical capacities (see for example Gustman and Steinmeier, 2005). Training could increase the enjoyableness of work as employees are/feel more proficient. The curvature of the CRRA utility function is determined by η, the risk aversion parameter. Individuals face random utility perturbations, represented by ε_t, which are extreme value type-1 distributed. The parameter α determines the importance of the preferences regarding earnings and effort relative to the random utility perturbations. Given that it is difficult to identify the risk aversion parameter,[20] I set the parameter η at a fixed level of 0.7, which is within the range Chetty (2006) finds.[21] The discount rate is set at 0.98

[16] Most people claim benefits as soon as they become available despite actuarial incentives (Gustman and Steinmeier, 2005). Also see Figure 1. I choose a common retirement age at age 63 for all employees to avoid inconsistencies with eligibility criteria that may arise in survey data (employment histories are based on retrospective surveys in NEPS).

[17] Age 85 roughly corresponds to the life-expectancy of someone who is today 62 years old. The life-expectancy varies by age but the difference between age 50 and 60 is small. It increases only by 0.7 year. Thus I generously use a horizon of 85 for everyone. As individuals only make choices up to age 62 life-expectancies for years beyond that age are irrelevant.

[18] Similar to the paper by Keane and Wolpin (1997) where individuals optimize over rewards instead of consumption. In contrast to Keane and Wolpin (1997) where individuals maximize their expected present value of their lifetime rewards, I assume that individuals maximize over the utilities of these rewards, allowing for decreasing marginal valuation of additional money.

[19] See Section 3.3 for a definition of human capital of training.

[20] As it is the case in many structural models even in papers, which attempt to estimate this parameter.

[21] Also Wakker (2008) implies that this is a reasonable assumption. Note that individuals' instantaneous utility is created by the income not consumption – individuals may have higher risk aversion with respect to the latter. For higher values of η, employees would care too little about disutilities of work

following Blundell et al. (2019) and Haan et al. (2018); all other parameters will be estimated.[22]

Utility costs of work and training The disutility parameters, which are defined as relative withdrawal from the reward,[23] have a rather broad interpretation. For example, disutility of work can also include (negatively) the joy of work or a benefit of not being unemployed. Likewise, the parameter ν reflects the sum of all sorts of immediate utility changes that are associated with training – this can be e.g. effort costs, monetary costs, or other frictions.

The **choice set** D_t is determined by the individual's age, exogenous training offers (TO) provided in the data (see section 2.1), and involuntary job separations that occur with probability JS (see section 3.5 for details). During working life the choice set can consist of up to three choice options $D_t = \{u, \ell, \ell t\}$ if the individual receives a training offer and does not face a job separation (TO=1, JS=0). If he does not receive a training offer he can only choose between unemployment and work $D_t = \{u, \ell\}$ (TO=0, JS=0), and if he loses his job (JS=1) he has no choice $D_t = \{u\}$ and becomes unemployed. Once employees become unemployed (due to choice or separations), they remain unemployed until retirement. This assumption is reasonable, as very few low-educated individuals return to employment once they become unemployed after age 50.[24] It further faciliates identification. At age 63, all individuals are assumed to retire and remain in retirement until the end of their life $D_t = \{r\}$ for $t \geq 63$. That is, individuals will not have any more choices to make once they become unemployed or retired.

and training.

[22] As the household context is arguably less important for employment decisions of this subsample which entered the labor market in the 1970s/1980s I will model the decisions independent of the presence of a partner or children. This further allows me to circumvent the problem that the NEPS does not provide income information of other household members. As I use only men aged above 50, who are relevant for my research question, I do not need to make strong assumptions about the equivalence of training and working conditions across decades, that are necessary in life-cycle models (e.g. Blundell et al., 2019).

[23] This implies larger withdrawals from the reward for higher wages. Yet, this effect is diminished in the respective utility due to the curvature of the utility function. For $\eta = 1$ (i.e. log-utility) the utility loss would be equivalent for different income (reward) levels. For $\eta < 1$ the utility loss would be higher for higher reward levels, for $\eta > 1$ it would be lower.

[24] Less than 8% of the low educated between 50 and 62 return to work once getting unemployed in my data, Etgeton (see also 2018).

3.3 Human capital

Training When individuals enter the model their human capital of training ($train$) is normalized to 0. Any human capital of experience and training that was acquired before is assumed to be reflected in the endowments of the wage-level w_{t0} and in the fact that they are in employment when entering the model. In the subsequent periods each time when the employee chooses to train ($tr_t = 1$) this adds to his human capital account but the human capital of training decays over time at the rate $\delta_{hc} := 0.93$.[25]

$$train_t = \sum_{s=t0}^{t} \delta_{hc}^{t-s} I(tr_t = 1) \quad (3)$$

The average level of human capital of training acquired since age 50 lies well below 1. It increases up to age 58 to an average level of 0.63 and then decreases again up until retirement.[26]

Training offers The availability of training, that is financing for on-the-job training by the employer (as observed in the data), is assumed to be exogenous.[27] For future time periods individuals expect their training offer to be equivalent to their current training offer: $E_t TO_{t+1} = E_t TO_{t+2} = TO_t$. Training offers are observed before the choice is made.

3.4 Rewards

When individuals are employed, their rewards equal their annual net wage $R_t(\ell) = R_t(\ell t) = w_t$, that is they receive the same salary when engaging in training.[28] The NEPS provides both stated gross and net wages. I use net wages, because these are

[25]This corresponds the average of the human capital depreciation rates in Blundell et al. (2019).

[26]See Figure 24 for average level of $train$ by age.

[27]Selection of older workers into firms with training offers at this age is unlikely. Section 2.2.2 confirmes that attitudes and ambitions of employees with and without training offers are similar.

[28]This model assumption is purposely different from typical human capital investment models, e.g. as applied in Fan et al. (2017), as in the context of on-the-job training of German employees in their late careers, where the company mostly finances and often even provides the training (Pischke, 2001; Görlitz, 2011), it would be an unreasonable assumption to model training costs as forgone earnings. Instead, potential training costs, monetary or non-monetary, are captured by the flexible parameter ν in the utility function. This way of modeling training costs is similar to Blundell et al. (2019).

closer to consumption and hence more useful for representing the individual's utility returns from work. I do not include savings in the model as savings are arguably of minor relevance for less-educated employees' late career choices due to the fact that savings would typically be close to zero for this group (see, for example, Börsch-Supan et al., 2015). When unemployed, the reward equals some unemployment benefit $R_t(u) = UB_t$ and, when retired, the reward equals the retirement benefit $R_t(r) = RB$. Details on the values of these rewards are provided below.

3.4.1 Wages w_t

As I include only individuals who are employed when entering the sample, I observe an initial wage w_{t0} for everyone. This wage reflects the market valuation of the employee's work when entering the model, including the human capital levels at this time and the general ambitions of the respective employee. In the subsequent periods, the development of this wage is assumed to depend on a general wage trend, human capital of training, and age.[29] Wages are assumed to emerge in the following way:

$$w_t = w_{t0} * (1 + \alpha_0 + \alpha_1(train_t) + \alpha_2 age_t)^{t-t0} \tag{4}$$

with the *age* relative to age 50, human capital of training $train$.[30] This definition allows wage level to decrease if the human capital of training decays. This assumption is in line with much of the literature (see eg. Blundell et al. (2016a)) and is reasonable, as the data reveals that a relevant share of the employees face negative wage growth at times. I assume that wages are deterministic from the worker's perspective. However, from the researcher's perspective they are not as I only observe wages that potentially include measurement error.

3.4.2 Unemployment benefit UB_t

I set the unemployment benefits to 60% of the previous wage, which is in line with the German rules.[31] They are paid for up to two years for employees aged 50 or older.

[29] The data does not indicate differences in relative wage growth across different wage levels for the group of less-educated, hence I removed this as control to save parameters.

[30] Log wages are assumed to follow a normal distribution. Measurement error follows a normal distribution with mean zero. Hence observed sample wages are assumed to be given by: $Log(w_t) + \varepsilon_t$ with $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$.

[31] For individuals with dependent children 67%. I will use 60% for everyone.

After this period they receive means tested transfers and housing benefits. This is also reflected in the reward function for unemployed people. The level of means tested transfers plus housing benefits is set at EUR 959 for everyone.[32]

Retirement benefit *RB*

In Germany, the retirement benefits depend on the time that an individual has contributed to the system, i.e. employment years, and on the contribution level in these years. More specifically, if the individual contributes less or more than the average person in a year, then the contribution year is scaled down or up, corresponding to the contribution level. The pension level is computed by adding up the scaled contribution years and multiplying it with the current “Rentenwert” (retirement benefit value) and penalties are deducted for early retirement (Deutsche Rentenversicherung, 2020). It turns out that for 45 contribution years, the gross retirement benefit amounts to roughly 45% of the average gross wage and that a missing contribution year leads to a deduction of roughly 1%. I reflect this in my model by considering the observed wage level and using penalties for years of unemployment prior to retirement during my observation period.[33] Each retiree receives this annuity (*RB*) for the rest of his life.

3.5 Job separation rate

The probability of becoming unemployed (job separation rate *JS*) depends on the employee’s age, the human capital of training, and the initial wage level $wage_{t0}$[34] when the individual entered the sample. The JS is assumed to follow a binomial-logit func-

[32]EUR 409 "ALGII"+550 housing benefit. As I don’t have precise information about individuals’ household context, savings, or housing costs an exact computation of means tested transfers is not possible.

[33]For the first two years of unemployment (ALG I) the contribution is reduced by 20%. As I do not observe all wage levels in the employment history and only very rough information about actual contribution years, I generously use the last wage level to calculate the retirement benefit and assume that the individual was fully employed in all years prior to my observation period. Therefor I do not account for the fact that the lower relative tax burden in retirement years improves the ratio with respect net-wages and benefits compared to the gross values.

[34]$wage_{t0}$ is divided by 1000 in this equation to avoid very small parameters in estimation.

tional form:[35]

$$JS(train, age, wage) = \Lambda\left(\beta_0 + \beta_1 train + \beta_2 age + \beta_3 wage_{t0}\right) \quad (5)$$

The parameter vector β captures the impact of the state variables on the job separations and hence the choice set. That means, the realization of JS and the resulting choice options that the individual has depend on his previous investments in human capital of training. The individual observes the realization of JS before making his decision.[36]

As I do not observe involuntary job-separations in the data, the relative magnitude of job-separations compared to voluntary transitions into unemployment, due to the individual's utility considerations, are determined by the functional form of my model. Similarly, the functional form determines the way the effect of human capital of training on employment is split between the job-separation rate and the utility function. However, the policy-relevant measure for my research question is the combination of the job-separations *and* the voluntary transitions: the employment persistence and the employment rate. The latter are identified with the data.[37]

3.6 Value functions

The resulting value functions of this dynamic-programming problem are defined as follows:

[35] $\Lambda(.) = exp(.)/(1 + exp(.))$

[36] The special set up in my model allows me to omit an experience variable. All employees who enter my model are employed and once they become unemployed they will remain unemployed until they are eligible for retirement. Any market valuation of experience that was gained prior to entering the model will be reflected in w_{t0} and the fact that the person is still employed. Any return to experience after entering the model is captured by the constant α_0 for wages and β_0 and β_2 (negatively) for the job separation rate.

[37] It would be interesting for further research to look more into the distinction between the effects of training on the job-separation rate and the utility function beyond the functional form, especially to validate the robustness of the reduction in the disutility of work due to training. This requires a data-set containing both detailed information about training participation and involuntary job separations. For a large enough data set also an exogenous shock to either channel would do.

The value function when choosing working and participating in on-the-job training:

$$
\begin{aligned}
V_t^{\ell t}(s_{it},\theta) &= U(R_s(\ell t),\ell t) \\
&\quad + \delta\Big[(1-Pr(JS=1|s_{i,t+1},\theta)) \\
&\quad \big[Pr(TO=1|s_{i,t+1})Emax\{V_{t+1}^u(s_{i,t+1},\theta),V_{t+1}^{\ell}(s_{i,t+1},\theta),V_{t+1}^{\ell t}(s_{i,t+1},\theta)\} \\
&\quad + (1-Pr(TO=1|s_{i,t+1}))Emax\{V_{t+1}^u(s_{i,t+1},\theta),V_{t+1}^{\ell}(s_{i,t+1},\theta)\}\big] \\
&\quad + Pr(JS=1|s_{it+1},\theta)E\{V_{t+1}^u(s_{i,t+1},\theta)\}\Big]
\end{aligned}
$$

The value function when choosing working

$$
\begin{aligned}
V_t^{\ell}(s_{it},\theta) &= U(R_s(\ell),\ell) \\
&\quad + \delta\Big[(1-Pr(JS=1|s_{i,t+1},\theta))\big[Pr(TO=1|s_{it})Emax\{V_{t+1}^u,V_{t+1}^{\ell},V_{t+1}^{\ell t}\} \\
&\quad + (1-Pr(TO=1|s_{i,t+1}))Emax\{V_{t+1}^u,V_{t+1}^{\ell}\}\big] \\
&\quad + Pr(JS=1|s_{i,t+1},\theta)E\{V_{t+1}^u\}\Big]
\end{aligned}
$$

and for unemployment

$$
V_t^u(s_{it},\theta) = U(R_s(u),u) + \delta E\{V_{t+1}^u(s_{i,t+1},\theta)\}.
$$

with parameters θ and the current state variables s_{it} reflecting wage, the human capital of training, age, and employment status. ($V_{t+1}^{\cdot}$ abbreviates $V_{t+1}^{\cdot}(s_{i,t+1},\theta)$.) $V_t^{\ell}(s_{it},\theta)$ and $V_t^{\ell t}(s_{it},\theta)$ differ in the instantaneous utility and in the value of the state variable "human capital of training" in $t+1$. The value function $V_t^u(s_{it},\theta)$ reflects that the individual will not make any further decisions. The value function is solved via backward induction, starting at the terminal period $\bar{T}$, which corresponds to age 85. See Appendix B for details on the estimation.

Potential effect of training on employment

If training reduces the job separation rate, it has a first and a second order effect on employment: First, job separations are reduced and hence the involuntary unemployment decreases. Second, voluntary unemployment becomes more costly relative to employment when the probability of future involuntary job loss is reduced. Similarly, if the effect of training on wages is positive, then voluntary unemployment becomes more costly, as the individual would miss out on increased future wage growth.

4 Results and model fit

In this section I present the model-parameters estimated with the maximum-likelihood estimation and provide information about the in-sample fit.

4.1 Estimation results

Table 3 shows the parameter estimates for the wage function, the job separation rate, and the utility function, with standard errors in parentheses.

Table 3: Parameter estimates

Parameter		Estimate	Std. Err.
Wage function			
α_0	(Intercept)	0.02091	(0.00203)***
α_1	(HC of training)	0.00015	(0.00011)
α_2	((Age-49)/10)	-0.00031	(0.00226)
Employment risk (JS)			
β_0	(Intercept)	-7.7561	(0.7361)***
β_1	(HC of training)	-0.1076	(0.0489)*
β_2	(Age-49)	0.3067	(0.0801)***
β_3	($wage_{t0}$/1000)	0.2990	(0.09767)**
Utility function			
ζ	(Disutility of employment)	0.40533	(0.16780)*
ζ^{age}	(Change in ζ by age)	0.00993	(0.01198)
τ	(Change in ζ by HC of training)	-0.00695	(0.00265)**
ν	(Disutility of training)	0.06370	(0.00828)***

Notes: Signif. codes: '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1. SD measurement error of wages $\sigma_\varepsilon = 0.1379496$ (SE 0.0003549). Utility scaling parameter $\alpha = 1.2034$. N=3050.
Source: NEPS data, less-educated male employees only.

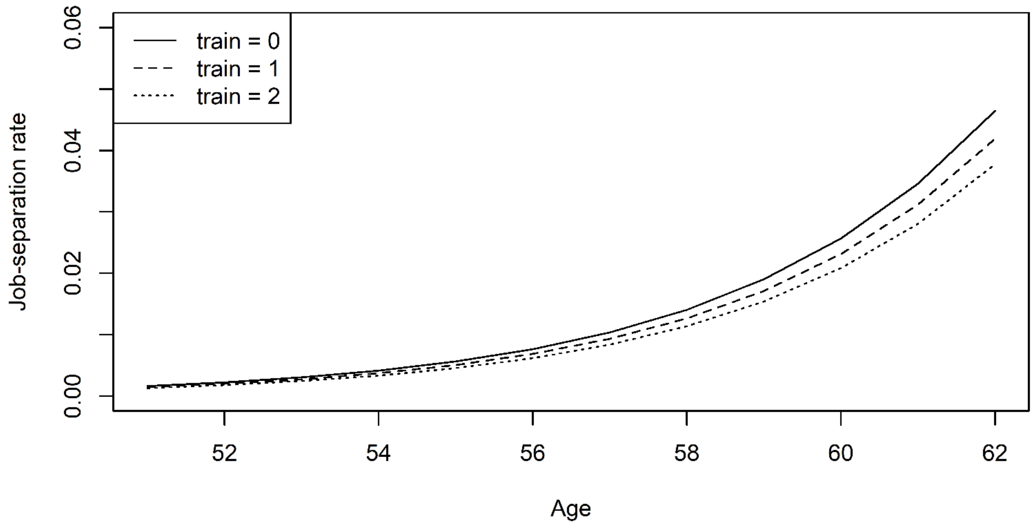
The results show a positive constant for the nominal wage growth rate of 2.1%. Parameters of wage-growth function indicate that wage growth increases with human capital of training and decreases with age on average but both coefficients are very small and not significant. This is in line with the results from most of the previous literature, which does not find significant effects of training on wages (see Section 1). The rate of job separations (JS) increases with age (significant at 10% level) and also significantly with the initial wage level, but it decreases significantly with training. Figure 5 shows the job-separation rate by age for an initial wage-level of EUR 2500 by different levels of human capital of training. One can see that the job-separation rate increases with age and consequently the largest percentage-point decrease due to human capital of training can be achieved near the retirement age. For example, at the age of 58, for an initial wage level of 2500 having one unit of human capital of training compared to having 0 units of human capital of training decreases the probability of job loss from 1.90% to 1.71%, i.e. by 0.19 percentage points. For age 62 one unit of human capital of training reduces the job-separation rate by 0.72 percentage points. The relative reduction is 10% for one unit of human capital of training and about 19% for two units compared to 0 units. Yet, even for two units of human capital of training, which lies far above the average of 0.42, training cannot fully counteract the age related increase in employment loss. Small effects of training on employment of older workers were also found by previous literature (see e.g. Picchio and Van Ours, 2013). The average job separation rate lies within the range of the findings of previous literature (see Haan et al., 2017).[38]

The parameters of the utility function indicate that working is associated with a disutility of 40% of the net-wages, which decreases slightly with human capital of training. The age trend of the disutility is positive but not significant. The disutility of training is 6.4% and significant.[39] That is, individuals face positive costs when participating in training despite the fact that their firm provides the financing and they continue to receive their regular salary. These costs could include effort costs, general taste, small time or monetary costs, or other frictions related to training participation. Despite these positive costs it can still be worthwhile for individuals to participate in training as it reduces the probability of becoming unemployed and it reduces the disutility of work for future periods. Due to the fact that the training costs are relatively low, some people still participate in training when they are near retirement and the remaining working periods where training could pay off are limited.

[38]This indicates that the distinction between job separations and chosen unemployment via the functional form of my model works sufficiently.

[39]This level of training costs is comparable to the paper by (Blundell et al., 2019) who fix the training costs at 2 hours forgone wage, which would be 5% in a 40 hrs week.

Figure 5: Job-separation rate by age and human capital of training



Notes: The black lines show the predicted job separation rate by age for individuals with no human capital of training (solid line), with human capital of 1 (dashed line), and 2 (dotted line) for an initial wage level of 2500 EUR. *Source:* NEPS; own calculations.

4.2 Goodness of fit

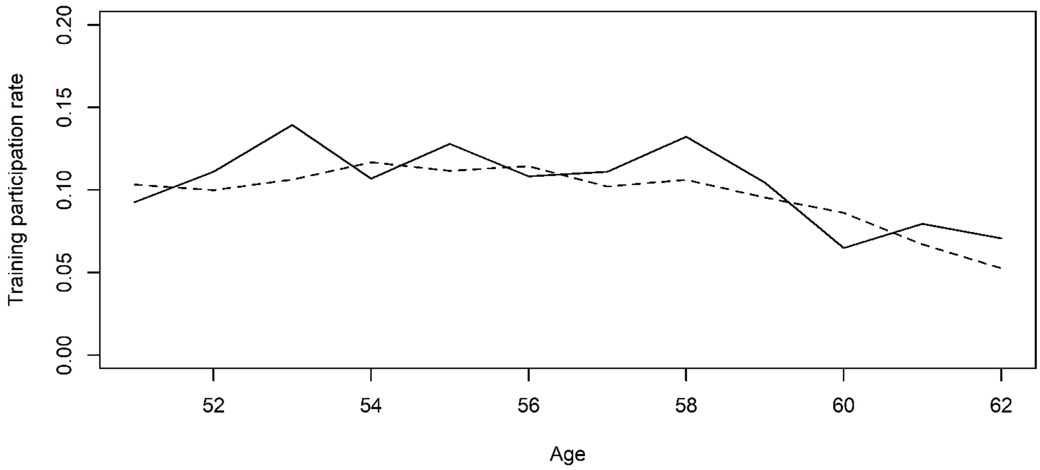
To evaluate the goodness of fit I compare the actual choices and wages in the data with the choices and wages that the model would predict by age.[40]

Overall the model fits well. Average training participation (Figure 6) fits well and deceases slightly with age, as it does in the original data. Also the average employment persistence, i.e. the probability to remain employed, by age fits well (Figure 7). It is very close to 1 in the early 50s and then first decreases slowly and then more sharply when approaching retirement age.

Figure 8 displays the original density of wages in solid black and the simulated in dashed green. The simulated wage density has a less pronounced spike around 2000 but nicely overlaps the original date. Figure 9 shows the mean and median wages by age. The simulated mean and median are slightly higher but roughly fit the data. The simulated median wage also reflects the decay close to retirement age as it can be

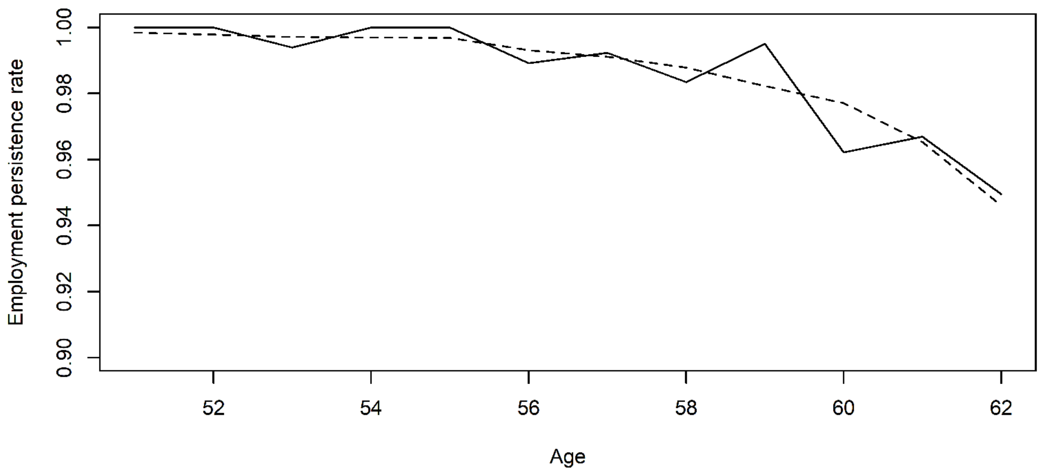
[40]For the simulated choices I replicate my sample 50 times to allow for different draws of random utility perturbations and measurement error.

Figure 6: Training rate by age original vs simulated



Notes: Training participation by age. Original values in solid line. Simulated values in dashed line. *Source:* NEPS data, subsample. Own calculations.

Figure 7: Employment persistence by age original vs simulated



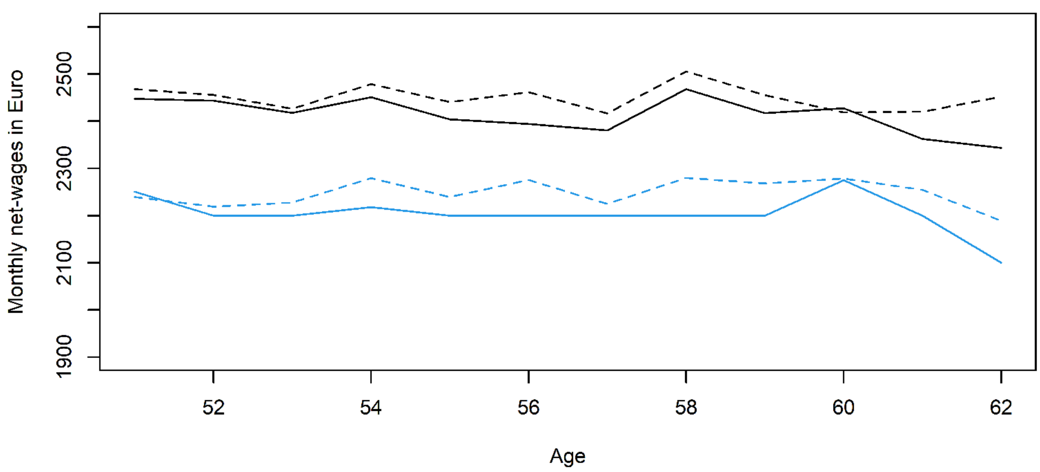
Notes: Share staying employed (employment persistence rate) by age. Original values in solid line. Simulated values in dashed line. *Source:* NEPS data, subsample. Own calculations.

observed in the original data.

Figure 8: Density of original wages and simulated wages

Notes: Density of monthly net-wages. Original values solid line, simulated wages in dashed line. *Source:* NEPS data, subsample. Own calculations.

Figure 9: Wages by age simulated vs. original



Notes: Mean and median monthly net-wages. Original values in solid line (black: mean wage; grey: median wage). Simulated values in dashed line. *Source:* NEPS data, subsample. Own calculations.

5 Counterfactual simulations

In the previous sections I have quantified parameters that determine the training decision of the less-educated men in my sample. I showed that on the one hand individuals have significant participation costs but on the other hand human capital of training has a positive impact on the employment prospects. My data set has allowed me to separate individual costs of training from the availability of training offers, that is the general availability of training funding in the firm. In this section I can now turn to the question whether a policy intervention should target training supply or individual participation incentives conditional on training supply, in order to increase employment near retirement. I investigate these two channels separately: First, the training offers (section 5.1) and second, the individual costs of training (section 5.2).

For the counterfactual simulations I randomly redraw 10,000 times from the sample of 51 year old in my data (less-educated male employees) and simulate their choices and corresponding human capital and wage measures till retirement age.[41] Afterwards, I calculate aggregate employment outcomes and compare the counterfactual to the baseline model.

5.1 Increasing training offers

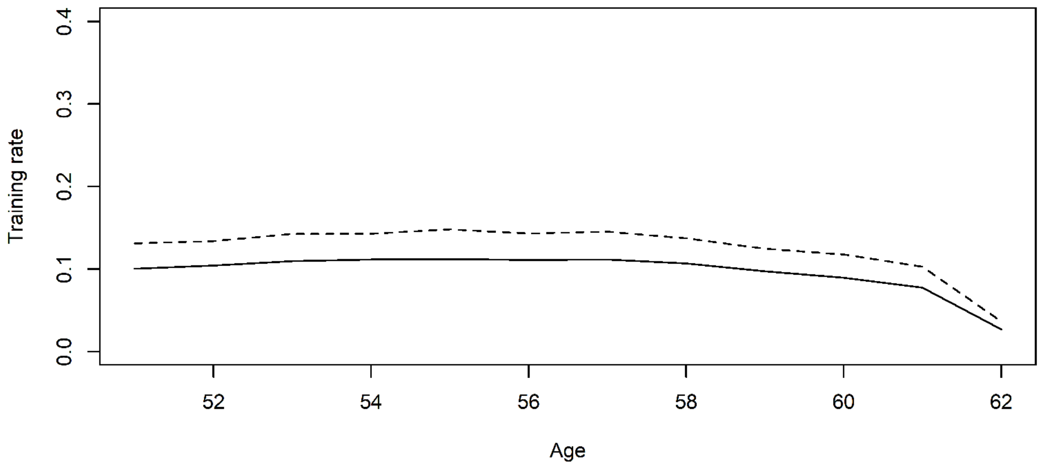
Scenario 1 An increase in training offers would enable more employees to choose training. Hence it could increase the incidence of training, and its beneficial effects on employment outcomes. Therefore, a policy intervention, like a subsidy of training costs for firms, might improve employment prospects of less-educated men in their late careers. I investigate the impact of an extreme policy intervention that would increase training offers to 100%: $TO = 1$ for everyone at any age. Note that the general willingness of the employers to invest in training (as reflected by the training offer rate) is with 80% quite high in my sample already, while the training participation rate is close to 10%. Consequently, I expect a relatively small impact of such an increase in the training-offer rate on employment outcomes.

This is exactly what I observe in my simulation. The change in the simulated training participation rate is very small (Figure 10). The average training participation in my

[41] I do this separately, for the baseline model (without intervention) and for the counterfactual scenarios.

sample increases from 9.7% to 12.6%. This small change in the training rate has hardly any impact on the employment persistence (Figure 11) and employment rate (Figure 12). At age 62, where the effect is largest, the employment rate rises from 83.5% to 84.0%, i.e. by 0.55 percent.

Figure 10: Scenario 1: Training rate – baseline model versus counterfactual



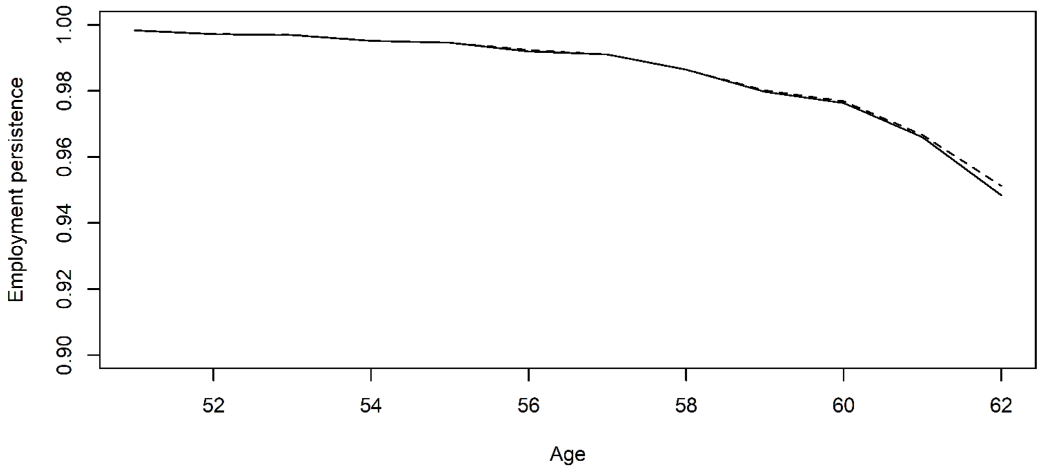
Notes: Baseline solid line, counterfactual scenario 1 dashed line. Simulation based on 10,000 randomly re-sampled individuals drawn from the original sample of 51 year old. *Source:* NEPS data. Own calculations.

In summary, the simulation shows that a policy intervention that targets only the willingness of firms to invest in training would not be effective in increasing the employment of less-educated men in their late careers, even if it achieves to increase the training offer rate to 100%. In practice an implementation would additionally be challenged by potential crowding out effects of firm provided further-education investments (see Görlitz, 2010). This risk is larger in a setting where general willingness to invest in training is already high - as in my data. Thus, any potential implementation of such a policy would need to be carefully deliberated. Given the negligible returns of a successful implementation, this intervention does not appear too promising.

5.2 Reducing individual training cost

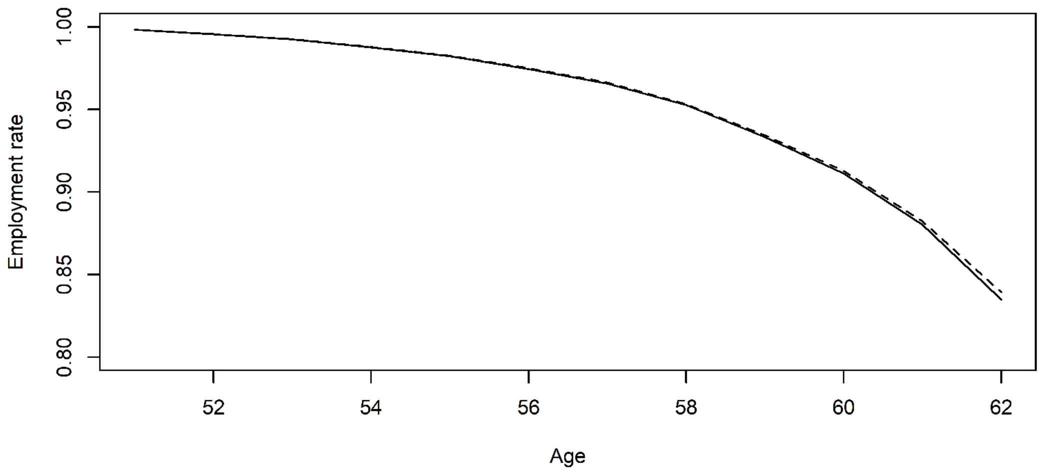
Scenario 2 Since the rate of training participation is much lower than the training-offer rate it might be more promising to think about a policy intervention that targets

Figure 11: Scenario 1: Employment persistence – baseline model versus counterfactual



Notes: Baseline solid line, counterfactual scenario 1 dashed line. Employment rate of individuals who have been employed in previous period. Simulation based on 10,000 randomly re-sampled individuals drawn from the original sample of 51 year old. *Source:* NEPS data. Own calculations.

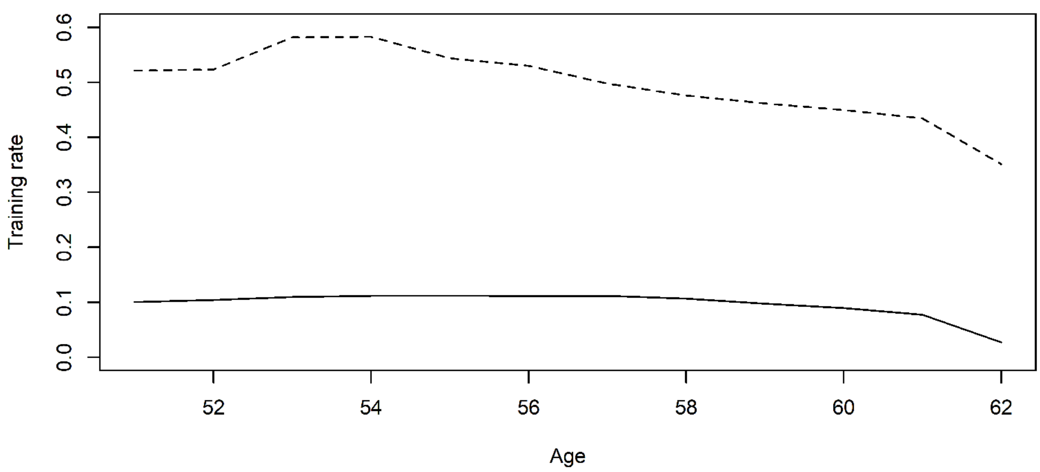
Figure 12: Scenario 1: Employment rate – baseline model versus counterfactual



Notes: Baseline solid line, counterfactual scenario 1 dashed line. Simulation based on 10,000 randomly re-sampled individuals drawn from the original sample of 51 year old. *Source:* NEPS data. Own calculations.

individual incentives, i.e. the utility costs of training (ν). Thus, I exogenously reduce the training costs in this second counterfactual analysis. To see the full potential of a policy that targets individual incentives for training participation I analyze the extreme case of $\nu = 0$. That is the employee's full disutility of training is compensated. Corresponding potential policies could pay fringe benefits or other compensation payments, which are payed in the year of training participation, or they could try to reduce the non-monetary costs of training participation. For instance, frictions like the effort to gather information about courses or to enroll in courses could create non-monetary costs. Easy access to information about training or default sign-up rules could reduce these costs.[42]

Figure 13: Scenario 2: Training rate – model vs counterfactual



Notes: Baseline solid line, counterfactual scenario 2 dashed line. Simulation based on 10,000 randomly re-sampled individuals drawn from the original sample of 51 year old. *Source:* NEPS data. Own calculations.

Reducing the utility costs of training to 0 would have a large impact on the training participation: It would increase to 50% on average, with the highest participation rate of 58.3% at age 54.[43] As a consequence the employment persistence and employment rate of older employees (Figures 14 and 15) would increase. The largest percentage point increase in employment would be achieved for the oldest employees: Assuming the same effect of training on involuntary separations under such a dramatic increase

[42]For the design of a precises policy intervention an additional analysis on the composition of the individual training costs would help.

[43]Note that less than 80% of the employees have a training offer, that is about two thirds participate in training on average.

in training participation, this would increase the employment rate of 62 year old less-educated males from 83.5% to 87.6%. This corresponds to a 4.9% increase in the employment rate of 62 year old. For age 60 the employment rate would increase from 91.1% to 92.8%, i.e. by 1.9%.

This second counterfactual simulation shows that a policy intervention which directly addresses the utility costs of employees could be more effective than an intervention that addresses the general provision of training from the firm side. Such a reduction could be achieved by different policy instruments: Besides compensation payments, which would amount to EUR 1911 for an employee with a monthly net-wage of EUR 2500 in this scenario, the reduction of non-monetary costs could be effective. For example, a study by Van den Berg et al. (2019) provides evidence that providing information about training programs can increase the training participation.

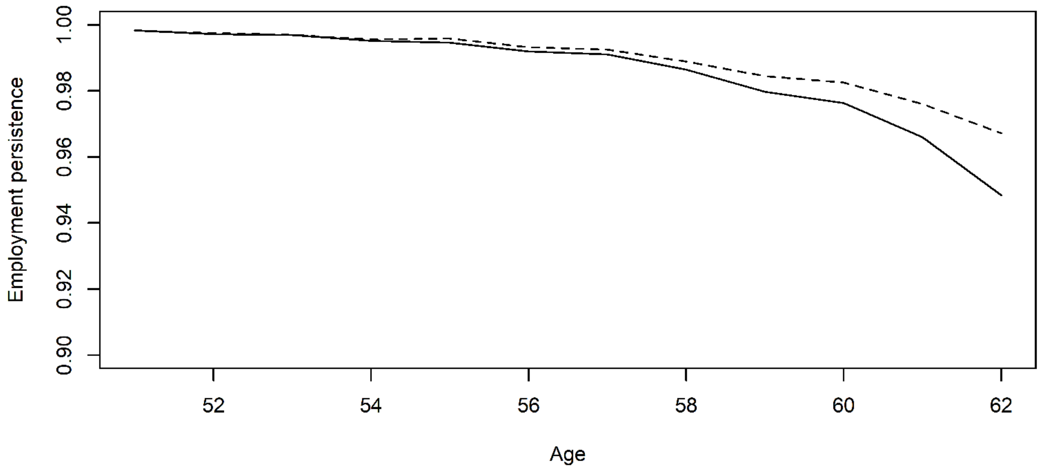
Yet, even a compensation of the entire training costs could not fully counteract decreasing employment rates of less-educated employees approaching retirement age. This is driven by two forces: First, the relative size of the absolute value of the training coefficient in the employment risk function is smaller than the age coefficient reflecting that training cannot fully compensate for advancing age. Second, some part of this unemployment is a result of the employee's trade-off between the utility of an additional full salary plus no penalties for the retirement benefit compared to the unemployment benefit without any disutility of work, despite the fact that the increased training activity would have decreased the disutility of work slightly.

Scenario 2b The intervention in scenario 2 may appear extreme as we look at a 100% reduction in training costs, while we looked at a 25% increase in training offers in scenario 1. That is why I add a counterfactual simulation where I consider a 25% reduction in training costs (scenario 2b). That is, if the chosen policy instrument is compensation payments to reduce the individual training costs, an employee with a monthly net-wage of EUR 2500 would receive a tax free compensation of EUR 478 for his training participation.[44]

As we see in Figure 16 the reduction of the individual training costs by 25% would lead to a less pronounced increase in the training rate compared to a 100% reduction: The average training participation would increase to 18.1% on average, with an increase

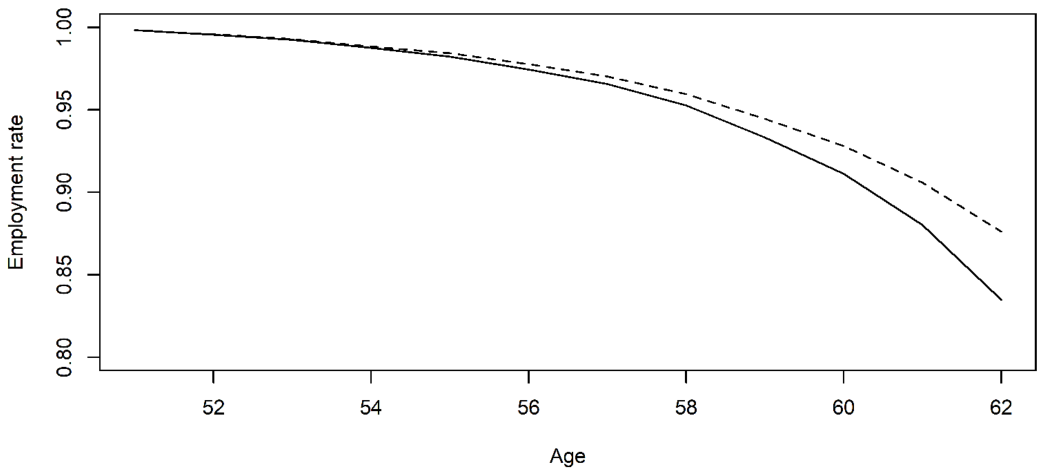
[44] This magnitude of the costs is still difficult to compare to scenario 1, as it is unclear what the firms would pay for the training provision and to what extend crowding-out would play a role.

Figure 14: Scenario 2: Employment persistence – baseline model vs counterfactual



Notes: Baseline solid line, counterfactual scenario 2 dashed line. Simulation based on 10,000 randomly re-sampled individuals drawn from the original sample of 51 year old. *Source:* NEPS data. Own calculations.

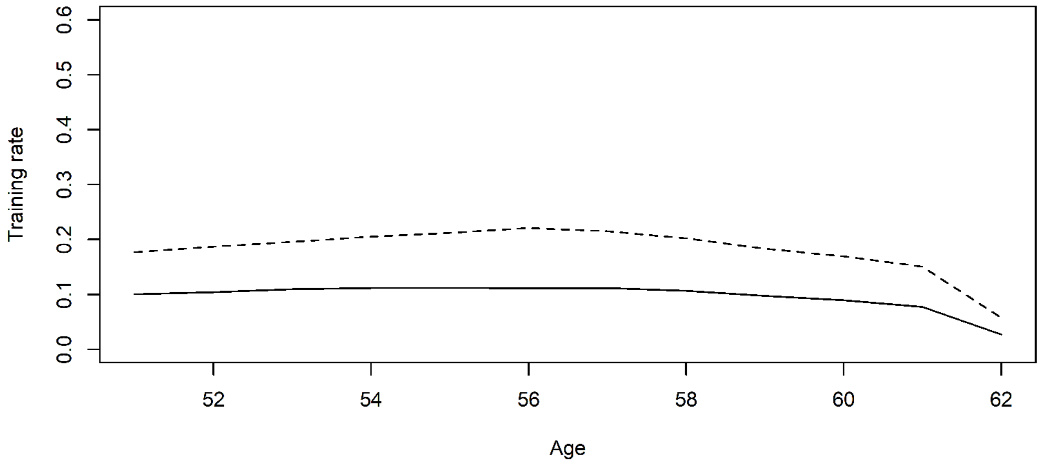
Figure 15: Scenario 2: Employment rate – baseline model versus counterfactual



Notes: Baseline solid line, counterfactual scenario 2 dashed line. Simulation based on 10,000 randomly re-sampled individuals drawn from the original sample of 51 year old. *Source:* NEPS data. Own calculations.

by 98% at age 57 from 11.2% to 21.5% (compared to a maximum increase of 33% in scenario 1). As a consequence the effects on the employment persistence and employ-

Figure 16: Scenario 2b: Training rate – model vs counterfactual

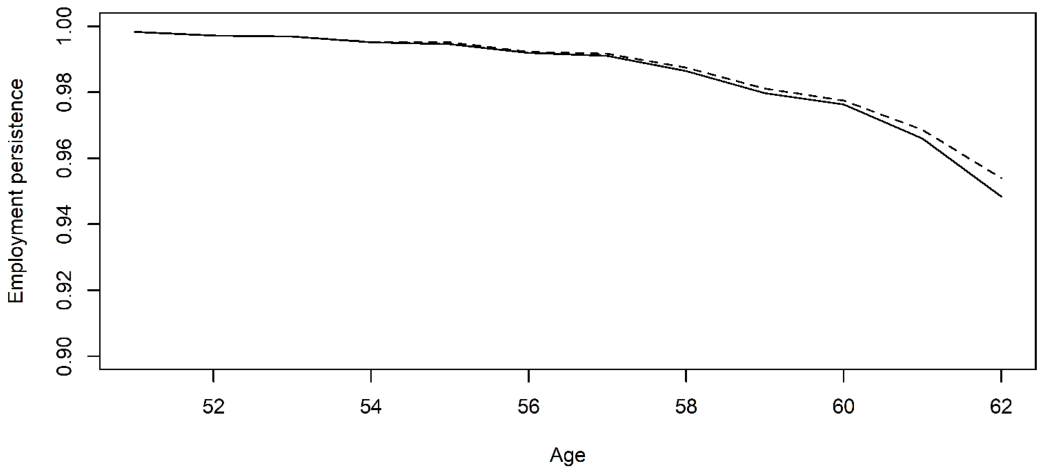


Notes: Baseline solid line, counterfactual scenario 2b dashed line. Simulation based on 10,000 randomly re-sampled individuals drawn from the original sample of 51 year old. *Source:* NEPS data. Own calculations.

ment rate (Figures 17 and 18) would also be more pronounced compared to scenario 1: At age 62 the employment rate would increase by 1.4% from 83.5% to 84.6%.

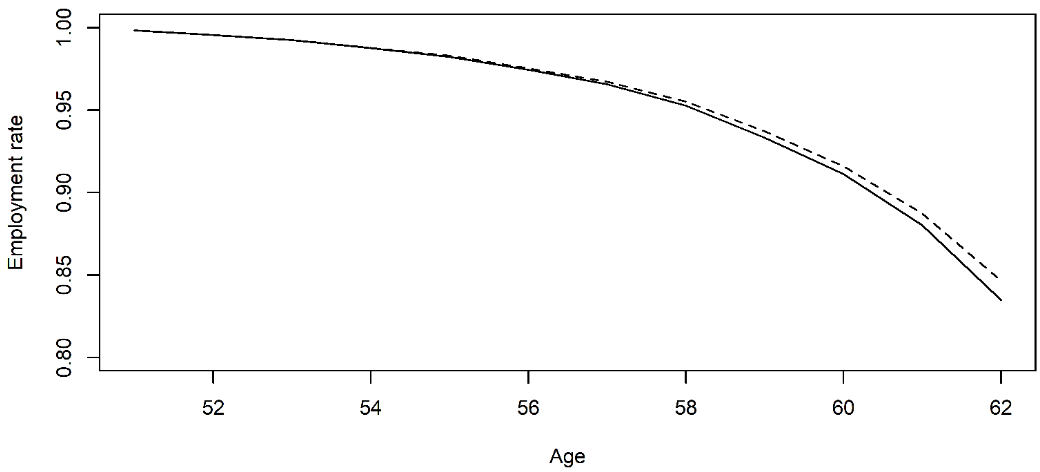
In conclusion, even a 25% reduction in the individual training costs would have a larger effect on employment than a 25% increase in training offers. Importantly in the simulation in scenario 1 I already reached the maximum possible intervention intensity for this channel, while the individual training cost could in principle even be overcompensated and could consequently achieve an even further increase in training participation in scenario 2. Hence, it reveals to be more promising to target individual incentives of training participation than the provision of training from the firms' side.

Figure 17: Scenario 2b: Employment persistence – baseline model vs counterfactual



Notes: Baseline solid line, counterfactual scenario 2b dashed line. Simulation based on 10,000 randomly re-sampled individuals drawn from the original sample of 51 year old. *Source:* NEPS data. Own calculations.

Figure 18: Scenario 2b: Employment rate – baseline model versus counterfactual



Notes: Baseline solid line, counterfactual scenario 2b dashed line. Simulation based on 10,000 randomly re-sampled individuals drawn from the original sample of 51 year old. *Source:* NEPS data. Own calculations.

6 Discussion and conclusion

In this paper, I investigate the role of on-the-job training for employment outcomes of less-educated men aged above 50 using a structural dynamic discrete-choice model. This model provides novel insights into the trade-offs that these employees face when they decide whether to participate in on-the-job training. An important feature of my data set, the NEPS, is that it provides the necessary information to distinguish between the general availability of training funding in the firm and the individual utility costs of training. Using this feature in my structural model I was able to quantify the benefits and costs of training for the employee and to simulate the effect of different policy interventions on employment outcomes. As a consequence I could answer the question whether policy makers should increase training supply or incentivize individual training take-up to foster overall training participation.

The estimated parameters support findings from the existing literature, which indicate that the human capital of training has little effect on wages but has an impact on the employment outcomes. Further, I find that training causes a small reduction in the disutility of work, which could be an interesting starting point for further research addressing work motivation of less-educated employees in their late careers.

The counterfactual simulations in the last section illustrated, that it is less the lack of training funding in firms that determines whether or not employees participate in training and more the individual training costs. Further, training participation is shown to have a positive impact on the employment rate. In an extreme case, where the individual training costs were reduced to zero, a small increase in elderly employees' employment persistence would be achieved and hence result in an increase in employment rates near retirement from 84% to 88%. Therefore, fostering on-the-job training could play a part in future policy interventions that seek to address unemployment of less-educated employees in their late careers. However, in its current form on-the-job training would not be able to fully counteract the fact that employees' employment persistence decreases with age. It would be interesting to see more research on the question of how on-the-job training could become more effective in improving employment outcomes among elderly less-educated employees; for example, on the quality of training or the fit to the needs of older less-educated employees.[45] Furthermore, research on the composition of the utility costs of training could help to design an

[45] A study by Bellmann et al. (2013) provides first descriptive evidence on this topic.

effective policy intervention targeting individual costs of training.

In conclusion, incentivizing on-the-job training participation for less-educated employees past their 50s could help to improve their employment outcomes near retirement.

References

Bassanini, A. (2006). Training, wages and employment security: an empirical analysis on european data. *Applied Economics Letters*, 13(8):523–527.

Bellmann, L., Dummert, S., and Leber, U. (2013). Betriebliche weiterbildung für ältere–eine längsschnittanalyse mit den daten des iab-betriebspanels. *Die Unternehmung*, 67(4):311–330.

Blossfeld, H.-P., Roßbach, H.-G., and von Maurice, J. H. (2011). Education as a lifelong process - the german national educational panel study (neps). *Zeitschrift für Erziehungswissenschaft: Sonderheft 14*.

Blundell, R., Costa Dias, M., Meghir, C., and Shaw, J. (2016a). Female labor supply, human capital, and welfare reform. *Econometrica*, 84(5):1705–1753.

Blundell, R., French, E., and Tetlow, G. (2016b). Retirement incentives and labor supply. In *Handbook of the economics of population aging*, volume 1, pages 457–566. Elsevier.

Blundell, R. W., Costa Dias, M., Goll, D., and Meghir, C. (2019). Wages, experience and training of women. *IFS Working Paper*, (W19/08).

Börsch-Supan, A., Bucher-Koenen, T., Coppola, M., and Lamla, B. (2015). Savings in times of demographic change: Lessons from the german experience. *Journal of Economic Surveys*, 29(4):807–829.

Börsch-Supan, A. and Ferrari, I. (2017). Old-age labor force participation in germany: What explains the trend reversal among older men? and what the steady increase among women? *National Bureau of Economic Research Working Paper*, (24044).

Cairo, I. and Cajner, T. (2018). Human capital and unemployment dynamics: Why more educated workers enjoy greater employment stability. *The Economic Journal*, 128(609):652–682.

Chetty, R. (2006). A new method of estimating risk aversion. *American Economic Review*, 96(5):1821–1834.

Conti, G. (2005). Training, productivity and wages in italy. *Labour economics*, 12(4):557–576.

Daniel, K. and Heywood, J. S. (2007). The determinants of hiring older workers: Uk evidence. *Labour Economics*, 14(1):35–51.

Dauth, C. (2020). Regional discontinuities and the effectiveness of further training subsidies for low-skilled employees. *ILR Review*, 73(5):1147–1184.

Deutsche Rentenversicherung (2020). Benefits. https://www.deutsche-rentenversicherung.de/DRV/EN/Leistungen/leistungen.html. (last access: July 9th 2020).

Ehlert, M. (2017). Who benefits from training courses in germany? monetary returns to non-formal further education on a segmented labour market. *European Sociological Review*, 33(3):436–448.

Etgeton, S. (2018). The effect of pension reforms on old-age income inequality. *Labour Economics*, 53:146–161.

Fan, X., Seshadri, A., and Taber, C. (2017). Understanding earnings, labor supply, and retirement decisions. *Michigan Retirement Research Center Working Paper*, (2017-367).

Fouarge, D., Schils, T., and De Grip, A. (2013). Why do low-educated workers invest less in further training? *Applied Economics*, 45(18):2587–2601.

Göbel, C. and Zwick, T. (2012). Age and productivity: sector differences. *De Economist*, 160(1):35–57.

Göbel, C. and Zwick, T. (2013). Are personnel measures effective in increasing productivity of old workers? *Labour Economics*, 22:80–93.

Gohl, N., Haan, P., Kurz, E., and Weinhardt, F. (2020). Working life and human capital investment: Causal evidence from pension reform. *IZA Discussion Paper*.

Görlitz, K. (2010). The effect of subsidizing continuous training investments—evidence from german establishment data. *Labour Economics*, 17(5):789–798.

Görlitz, K. (2011). Continuous training and wages: An empirical analysis using a comparison-group approach. *Economics of education review*, 30(4):691–701.

Gustman, A. L. and Steinmeier, T. L. (2005). The social security early entitlement age in a structural model of retirement and wealth. *Journal of public Economics*, 89(2-3):441–463.

Haan, P., Kemptner, D., and Prowse, V. L. (2017). Insurance, redistribution, and the inequality of lifetime income. *DIW Berlin Discussion Paper*, (1716).

Haan, P., Kemptner, D., and Prowse, V. L. (2018). Insurance, Redistribution, and the Inequality of Lifetime Income. *IZA Discussion Papers*, (11275).

Haan, P. and Prowse, V. (2014). Longevity, life-cycle behavior and pension reform. *Journal of Econometrics*, 178:582–601.

Henningsen, A. and Toomet, O. (2011). maxlik: A package for maximum likelihood estimation in r. *Computational Statistics*, 26(3):443–458.

Jürges, H. and Schneider, K. (2004). Dynamische lohneffekte beruflicher weiterbildung: Eine längsschnittanalyse mit den daten des soep. *MEA Workingpaper*.

Keane, M. P. and Wolpin, K. I. (1997). The career decisions of young men. *Journal of political Economy*, 105(3):473–522.

Kluve, J. (2010). The effectiveness of european active labor market programs. *Labour economics*, 17(6):904–918.

Kruppe, T. and Trepesch, M. (2017). Weiterbildungsbeteiligung in deutschland: Auswertungen mit den daten der erwachsenenbefragung des nationalen bildungspanels" bildung im erwachsenenalter und lebenslanges lernen". Technical report, IAB-Discussion Paper.

Kuruscu, B. (2006). Training and lifetime income. *American Economic Review*, 96(3):832–846.

Leuven, E. (2005). The economics of private sector training: a survey of the literature. *Journal of economic surveys*, 19(1):91–111.

Low, H., Meghir, C., and Pistaferri, L. (2010). Wage risk and employment risk over the life cycle. *American Economic Review*, 100(4):1432–67.

Picchio, M. and Van Ours, J. C. (2013). Retaining through training even for older workers. *Economics of Education Review*, 32:29–48.

Pischke, J.-S. (2001). Continuous training in germany. *Journal of population economics*, 14(3):523–548.

Sanders, J., Oomens, S., Blonk, R. W., and Hazelzet, A. (2011). Explaining lower educated workers' training intentions. *Journal of Workplace Learning*.

Staubli, S. and Zweimüller, J. (2013). Does raising the early retirement age increase employment of older workers? *Journal of public economics*, 108:17–32.

Van den Berg, G. J., Dauth, C., Homrighausen, P., and Stephan, G. (2019). Informing employees in small and medium sized firms about training: Results of a randomized field experiment. *IAB Discussion Paper 22*.

Wakker, P. P. (2008). Explaining the characteristics of the power (crra) utility family. *Health economics*, 17(12):1329–1344.

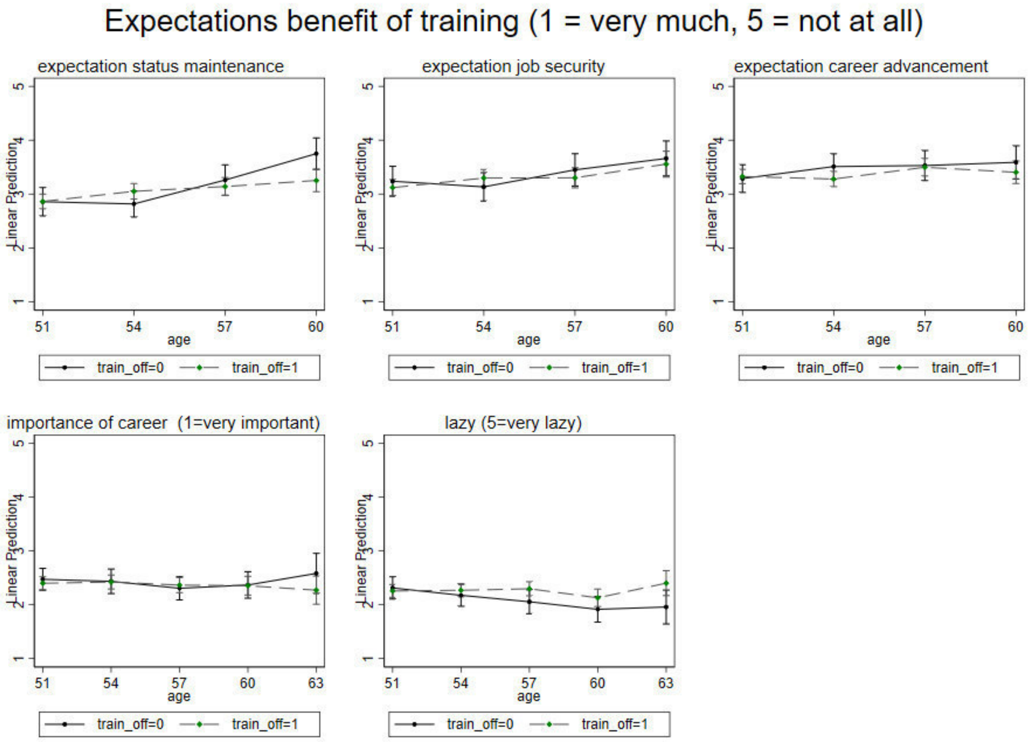
Zwick, T. (2002). Continuous training and firm productivity in germany. Technical report, ZEW Discussion Papers.

Zwick, T. (2011). Why training older employees is less effective. *ZEW-Centre for European Economic Research Discussion Paper*, (11-046).

This paper uses data from the National Educational Panel Study (NEPS): Starting Cohort Adults, doi:10.5157/NEPS:SC6:11.0.0. From 2008 to 2019, NEPS data was collected as part of the Framework Program for the Promotion of Empirical Educational Research funded by the German Federal Ministry of Education and Research (BMBF). As of 2014, NEPS is carried out by the Leibniz Institute for Educational Trajectories (LIfBi) at the University of Bamberg in cooperation with a nationwide network.

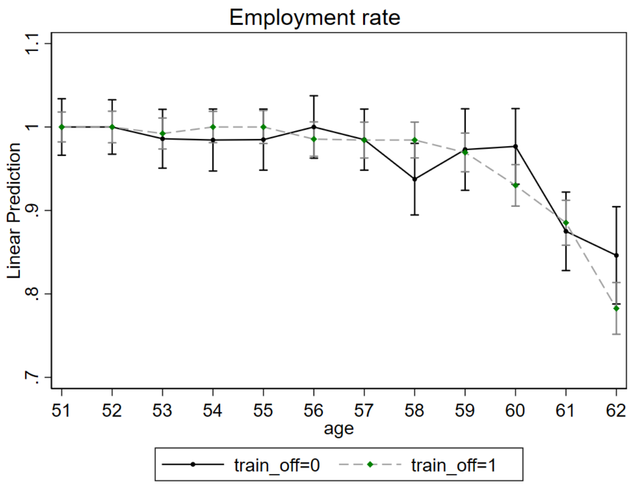
A Descriptives

Figure 19: Expectations and ambitions by training offers



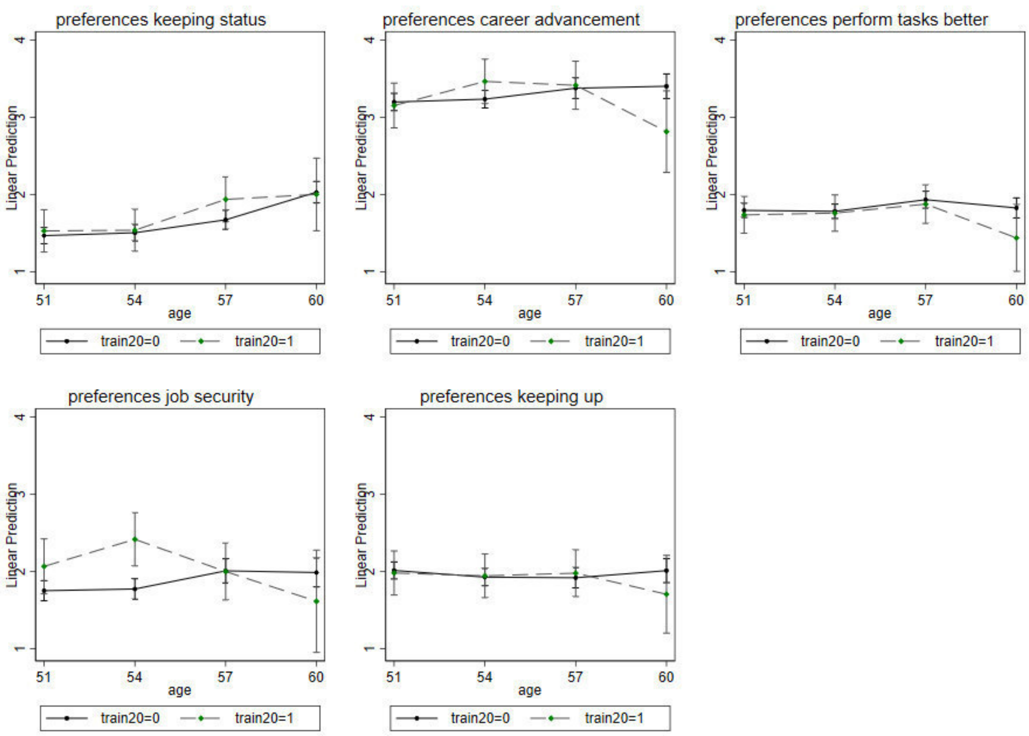
Source: NEPS; own calculations bases on estimation sample. Male low educated employees only. Top row expectations; bottom row ambitions.

Figure 20: Employment by training offers



Source: NEPS; own calculations bases on estimation sample. Male employees, low educated only.

Figure 21: Career ambitions by training participation



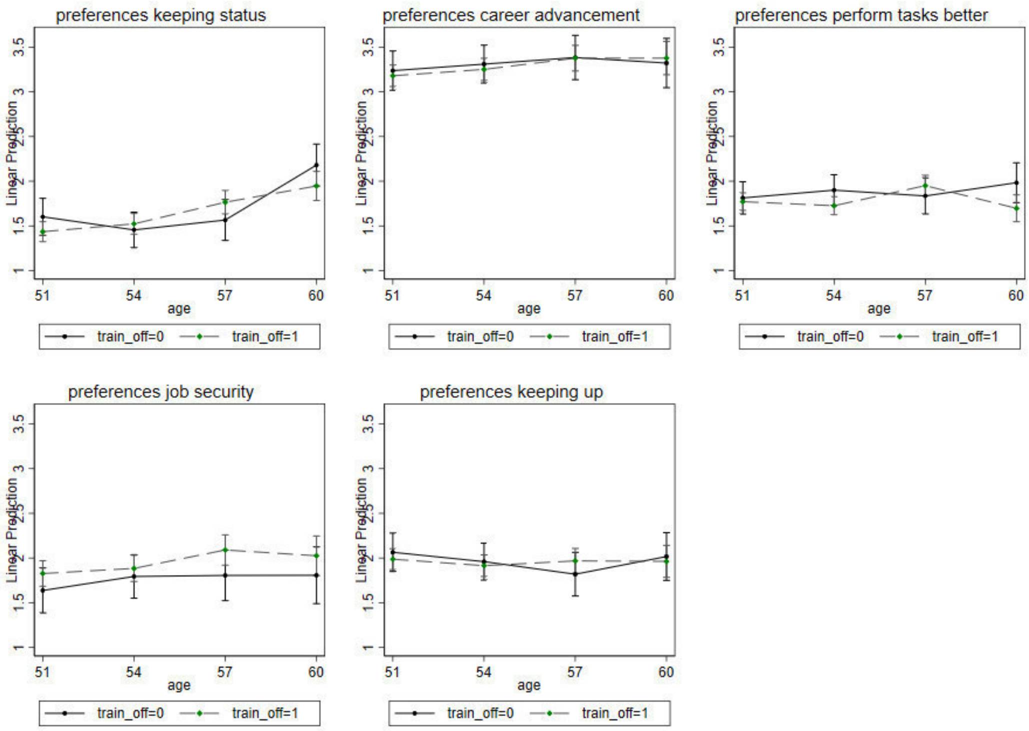
Notes: Scale ranges from 1 ="very important" to 5 ="very unimportant". Low educated male employees only. Training participation defined as having participated at least in 20hrs of training. Age is grouped into three year intervals, i.e. the data point at age 51 refers to ages 50-52. *Source:* NEPS; own calculations.

Table 4: Individual characteristics by training offers

Variable	**No training offer**	**Training offer**
Ambitions		
Importance of status maintenance	1.994	1.971
	(1.303)	(1.292)
Importance career advancement	3.528	3.331
	(1.122)	(1.105)
Importance perform tasks better	1.977	1.891
	(1.018)	(0.883)
Importance job security	2.038	2.087
	(1.312)	(1.341)
Importance of keeping up	2.109	2.038
with colleges	(1.177)	(1.065)
Attitudes		
Lazy	2.132	2.248
	(1.072)	(1.131)
Importance of career	2.593	2.583
	(1.091)	(1.105)
Wages		
Gross wage	2990.9	3797.9
	(1235.7)	(1502.8)
Net wage	2034.4	2529.3
	(1035.7)	(949.1)

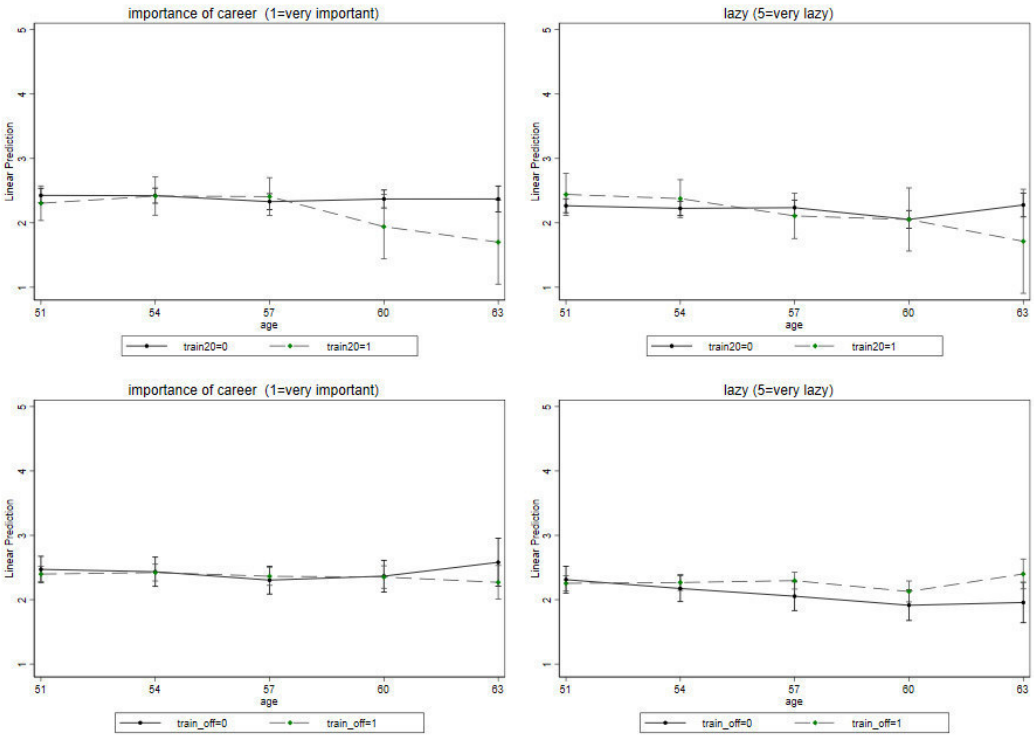
Means, sd in parentheses. Ambitions: 1 = very important, 5= very unimportant. Lazyness: 1= not lazy at all, 5 very lazy. For a brake down by age see Figures 22 and 23 *Source*: NEPS data, low educated male employees in full-time employment only.

Figure 22: Career Ambitions by training offers (TO)



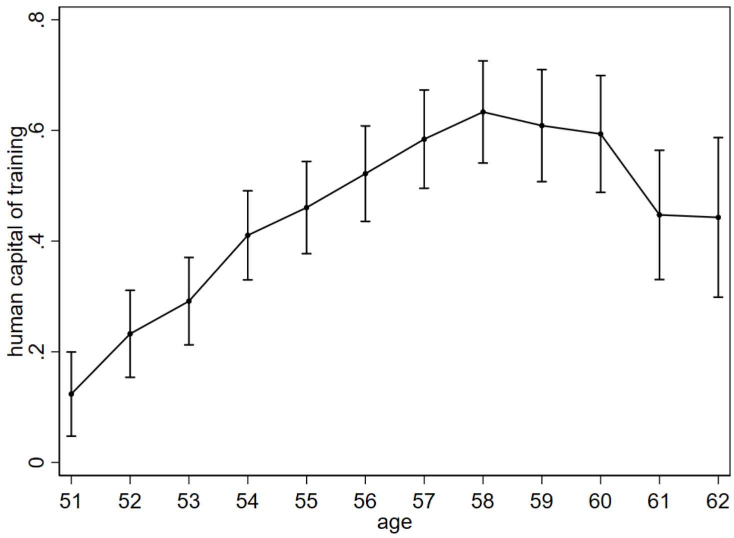
Notes: Scale ranges from 1 ="very important" to 5 ="very unimportant". Low educated male employees only. *Source:* NEPS; own calculations.

Figure 23: Ambitions by training offers and training participation



Source: NEPS; own calculations bases on estimation sample. Male low educated employees only. Top row training participation; bottom row training offers.

Figure 24: Pre-choice human capital of training by age



Source: NEPS; own calculations base. Human capital of training as defined in section 3.3.

B Estimation

I use a multinomial-logit model to approximate the choices. First, the dynamic programming problem is solved using backward induction. I use linear interpolation for expected value functions of periods that are more than two periods ahead from the decision period. Interpolation is used for the state variables of wages and human capital of training only.[46]

The log-likelihood function is defined as follows:

$$L = \sum_{i=1}^{N} log\big[\prod_{t=t0}^{\bar{t}} P(d_{it}|\theta, s_{it}) f(w_{it}^{obs}|\theta, s_{it})\big] \tag{6}$$

with parameter vector θ and state-variables $s_{it} = \{train_{it}, TO_{it}, wage_{it}, JS_{it}, age_{it}\}$

Individual likelihood contributions for parameters θ and state variables s_{it}:

$$Pr(d = \ell t|s_{it}, \theta) = P(D_{it} = \{u, \ell, \ell t\}|s_{it}, \theta)\frac{exp(\bar{V}_{\ell t}(s_{it}, \theta))}{\sum_{c\in\{\ell t, \ell, u\}} exp(\bar{V}_c(s_{it}, \theta))} f(wage^{obs}|s_{it}, \theta)$$

$$\begin{aligned}
&Pr(d = \ell|s_{it}, \theta) = \\
&\big[Pr(D_{it} = \{u, \ell, \ell t\}|s_{it}, \theta)\frac{exp(\bar{V}_{\ell}(s_{it}, \theta))}{\sum_{c\in\{\ell t, \ell, u\}} exp(\bar{V}_c(s_{it}, \theta))} \\
&+ Pr(D_{it} = \{u, \ell\}|s_{it}, \theta)\frac{exp(\bar{V}_{\ell}(s_{it,\theta}))}{\sum_{c\in\{\ell, u\}} exp(\bar{V}_c(s_{it}, \theta))}\big] f(wage^{obs}|s_{it}, \theta)
\end{aligned}$$

$$\begin{aligned}
&Pr(d = u|s_{it}, \theta) = \\
&Pr(D_{it} = \{u, \ell, \ell t\}|s_{it}, \theta)\frac{exp(\bar{V}_u(s_{it}, \theta))}{\sum_{c\in\{\ell t, \ell, u\}} exp(\bar{V}_c(s_{it}, \theta))} \\
&+ Pr(D_{it} = \{u, \ell\}|s_{it}, \theta)\frac{exp(\bar{V}_u(s_{it}, \theta))}{\sum_{c\in\{\ell, u\}} exp(\bar{V}_c(s_{it}, \theta))} + Pr(D_{it} = \{u\}|s_{it}, \theta)
\end{aligned}$$

where $\bar{V}()$ is the systematic component of the value function without the preference shock. I estimate the parameters of the model with the maximum-likelihood method using nonlinear minimization with a Newton-type method in R (nlm). Several starting values were tested to ensure the parameters represent the global optimum. The stan-

[46]The grid for net-monthly wages is $\{1500, 2000, 2500, 3000, 3500, 4200, 5000, 10000\}$, for training $\{0, 0.8, 1, 1.8, 3\}$.

dard errors at the optimum are derived using the information matrix equality (BHHH method) with numerical gradient (see Henningsen and Toomet, 2011).