Participation in training at older ages

A European perspective on path dependency in life course trajectories

Turek, K.; Henkens, K.

DOI
10.1016/j.alcr.2020.100396

Publication date
2021

Document Version
Final published version

Published in
Advances in Life Course Research

Advances in Life Course Research 48 (2021) 100396

Contents lists available at ScienceDirect

Advances in Life Course Research

journal homepage: www.elsevier.com/locate/alcr

Participation in training at older ages: A European perspective on path dependency in life course trajectories

Konrad Turek [a,b,c,*], Kène Henkens [a,b,c,d]

[a] Netherlands Interdisciplinary Demographic Institute (NIDI-KNAW), Lange Houtstraat 19, NL-2511 CV The Hague, The Netherlands
[b] University of Groningen, PO Box 72, 9700 AB Groningen, The Netherlands
[c] University of Amsterdam, Spui 21, 1012 WX Amsterdam The Netherlands
[d] University Medical Center Groningen (UMCG-RUG), PO box 30.001, 9700 RB Groningen, The Netherlands

ARTICLE INFO

Keywords:
Lifelong learning
Training
Path dependency
Older workers
Population ageing
SHARE

ABSTRACT

Investments in lifelong learning may create unsatisfactory results, and this could potentially contribute to the reproduction of inequalities. We argue that the process is related to the accumulation of opportunities and barriers for participation in training, which can lock individuals in disadvantageous path-dependent trajectories. We take a longitudinal approach to analyse whether participation in training in older age is path-dependent, and whether this path dependency is related to institutional contexts. Using data from the Survey of Health, Ageing, and Retirement in Europe (SHARE), we trace individual training trajectories in the population aged 50+ in twelve European countries between 2010 and 2015 (27 370 respondents). Hierarchical Bayesian logit models serve to assess the probability of training during the sixth wave, with a lagged dependent variable as a predictor. Results suggest that training participation is path-dependent and participation in training is limited for people who have not trained previously. It is also related to macrostructural context: path dependency is lower in countries with stronger knowledge economies, stronger emphasis on education, and a proactive ageing climate. Recognising path dependency can help to improve access to training and design policies that address problems of cohesion, active ageing and adult learning.

1. Introduction

Continuous acquisition and adjustment of skills are deemed necessary to extend working lives and increase employability at older ages. Lifelong learning (LLL) is important to counteract skill obsolescence, stimulate active ageing, enhance social capital, and empower political inclusion (Cedefop, 2012; Evans, Schoon, & Weale, 2013; Groot & Van den Brink, 2000; Picchio & van Ours, 2013). From a broader policy perspective, LLL might also reduce socioeconomic inequalities and disparities in health, or quality of life (EC, 2010; Green, 2006). Building on these arguments, nearly all strategic policy documents in the European Union in the past decades refer to LLL as a priority (Holford & Mleczko, 2013). Despite large budgets, investments in LLL in the European Cohesion Policy 2007–2013 were inefficient and did not reach expected targets, especially for older age groups. Instead of social and economic cohesion, they often contributed to existing disparities through accumulation of advantages and disadvantages based on unequal access to education and selective approaches to training in companies (Cedefop, 2015; EC, 2010, 2013; Formosa, 2012). At the individual level, accumulation of opportunities and barriers for training results in path dependency – a process in which the likelihood of participation depends on previous participation. The "shadow of the past" can lock individuals in progressing chains of risks and disadvantages, limiting their potential and opportunities for switching into more beneficial pathways (Ben-Schlomo & Kuh, 2002; Bernardi, Huinink, & Settersten, 2019; Kuh, Ben-Shlomo, Lynch, Hallqvist, & Power, 2003). Previous training can improve abilities and motivation for further learning; it may also extend access to educational opportunities at work (Froehlich, Beausaert, & Segers, 2015; Hansson, 2008; Lazazzara, Karpinska, & Henkens, 2013; Pak, Kooij, De Lange, & Van Veldhoven, 2018; Van der Heijden, Gorgievski, & De Lange, 2016). As such, path dependency is an important causal pathway in analysing lifelong learning.

In this article, we ask whether participation in training in older age is path-dependent, and whether this path dependency is related to institutional contexts. To assess path dependency, we analyse how previous attendance in training affects the probability of further attendance. In

* Corresponding author at: NIDI, Lange Houtstraat 19, NL-2511 CV The Hague, The Netherlands.
E-mail addresses: turek@nidi.nl (K. Turek), henkens@nidi.nl (K. Henkens).

https://doi.org/10.1016/j.alcr.2020.100396
Received 8 July 2019; Received in revised form 21 September 2020; Accepted 4 December 2020
Available online 23 December 2020

particular, we focus on people who enter training for the first time for at least a few years. Low participation in LLL is shown to derive primarily from lack of possibilities for training rather than lack of personal motivation to participate (Kilpi-Jakonen, Vono de Vilhena, & Blossfeld, 2015; Leuven & Oosterbeek, 1999; Picchio & van Ours, 2013; Roosmaa & Saar, 2010; Rubenson & Desjardins, 2009). Opportunities to learn are distributed unevenly, but institutional arrangements and policies may help to overcome external and individual barriers to participation, creating more equitable conditions. Reduction of social inequalities through LLL policies is impossible unless LLL is more accessible to groups that are less likely to participate. Therefore, we take a comparative approach to assess disparities in training participation between institutional contexts and analyse the factors that might explain them. Using data from the Survey of Health, Ageing, and Retirement in Europe (SHARE), we observe individuals in twelve countries over five years. A multilevel data structure allowed us to study the role of macro-level predictors related to the economic, welfare and sociocultural contexts.

With this study, we contribute to the literature in three ways. First, by using a longitudinal perspective on training participation, we add to growing literature on the analysis of life courses (Piccarreta & Studer, 2018). Extant research commonly focuses on cross-sectional views, analysing supply and demand factors that drive educational attainment in old age (Roosmaa & Saar, 2010; Saar & Räis, 2017). This perspective is unable to show temporal dynamics in participation at the individual level, contrary to panel data, which allows tracing individual learning trajectories and viewing them in terms of continuity. To our knowledge, this study is first to empirically assess path dependency in training participation.

Second, this article adds to evidence regarding LLL at older ages by combining a life course approach with an international comparative analysis. Although cross-national life course perspective has been considered before, for example by Blossfeld, Kilpi-Jakonen, De Vilhena, and Buchholz (2014), the analyses have not linked the micro-level pathways and institutional contexts directly. We test whether the variability in training pathways across countries relates to macro-characteristics, such as development of knowledge economy, generosity of the welfare state, public support for education, and active ageing culture. The results suggest that some contextual factors can foster the potential for breaking down path dependency.

Third, the results have implications for policies that address problems of cohesion, active ageing, and improvement of adult-learning attendance. We argue that the life course perspective is necessary to recognise path dependency and address measures to improve opportunities for training. Limited access to training for disadvantaged groups, such as older and less-skilled people who are overlooked in market-based systems, might further drive accumulation of inequalities. Path dependency hampers policies that address cohesion and potentially lead to their failure. Although emphasised in some studies (Kilpi-Jakonen et al. 2015; Leuven & Oosterbeek, 1999; Roosmaa & Saar, 2010; Rubenson & Desjardins, 2009), these arguments do not have sufficient empirical evidence.

This article is structured as follows: first, we discuss the mechanisms which shape path dependency in training at older ages and develop our hypotheses. Second, in the methods section, we describe the data and statistical approach. In the third part, we present results of the analysis. In the last section, we discuss the results and main conclusions.

2. Path dependency in training participation: possible underlying mechanisms

2.1. Life course perspective and path dependency

From a life course perspective, participation in training can be considered from a longitudinal perspective as a trajectory of behaviours, where earlier decisions and previous biography shape an individual's situation later in life. As such, it is subject to life course mechanisms of accumulation of advantages and disadvantages (Crystal & Shea, 1990; O'Rand, 1996) or accumulation of inequalities (Ferraro, Schippee, & Schafer, 2009), which drive progressive processes of differentiation between individuals. The central assumption is that early advantages reduce the risk of exposure to adverse transitions and increase the access to beneficial opportunity structures, whereas early disadvantages may increase the likelihood of persistent disadvantage (Crystal & Shea, 1990). This is directly related to the idea of path dependency, i.e. a process in which the probability of a biographical event in time t depends on the longer life history reaching back at least several time units (Bernardi et al., 2019; Kuh et al., 2003). In this perspective, path dependency in training can be seen as resulting from the accumulation of opportunities and barriers for participation in educational activities.

Following Ben-Schlomo and Kuh (2002) and Kuh et al. (2003), we can refer to two life course mechanisms that may lead to path dependency. The first one is *accumulation with risk clustering* in which exposures are clustered due to a common underlying factor. In case of training, education and learning abilities – if assumed to be time-constant – affect the likelihood of participation similarly at time t and $t + 1$, and result in accumulation at the individual level. The second mechanism is a *chain of risks with additive effect* where exposures (beneficial or adverse) are linked to each other in a sequence, and one tends to lead to another. Both of them are relevant for this study. Participation in training can increase skills that improve learning abilities, awareness of benefits of learning, and motivation for further participation (Froehlich et al., 2015; Hansson, 2008; Pak et al., 2018; Zwick, 2012). Previous training attendance might also ease further access by signalling a worker's potential for development to employers and affecting their training-related decisions (Lazazzara et al., 2013; Spence, 1973; Van der Heijden et al., 2016). As a result, individuals who do not train – for whatever individual or external reasons – can be locked in path-dependent trajectories, where chances for leaving the disadvantaged position decrease with time. On the contrary, individuals who participate in training may gain easier access to opportunities and increase individual resources for further participation. The life course perspective suggests that such long-lasting processes of differentiation stimulate development of intracohort inequalities at older ages.

The idea of accumulation found a fruitful ground in education research, where it has been extensively applied to study the outcomes of early-life educational disadvantages (Kerckhoff, 1993; Kerckhoff & Glennie, 1999; Pallas, Natriello, & McDill, 1989). It was also adapted to the studies of adult education and LLL (Blossfeld et al., 2014; Bukodi, 2016; Elman & O'Rand, 2004; Kilpi-Jakonen et al., 2015), though shortages in appropriate data sources limit empirical longitudinal evidence. Life course studies commonly characterise LLL as a tool for stratification that can stimulate the growth or decrease of inequalities. In one of only a few longitudinal studies, Bukodi (2016) provides evidence that LLL increases inequalities over the life course. The author traces individuals from the United Kingdom from their teenage years to age 38, finding that training is more beneficial to individuals with high initial socioeconomic positions than for those less-advantaged. Blossfeld et al. (2014) and Kilpi-Jakonen et al. (2015) use comparative evidence, suggesting that participation in training in adult age depends on socioeconomic positions, which reflects the accumulation hypothesis.

Following the life course accumulation theory and general education research, we formulate the first hypothesis that:

H1. Training participation in older age is path-dependent: the probability of training is greater for individuals who participated previously.

2.2. Institutional contexts

The next question is how and why path dependency in training might differ across countries. Accumulation mechanisms operate at the individual level (George, 1993), but they are affected by institutional, economic and socio-cultural contexts (Dannefer & Uhlenberg, 1999; Ferraro et al., 2009; Kohli, 2007). This social embeddedness of life courses (Elder, 1994) points to the structural origins of later-life inequalities. In Europe, training at older ages is most common in the Nordic states, the United Kingdom, the Netherlands, and Switzerland, and low participation occurs in Southern and Central Europe and the Baltic states (Beblavy, Thum, & Potjagailo, 2014; Brunello et al., 2007; Dämmrich., De Vilhena, & Reichart, 2014). Research suggests a negative correlation between general participation and inequality of participation; differences between low-and high-skilled adults are greater in countries with low general participation (southern Europe and the Baltic countries), and lower in countries with high participation (Nordic countries) (Roosmaa & Saar, 2010). The variation between countries is usually explained by referring to economic situation, welfare regimes, education system, and sociocultural context (Boeren, Nicaise, & Baert, 2010).

Strong and competitive economies show, in general, higher demand for qualifications and knowledge (Beblavy et al., 2014; Dämmrich. et al., 2014; Saar & Räis, 2017). The need is driven by innovations, knowledge-based development, and competition between companies (Coulombe & Tremblay, 2007). For example, Dämmrich. et al. (2014) show that countries with higher public expenditures on research and development, such as Finland or Sweden, report higher participation in adult learning. Saar and Räis (2017) relate higher participation rates in countries such as Denmark to their higher innovation performance. Investments in human capital also tend to be more intensive in prosperous economic conditions, while in a recessive context, the primary concern of organisation is a short-term stabilisation and reduction of unnecessary expenditures, which usually include training (Brunello et al., 2007; EC, 2013; Munnell & Rutledge, 2013). We expect then the opportunities to participate in training, and as a consequence the path dependency to be affected by the economic factors, such as country's degree of innovation, investments in research and development, and the overall development of knowledge economy:

H2. Path dependency in training participation is lower in countries with more developed knowledge economies.

Institutional arrangements, welfare regimes, and public policy can improve persons' capability of overcoming a variety of barriers to training participation. By supporting disadvantaged groups, they create fairer conditions and reshape unequal distributions of opportunities to participation (Roosmaa & Saar, 2010; Rubenson & Desjardins, 2009). An institutional context may also influence employers' investment in staff training. Elements such as retirement regulations, taxes, and incentives affect calculation of costs, benefits, and the expected period of return on investment (Lazazzara et al., 2013). Dämmrich. et al. (2014) find higher participation in adult learning in countries with higher public expenditures on social protection. We assume that more extensive welfare states – defined in terms of generosity of social expenditure (Razin, Sadka, & Swagel, 2002) – are more likely to reduce intracohort inequalities in old-age training (Beblavy et al., 2014; Green, 2006; Rubenson, 2006). For example, liberal and market-oriented regimes, such as in the United States and the United Kingdom, are less oriented at social cohesion and provide less support for educational activity. The main determinant of training is then the labour market demand for human capital, what results in lower and more unequally distributed LLL participation (Green, 2006, 2011; Riddell & Weedon, 2012; Verdier et al., 2013). In contrast, the Scandinavian welfare model with a high level of social expenditures and more active labour market policies provides stronger support for social integration, individual aspirations, and empowerment (Riddell & Weedon, 2012). Here, participation in adult education is less market-driven and relies more on public support, and consequently, rates are higher and inequalities lower (Verdier et al., 2013). To account for the welfare state differences, we consider the size of countries' social expenditure as an indicator of the welfare generosity and supportive power (Razin et al., 2002):

H3. Path dependency in training participation is lower in generous and

Table 1
Descriptive statistics of control variables (for wave 6).

	Austria	Germany	Sweden	Spain	Italy	France	Denmark	Switzerland	Belgium	Czech Rep.	Slovenia	Estonia	Total	Average training rate (%)
Male (%)	42.2	47.1	43.9	43.9	44.4	42.7	46	45.4	44.2	39.5	41.8	37.7	42.5	10.2
Female (%)	57.8	52.9	56.1	56.1	55.6	57.3	54.0	54.6	55.8	60.5	58.2	62.3	57.5	12.8
Age (at wave 4)														
50−60	34.0	28.4	16.9	30.6	28.5	38.0	43.0	38.1	40.9	33.3	38.5	34.1	34.6	19.4
61−70	37.0	40.9	44.6	31.4	40.0	31.2	33.6	34.8	31.4	41.6	32.8	33.2	35.4	10.2
71−80	22.1	25.5	28.6	27.6	25.6	21.9	17.5	20.8	19.9	21.1	21.7	26.7	23.1	4.9
81+	6.9	5.2	9.9	10.4	5.9	8.9	5.8	6.3	7.7	4.1	7.1	6.1	6.9	1.7
Education (%)														
Primary	24.4	11.1	41.5	83.1	70.8	41.8	15.6	19.2	40.2	45.7	33.6	29.1	39.1	3.7
Secondary	49.7	54.2	29.0	9.0	23.3	35.2	40.7	64.6	26.9	41.3	48.6	48.9	39.3	11.8
Tertiary	25.9	34.7	29.6	7.9	5.9	22.9	43.7	16.2	32.9	13.0	17.8	22.1	21.6	25.6
Employment pattern (%)														
Not working	74.1	67.9	67.0	69.1	73.4	68.5	46.0	51.1	64.1	73.8	78.1	56.5	65.5	6.6
Employed	12.7	14.2	14.5	10.3	11.7	15.5	31.2	28.3	17.8	12.0	10.8	20.4	16.7	29.4
Deactivation	8.7	11.4	13.2	7.5	5.7	9.4	11.8	11.2	8.5	8.5	5.7	11.1	9.3	14.6
Reactivation	0.6	0.8	0.7	1.1	1.0	0.8	1.3	2.4	0.5	1.1	0.4	2.4	1.2	16.7
Other	3.9	5.8	4.5	12.1	8.3	5.9	9.7	7.1	9.1	4.6	5.0	9.7	7.4	11.5
Total N	2829	968	1261	2398	2172	2835	1645	2459	3146	3224	1606	4356	28,899	11.7

Note: Statistics for the analytical sample. Unweighted.
Source: SHARE data 2010–2015 (own estimates).

supportive welfare states.

The third country-specific element that might influence training participation at older ages is the education system (Boeren et al., 2010; Rubenson, 2006). Desjardins (2013) connects adult education with the country's emphasis on education. He argues that countries with strong state support for learning systems, such as Finland, Sweden and Denmark, reveal comparatively better coordination, lower barriers to participation and more opportunities for adult learning. Dämmrich. et al. (2014) find a positive effect of public expenditures in education on non-formal employer-sponsored learning activities. We expect then that country's level of support for education – defined in terms of public expenditure on education – affects training participation:

H4. Path dependency in training participation is lower in countries with stronger public support for education.

From the sociocultural perspective, the level of training participation is often related to the culture of learning, culture of age and age-related norms (Boeren et al., 2010). These factors may influence attitudes and motivations toward learning at older ages by shaping a more or less active image of older workers. By the same token, old-age stereotypes regarding trainability can create barriers to education, especially when they stimulate employers' age-discriminatory decisions regarding provision of organisation-sponsored training (Harper, Khan, Saxena, & Leeson, 2006; Posthuma & Campion, 2009; Turek & Henkens, 2019; Van Dalen, Henkens, & Schippers, 2009). Continuous learning also constitutes a core element of the active ageing culture which focuses on improving quality of life, health, and wellbeing through social and intellectual activity (Cedefop, 2012; Narushima, Liu, & Diestelkamp, 2018). From the sociocultural perspective, Nordic countries foster participation in training through their active attitude towards ageing, emphasis on lifelong learning, and relatively low level of ageism. Low-training countries, on the contrary, such as Southern European countries (e.g. Italy, Portugal and Greece) and also some Eastern European (e.g. Poland and Hungary) exhibit features of early exit cultures with a more passive image of old age (Boeren et al., 2010; Formosa, 2012; Rubenson & Desjardins, 2009; Rubenson, 2006). With the last hypothesis, we expect to observe country differences in training participation patterns due to sociocultural factors, namely the culture of active ageing:

H5. Path dependency in training participation is lower in countries with a more proactive culture of ageing.

3. Methods

3.1. Data

Data came from the Survey of Health, Ageing, and Retirement in Europe (SHARE, Release 6.0). SHARE is a cross-national, longitudinal research program that collects data on nationally representative samples of adults aged 50 and older from 27 countries (Börsch-Supan, Brandt, Hunkler, Kneip, & Korbmacher, 2013). The present study is restricted to 12 countries that participated during waves 4 (2010/11), 5 (2013), and 6 (2015). We include only those respondents who were 50 years or older during wave 4 and who were interviewed during all three waves, resulting in a sample of 28 899 individuals (Table 1). There were no missing values for the control variables. In the final analysis, we included cases with full information on training participation in all waves (95 % of the sample = 27 370 observations).

3.2. Dependent variable

For a dependent variable, we use a question that indicated attendance in educational or training courses during the last 12 months (0 = no, 1 = yes). The question was asked in the same form during waves 4 through 6: "Which of the activities listed on this card - if any - have you done in the last twelve months?", with a multiple selection item list which included "Attended an educational or training course" (a different question was used during waves 1 and 2, and wave 3 did not include a similar item). The average training participation rate in wave 6, in the analytical sample, was 11.7 %.

3.3. Control variables

Control variables (Table 1) include gender, age (included as a continuous variable)[1], and education (grouped into three categories according to ISCED-97 levels: 0–2=primary, 3–4=secondary, and 5–6=tertiary).. We also include employment pattern which was created based on employment during three waves and included five categories—not working (i.e., unemployed, retired, or inactive during all waves), continuously employed (i.e., employed or self-employed during each wave), deactivation (i.e., first working then retired/unemployed/inactive), reactivation (i.e., first retired/unemployed/inactive and then working), and other.

3.4. Country-level predictors

Based on the previous literature (see: Dämmrich. et al., 2014; Riddell & Weedon, 2012; Saar & Räis, 2017), we selected four macro-predictors to test hypotheses related to the macro-context (Table 3). Degree of knowledge economy (H2) is represented by the Knowledge Economy Index (KEI), which indicates a country's overall degree of development regarding a knowledge economy. KEI was computed based on the World Bank's method (Chen & Dahlman, 2006), which uses the degree of economic and institutional incentives for efficient use of human capital, education and human resources, innovation potential, and the quality of information and communication technologies infrastructure. The generosity of the welfare state (H3) is indicated by total social welfare expenditures as a percentage of GDP (SWE), comprising total social spending toward old age, survivors, incapacity-related benefits, health, family, active labour market programmes, unemployment, housing, and other social policies. Public support for education (H4) is represented by government expenditures on education as a percentage of GDP (EDU). As a measure of the culture of active ageing (H5), we use the Active Ageing Index (AAI), which measures the degree to which older people live independent lives and participate in paid employment and social activities, and their capacity to remain active into old age. It is calculated using 22 indicators grouped into four domains—employment, participation in society, independent living, and capacity for active ageing (AAI, 2013). Detailed statistics for countries appear in Supplementary Table A1. Values for the macro-predictors are provided for years 2014–2015, to which the training question refers, or the closes possible time. Values of these indicators do not vary significantly over time, thus using alternative reference years provides similar results.[2]

KEI correlates strongly with EDU and AAI at almost 0.8 (Supplementary Table A2). The correlation between EDU and AAI had a medium-level strength of 0.55, and correlations between SWE and other predictors were weak or close to zero. For comparisons between models, all macro-predictors were z-standardized (mean = 0, SD = 1).

[1] We do not impose any upper age limit since we are interested in training in old age in general, both in the working-age and beyond it, although the training rate above the age of 70 drops drastically

[2] Following other studies (Dämmrich et al., 2014; Green & Janmaat, 2011; Riddell & Weedon, 2012; Saar & Räis, 2017), we compared alternative macro-predictors, such as employment rate of people aged 50–74, socioeconomic inequalities (Gini coefficient), expenditures on education as a portion of public expenditures, GDP per capita, and GDP growth 2010–2015 (Supplementary Table A5). KEI, SWE, EDU, and AAI were chosen because they were reliable, were grounded in the theoretical framework, and provided the clearest interpretation.

3.5. Analytical approach

We estimate the probability of participation in training during wave 6 based on information regarding training trajectory during earlier waves; a lagged dependent variable (LDV) approach.[3] To account for the hierarchical structure of the data, we build a multilevel (mixed-effects) model with individuals clustered in countries, as presented in Eq. 1. It includes random intercept to account for differences in the average training participation between countries. The LDV measures path dependency in training participation. Lags 1 (i.e., participation in training one wave before) and 2 (i.e., two waves before) of the dependent variable are included as predictors, with a random slope allowing for the effect to vary across countries To test hypotheses H2–H5, we fit models with country-level predictors included one at a time. These models contain an interaction between macro-predictors and the LDV. The interaction is necessary to verify the relationship between the country-level predictors and path dependency, as hypothesised.

$$
\begin{aligned}
\operatorname{logit}\{\Pr(\text{train}_{ij} = 1 | x_{ij}, u_{0j}, u_{1j})\} &= \beta_0 + \beta_1 \text{lag}_{\text{train}\,ij} + \beta_2 \text{macro}_j + \beta_3 \text{lag}_{\text{train}\,ij} \times \text{macro}_j \\
&\quad + \beta_c \text{control}_{ij} + \left(u_{0j} + u_{1j}\text{lag}_{\text{train}\,j}\right), \text{ for } i \\
&= 1, \ldots, n;\ j = 1, \ldots, k \qquad (1)
\end{aligned}
$$

$$
\begin{bmatrix} u_{0j} \\ u_{1j} \end{bmatrix} \sim \mathrm{N}(0, \Omega_u), \ \Omega_u = \begin{bmatrix} \sigma^2_{u0} & \\ \sigma_{u01} & \sigma^2_{u1} \end{bmatrix}
$$

We apply a Bayesian estimation and fit a mixed-effects Hierarchical Bayes Logit Model (HBLM). Bayesian approach is superior versus the frequentist approaches in many aspects, including reliability and clarity of the results, validation of the model and flexibility during post-estimation (McElreath, 2016). There are two major statistical reasons for which we apply it in this study. First, a sample of $k = 12$ countries is too small for modelling based on maximum likelihood (ML) and might lead to biased results (Bryan & Jenkins, 2016; Maas & Hox, 2005). In such cases, HBLM is recommended (Gelman & Hill, 2007)[4] . Second, the model has a complex design with a multilevel structure, LDV with a random slope, and cross-level interaction. Contrary to HBLM, ML cannot handle this degree of complexity.

Estimation was conducted using MCMC sampling with the Hamiltonian Monte Carlo algorithm (4 chains with 4 000 iterations, 1 000 for warmup, and total post-warmup sample = 12 000) using the brms package (Bürkner, 2017) based on Stan computational framework (http://mc-stan.org/) in R ver. 3.5.1. All models converged with large effective posteriori sample sizes. Trace plots were inspected visually showing no signs of nonconvergence, and the $\widehat{R}$ values equal to 1 suggesting proper chains convergence (Bürkner, 2017).[5],[6]

Table 2
Categories of the lagged dependent variable.

wave 4	wave 5	LDV category
No	No	LagA
Yes	No	LagB
No	Yes	LagC
Yes	Yes	LagD

In the results section, we present the mean and 95 % credible intervals (CI) of the posterior distribution (a range in which the true value of a parameter lies with a 95 % probability). For interpretation and visualisation of the effects of predictors, especially in the case of interactions, we use predicted mean values of the response distribution (i.e., predicted probabilities). Assessment of model fit was conducted using WAIC and a median of Bayesian R^2 (Gelman, Goodrich, Gabry, & Vehtari, 2019; Vehtari, Gelman, & Gabry, 2016). Lower values of WAIC indicate better fit. Bayesian R^2 is a posteriori ratio of predicted variance and variance plus error variance, showing a data-based estimate of the proportion of variance explained by new data. To verify research hypotheses, we test the existence of cross-level interaction effects. We do it by comparing slopes of the interaction term using Cohen's (1988) measure of effect size and a corresponding Cohen's U_3 nonoverlap measure. The effect size, computed as $(\mu_1 - \mu_2)/\sqrt{(\sigma_1^2 + \sigma_2^2)/2}$ and adapted to the Bayesian framework (Kruschke, 2012), measures the difference between mean values of two coefficients relative to the pooled variability of these coefficients. A greater value indicates a greater effect, with values higher than 1.6 corresponding to a nonoverlap value higher than 0.95. The nonoverlap measure informs about credibility of a difference between slopes (computed based on *a posteriori* samples of coefficients) as a share of scenarios in which slope A is larger than slope B.

Given three waves, there are four possible combinations of LDV (Table 2). The point of interest is the category LagA, which shows the probability of training during wave 6 after non-participation and indicates the openness of training systems for new participants. With random slopes we allow coefficients to vary between countries and can predict specific effects for each of them.

[3] This approach allows us to omit potential problems related to application of fully dynamic panel models with LDV (Keele & Kelly, 2006; Wilkins, 2017).

[4] HBLM is recommended for multilevel analysis with a small number of clusters due to use of priors and Markov Chain Monte Carlo (MCMC) sampling, which improve the reliability of estimates. MCMC iteratively samples parameter estimates, compares them to observed data, and updates the estimates. At the convergence point, an a posteriori distribution of all model parameters is given, meaning that each coefficient has its own distribution with an average that corresponds to the standard logit model's coefficient. Hierarchical models shrink varying coefficients toward the grand mean, borrowing information from other clusters and providing more conservative and reliable estimates which are less affected by extreme values and less likely to suggest a false positive result (see e.g. Gelman & Hill, 2007).

[5] In the model, we use weakly informative, regularizing priors. For intercepts and coefficients, we use weekly informative normal (0, 10) priors that imply no strong expectations for the parameters' values but conservatively guard against overestimating associations between variables (Bürkner, 2017; McElreath, 2016). For the variance part, we use half-Cauchy priors (0, 1), which is a special case of *t* family priors that is most suitable for hierarchical models with a small number of groups (Bürkner, 2017; Gelman, 2006; McElreath, 2016). With a broad peak in distribution at zero, Cauchy prior shifts group-level parameters towards the grand mean to reduced influence of outliers but allows for larger deviations in areas of high likelihood. This shrinkage ability supports more stable estimates for the country-level parameters that we are most interested in this study.

[6] Simplified versions of models (e.g., without random slopes) produced nearly the same results both with ML and Bayesian frameworks (Supplementary Table A4). As a robustness check, we tested alternative specifications of HBLM models. Results were stable across specifications of priors (e.g., cauchy [0, 5] and half-student-t [4, 0, 1] for the variance part). Adding random slopes for control variables produced nearly the same results.

Table 3
Macro-level predictors.

Hypothesis	Predictor	Code	Ref. year	Source	Values: min (L), average (M), max (H)
H2: Knowledge Economy	Knowledge Economy Index	KEI	2012	World Bank methodology (Chen & Dahlman, 2006); Retrieved from DICE database[1].	L = 7.9 M = 8.6 H = 9.4
H3: Size of welfare state	Social welfare expenditure as a % of GDP	SWE	2015	OECD (2015) online database[2]	L = 15.9 M = 24.8 H = 32.0
H4: Public support for education	Expenditure on education as a % of GDP	EDU	2014	World Bank online database (WB, 2014) [3]	L = 5.1 M = 5.5 H = 7.7
H5: Culture of active ageing	Active Ageing Index	AAI	2014	DG EMPL & UNECE methodology (AAI, 2013). For UE countries retrieved from Active Ageing Index Portal[4]. For Switzerland calculated by the Swiss Federal Statistical Office (FSO, 2018).	L = 29.8 M = 36.6 H =44.9

[1] DICE Database "Knowledge Economy Index, 1995–2012". ifo Institute, Munich, 2013. Available online: www.cesifo-group.de/DICE/fb/ziuXgj7S.
[2] OECD (2015). The OECD Social Expenditure Database. Available online: www.stats.oecd.org.
[3] World Bank Open Database. Available online: www.data.worldbank.org.
[4] Active Ageing Index Portal. Available online: www.statswiki.unece.org/display/AAI/Active+Ageing+Index+Home.

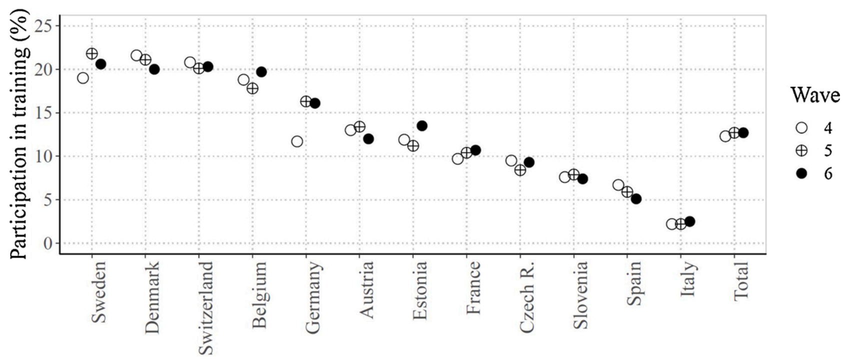

Fig. 1. Participation in training by country and wave.
Source: SHARE data 2010–2015 (own estimates). Based on a full sample without analytical inclusion criteria.

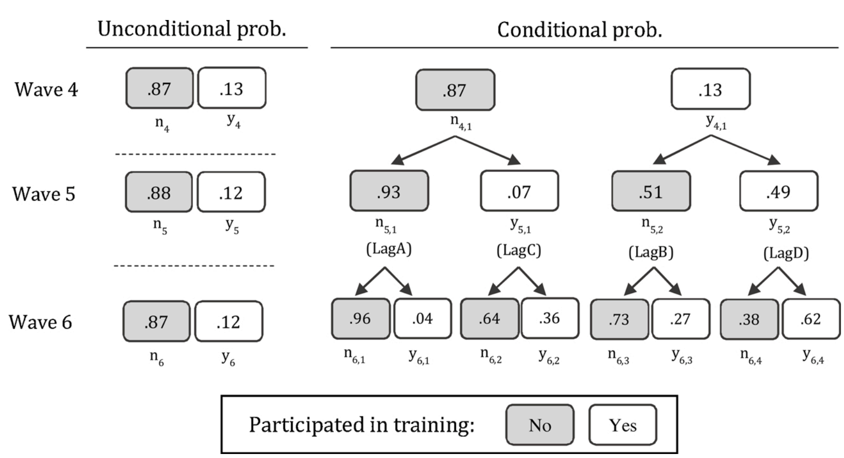

Fig. 2. A descriptive overview of the probability of training unconditional and conditional on previous participation.
Notes: Based on frequencies calculated for individuals who participated in all three waves.
Source: SHARE data 2010–2015 (own estimates).

4. Results

4.1. Descriptive overview

Average participation in training differed greatly among countries, ranging from ca. 20 % in Sweden, Denmark, Switzerland, and Belgium to 2% in Italy (Fig. 1), but participation across waves was stable within countries.

The average participation rate in the general pooled sample was around 12.5 % during each wave. This overall probability of participation can be compared with probabilities conditional on previous participation paths. Fig. 2 shows the flow of individuals through categories of participation and non-participation during three waves. It visualises that the likelihood of training is highly dependent on whether

Table 4
Results from Bayesian hierarchical logit models for the probability of training during wave 6 based on participation in waves 4 and 5 (lags of training) and individuals characteristics.

	Model 1			Model 2		
	Log-odds	95 % CI	OR	Log-odds	95 % CI	OR
Intercept	−1.07	[-1.26; -0.89]	0.34	−1.69	[-1.97; -1.41]	0.18
Lags of training (*Ref.: LagB - training in w4, but not in w5*)						
LagA (no training in w4 and w5)	−2.01	[-2.33; -1.71]	0.13	−1.58	[-1.85; -1.33]	0.21
LagC (training in w5, but not in w4)	0.41	[0.23; 0.58]	1.50	0.42	[0.24; 0.60]	1.53
LagD (training in w4 and w5)	1.42	[1.21; 1.60]	4.13	1.30	[1.09; 1.48]	3.67
Female	–	–	–	0.26	[0.17; 0.35]	1.29
Age (0 = 50 y.o.)	–	–	–	−0.32	[-0.40; -0.24]	0.73
Education (*Ref.: Primary*)	–	–	–			
Secondary	–	–	–	0.58	[0.45; 0.72]	1.79
Tertiary	–	–	–	1.11	[0.97; 1.24]	3.03
Employment pattern (*Ref = Not working*)	–	–	–			
Employed continuously	–	–	–	0.53	[0.39; 0.68]	1.70
Deactivation	–	–	–	0.04	[-0.12; 0.19]	1.04
Reactivation	–	–	–	0.25	[-0.11; 0.59]	1.28
Other	–	–	–	0.14	[-0.06; 0.34]	1.15
Variance part						
sd(Interecept)	0.24	[0.08; 0.46]	–	0.22	[0.07; 0.44]	–
sd(LagA)	0.47	[0.26; 0.78]	–	0.37	[0.18; 0.65]	–
sd(LagC)	0.11	[0.00; 0.32]	–	0.13	[0.01; 0.36]	–
sd(LagD)	0.18	[0.01; 0.43]	–	0.17	[0.01; 0.42]	–
cor(Intercept, LagA)	0.33	[-0.32; 0.85]	–	0.24	[-0.43; 0.82]	–
cor(Intercept, LagC)	0.08	[-0.76; 0.84]	–	0.17	[-0.70; 0.86]	–
cor(Intercept, LagD)	0.25	[-0.59; 0.88]	–	0.15	[-0.65; 0.84]	–
cor(LagA, LagC)	0.10	[-0.75; 0.84]	–	0.11	[-0.73; 0.83]	–
cor(LagA, LagD)	0.25	[-0.60; 0.85]	–	0.29	[-0.60; 0.88]	–
cor(LagC, LagD)	0.11	[-0.76; 0.86]	–	0.12	[-0.74; 0.85]	–
N	27,370			27,370		
WAIC	14646.3			13914.5		
Bayes R2 (median)	0.252			0.285		

Notes: The model accounts for the multilevel structure by clustering results from 12 countries and allowing for random intercept and random slopes for lags. OR – odds ratio. Effective sample sizes from MCMC estimation between: (M1) 4694–13214, (M2) 3926–25620.
Source: SHARE data 2010–2015 (own estimates).

respondents participated in training during earlier waves. For example, respondents that participated in training in wave 4 had a 51 % chance of participating again in training in wave 5. Those who did not participate in training at wave 4 had only a 7% chance of participating in training in wave 5. Diversification of patterns continued into wave 6, resulting in four conditional probabilities of training along four paths (LagA, LagB, LagC, and LagD), which ranged from 0.04 to 0.62. The scheme depicts in a descriptive way the expected path dependency in the form of an increasing probability to remain on non-training (0.87 → 0.93→0.96) and training paths (0.13→0.49→0.62). The chance of starting training during wave 6 after a non-participation pattern was only 4%.

4.2. Path dependency in training participation

Table 4 presents the results of the estimation of the probability of training in wave 6 based on earlier participation in training (Lags of training). Two models are presented: Model 1 with LDV only and Model 2 with additional sociodemographic control variables.

Adding control variables in Model 2 improved the model fit (WAIC decreased by 732, se=57.2). Females, on average, had a greater probability of training (OR=1.29), and the probability of training decreased with age (OR = 0.73) and increased with education level ($OR_{secondary}$ = 1.79, $OR_{tertiary}$ = 3.03). Higher coefficients were also found for the employed group (OR = 1.79) in comparison to the not-working group. The results show that also after controlling for sociodemographic characteristics, there is a strong path dependency in training participation. Individuals who did not participate in training earlier (Lag A) are much less likely to be trained in wave 6 (OR_{LagA} = 0.13) than those who participated in one of the earlier waves (OR_{LagB} = 1.50; OR_{LagC} = 4.13). Interpretation of the effects of LDV is easier when presented as probabilities (last row of Table 5). The average probability that a person who did not attend training during waves 4 and 5 will train during wave 6 (LagA) was only 5%, less than half of the average probability of training (12 %). The probability was much higher for people who trained during a previous wave (between 27 % for LagB and 36 % for LagC) or both waves (62 % for LagD).

These results provide support for hypothesis H1, which is derived from the theoretical assumption regarding the life course accumulation mechanisms. In particular, it stays in line with the expectation that accumulation of opportunities and barriers can lock individuals in disadvantageous path-dependent trajectories. This dependency is established on mechanisms of accumulation with risk clustering and chains of risks with additive effect (Ben-Schlomo & Kuh, 2002; Kuh et al., 2003). Accordingly, propensity and opportunities for training participation are affected by factors which can be stable over time (e.g., education, learning abilities, and company training policy) or can be reinforced by previous participation (e.g., motivation, specific human capital, employers' recognition of learning potential), resulting in path dependency.

4.3. The role of country context for path dependency

Table 5 also gives insight into the differences across countries. All countries exhibit a clear path dependency, as the pattern of conditional probabilities similarly increases from LagA to LagD. There are, however, substantial differences. For example, the likelihood of training, without any earlier participation in training in the observed period (LagA) was highest in Sweden (0.08 [0.06;0.10]) and Switzerland (0.08 [0.06;0.09]), and lowest in Italy (0.02 [0.01;0.03]) and Spain (0.02 [0.01;0.02]). Since this is a probability model, predictions differ based on specification of controls. For example, when predicted for high-training groups, such as employed women, aged 50, and with tertiary education, differences between countries became even larger, ranging from 0.11 [0.08;0.15] in Italy to 0.34 [0.28;0.40] in Sweden for LagA.

As argued in the theoretical section, these differences between countries can be related to their macro-level characteristics. Specifically,

Table 5
Probability of training during wave 6 conditional on training during waves 4 and 5 by country.

	LagA (no training in w4 and w5)		LagB (training in w4, but not in w5)		LagC (training in w5, but not in w4)		LagD (training in w4 and w5)	
Austria	0.04	[0.04;0.05]	0.27	[0.23;0.31]	0.34	[0.28;0.40]	0.60	[0.55;0.66]
Belgium	0.07	[0.06;0.08]	0.30	[0.26;0.35]	0.42	[0.36;0.48]	0.68	[0.63;0.73]
Czech R.	0.04	[0.03;0.05]	0.24	[0.20;0.28]	0.34	[0.28;0.40]	0.57	[0.50;0.63]
Denmark	0.07	[0.06;0.09]	0.27	[0.23;0.32]	0.35	[0.28;0.41]	0.64	[0.59;0.70]
Estonia	0.04	[0.03;0.04]	0.27	[0.23;0.31]	0.35	[0.30;0.41]	0.63	[0.57;0.68]
France	0.05	[0.04;0.06]	0.27	[0.22;0.31]	0.33	[0.27;0.39]	0.56	[0.50;0.62]
Germany	0.05	[0.04;0.07]	0.28	[0.23;0.34]	0.40	[0.33;0.49]	0.63	[0.55;0.70]
Italy	0.02	[0.01;0.03]	0.18	[0.13;0.24]	0.27	[0.19;0.36]	0.47	[0.34;0.58]
Slovenia	0.04	[0.03;0.05]	0.22	[0.17;0.27]	0.30	[0.23;0.37]	0.55	[0.47;0.63]
Spain	0.02	[0.01;0.02]	0.19	[0.14;0.23]	0.24	[0.18;0.30]	0.48	[0.38;0.57]
Sweden	0.08	[0.06;0.10]	0.26	[0.22;0.31]	0.34	[0.28;0.41]	0.55	[0.48;0.62]
Switzerland	0.08	[0.06;0.09]	0.30	[0.26;0.35]	0.41	[0.35;0.47]	0.65	[0.60;0.70]
Total	0.05	[0.04;0.06]	0.27	[0.23;0.32]	0.36	[0.30;0.42]	0.62	[0.56;0.67]

Note: Prediction for the observed values based on Model 2 (Table 2). 95 % CI in brackets.
Source: SHARE data 2010–2015 (own estimates).

Table 6
Results from Bayesian hierarchical logit models for the probability of training during wave 6based on participation in waves 4 and 5 (lags of training), individuals characteristics, and macro-level predictors. Only cross-level interaction-term shown: effects of macro-predictor at the levels of LDV.

	Model 3		Model 4		Model 5		Model 6	
	KEI		SWE		EDU		AAI	
(A) Regression results (log-odds and 95 % CI)								
Interaction effect (slope) for macro-predictor								
LagA	0.36	[0.15;0.57]	0.00	[-0.33;0.33]	0.31	[0.07;0.56]	0.38	[020.;0.56]
LagB	0.07	[-0.12; 0.26]	−0.04	[-0.24; 0.14]	0.03	[-0.17; 0.23]	0.07	[-0.10; 0.25]
LagC	0.15	[-0.07;0.38]	−0.09	[-0.31;0.13]	0.01	[-0.22;0.25]	0.19	[-0.02;0.40]
LagD	0.20	[-0.01;0.42]	−0.05	[-0.17;0.27]	0.17	[-0.06;0.39]	0.17	[-0.05;0.39]
(B) Test of differences between slopes of LDV								
Probability of the difference between slopes (nonoverlap)								
$\Delta_{AB}>0$	0.99		0.63		1.00		1.00	
$\Delta_{AC}>0$	0.96		0.75		0.99		0.96	
$\Delta_{AD}>0$	0.92		0.67		0.90		0.98	
Effect size of the difference between slopes								
Δ_{AB}	2.92		0.31		2.54		3.48	
Δ_{AC}	1.93		0.63		2.50		1.92	
Δ_{AD}	1.51		0.40		1.23		2.07	
N	27,370		27,370		27,370		27,370	
WAIC	13917.8		13915.6		13914.0		13914.7	
Bayes R2 (median)	0.29		0.29		0.29		0.29	

Note: The model accounts for the multilevel structure by clustering results from 12 countries and allowing for random intercept and random slopes for lags. All models additionally control for gender, age, education and employment pattern and are clustered by country with a random slope for LDV. Effective sample sizes from MCMC estimation for presented coefficients between: (M3) 5505– 9398, (M4) 5428– 12,614, (M5) 5821– 10149, (M6) 4715– 8506. Estimated in part B based on a posteriori samples of coefficients (n = 12.000). Nonoverlap measure indicates credible results with values get closer to 1. Higher values of the corresponding effect sizes indicate a higher effect, and values > 1.6 are equivalent to the nonoverlap value >0.95.
KEI - Knowledge Economy Index; SWE - Social welfare expenditure as a % of GDP; EDU - Expenditure on education as a % of GDP; AAI - Active Ageing Index. LagA – no training in w4 and w5; LagB - training in w4, but not in w5; LagC – training in w5, but not in w4; LagD – training in w4 and w5.
Source: SHARE data 2010–2015 (own estimates).

we expected that path dependency in training participation would be less prevalent in countries with more developed knowledge economies (H2), supportive welfare states (H3), stronger public support for education (H4), and a more proactive culture of ageing (H5). To verify these hypotheses, we test cross-level interactions between each of the macro-level predictors and LDV representing path dependency. We fit four models with four macro-level predictors included one at a time. Results from these analyses, limited only to the interaction term, appear in Table 6, and the general trend is visualised in Fig. 3 (full model estimates are presented in Supplementary Table A3).

Part A of Table 6 shows interaction effects (slopes) of macro-predictors for the four categories of LDV (with controls set to zero). The key is the interaction with the category of people who have not trained before (i.e., LagA), which indicates a level of openness of a training system. The results show that in countries with a higher Knowledge Economy Index (KEI), expenditure on education (EDU), and Active Ageing Index (AAI), the likelihood of training after a non-participation trajectory (LagA) is higher than in countries with lower values of these measures (the interaction effect for LagA ranges between 0.31 and 0.38 and is credibly higher than for other lags). The interaction effect for social welfare expenditure (SWE) was not significant. Formal tests presented in part B of the table confirm differences between the slopes in the interaction term (with the strongest interaction effect for LagA).

The final results are shown in Fig. 3 as a predicted probability of training. The cross-level interaction is visualised by the difference in the gradient of change between the lower set of lines that represents entering training after non-participation pattern, which can be called accessibility (LagA) and the upper set that represents the probability of continued training (LagD).

The probability of training increased, on average, with values for KEI, EDU, and AAI both for LagA and LagD, but the increase was significantly higher for slopes of LagA. A substantially higher value of

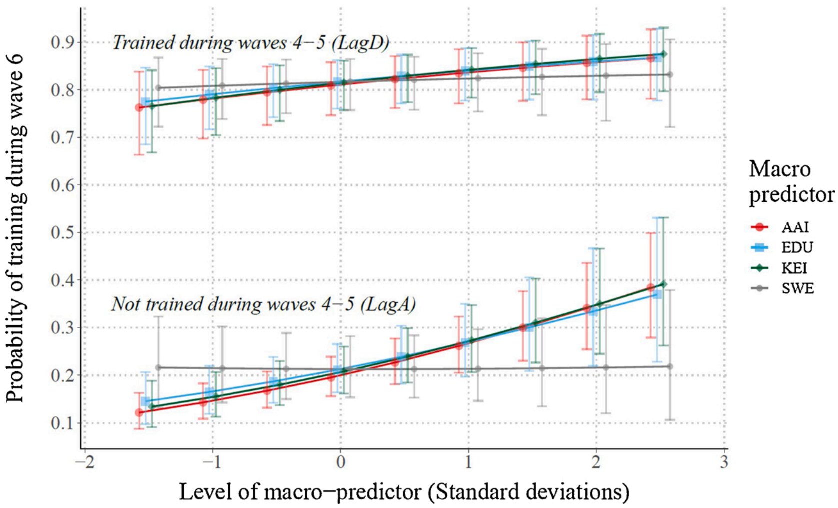

Fig. 3. Predicted probability of participation in training during wave 6 for different levels of macro-level predictors and two patterns of previous participation: people who trained (LagD) and not trained (LagA) in waves 4 and 5.
Note: For readability, lines for LagB and LagC are not shown. They would be located between LagD and LagA. Predicted probability for employed females with higher education, aged 50.
Source: SHARE data 2010–2015 (own estimates).

the interaction effect for LagA signals that these country-level characteristics particularly strongly correlate with the probability of entering training in case of people who have not participated before (accessibility).[7] No such relationship was observed for SWE, for which both lines were nearly parallel and the entire interaction term was invalid, thus not supporting H3 about the role of supportive welfare states. Our results support hypothesis H2 that stronger and more innovative economies provide greater opportunities to train at older ages. Economic and technological development encourages employers and employees to invest in skills and knowledge (Descy & Tessaring, 2005; Hanushek & Kimko, 2000). As a result, there are more opportunities for adult learning, also at older ages due to the general trend to extend working lives (Turek, Oude Mulders, & Henkens, 2020). Hypothesis H4 also found support in the data: path dependency in training participation is lower in countries with stronger public support for education. Larger investments in education reduce barriers to participation and ensure more opportunities for adult learning (Boeren et al., 2010; Desjardins, 2013; Rubenson, 2006). In line with hypothesis H5, proactive culture of ageing correlates with weakening of the path dependency. An active image of old age may influence older people's attitudes and motivations toward learning (Formosa, 2012; Rubenson & Desjardins, 2009), whereas negative stereotypes regarding trainability in older age can create reduce opportunities for training at work (Posthuma & Campion, 2009; Turek & Henkens, 2019; Van Dalen et al., 2009).

5. Conclusions

Low training attendance of older people is a prominent challenge for EU policies, and many countries have dedicated strategic programs to improve it, though the efficiency of the measures is disputable. We assess the path dependency of training to provide a new perspective on LLL at older ages. The literature offers some insights into factors that shape differences in training attendance across countries, such as demand for human capital and characteristics of welfare regimes (Riddell & Weedon, 2012; Roosmaa & Saar, 2010; Saar & Räis, 2017). However, cross-sectional data cannot reveal how LLL-related inequalities are shaped over time. This study treats lifelong learning literally as a process that occurs longitudinally. Based on three waves of SHARE panel data for a population 50+ in twelve European countries, we trace individual trajectories of training and analyse them in terms of conditionality on previous attendance.

During each wave, about 12.5 % of the population declared participation in training during the previous 12 months. Although differences across countries were large, the attendance rate was stable across waves for each country. After switching from a cross-sectional to longitudinal perspective, we find that dynamics at the individual level is considerable. Results show that training participation is path-dependent, which suggests that previous biography and decisions have a causal impact on future events (Bernardi et al., 2019; Liebowitz & Margolis, 1995). On the one hand, previous participation strongly improves the chances of further participation, which means that once they start learning, people are much more inclined to continue. On the other hand, participation in training is limited for people who have not trained previously. Path dependency can be responsible for "locking-in" the pathway and making it difficult for an individual to escape from the progressing chain of risks (Liebowitz & Margolis, 1995). Such a life course stagnation (Bernardi et al., 2019) can affect individual development by limiting opportunities and individual potential for participation in training.

What is surprising, and what provides an original contribution to LLL research is that the strength of path dependency differs across countries. In Spain and Italy, participation is more selective and conditioned by past activity than in Sweden, Switzerland, and Belgium. We investigated whether these differences relate to the degree of countries' contextual factors by testing an interaction between selected macro-characteristics and path dependency in individuals' training participation. The most important was interaction with the category of people who have not trained before, indicating how does the macro-context correlate with the accessibility of a training system. Accessibility, in this perspective, indicates the potential for breaking down the path dependency. As we have found, the level of path dependency is related to the level of development of knowledge economy, public support for education, and active ageing culture. With an increase in these macro-predictors, path dependency is lower. General welfare generosity and supportive power did not correlate with path dependency.

More specifically, the results suggest that stronger and more innovative economies provide greater opportunities to train at older ages. Investing in human capital is important for meeting the challenges of economic and technological progress in general (Descy & Tessaring, 2005; Hanushek & Kimko, 2000). In addition to this, the process of population ageing requires organisation to increase their investments in training of older workers in particular. As some studies indicate, a pro-active shift in management of older workers can be observed, and organisations are not only providing more opportunities to train in older age but also implement more comprehensive approaches to actively

[7] It should be noted that it is impossible to conclude whch of the macro-predictors had a stronger effect. Differences in corresponding estimates between models are small. E.g., the probability that the effect size for Δ_{AD} for AAI=2.07 was higher than EDU=1.23 is only 0.74.

involve their older staff (Moen, Kojola, & Schaefers, 2017; Turek, Oude Mulders, & Henkens, 2020).

From the structural perspective – which focuses on welfare state and public policies and their ability to remove barriers to participation (Boeren et al., 2010; Rubenson & Desjardins, 2009) – one recommendation for future research is to use more refined and differential measures of welfare generosity and supportive power. A measure of social expenditure that combines expenditures in very disparate areas, such as social protection and pro-active policies, is too general to account for mechanisms that drive educational attainment. This might also be true for the general typologies of welfare regimes used in many studies of LLL (Green, 2006; Rubenson & Desjardins, 2009; Rubenson, 2006; Verdier et al., 2013). Only part of a welfare state—the weight ascribed to education—appears to be a relevant indicator of training behaviours that correlates positively with access to training. This result should be considered in relation to policies that address socioeconomic inequalities. Larger investments in education indicate that a state is increasing emphasis on social cohesion (Putnam, 2004), but cohesion cannot be achieved if a training system is closed. Availability of opportunities for education, especially for people who did not attend it previously, constitutes the fundament of an efficient cohesion approach. Low accessibility and strong accumulation of training may be one reason LLL policies fail since they do not reach the target population. This conclusion corroborates other studies that suggest that a reduction of social inequalities through LLL policies is impossible unless LLL is more accessible to groups that are less likely to participate (Kilpi-Jakonen, et al. 2015; Picchio & van Ours, 2013; Roosmaa & Saar, 2010). LLL shapes individual life trajectories and affects the socioeconomic inequalities, stimulating their growth if there is a strong path dependency and decreasing them if participation is more accessible. Additionally, improvements to accessibility relate directly to average participation rates. Countries with high training attendance, such as Sweden, Denmark, Belgium, and Switzerland, are characterised by lower path dependency. Countries with low training attendance, such as Italy and Spain, have higher path dependency. Policy programmes related to LLL often use average training rates as target indicators to measure the efficiency of public interventions. This study supports the argument that the way to increase participation leads through increasing access to training (Roosmaa & Saar, 2010).

Finally, we found that path dependency in training is lower in countries with proactive ageing cultures. Cultures of old age, age roles, norms, and stereotypes affect attitudes toward learning at older ages (Formosa, 2012; Rubenson & Desjardins, 2009). Countries such as Sweden and Switzerland have excessively proactive cultures of old age, with a strong emphasis on education. Recognition of LLL's role in successful ageing creates a foundation for active attitudes of individuals (Withnall, 2010). Employers' decisions regarding training are also affected by old-age norms and culture-based expectations (Posthuma & Campion, 2009; Turek & Henkens, 2019). Results from the current study corroborate the argument that access to training at older ages is a necessary condition for active ageing.

This study has some limitations. Effects of macro-level indicators, such as specific type of expenditures as a percentage of GDP, should not be interpreted causally in connection with micro-level behaviours. As contextual factors, they reflect complex economic, structural, and sociocultural mechanisms, and interpretation should be embedded in a theoretical framework. Due to model complexity, we cannot include all macro-factors in a single model, separating effects. Alternative specification of macro-predictors is possible, e.g. taking into account nominal, not relative values of expenditures. Many of them were considered (see Appendix Table A5), but KEI, SWE, EDU, and AAI provided the clearest interpretation and referred to approaches used in other studies (Dämmrich. et al., 2014; Green & Janmaat, 2011; Riddell & Weedon, 2012; Saar & Räis, 2017). Data from 12 countries is also insufficient for drawing causal conclusions, but contrary to the frequentist approach, Bayesian modelling provides reliable estimates of models with one macro-predictor. Although SHARE is a unique source of comparative and longitudinal data, the available indicator for training participation is limited, and we lack detailed information about type, form and duration of the educational course, as well as motivation of the learner. This is not a retrospective study, and we do not control for training behaviours which happened before entering the three-wave observation window. Similarly, we cannot control for what individuals anticipate in the future, e.g. expected retirement age, what could affect their current training behaviours (Bernardi et al., 2019). SHARE data is only for the population aged 50+, and since this is the first study of path dependency in training, we can hypothesise only whether similar patterns would occur for younger groups. The analysis of training participation is limited to the structural barriers, omitting individual dispositions for participation (Froehlich et al., 2015; Zwick, 2012) and organisation-level factors, such as organisational culture or age management policies (Armstrong-Stassen & Schlosser, 2008; Pak et al., 2018). Analyses covered the period 2010–2015, when most European countries were experiencing economic slowdowns that likely resulted in reduction to investment in human capital, especially among older generations (EC, 2013; Munnell & Rutledge, 2013).

Despite these limitations, this study provides novel insights into the nature of LLL at older ages. Accumulation of advantages and disadvantages shapes development of socioeconomic structures and stimulates divergence in which initial differences enlarge over time (Crystal & Shea, 1990; Dannefer, 2003; Ferraro et al., 2009; O'Rand, 1996). An ageing population magnifies the roles of these mechanisms; increasing lifespans and longer working careers provide more time for accumulation-driven inequalities to develop, both within older generations and between younger and older cohorts. Consequently, the role of investments in adult education increases. LLL is not merely an effect of accumulated life course inequalities, but a tool for their further development. Strong path dependency and low access to training might only petrify or reinforce socioeconomic disparities, having a more profound influence on the lives of older people. If we want active, productive, and more equal societies, LLL policies must be efficient at encouraging participation of disadvantaged individuals, especially those at older ages. We argue that policies that address life course developments should include a life course perspective. Only then can the potential path dependencies be broken by adequate measures.

Founding

This work has received founding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No 748671 – LEEP – H2020-MSCA-IF-2016/H2020-MSCA-IF-2016. This paper uses data from SHARE Waves 4, 5, and 6 (Börsch-Supan et al. 2013). The SHARE data collection has been funded by the European Commission through FP5 (QLK6-CT-2001-00360), FP6 (SHARE-I3: RII-CT-2006-062193, COMPARE: CIT5-CT-2005-028857, SHARELIFE: CIT4-CT-2006-028812), FP7 (SHARE-PREP: GA N°211909, SHARE-LEAP: GA N°227822, SHARE M4: GA N°261982, DASISH: GA N°283646) and Horizon 2020 (SHARE-DEV3: GA N°676536, SHARE-COHESION: GA N°870628, SERISS: GA N°654221, SSHOC: GA N°823782) and by DG Employment, Social Affairs & Inclusion. Additional funding from the German Ministry of Education and Research, the Max Planck Society for the Advancement of Science, the U.S. National Institute on Aging (U01_AG09740-13S2, P01_AG005842, P01_AG08291, P30_AG12815, R21_AG025169, Y1-AG-4553-01, IAG_BSR06-11, OGHA_04-064, HHSN271201300071C) and from various national funding sources is gratefully acknowledged (see www.share-project.org).

Appendix A

Table A1
Statistics of macro-predictors for countries.

	KEI[1]	SWE[2]	EDU[3]	AAI[4]
Austria	8.6	27.7	5.4	34.1
Germany	8.9	24.9	4.9	37.4
Sweden	9.4	26.3	7.7	44.9
Spain	8.4	24.7	4.3	32.6
Italy	7.9	28.5	4.1	34.0
France	8.2	32.0	5.5	35.8
Denmark	9.2	29.0	7.6	40.3
Switzerland	8.9	15.9	5.1	44.0
Belgium	8.7	29.2	6.6	37.7
Czech Republic	8.1	19.4	4.0	34.4
Slovenia	8.0	22.6	5.3	29.8
Estonia	8.4	17.7	5.5	34.6
Avarage	8.6	24.8	5.5	36.6
SD	0.5	4.8	1.2	4.4

Source:
[1] DICE Database "Knowledge Economy Index, 1995−2012". ifo Institute, Munich, 2013. Available online: www.cesifo-group.de/DICE/fb/ziuXgj7S.
[2] OECD (2015). The OECD Social Expenditure Database. Available online: www.stats.oecd.org.
[3] World Bank Open Database. Available online: www.data.worldbank.org.
[4] Active Ageing Index Portal. Available online: www.statswiki.unece.org/display/AAI/Active+Ageing+Index+Home.

Table A2
Correlation between macro-predictors.

	KEI correlation	SWE	EDU	AAI
KEI	1			
SWE	0.06	1		
EDU	0.77	0.35	1	
AAI	0.84	−0.06	0.61	1

Source: own estimates.

Table A3
Bayesian hierarchical logit models for the probability of training during wave 6, including macro-level predictors. Full model (log-odds and 95 % CI in brackets).

	(M3)		(M4)		(M5)		(M6)	
	KEI		SWE		EDU		AAI	
Intercept	−1.67	[-1.94; -1.40]	−1.69	[-1.98; -1.41]	−1.68	[-1.96; -1.9]	−1.67	[-1.95; -1.40]
Lags of training (main ef.) (*Ref.: LagB - training in w4, but not in w5*)								
LagA (no training in w4 and w5)	−1.57	[-1.77;-1.38]	−1.57	[-1.85; -1.31]	−1.58	[-1.79;-1.38]	−1.58	[-1.76;-1.41]
LagC (training in w5, but not in w4)	0.41	[0.22; 0.60]	0.42	[0.24; 0.60]	0.42	[0.24; 0.60]	0.41	[0.22; 0.60]
LagD (training in w4 and w5)	1.29	[1.09; 1.47]	1.32	[1.12; 1.50]	1.28	[1.09; 1.46]	1.28	[1.08; 1.47]
Macro (main ef.)	0.07	[-0.12; 0.26]	−0.04	[-0.24; 0.14]	0.03	[-0.17; 0.23]	0.07	[-0.10; 0.25]
Macro#Lag (interact.)								
Macro#LagA	0.29	[0.08; 0.51]	0.04	[-0.24; 0.32]	0.28	[0.07; 0.51]	0.31	[0.13; 0.49]
Macro#LagC	0.08	[-0.13; 0.30]	−0.05	[-0.21; 0.12]	−0.02	[-0.21; 0.19]	0.12	[-0.07; 0.32]
Macro#LagD	0.13	[-0.09; 0.35]	0.1	[-0.08; 0.26]	0.14	[-0.06; 0.34]	0.10	[-0.10; 0.31]
Female	0.26	[0.17; 0.35]	0.26	[0.17; 0.35]	0.26	[0.17; 0.35]	0.26	[0.17; 0.35]
Age (0 = 50 y.o.)	−0.33	[-0.41; -0.25]	−0.32	[-0.4; -0.25]	−0.32	[-0.40; -0.25]	−0.33	[-0.40; -0.25]
Education (*Ref.: Primary*)								
Secondary	0.58	[0.45; 0.72]	0.58	[0.45; 0.72]	0.58	[0.45; 0.71]	0.59	[0.46; 0.72]
Tertiary	1.10	[0.97; 1.25]	1.11	[0.97; 1.25]	1.1	[0.97; 1.24]	1.11	[0.98; 1.25]
Employment pattern (*Ref = Not working*)								
Employed continuously	0.52	[0.38; 0.67]	0.53	[0.38; 0.67]	0.53	[0.39; 0.67]	0.52	[0.38; 0.66]
Deactivation	0.03	[-0.12; 0.19]	0.04	[-0.12; 0.19]	0.04	[-0.12; 0.19]	0.03	[-0.12; 0.19]
Reactivation	0.24	[-0.12; 0.57]	0.24	[-0.11; 0.59]	0.25	[-0.10; 0.59]	0.23	[-0.13; 0.57]
Other	0.13	[-0.07; 0.33]	0.14	[-0.05; 0.34]	0.14	[-0.06; 0.33]	0.13	[-0.07; 0.33]
Variance part								
sd(Interecept)	0.19	[0.05; 0.38]	0.25	[0.08; 0.48]	0.23	[0.08; 0.46]	0.19	[0.05; 0.39]
sd(LagA)	0.22	[0.04; 0.47]	0.39	[0.18; 0.71]	0.24	[0.08; 0.47]	0.16	[0.01; 0.39]
sd(LagC)	0.14	[0.01; 0.39]	0.13	[0.01; 0.36]	0.13	[0.01; 0.36]	0.15	[0.01; 0.40]
sd(LagD)	0.16	[0.01; 0.41]	0.14	[0.01; 0.40]	0.14	[0.01; 0.39]	0.17	[0.01; 0.43]
N	27,370		27,370		27,370		27,370	
WAIC	13917.8		13915.6		13914.0		13914.7	
Bayes R^2 (median)	0.29		0.29		0.29		0.29	

Note: Correlations in the variance part not shown. Effective sample sizes between: (M3) 3176–17829, (M4) 4458–23132, (M5) 4327–19849, (M6) 2352–17581.
Source: SHARE data 2010–2015 (own estimates).

Table A4
Comparison of Bayesian models presented in the article with simplified versions estimated with Maximum Likelihood.

	Model with lags and controls			Model with lags, controls and macro-predictor (AAI)		
	B1	ML1a	ML1b	B2	ML2a	ML2b
Specification:	B; RS+	ML; RS–	ML; RS+ ; Unreliable	B; RS+; Inter +	ML; RS–; Inter–	ML; RS+; Inter+; Unreliable
Intercept	−1.69 ***	−1.76 ***	−1.67 ***	−1.67 ***	−1.74 ***	−1.66 ***
Lags of training (main ef.) (*Ref.: LagB - training in w4, but not in w5*)						
LagA (no training in w4 and w5)	−1.58 ***	−1.51 ***	−1.59 ***	−1.58 ***	−1.51 ***	−1.58 ***
LagC (training in w5, but not in w4)	0.42 ***	0.44 ***	0.40 ***	0.41 ***	0.44 ***	0.41 ***
LagD (training in w4 and w5)	1.30 ***	1.32 ***	1.27 ***	1.29 ***	1.32 ***	1.26 ***
Female	0.26 ***	0.26 ***	0.26 ***	0.26 ***	0.26 ***	0.26 ***
Age (0 = 50 y.o.)	−0.32 ***	−0.32 ***	−0.32 ***	−0.33 ***	−0.32 ***	−0.33 ***
Education (*Ref.: Primary*)						
Secondary	0.58 ***	0.60 ***	0.58 ***	0.58 ***	0.60 ***	0.59 ***
Tertiary	1.11 ***	1.13 ***	1.10 ***	1.11 ***	1.13 ***	1.11 ***
Employment pattern (*Ref = Not working*)						
Employed continuously	0.53 ***	0.54 ***	0.53 ***	0.52 ***	0.53 ***	0.52 ***
Deactivation	0.04 ***	0.05	0.04	0.03 ***	0.04	0.03
Reactivation	0.25 ***	0.26	0.25	0.23 ***	0.24	0.23
Other	0.14 ***	0.15	0.14	0.13 ***	0.14	0.13
Macro (main ef.) [AAI]	–	–	–	0.06 ***	0.24 ***	0.23 ***
Macro#LagA	–	–	–	0.31 ***	–	–
Macro#LagC	–	–	–	0.12 ***	–	–
Macro#LagD	–	–	–	0.10 ***	–	–
Variance part						
sd(Interecept)	0.22	0.30	0.41	0.23	0.24	0.14
sd(LagA)	0.37	–	0.35	0.24	–	0.33
sd(LagC)	0.13	–	0.27	0.13	–	0.16
sd(LagD)	0.17	–	0.22	0.14	–	0.16
N	27,370	27,370	27,370	27,370	27,370	27,370

Notes: Only model with AAI presented as an example. * $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$.
Specification of the model: Estimation method: Bayesian (*B*), Maximum Likelihood (*ML*). Random slopes for lags: included (*RS+*), not included (*RS–*). Cross-level interaction between macro-predictor and LDV: included (*Inter+*), not included (*Inter–*).
Unreliable - model does not converge. Results shown for a model estimated with a simpler method of optimisation (*R::lme4* option "nAGQ" = 0), which does not guarantee correct results.

Table A5
Alternative macro variables: Bayesian hierarchical logit models for the probability of training during wave 6, including macro-level predictors. Only cross-level interaction-term shown: effects of macro-predictor at the levels of LDV.

	MA1		MA2		MA3		MA4		MA5		MA6	
	Empl.rate 50–74		Gini		EDU_publ		GDPpcap		GDP 10–15		AAI 2016	
(A) Regression results (log-odds and 95 % CI)												
Intercept for macro-predictor												
LagA	−3.25	[-3.58; -2.92]	−3.27	[-3.61; -2.93]	−3.26	[-3.57; -2.95]	−3.26	[-3.58; -2.95]	−3.26	[-3.61; -2.9]	−3.32	[-3.6; -3.04]
LagB	−1.67	[-1.96; -1.40]	−1.70	[-1.98; -1.42]	−1.67	[-1.95; -1.40]	−1.68	[-1.97; -1.40]	−1.68	[-1.97; -1.41]	−1.71	[-2,0; -1.41]
LagC	−1.26	[-1.56; -0.96]	−1.27	[-1.58; -0.96]	−1.26	[-1.56; -0.96]	−1.27	[-1.57; -0.96]	−1.26	[-1.58; -0.97]	−1.32	[-1.64; -1,0]
LagD	−0.37	[-0.68; -0.07]	−0.39	[-0.70; -0.09]	−0.38	[-0.69; -0.07]	−0.39	[-0.70; -0.09]	−0.37	[-0.70; -0.06]	−0.43	[-0.76; -0.1]
Slopes for macro-predictor												
LagA	0.25	[-0.03; 0.53]	−0.24	[-0.53; 0.05]	0.33	[0.10; 0.57]	0.31	[0.06; 0.55]	0.20	[-0.10; 0.51]	0.37	[0.18; 0.56]
LagB	0.07	[-0.11; 0.27]	−0.01	[-0.21; 0.19]	0.08	[-0.10; 0.27]	0.05	[-0.12; 0.23]	0.09	[-0.10; 0.30]	0.03	[-0.17; 0.24]
LagC	0.15	[-0.08; 0.38]	−0.06	[-0.30; 0.18]	0.12	[-0.09; 0.35]	0.15	[-0.06; 0.36]	0.11	[-0.12; 0.36]	0.09	[-0.14; 0.33]
LagD	0.07	[-0.15; 0.32]	−0.18	[-0.41; 0.05]	0.12	[-0.10; 0.37]	0.17	[-0.03; 0.39]	0.07	[-0.16; 0.33]	0.14	[-0.10; 0.40]
(B) Effects for slopes of LDV												
Probability of the difference between slopes (nonoverlap)												
$\Delta_{AB}>0$	0.93		0.97		0.99		0.99		0.79		1.00	
$\Delta_{AC}>0$	0.79		0.91		0.95		0.93		0.72		0.99	
$\Delta_{AD}>0$	0.93		0.69		0.96		0.90		0.83		0.98	
Effect size of the difference between slopes												
Δ_{AB}	1.57		1.84		2.30		2.41		0.83		3.43	
Δ_{AC}	0.83		1.33		1.75		1.41		0.61		2.60	
Δ_{AD}	1.44		0.47		1.70		1.19		0.93		2.08	
N	27,370		27,370		27,370		27,370		27,370		24,964	

Note: All models additionally control for gender, age, education and employment pattern and are clustered by country with a random slope for LDV. For interpretation, see description of Table 6 in the article.
Empl. rate 50−74- Employment rate for group aged 50−74 (2015). Gini – Gini coefficient, measure of socioeconomic inequalities (2015). EDU_publ – Expenditures on education as a portion of public expenditures (2014). GDPpcap - GDP per capita (2015). GDP 10−15 - GDP growth 2010–2015. AAI 2016 – Active Ageing Index 2016 (without Switzerland). LagA – no training in w4 and w5; LagB - training in w4, but not in w5; LagC – training in w5, but not in w4; LagD – training in w4 and w5. Source: SHARE data 2010–2015 (own estimates).

References

AAI. (2013). *Active ageing index 2012: Concept, methodology and final results*. Vienna: European Centre for Social Welfare Policy and Research.
Armstrong-Stassen, M., & Schlosser, F. (2008). Benefits of a supportive development climate for older workers. *Journal of Managerial Psychology, 23*(4), 419–437.
Beblavy, M., Thum, A. E., & Potjagailo, G. (2014). Learning at every age? Life cycle dynamics of adult education in Europe. *Studies in Continuing Education, 36*(3), 358–379.
Ben-Schlomo, Y., & Kuh, D. (2002). A life course approach to chronic disease epidemiology: Conceptual models, empirical challenges and interdisciplinary perspectives. *International Journal of Epidemiology, 31*(2), 285–293.
Bernardi, L., Huinink, J., & Settersten, S. (2019). The life course cube: A tool for studying lives. *Advances in Life Course Research, 41*, 1–13.
Blossfeld, H. P., Kilpi-Jakonen, E., De Vilhena, D., & Buchholz, S. (Eds.). (2014). *Adult learning in modern societies: An International comparison from a life-course perspective*. Cheltenham: Edward Elgar.
Boeren, E., Nicaise, I., & Baert, H. (2010). Theoretical models of participation in adult education: The need for an integrated model. *International Journal of Lifelong Education, 29*(1), 45–61.
Börsch-Supan, A., Brandt, M., Hunkler, C., Kneip, T., Korbmacher, J., et al. (2013). Data resource profile: The survey of health, ageing and retirement in Europe (SHARE). *International Journal of Epidemiology, 42*(4), 992–1001.
Brunello, G., Garibaldi, P., & Wasmer, E. (Eds.). (2007). *Education and training in Europe*. Oxford: Oxford University Press.
Bryan, M. L., & Jenkins, S. P. (2016). Multilevel modelling of country effects: A cautionary tale. *European Sociological Review, 32*(1), 3–22.
Bukodi, E. (2016). Cumulative inequalities over the life-course: Life-long learning and social mobility in Britain. *Journal of Social Policy, 46*(2), 367–404.
Bürkner, P. C. (2017). Brms: An r package for Bayesian multilevel models using Stan. *Journal of Statistical Software, 80*(1), 1–28.
Cedefop. (2012). *Working and ageing. The benefits of investing in an ageing workforce*. Luxembourg.
Cedefop. (2015). *Unequal access to job-related learning: Evidence from the adult education survey*. Luxembourg.
Chen, D. H. C., & Dahlman, C. J. (2006). *The knowledge economy, the KAM methodology and World Bank operations*. Washington, DC: World Bank.
Cohen, J. (1988). *Statistical power analysis for the behavioral sciencies* (second edition). Hillsdale: Lawrence Erlbaum.
Coulombe, S., & Tremblay, J. (2007). Explaining cross-country differences in job-related training: Macroeconomic evidence from OECD countries. *Économie Internationale, 110*, 5–29.
Crystal, S., & Shea, D. (1990). Cumulative advantage, cumulative disadvantage, and inequality among elderly people. *The Gerontologist, 30*(4), 437–443.
Dämmrich, J., De Vilhena, D., Reichart, E., et al. (2014). Participation in adult learning in Europe: The impact of Country-level and individual characteristics. In H.-P. Blossfeld (Ed.), *Adult learning in modern societies an International comparison from a life-Course perspective* (pp. 29–55). Cheltenham: Edward Elgar.
Dannefer, D. (2003). Cumulative advantage, disadvantage and the life course. *The Journals of Gerontology: Series B, 58*(6), 327–337.
Dannefer, D., & Uhlenberg, P. (1999). Paths of the life course: A typology. In V. L. Bengtson, & K. W. Schaie (Eds.), *Handbook of theories of aging* (pp. 306–326). New York: Springer Publishing.
Descy, P., & Tessaring, M. (2005). *The value of learning Evaluation and impact of education and training*. Luxemburg: Cedefop.
Desjardins, R. (2013). Considerations of the impact of neoliberalism and alternative regimes on learning and its outcomes: An empirical example based on the level and distribution of adult learning. *International Studies in Sociology of Education, 23*(3), 182–203.
EC. (2010). *Why socioeconomic inequalities increase? Facts and policy responses in Europe*. Brussels: European Commission Directorate-General for Research.
EC. (2013). *Funding of education in Europe 2000-2012: The impact of the economic crisis*. Eurydice Report. Luxembourg: Publications Office of the European Union.
Elder, G. H. (1994). Time, human agency, and social change: Perspectives on the life course. *Social Psychology Quarterly, 57*(1), 4–15.
Elman, C., & O'Rand, A. M. (2004). The race is to the swift: Socioeconomic origins, adult education, and wage attainment. *The American Journal of Sociology, 110*(1), 123–160.
Evans, K., Schoon, I., & Weale, M. (2013). Can Lifelong Learning Reshape Life Chances? *British Journal of Educational Studies, 61*(1), 25–47.
Ferraro, K. P., Schippee, T. P., Schafer, M. H., et al. (2009). Cumulative inequality theory for research on aging and the life course. In V. L. Bengtson (Ed.), *Handbook of theories of aging* (pp. 413–433). New York: Springer.
Formosa, M. (2012). European Union policy on older adult learning: A critical commentary. *Journal of Aging & Social Policy, 24*(4), 384–399.
Froehlich, D., Beausaert, S., & Segers, M. (2015). Age, employability and the role of learning activities and their motivational antecedents: A conceptual model. *International Journal of Human Resource Management, 26*(16), 2087–2101.
FSO. (2018). *Active ageing*. Neuchâtel: Demos 1/2018 - Federal Statistical Office.
Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models. *Bayesian Analysis, 1*(3), 515–534.
Gelman, A., & Hill, J. (2007). *Data analysis using regression and Multilevel/Hierarchical models*. Cambridge: Cambridge University Press.
Gelman, A., Goodrich, B., Gabry, J., & Vehtari, A. (2019). R-squared for Bayesian regression models. *The American Statistician* (Published online).
George, L. K. (1993). Sociological perspectives on life transitions. *Annual Review of Sociology, 19*(1), 353–373.
Green, A. (2006). Models of Lifelong Learning and the' knowledge society'. *Compare A Journal of Comparative Education, 36*(3), 307–325.
Green, A. (2011). Lifelong learning, equality and social cohesion. *European Journal of Education, 46*(2), 228–243.
Green, A., & Janmaat, J. G. (2011). *Regimes of social cohesion. Societies and the crisis of globalization*. New York: Palgrave Macmillan.
Groot, W., & Van den Brink, H. M. (2000). Education, training and employability. *Applied Economics, 32*(5), 573–581.
Hansson, B. (2008). *Job-related training and benefits for individuals. A review of evidence and explanations*. Paris: OECD.
Hanushek, E. A., & Kimko, D. D. (2000). Schooling, labor-force quality, and the growth of nations. *The American Economic Review, 90*(5), 1184–1208.
Harper, S., Khan, H., Saxena, A., & Leeson, G. W. (2006). Attitudes and practices of employers towards ageing workers: Evidence from a global survey on the future of retirement. *Ageing Horizons, 5*, 31–41.
Holford, J., & Mleczko, A. (2013). Lifelong learning: National policies in the European perspective. In E. Saar, O. B. Ure, & J. Holford (Eds.), *Lifelong learning in Europe: National patterns and challenges* (pp. 25–45). Cheltenham: Edward Elgar.
Keele, L., & Kelly, N. (2006). Dynamic models for dynamic theories: The ins and outs of lagged dependent variables. *Political Analysis, 14*(02), 186–205.
Kerckhoff, A. C. (1993). *Diverging pathways: Social structure and career deflections*. New York: Cambridge University Press.
Kerckhoff, A. C., & Glennie, E. (1999). The Matthew effect in American education. *Research in Sociology of Education and Socialization, 12*, 35–66.
Kilpi-Jakonen, E., Vono de Vilhena, D., & Blossfeld, H. P. (2015). Adult learning and social inequalities: Processes of equalisation or cumulative disadvantage? *International Review of Education, 61*(4), 529–546.
Kohli, M. (2007). The institutionalisation of the life course: Looking back to look ahead. *Research in Human Development, 4*(3–4), 253–271.
Kruschke, J. (2012). Bayesian estimation supersedes the t test. *Journal of Experimental Psychology, 142*, 573–603.
Kuh, D., Ben-Shlomo, Y., Lynch, J., Hallqvist, J., & Power, C. (2003). Life course epidemiology. *Journal of Epidemiology and Community Health, 57*(11), 778–783.
Lazazzara, A., Karpinska, K., & Henkens, K. (2013). What factors influence training opportunities for older workers? Three factorial surveys exploring the attitudes of HR professionals. *International Journal of Human Resource Management, 23*(11), 2154–2172.
Leuven, E., & Oosterbeek, H. (1999). The demand and supply of work-related training. *Research in Labor Economics, 18*, 303–330.
Liebowitz, S. T., & Margolis, S. (1995). Path dependence, lock-in, and history. *Journal of Law, Economics, & Organization, 11*(1), 205–226.
Maas, C., & Hox, J. (2005). Sufficient sample sizes for multilevel modeling. *Methodology, 1*(3), 86–92.
McElreath, R. (2016). *Statistical rethinking: A Bayesian course with examples in R and Stan*. Boca Raton, FL: CRC Press.
Moen, P., Kojola, E., & Schaefers, K. (2017). Organizational change around an older workforce. *The Gerontologist, 57*, 847–856.
Munnell, A. H., & Rutledge, M. S. (2013). The effects of the great recession on the retirement security of older workers. *The Annals of the American Academy of Political and Social Science, 650*(1), 124–142.
Narushima, M., Liu, J., & Diestelkamp, N. (2018). I learn, therefore I am: A phenomenological analysis of meanings of lifelong learning for vulnerable older adults. *The Gerontologist, 58*(4), 696–705.
O'Rand, A. M. (1996). The precious and the precocious: Understanding cumulative disadvantage and cumulative advantage over the life course. *The Gerontologist, 36*(2), 230–238.
Pak, K., Kooij, D., De Lange, A., & Van Veldhoven, M. (2018). Human Resource Management and the ability, motivation and opportunity to continue working: A review of quantitative studies. *Human Resource Management Review, 29*(3), 336–352.
Pallas, A. M., Natriello, G., & McDill, E. L. (1989). The changing nature of the disadvantaged population current dimensions and future trends. *Educational Researcher, 18*(5), 16–22.
Piccarreta, R., & Studer, M. (2018). Holistic analysis of the life course: Methodological challenges and new perspectives. *Advances in Life Course Research* (Published online).
Picchio, M., & van Ours, J. C. (2013). Retaining through training even for older workers. *Economics of Education Review, 32*, 29–48.
Posthuma, R., & Campion, M. (2009). Age stereotypes in the workplace: Common stereotypes, moderators, and future research directions. *Journal of Management, 35*(158), 158–188.
Putnam, R. (2004). Education, diversity, social cohesion and "social Capital". *Note for discussion presented to meeting of OECD education ministers, Dublin, March 2004*.
Razin, A., Sadka, E., & Swagel, P. (2002). The aging population and the size of the welfare state. *The Journal of Political Economy, 110*(4), 900–918.
Riddell, S., & Weedon, E. (2012). Lifelong learning and the wider European socioeconomic context. In S. Riddell, J. Markowitsch, & E. Weedon (Eds.), *Lifelong learning in Europe. Equity and efficiency in the balance* (pp. 17–38). Bristol: The Policy Press.
Roosmaa, E. L., & Saar, E. (2010). Participating in non-formal learning: Patterns of inequality in EU-15 and the new EU-8 member countries. *Journal of Education and Work, 23*(3), 179–206.
Rubenson, K. (2006). The nordic model of lifelong learning. *Compare A Journal of Comparative Education, 36*(3), 327–341.

Rubenson, K., & Desjardins, R. (2009). The impact of welfare state regimes on barriers to participation in adult education: A bounded agency model. *Adult Education Quarterly, 59*(3), 187–207.

Saar, E., & Räis, M. L. (2017). Participation in job-related training in European countries: The impact of skill supply and demand characteristics. *Journal of Education and Work, 30*(5), 531–551.

Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics, 87*(3), 355–374.

Turek, K., & Henkens, K. (2019). How skill requirements affect the likelihood of recruitment of older workers in Poland: The indirect role of age stereotypes. *Work, employment and society*. First published online.

Turek, K., Oude Mulders, J., & Henkens, K. (2020). The proactive shift in managing an older workforce 2009–2017: A latent class analysis of organizational policies. *The Gerontologist, 60*(8), 1515–1526.

Van Dalen, H. P., Henkens, K., & Schippers, J. (2009). Dealing with older workers in Europe: A comparative survey of employers' attitudes and actions. *Journal of European Social Policy, 19*(1), 47–60.

Van der Heijden, B., Gorgievski, M., & De Lange, A. (2016). Learning at the workplace and sustainable employability: A multi-source model moderated by age. *European Journal of Work and Organisational Psychology, 25*(1), 13–30.

Vehtari, A., Gelman, A., & Gabry, J. (2016). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing, 27*(5), 1413–1432.

Verdier, E., et al. (2013). Lifelong learning regimes versus vocational education and training systems in Europe: The growing hybridisation of national models. In J. G. Janmaat (Ed.), *The dynamics and social outcomes of education systems* (pp. 70–93). New York: Palgrave Macmillan.

Wilkins, A. S. (2017). To lag or not to lag?: Re-Evaluating the use of lagged dependent variables in regression analysis. *Political Science Research and Methods, 6*(02), 393–411.

Withnall, A. (2010). *Improving learning in later life*. London, New York: Routledge.

Zwick, T. (2012). Training effectiveness – Differences between younger and older employees. *Working and ageing. The benefits of investing in an ageing workforce* (pp. 36–54). Luxembourg: Cedefop.