

Training and wages of older workers in Europe

Michele Belloni · Claudia Villosio

Published online: 15 November 2014
© Springer-Verlag Berlin Heidelberg 2014

Abstract The financial deficits of many social security systems caused by ageing populations and stagnating economies are forcing workers to retire later from the labour market. An extended working life, combined with rapid technological progress in many sectors, is likely making older workers' skills obtained in school obsolete. In this context, lifelong investment in training is widely recognised among the international research and policy community as a key element to increase or at least limit the decline in productivity of older workers. This paper investigates the relationship between training undertaken by European older workers and their wages, relying on the Survey of Health, Ageing and Retirement in Europe. There is no evidence of training wage premium for older workers residing in many European countries including Denmark, Sweden, Belgium, the Netherlands, and Switzerland. Very high premiums are instead found for Austria, Germany, Greece and Italy. It is, however, likely that these high premiums are overestimated due to training endogeneity and sample selection bias.

Keywords Older workers · Training · Wages · Cognitive abilities · Sample selection bias · Attrition

Responsible Editor: Michal Myck (guest editor) and Howard Litwin.

M. Belloni
Economics Department University Ca' Foscari of Venice, CeRP-
Collegio Carlo Alberto, and Netspar, San Giobbe 873,
30121 Venice, Italy
e-mail: michele.belloni@unive.it

C. Villosio (✉)
LABORatorio R. Revelli-Collegio Carlo Alberto, Via Real
Collegio 30, 10024 Moncalieri, TO, Italy
e-mail: claudia.villosio@laboratoriorevelli.it

JEL Classification J14 · J24 · J31

Introduction

The long-term increase in longevity that many countries are experiencing is a direct reflection of the success of our societies in securing higher living standards. Such increasing life expectancy, together with low economic growth, has, however, put pressure on the financial equilibrium of many pay-as-you-go pension systems in Europe and in other industrialised countries. Governments have therefore often been forced to increase the average retirement age out of the labour force by either limiting access to early retirement or making early exit routes less attractive (e.g. Gruber and Wise 2004). An extended working life—combined with the rapid technological progress in many sectors—is likely making older workers' skills obtained in school obsolete. In this context, lifelong investment in training is widely recognised among the international research and policy community as a key element to increase or at least limit the decline in productivity of older workers (e.g. OECD 2006).

The theoretical foundations of the effect of training on worker productivity lie in human capital theory (Becker 1964, 1993). According to this theory, training activities (and education) are investments, since they are undertaken to increase individuals' stock of knowledge, skills, competencies, and abilities, which are important elements of human capital. An individual's decision to invest in training is based upon a comparison of the net present value of costs (i.e. direct training costs and lower wages while in training) and the net present value of benefits (i.e. higher marginal product and higher wages) of such an investment. Individuals invest in training during a period and obtain

returns to the investment in subsequent periods: The longer the residual working life, the more advantageous the investment in training. One of the main predictions of this model is, therefore, that both participation and returns to training decline with increasing age and approaching retirement. Generous early retirement schemes may therefore discourage participation in training aimed at older workers (Fouarge and Schils 2009). When analysing older workers, it is particularly relevant to distinguish between general and specific training. General training develops transferable skills that the trainee can bring to another firm, while specific training improves skills targeted to the firm in which the trainee works. Human capital models predict that while the workers receive all the returns to general training (and therefore they pay for it), employees and employers share the returns (and costs) to specific training.

Great research efforts have been dedicated to empirically test the predictions of human capital theory. A branch of this literature evaluates the effect of training on productivity indirectly, by means of its effect on workers' wages. It assumes that wages are a sufficient statistic for productivity (Dearden et al. 2006) and often relies on the traditional neoclassical labour market model with perfectly competitive wages (noticeable exceptions are Acemoglu and Pischke 1999 and Dostie and Léger 2011). Positive effects of training on wages are commonly found for most European countries, as well as for the United States. Nevertheless, the size of the estimated effect varies widely across countries and is very sensitive to the data and analytical methods used. Extended empirical literature on this topic exists for the United Kingdom, where estimated returns to training range from 1 (Booth 1993) to 18 % (Booth 1991). Positive effects of training on wages are found for Norway (1 % increase, Schøne 2004), Switzerland (2 %, Gerfin 2004), and Portugal (30 %, Budria and Pereira 2007). The results for Germany (cf. no return in Pischke 2001, with 5 % return in Mühler et al. 2007) and France (cf. no return in Goux and Maurin 2000, with positive return for job switchers in Fougère et al. 2001) are less clear.

Due to different concepts and definitions of training across countries and datasets, the comparability of results across countries is limited. A few studies circumvent some of these limitations by exploiting cross-country data (see OECD 1999, 2004). Bassanini et al. (2007) estimate a positive impact of training on earnings for most of the analysed countries; this return ranges from 3.7 % for the Netherlands to 21.6 % for Greece and is sensitive to the statistical method employed. Some papers investigate the differentials in training returns (and incidence) between younger and older workers (Booth 1991; Warr and Fay 2001; Bassanini 2006; Bassanini et al. 2007; Lang 2012; Dostie and Léger 2011). These studies mostly confirm human capital theory predictions, showing lower returns

for older workers compared to younger ones. To our knowledge, there are no studies focusing on older adults from an international perspective.

In this paper, we investigate the relationship between training undertaken by workers aged 50 and above and their wages. We rely on the Survey of Health, Ageing and Retirement in Europe (SHARE), which provides information on wages and training for European individuals aged 50+. Besides its specific focus on older adults, this source of data is particularly suitable for international studies due to its wide country coverage and harmonisation in the definition of training in the country questionnaires. We test whether our results are affected by unobserved ability, attrition, and sample selection biases.

Data and empirical strategy

This study relies on data from SHARE (www.share-project.org), a multidisciplinary and cross-national panel database of micro data on health, socioeconomic status (SES), and social and family networks. The target population consists of individuals aged 50+ who speak the official language of each country and do not live abroad or in an institution, as well as their spouses or partners, irrespective of age. Four waves of SHARE are currently available. The baseline study (interview years 2004–2005) covers 11 European countries representing different regions of continental Europe, from Scandinavia (Denmark and Sweden) through Central Europe (Austria, Belgium, France, Germany, the Netherlands, and Switzerland) to the Mediterranean countries (Greece, Italy, and Spain). Additional countries joined the survey in the following waves. The third wave of SHARE (conducted in 2008–2009), called SHARELIFE, was a retrospective life history survey.

This study exploits two key variables reconstructed from the first two waves of SHARE: hourly wages and participation in training activities. Hourly wages are measured from the second wave of the survey. A set of variables concerning employees' current main job is included in the 'Employment and Pension' module of this wave, which comprises the following: (i) last payment taken home from work, (ii) frequency of payment ('How often do you get paid?': 'Every week', 'every two weeks', and 'every year'), and (iii) total hours worked per week. We reconstruct hourly wages, net of taxes (Flores and Kalwij 2013), dividing net wages—obtained by combining (i) and (ii)—by the number of hours worked—from (iii). Unfolding bracket values for last payment taken home from work are used to reduce the number of missing values.

Participation in training activities is retrieved from the first wave of SHARE for those employees whose hourly wages in the second wave of the survey (i.e. one to three

years later) could be reconstructed. To measure training participation, we exploit a question (part of the module ‘Activities’) that asks interviewees whether they attended any educational or training courses in the last four weeks. We note, however, that this question does not permit differentiation between general and specific training. No information on the length of the training course is gathered.

As mentioned above, we consider individuals surveyed in both Waves 1 and 2, that is, about 18,000 respondents. About one-third of the baseline sample is lost due to attrition between the first and second waves of the survey. Further focusing on employees in both waves leads to a selection of 4,304 individuals, since retired individuals represent roughly half of the samples in both waves. Note that about 13 % of employees in Wave 1 who were re-interviewed in Wave 2 retired between the two waves. Moreover, an additional 6 % became unemployed, permanently sick, or homemakers. We also note that transition out of employment is not random with respect to training: 14 % (3 %) of untrained workers in Wave 1 moved to retirement (unemployment), compared to only 8 % (2 %) of trained workers. We return to these issues later.

Information on wages is available for 3,824 individuals, including 136 whose wages were imputed using bracket values. To account for outliers, we exclude the top and bottom 1 % of the hourly wage distribution obtained, as well as individuals working fewer than 15 h or more than 70 h per week (i.e. about 10 % of the residual sample). Finally, excluding observations for which either weights or control variables are missing leads to a final estimation sample of 2,312 individuals.

Based on these data, we estimate the following equation for (the logarithm of) wages:

$$\log(w_{i2}) = x_{i2}\beta + \tau_{i1}\gamma + \varepsilon_{i2}, \quad (1)$$

where $\log(w_{i2})$ is the logarithm of net hourly wages of individual i at the time of Wave 2 (i.e. in either 2006 or 2007), x is a vector of exogenous demographic and job-related individual characteristics, τ is a dummy variable equal to one if individual i participated in any training activity at the time of Wave 1 (i.e. in either 2004 or 2005), and ε is a random term that is assumed to satisfy the standard assumptions. The key parameter to be estimated is γ , which measures the causal effect of training on wages if the above assumptions are satisfied. In the main empirical specification of Eq. (1), x includes a second-order polynomial for age, (log of) tenure, dummy variables for the main International Standard Classification of Education (ISCED) educational groups, gender, whether working in the public or private sector, the country of residence, work sector, occupation—by one-digit International Standard Classification of Occupations (ISCO) code—and year of interview.

We estimate Eq. (1) with the method of ordinary least squares (OLS), thus the estimate of γ in Eq. (1) is consistent only if training is uncorrelated with the error term, conditional on the other covariates. This assumption does not hold if unobserved individual characteristics, such as ability, affect wages and are correlated with training. Since ability, training participation, and wages are expected to be positively correlated, neglecting such an omitted variable/selectivity problem can lead to overestimation of the return to training. The state of the art in this literature is to estimate fixed effects or first-differences versions of Eq. (1), where a time-invariant component is included in the error term. With these methods, one gets rid of individual unobserved characteristics as long as they are constant over time (Wooldridge 2002). However, such models cannot be estimated with the data at our disposal, because we are not able to reconstruct hourly wages from the fourth wave of SHARE. Having had both hourly wages and delayed training participation at two points in time—that is, wages in Waves 2 and 4 and training in Waves 1 and 2—would have allowed estimation of a first-differences version of Eq. (1). To partly overcome this lack of data, we also present the results of augmented models that include proxy variables for ability taken from SHARELIFE.

Results

Figure 1 reports the distribution (of the logarithm) of hourly wages, by training status. It highlights that the distribution of trained workers’ hourly wages is somewhat shifted to the right with respect to that of untrained workers. Table 1 shows descriptive statistics, by training status. Trained workers represent 17.8 % (411 individuals) of the estimation sample. On average, they receive higher wages than untrained workers; the difference between trained and untrained workers in terms of the mean of (the logarithm of) hourly wages (trained 2.68, s.d. 0.42; untrained 2.48, s.d. 0.45) is statistically significant at the 1 % level. Females (55 vs. 43 %) and public sector workers (28 vs. 22 %) represent a significantly higher percentage of trained workers than of untrained workers. More than 50 % of trained workers have a tertiary education degree (ISCED 5–6), while only 10 % have either no education (ISCED 0) or a primary or lower-secondary degree (ISCED 1–2). Among untrained workers, most (40 %) of them have an upper or post-secondary education (ISCED 3–4), while 30 % of this sample has a lower level of education (ISCED 0–2). Figure 2 compares the training incidence and the difference between trained and untrained workers in the mean of the logarithm of hourly wages, by country. The latter variable can be seen as an uncorrected training wage premium at the country level. Two facts emerge from



Fig. 1 Distribution of log hourly wages, by training status. Source: SHARE Wave 1–2

Table 1 Estimation sample: descriptive statistics, by training status

Variables	Untrained workers		Trained workers	
	Mean	Std. Dev.	Mean	Std. Dev.
Log hourly wage	2.48	0.45	2.68	0.42
Log tenure	2.84	0.91	2.93	0.88
Female	0.43	0.50	0.55	0.50
Age	54.89	3.63	54.45	3.30
age ² (/100)	30.26	4.12	29.75	3.67
public sector	0.22	0.42	0.28	0.45
<i>Education (ISCED 97)</i>				
No education (ISCED 0)	0.02	0.15	0.01	0.07
Primary and lower-secondary education (ISCED 1–2)	0.28	0.45	0.09	0.29
Upper and post-secondary education (ISCED 3–4)	0.40	0.49	0.39	0.49
Tertiary education (ISCED 5–6)	0.29	0.46	0.51	0.50
<i>n. of observations</i>	1,901		411	

Source: SHARE Wave 1–2

Note: Variables not reported: country of residence, sector, occupation, and year of interview

Fig. 2. First, both variables are characterised by wide dispersion (thus, later we investigate to some extent the issue of country heterogeneity in training premium). Second, there is a negative association between them. This is in line with previous studies (e.g. Bassanini et al. 2007) that report evidence of a negative association between training incidence and return to training across countries at all working ages.

Table 2 shows the OLS parameter estimates of Eq. (1) for two specifications. Specification (a) assumes that the training wage premium is the same in all the analysed countries. Specification (b) relaxes this assumption of

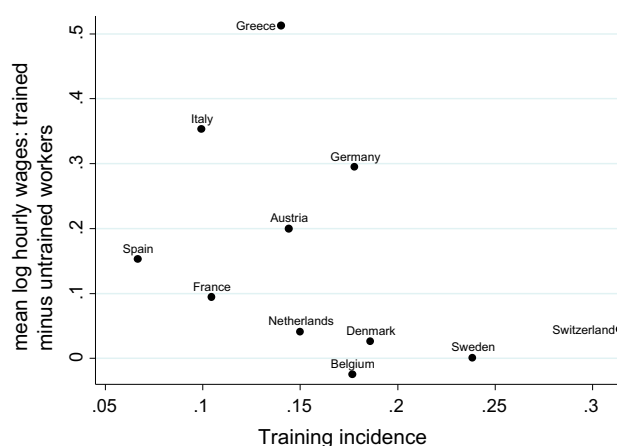


Fig. 2 Training incidence and difference between trained and untrained workers in the mean of log hourly wages, by country. Source: SHARE Wave 1–2

homogeneity across countries. Obviously, it would have been very informative to estimate country-specific premiums. However, this last exercise is not feasible given the small sample size and low training incidence observed in various countries included in our dataset. Therefore, starting from a flexible specification where the training variable is interacted with country dummies and relying on the descriptive evidence reported in Fig. 2, we progressively combine countries showing similar results, obtaining groups of countries characterised by increasingly different training premiums. More parsimonious specifications were tested against more flexible ones. This procedure results in three country groups: (1) Austria, Germany, Greece, and Italy; (2) France and Spain; and (3) Denmark, Sweden, Belgium, the Netherlands, and Switzerland.

Specification (a) shows that the estimate for the training variable is equal to 0.0634 and is significant at the 1 % level: for employees aged 50 and above residing in one of the 11 analysed European countries, undertaking training activities is associated with 6.3 % higher wages. This premium is sizeable and is similar to that of attaining an upper or post-secondary education degree (ISCED 3–4) instead of a primary or lower-secondary (ISCED 1–2) education degree ($0.23 - 0.16 = 7\%$). All the control variables have the expected sign and are highly significant: Hourly wages increase with education attainment, tenure, and age (at a decreasing rate). Specification (b) shows that the training wage premium is highly heterogeneous across countries: The incremental effects for country groups 2 and 3 are statistically significant at the 10 and 1 % levels, respectively. Training premiums are highest in Austria, Germany, Greece, and Italy (19 %) and are reduced by about half in France and Spain ($19.3 - 10.4 = 8.9\%$). No training premium is found for Denmark, Sweden, Belgium, the Netherlands, and Switzerland. For this last group of

Table 2 Training wage premiums

Variables	(a)	(b)
Training (country group 1: AT, DE, GR, IT)	0.0634*** (0.0215)	0.193*** (0.0293)
Log tenure	0.0720*** (0.00875)	0.0685*** (0.00854)
Female	-0.167*** (0.0164)	-0.147*** (0.0161)
Public sector	-0.0561*** (0.0196)	-0.0580*** (0.0196)
Age - 50	0.0170*** (0.00571)	0.0142** (0.00575)
Age- 50 squared	-0.00169*** (0.000428)	-0.00128*** (0.000432)
No education (ISCED 0)	-0.305*** (0.0554)	-0.245*** (0.0615)
Primary and lower-secondary education (ISCED 1-2)	-0.234*** (0.0261)	-0.238*** (0.0259)
Upper and post-secondary education (ISCED 3-4)	-0.162*** (0.0202)	-0.158*** (0.0194)
Training * (country group 2: FR, ES)		-0.104* (0.0602)
Training * (country group 3 : DK, SE, BE, NL, CH)		-0.256*** (0.0441)
Wald test H_0 : no effect of training in country group 2 (p value)		0.08
Wald test H_0 : no effect of training in country group 3 (p value)		0.17
Observations	2,312	2,312
R-squared	0.421	0.425

Source SHARE Waves 1-2

OLS estimates; Standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$; Additional controls: dummy variables for country of residence, sector, occupation, and year of interview; Reference categories: private sector, tertiary education (ISCED 5-6), training in country group 1: AT, DE, GR, IT

countries, the point estimate for the training variable is negative (19.3–25.6 = -6.3 %); however, the Wald test reported at the bottom of Table 2 on the joint significance of the training variable plus the incremental effect for group 3 suggests that this negative coefficient is not significantly different from zero (p value 0.17).

Sensitivity analysis

Attrition, transition out of employment, and selection into training

Our sample selection criteria require individuals to be employed in both waves. Earlier, we underlined that attrition

between the first and second wave of SHARE is high; moreover, we showed that a non-negligible percentage of individuals employed in the first wave retires or becomes unemployed between the two waves. Attrition and transitions out of employment might thus severely affect our findings. To shed light on these issues, we perform a test for attrition bias (see Beckett et al. 1988). We select individuals employed in the first wave of SHARE (excluding those whose hourly wages could not be reconstructed and outliers). We split this sample into three groups including, respectively, individuals (a) not present in Wave 2 due to attrition, (b) present in Wave 2 but not employed (retired or became unemployed between the two waves), and (c) employed in Wave 2, i.e. included in our estimation sample described previously. We then run a regression which explains (log of) hourly wages in Wave 1 by means of four “attrition” variables (see later), a dummy variable for having undertaken training in Wave 1 and the same set of control variables as in Eq. 1. The four “attrition” variables are a dummy variable equal to one if the individual belongs to group (a) (“Not present in Wave 2”), a dummy variable equal to one if the individual belongs to group (b) (“Not employed in Wave 2”) and two interaction terms between each of these dummy variables and the training dummy. Issues of attrition/sample selection bias are signalled by significance of the estimated parameters for these interaction terms, as well as by the joint significance of these interaction terms with the dummy variables “Not present in Wave 2”/“Not employed in Wave 2” (see Beckett et al. 1988).

Table 3 shows the results of this exercise, for the full sample and by group of countries. No clear evidence of problems stemming from attrition and/or transitions out of employment emerges for the full sample and for country groups 2 and 3. Results for country group 1, however, signal evidence of problems from both. For this country group, in fact, both interaction terms of the “attrition” variables are significant at 5 % level; the p value of the F-test for no joint significance of the interaction term plus constant for group “Not employed in Wave 2” is equal to 5.68 % (while that for group “Not present in Wave 2” is equal to 10.3 %). In other words, among those who remain employed, trained workers had a 10.2 % higher wage than untrained workers (point estimate of the training coefficient); conversely, among those who leave the sample due to attrition, the point estimate of the difference in wages between trained and untrained workers was equal to about -5.4 % (= 0.102–0.156). Among those who leave the sample due to exit from employment, such a difference was even larger: about -21.1 % (= 0.102–0.313). A comparison between these two differences shows that the problem of transitions out of employment is more severe than that of attrition. Note that the group “Not present in Wave 2” is a mixed one—it may also include individuals who left employment.

Table 3 Tests for attrition/sample selection bias-dependent variable: log of hourly wages in Wave 1

Variables	Full sample	By country group:		
		1 (AT, DE, GR, IT)	2 (FR, ES)	3 (DK, SE, BE, NL, CH)
Training	0.0134 (0.0254)	0.102* (0.0528)	−0.138 (0.103)	−0.0392* (0.0234)
<i>“Attrition” variables:</i>				
Not present in Wave 2	0.0180 (0.0148)	−0.000672 (0.0282)	0.0477 (0.0480)	0.0163 (0.0184)
Not employed in Wave 2	−0.00453 (0.0204)	0.0225 (0.0379)	−0.0414 (0.0705)	0.00667 (0.0262)
Training * not present in Wave 2	−0.0599 (0.0401)	−0.156** (0.0784)	0.0375 (0.154)	0.0462 (0.0426)
Training * not employed in Wave 2	−0.0853 (0.0671)	−0.313** (0.130)	0.482 (0.294)	0.0272 (0.0654)
<i>F-test H_0: no difference including constant between employed in Wave 2 and group:</i>				
Not present in Wave 2 (<i>p</i> value)	0.250	0.103	0.523	0.181
Not employed in Wave 2 (<i>p</i> value)	0.375	0.0568	0.255	0.836
Observations	3,742	1,022	341	2,379
R-squared	0.422	0.436	0.537	0.347

Note: OLS estimates; Standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$; Additional controls: Log tenure, age, age squared, dummy variables for female, public sector, ISCED educational groups, country of residence, sector, occupation, and year of interview; Reference categories: employed in Wave 2, not undertaking training in Wave 1

This analysis points out that our estimation sample excludes trained individuals with lower wages at the time of training. The sign of the bias created by this exclusion is not certain. In the extreme case of equal return to training across workers regardless of their group (employed in Wave 2, Not present/employed in Wave 2), such exclusion would not create any bias. However, one may expect that individuals undertaking training and retiring would have (on average) had returns to training lower than those remaining employed. In the last case, exclusion of these individuals from the sample biases our results upward.

The parameter estimates for the training dummy variable in Table 4 also suggest evidence of training endogeneity for country group 1. In Austria, Germany, Greece, and Italy—for which we find a training premium equal to 19.3 %, see Table 2—in fact, wages of trained workers were 10 % higher (estimate significant at 10 % level) than those of similar untrained workers at the time they undertake training. This indicates that training premiums found for this group of countries can be overestimated. The next subsection presents results of further checks on training endogeneity.

Selection into training: cognitive abilities and childhood circumstances

Earlier, we pointed out that the OLS estimates for training reported in Table 2 could be biased upwards due to

unobserved ability. Cognitive abilities are an important component of abilities and are rewarded in the labour market with higher earnings (Case and Paxson 2008). In this subsection, we control for cognitive abilities in the econometric model and check to what extent the estimated training premium is sensitive to these changes.

A ‘Cognitive Functioning’ module is included in SHARE that reports the results of simple tests of orientation in time, such as asking the interview date; verbal fluency, as in counting the number of words that can be named in a minute; memory, recalling as many words as possible from a 10-word list; and numeracy, testing daily life numerical calculations (for more details, see, e.g. Christelis et al. 2010). Starting from these tests, we build an index of cognitive abilities (Leist et al. 2013). This index is given by the average value of the (standardised) results of the verbal fluency, memory, and numeracy tests in SHARE Wave 1. We exclude tests of orientation in time since they show too little variation in our sample of workers. The higher the value of the index, the higher the cognitive abilities.

Cognitive abilities, such as those measured in SHARE Wave 1, may be innate and partly the result of an ageing process. Many empirical studies indicate the important role of lifecycle decisions and factors (education, occupational choices, lifestyles, and chronic diseases) in the process of skill formation (for a survey, see Mazzonna and Peracchi 2012). According to models of skill formation *à la* Cunha

Table 4 Training wage premiums—Robustness checks: controlling for cognitive abilities

Variables	All countries				Country groups 2 & 3			
	(a)	(b)	(c)	(d)	(a)	(b)	(c)	(d)
Training	0.069*** (0.024)	0.065*** (0.024)	0.056** (0.024)	0.055** (0.024)	−0.029 (0.024)	−0.031 (0.024)	−0.027 (0.024)	−0.027 (0.024)
Cognitive abilities Index		0.026* (0.014)		0.008 (0.015)		0.052*** (0.013)		0.026* (0.014)
No education (ISCED 0)	−0.382*** (0.064)	−0.365*** (0.067)	−0.340*** (0.078)	−0.336*** (0.080)	−0.418*** (0.053)	−0.373*** (0.055)	−0.357*** (0.060)	−0.338*** (0.061)
Primary and lower-secondary education (ISCED 1-2)	−0.234*** (0.029)	−0.229*** (0.030)	−0.198*** (0.031)	−0.198*** (0.031)	−0.237*** (0.028)	−0.220*** (0.028)	−0.198*** (0.030)	−0.192*** (0.030)
Upper and post-secondary education (ISCED 3-4)	−0.181*** (0.023)	−0.177*** (0.023)	−0.156*** (0.024)	−0.154*** (0.024)	−0.172*** (0.023)	−0.159*** (0.023)	−0.136*** (0.024)	−0.129*** (0.024)
<i>Initial conditions</i>								
Relative school performance index			0.022* (0.012)	0.021* (0.012)			0.046*** (0.012)	0.042*** (0.012)
Few books at home			−0.067*** (0.021)	−0.066*** (0.021)			−0.033 (0.020)	−0.030 (0.020)
Breadwinner blue collar			−0.055*** (0.020)	−0.054*** (0.020)			−0.058*** (0.020)	−0.058*** (0.020)
Observations	1,788	1,783	1,702	1,698	1,314	1,309	1,249	1,245
R-squared	0.456	0.457	0.463	0.463	0.383	0.390	0.397	0.399

Source: (a) SHARE Wave 1-2, (b-e) SHARE Wave 1-2-3

Note: OLS estimates; Standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$; Country groups 2&3: FR, ES, DK, SE, BE, NL, CH; Additional controls: Log tenure, age, age squared, dummy variables for female, public sector country of residence, sector, occupation, and year of interview; Reference category: tertiary education (ISCED 5-6)

and Heckman (2007), however, the roots of cognitive abilities are found in the home environment and parental influences in childhood. In the ‘Childhood Circumstances’ module, SHARELIFE includes a set of variables concerning individuals’ early-life conditions. It has been postulated that these variables may suffer from recall bias, that is, individuals may not remember events far in the past so well. Havari and Mazzonna (2014) show that these variables are characterised by a good level of internal and external consistency. From these, we select the following four: (i) performance in math relative to other children in the same class (from one, denoting ‘much better’, to five, for ‘much worse’), (ii) performance in language (same answer format as in (i)), (iii) the number of books at home (from one, for ‘none or very few’—i.e. zero to 10 books—to five, ‘enough to fill two or more bookcases’—i.e. more than 200 books), and iv) the occupation of the main breadwinner in the household (ISCO one-digit groups). All these variables refer to the time when the individual was aged 10 (“initial conditions”). From (i) and (ii), we build an index for self-reported relative school performance (*RSP index* henceforth). The literature (e.g. Blackburn and Neumark 1992) has often used intelligence quotient scores

as a proxy for unobserved ability when estimating wage equations in the absence of panel data (a so-called plug-in solution to the omitted variable problem in OLS regressions). Our *RSP index* takes on three possible values: two, if individuals report having performed better or much better than their schoolmates in both math and in their own language; one if individuals report having performed better or much better in either math or language; and zero otherwise. From (iii) and (iv), we build two dummy indicators of SES in childhood: *few books at home* (equal to one if the number of books was ‘none or very few’ to ‘enough to fill one shelf’ and zero otherwise) and *breadwinner blue collar* (equal to one for ISCO codes 6–9 and zero otherwise).

We already documented sizeable attrition between SHARE Waves 1 and 2. Making use of SHARELIFE introduces additional attrition, since we need respondents to be interviewed in both Waves 1 and 2 and in Wave 3. Due to this, the sample size reduces from 2,312 (Waves 1 and 2) to about 1,700 observations (Waves 1–3). In the resulting sample, the *cognitive abilities index* varies quite substantially: It has a mean equal to about zero and a standard deviation of 0.71. It is significantly higher for

trained workers (mean 0.23, s.e. 0.04) than for untrained workers (mean -0.04 , s.e. 0.02). The *RSP index* is also much higher for trained workers (mean 1.02, s.e. 0.04) than for untrained workers (mean 0.80, s.e. 0.02). A total of 32 % of trained workers had few books at home in childhood, compared to 53 % for untrained workers. Smaller differences between trained and untrained workers are observed when looking at whether the breadwinner at home was a blue-collar worker (61 vs. 65 %).

Table 4 reports OLS estimates of alternative specifications estimated on the subsample of SHARE Waves 1–3. It separately displays results for all countries (left panel) and for the subsample of countries for which the sensitivity analysis reported in the previous subsection does not indicate problems of training endogeneity or attrition/sample selection bias, i.e. country groups 2 and 3 (right panel). Given the above-described problem of attrition stemming from the use of SHARELIFE, we re-estimate our base specification without cognitive abilities measures on the resulting sample (specifications (a)). We then add to the model proxy variables for cognitive abilities and/or childhood circumstances (specifications (b–d)).

We first comment results for all countries. The coefficient for training increases somewhat when the base specification is estimated on this smaller sample: 0.069 (cf. with 0.063 in Table 2). Specification (b) includes the *cognitive abilities index* as an additional explanatory variable (5 observations have missing values for the cognitive abilities index). The estimated parameter for training decreases only negligibly (0.065). Specification (c) includes “initial conditions” in place of the *cognitive abilities index* variable. The estimated coefficient for the *RSP index* (0.022) implies that individuals with the highest index value (= 2) earn about 5 % higher hourly wages than those with the lowest one (= 0). Having had few book at home is associated with much lower hourly wages (-6.7 %) than having had more books. A similar result is found for being the child of a blue-collar worker (-5.5 %) instead of a white-collar worker. These findings are in line with those of Flores et al. (2013b), who quantify the association between childhood SES and annual earnings over the lifecycle, using SHARELIFE. The estimated training premium decreases to some extent, but not dramatically (from 6.9 to 5.6 %), once initial conditions are controlled for. Interestingly, the estimated formal education premiums—measured by the coefficients of the ISCED educational level dummy variables—also decrease once childhood SES variables and *RSP index* are controlled for. Finally, specification (d) includes both the *cognitive abilities index* and initial conditions. As expected, the coefficient for the *cognitive abilities index* loses significance (cf. specifications (b) and (d)); the coefficient for training is almost the same as in specification (c).

Results for country groups 2 and 3 (Table 4; right panel) show that the coefficient for training is not significantly different from zero, regardless of whether or not cognitive abilities measures and/or childhood circumstances are included in the model (cf. specifications (a) to (d)). A similar result of no training wage premium for these country groups is obtained by estimating on SHARE Waves 1 and 2 a model—alike specification (b) in Table 2—in which the training dummy is interacted with a dummy for belonging to country group 2 or 3 (results available upon requests).

On the one hand, the above-illustrated findings provide credibility to the OLS estimate of the average training premium (6.3 %) reported in Table 2. Nonetheless, many unobserved factors may easily lead to biased OLS estimates. Noncognitive abilities such as perseverance, motivation, time preference, risk aversion, self-control, and preference for leisure, for instance, are known to have a direct impact on wages, even after controlling for education (see Cunha and Heckman 2007 and references therein). We do not have any measure of noncognitive abilities in SHARE, and we are not able to estimate a panel data model that would allow us to get rid of time-invariant unobserved factors. Empirical studies such as that of Basanini et al. (2007), who estimate both OLS and fixed effects models, typically show that the latter result in much lower (or even no) returns to training than the former models do. On the other hand, findings of Table 4 highlight that such average training premium of 6.3 % is determined by the high premium found for Austria, Germany, Greece, and Italy for which we previously reported evidence of possible endogeneity and attrition/self-selection biases.

Conclusions

There is a growing consensus among the research and policy community on the importance of training investments in keeping older workers' skills updated. Despite the interest in these issues, the empirical evidence on the effect of training on older workers' productivity is scarce, as are suitable data. Using wages as a proxy for productivity, this paper investigates the relationship between training undertaken by European older workers and their wages, adopting an international perspective by relying on SHARE data.

We find that, for employees aged 50 and above residing in one of the 11 European countries analysed, undertaking training activities is associated with 6.3 % higher wages. Training wage premiums are highly heterogeneous across countries: they are highest in Austria, Germany, Greece, and Italy (19 %) and are about half that in France and

Spain (9 %). No premium is found for Denmark, Sweden, Belgium, the Netherlands, and Switzerland. Furthermore, we provide descriptive evidence that older workers undertaking training have lower probabilities to become unemployed or to retire.

We tested whether our results are affected by attrition, sample selection (exclusion of individuals retiring or becoming unemployed between waves), and unobserved ability biases. In terms of attrition and sample selection, we found that our estimation sample for the highest training premium countries excludes trained individuals with lower wages at the time of training. The problem of transition out of employment is more severe than that of attrition. The sign of the bias created by this exclusion is not certain. One may expect that individuals undertaking training and retiring would have had returns to training lower than those remaining employed. In this case, exclusion of these individuals from the sample biases our results upward.

The importance of unobserved ability bias is not clear. On the one hand, results seem to be robust to controlling for cognitive abilities (as measured at old age) and for their early-life determinants. On the other hand, for the highest training premium countries, we found evidence of training endogeneity: individuals undertaking training had on average higher wages at the time of the training. It is likely that unobserved ability is determining both training participation and wages. This indicates that training premiums found for this group of countries can be overestimated. Unfortunately, we are not able to handle the issue of training endogeneity with the data at our disposal: SHARE data neither provide information on noncognitive ability, nor allow us to estimate panel data models.

Our results indicate that for older workers residing in many European countries including Denmark, Sweden, Belgium, the Netherlands, and Switzerland there is no training wage premium. Very high premiums are instead found for Austria, Germany, Greece, and Italy. It is however likely that these high premiums are overestimated, and therefore they should be re-evaluated once more informative data becomes available.

Acknowledgments This research has been supported by the European Commission, Seventh Framework Programme, ‘Employment 2025: How will multiple transitions affect the European labour market’ (NEUJOBS, grant agreement no: 266833). The authors would like to thank the Guest Editor M. Myck, the Editor H. Litwin, two anonymous Reviewers, R. Alessie, A. Brugiavini, M. Flores, E. Fontaíña, E. Havari, K. Henkens, R. Mendes, E. Meschi, G. Pasini, A. Ruzik, I. Styczyńska, A. Tisch, and participants of the Third International Workshop on the Socio-Economics of Ageing (IWSEA), the Health, Education and Retirement over the Life Cycle (HERO13) conference and the Amsterdam Institute for Advanced Labour Studies seminar (AIAS-UvA) for their helpful comments. Michele Belloni is also affiliated with CERP-Collegio Carlo Alberto. The usual disclaimers apply.

This paper uses data from SHARE Wave 1 and 2 release 2.5.0, as of May 24th 2011 and SHARELIFE release 1, as of November 24th 2010. The SHARE data collection has been primarily funded by the European Commission through the 5th Framework Programme (project QLK6-CT-2001-00360 in the thematic programme Quality of Life), through the 6th Framework Programme (projects SHARE-13, RII-CT-2006-062193, COMPARE, CIT5-CT-2005-028857, and SHARELIFE, CIT4-CT-2006-028812) and through the 7th Framework Programme (SHARE-PREP, N° 211909, SHARE-LEAP, N° 227822 and SHARE M4, N° 261982). Additional funding from 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 and OGHA 04-064) and the German Ministry of Education and Research as well as from various national sources is gratefully acknowledged (see www.share-project.org for a full list of funding institutions).

References

- Acemoglu D, Pischke S (1999) Beyond becker: training in imperfect labor markets. *Econ J* 119:F112–F142
- Bassanini A (2006) Training, wages and employment security: an empirical analysis on European data. *Appl Econ Lett* 13(8):523–527
- Bassanini A, Booth A, Brunello G, De Paola M, Leuven E (2007) Workplace training in Europe. In: Brunello G, Garibaldi P, Wasmer E (eds) *Education and training in Europe*. Oxford University Press, Oxford, pp 8–13
- Becker G (1964) *Human capital: A theoretical and empirical analysis, with special reference to education*. The University of Chicago Press, Chicago
- Becker G (1993) *Human capital: A theoretical and empirical analysis, with special reference to education*, 3rd edn. The University of Chicago Press, Chicago
- Beckett S, Gould W, Lillard L, Welch F (1988) The panel study of income dynamics after fourteen years: an evaluation. *J Labor Econ* 6(4):472–492
- Blackburn MK, Neumark D (1992) Unobserved ability, efficiency wages, and interindustry wage differentials. *Q J Econ* 107(4):1421–1436
- Booth A (1991) Job-related formal training: who receives it and what is it worth? *Oxford B Econ Stat* 53:281–294
- Booth A (1993) Private sector training and graduate earnings. *Rev Econ Stat* 75(1):164–170
- Budria S, Pereira PT (2007) The wage effects of training in Portugal: differences across skill groups, genders, sectors and training types. *Appl Econ* 39(6):787–807
- Case A, Paxson C (2008) Stature and status: height, ability, and labor market outcomes. *J Polit Econ* 116:499–532
- Christelis D, Jappelli T, Padula M (2010) Cognitive abilities and portfolio choice. *Eur Econ Rev* 54:18–38
- Cunha F, Heckman J (2007) The technology of Skill Formation. *Am Econ Rev* 97:31–47
- Dearden L, Reed H, Van Reenen J (2006) The Impact of training on productivity and wages: evidence from British panel data. *Oxford B Econ Stat* 68(4):397–421
- Dostie B, Léger PT (2011) Firm-sponsored classroom training: is it worth it for older workers? CIRPEE Working Paper 11–36
- Flores M, Kalwij A (2013) What do wages add to the health-employment nexus? Evidence from older European workers. Documentos de trabajo—Análise Económica 0054, IDEGA—Instituto Universitario de Estudios e Desenvolvimento de Galicia

- Flores M, García-Gómez P, Kalwij A (2013b) Early life circumstances and life-cycle labor market outcomes. In: Flores M essays on early life circumstances, health and labor market outcomes in Europe. Dissertation, chapter 4
- Fouarge D, Schils T (2009) The effect of early retirement incentives on the training participation of older workers. *Labour* 23:85–109
- Fougère D, Goux D, Maurin E (2001) Formation continue et carrières salariales. une évaluation sur données individuelles. *Annales d'Économie et de Statistique* 62:49–69
- Gerfin M (2004) Work-Related Training and Wages: an empirical analysis for male workers in Switzerland. IZA DP 1078
- Goux D, Maurin E (2000) Returns to firm-provided training: evidence from French worker-firm matched data. *Labour Econ* 7(1):1–19
- Gruber J, Wise DA (2004) Introduction and summary. In: Gruber J, Wise DA (eds) *Social security programs and retirement around the world: Micro-estimation*. University of Chicago Press, Chicago, pp 1–40
- Havari E, Mazzonna F (2014) Can we trust older people's statements on their childhood circumstances? Evidence from SHARELIFE. *Eur J Pop* (forthcoming)
- Lang J (2012) The aims of lifelong learning: Age-related effects of training on wages and job security. SOEPpapers on multidisciplinary panel data research, DIW Berlin, 478/2012
- Leist AK, Glymour MM, Mackenbach JP, van Lenthe FJ, Avendano M (2013) Time away from work predicts later cognitive function: differences by activity during leave. *Ann Epidemiol* 23(8):455–462
- Mazzonna F, Peracchi F (2012) Ageing, cognitive abilities and retirement. *Eur Econ Rev* 56(4):691–710
- Mühler G, Beckmann M, Schauenberg B (2007) The returns to continuous training in Germany: new evidence from propensity score matching estimators. *RMS* 1:209–235
- OECD (1999) Training of adult workers in OECD countries: Measurement and analysis. *Employment Outlook*, OECD, Paris, chapter 3
- OECD (2004) Improving skills for more and better jobs: does training make a difference? *Employment outlook*, OECD, Paris, Chapter 4
- OECD (2006) *Live Longer. Work longer*, Paris
- Pischke JS (2001) Continuous training in Germany. *J Pop Econ* 14:523–548
- Schöne P (2004) Firm-financed training: firm specific or general skills? *Empir Econ* 29:885–900
- Warr P, Fay D (2001) Short report: age and personal initiative at work. *Eur J Work Organ Psy* 10(3):343–353
- Wooldridge JM (2002) *Econometrics Analysis of Cross Section and Panel Data*. The MIT Press, Cambridge, Massachusetts