

Discussion Paper No. 842

**BOOKS ARE FOREVER:
EARLY LIFE CONDITIONS,
EDUCATION AND LIFETIME EARNINGS
IN EUROPE**

Giorgio Brunello
Guglielmo Weber
Christoph T. Weiss

May 2012

The Institute of Social and Economic Research
Osaka University
6-1 Mihogaoka, Ibaraki, Osaka 567-0047, Japan

Books are forever: Early life conditions, education and lifetime earnings in Europe¹

by

Giorgio Brunello
Guglielmo Weber
Christoph T. Weiss

We estimate the effect of education on lifetime income in Europe, by distinguishing between individuals who lived in rural or urban areas during childhood and between individuals who had access to many or few books at age ten. We instrument years of education using reforms of compulsory education in nine different countries, and find that individuals in rural areas were most affected by the reforms. Among those affected, individuals with many books at home at age ten enjoyed substantially higher returns to their additional education. We argue that the long – lasting beneficial effects of having books at home are due to the cultural environment in the household and the development of cognitive skills rather than to the presence of short - term liquidity constraints.

JEL Code: J24

Keywords: lifetime earnings, returns to education, Europe

¹ Brunello: University of Padua, CESifo and IZA; Weber: University of Padua, CEPR and IFS; Weiss: Antonveneta Fellow, University of Padua. We thank Tom Crossley, Margherita Fort, Richard Freeman, Lisa Kahn, Peter Kuhn, Mario Padula, Luigi Pistaferri, Alfonso Rosolia, Richard Spady and participants at seminars in Bologna, Cambridge (RES), Kyoto (TPLS), Koç University, Groningen and Rome (Brucchi Luchino) for comments and suggestions. A previous version of this paper has been circulated with the title “The effects of schooling on lifetime earnings”. The authors gratefully acknowledge financial support from the Antonveneta Centre for Economic Research (CSEA) and Fondazione Cariparo. This paper uses data from SHARELIFE release 1, as of November 24th 2010 and SHARE release 2.5.0, as of May 24th 2011. 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-I3, RII-CT-2006-062193, COMPARE, CIT5-CT-2005-028857, and SHARELIFE, CIT4-CT-2006-028812) and through the 7th framework programme (SHARE-PREP, 211909 and SHARE-LEAP, 227822). Additional funding from the U.S. National Institute on Aging (U01 AG09740-13S2, P01 AG005842, P01 AG08291, P30 AG12815, Y1-AG-4553-01 and OGHA 04-064, IAG BSR06-11, R21 AG025169) as well as from various national sources is gratefully acknowledged. The usual disclaimer applies.

1. Introduction

At least since Mincer (1974) many economists have estimated the returns to education. The vast and ever growing literature in this area has been recently reviewed by Card (2001) and Heckman, Lochner and Todd (2006). In this paper, we depart from standard practice in various ways. First of all, we estimate the effect of education on lifetime earnings, not just current earnings. Secondly, we distinguish between individuals who lived in rural or urban areas during their childhood (at age ten) – an important distinction given that costs and foregone earnings of attending school were (maybe still are) likely to be higher for children living on farms or remote agricultural villages. Last, but by no means least, we also distinguish between individuals who had access to many or few books at age ten.

We provide evidence that a sizeable fraction of 50+ Europeans grew up in a household with less than a shelf of (non-school) books, and we show that the returns to education for individuals brought up in such households were much lower than for the luckier ones who had more direct access to books. In this sense we claim that books – like diamonds – are forever. Finally, we acknowledge that books at age ten could capture parental family economic resources or parental care early in life, but provide evidence that the latter is the more likely reason why books matter (in rural areas).

The empirical literature in labour economics typically estimates returns to education by using current rather than lifetime income. This practice has been challenged on the grounds that – when age-earnings profiles are not parallel with respect to educational attainment – a better measure of economic success is lifetime earnings or lifetime income. Figure 1 shows the age-earnings profiles of males from age 25 to 55 in nine European countries for which we have data, using the residuals from regressions on country and cohort dummies. The profiles exhibit the familiar concave shape. For each age in the relevant range, we also plot in Figure 2 the vertical distance of log earnings for individuals with education above (or equal to) their country-specific median education and individuals with education below the median.² This distance declines with age, with the possible exception of the final 5-6 years. We infer from this that age-earnings profiles are not parallel in education but converge over time.

This visual evidence suggests that estimates of the returns to education should be based on lifetime rather than on current income. Recent research confirms our visual inspection and shows that the returns to education based on annual earnings are significantly biased compared to those based on lifetime earnings, particularly if the sample includes many older workers (see Haider and Solon, 2006 and Bhuller Mogstad and Salvanes, 2011). This evidence, however, relies on administrative data from

² As in Figure 1, we use the residuals from regressions on country and cohort dummies.

only two countries – the US and Norway. We show that similar results hold in a broader context.

We estimate the returns to education in a number of European countries using a rich data set which contains detailed retrospective information on earnings, pensions and several variables of interest, including childhood characteristics. The European countries covered in our study are: Austria, Belgium, the Czech Republic, Denmark, France, Germany, Italy, the Netherlands and Sweden. The data are drawn from the third wave of the Survey of Health, Ageing and Retirement in Europe (SHARE).

In line with previous literature (see for instance Acemoglu and Angrist, 2001, Oreopoulos, 2006, and Pischke and van Wachter, 2005), we recognize that education is a choice variable and use the exogenous variation of compulsory years of education across countries and cohorts to identify the causal effect of years of education on current and lifetime earnings. We contribute to this literature by allowing compulsory education to have heterogeneous effects on current education, which vary with whether affected individuals lived in a rural or in an urban area at age ten.

We estimate returns to education using lifetime earnings, current (or last) earnings and earnings in the first job. In line with the evidence from Norway presented in Bhuller et al. (2011), we show that - on average - an additional year of education increases lifetime income by almost 9 percent. This is reassuring, given that the Norwegian earnings data are less affected by measurement error than survey data. We also find that the life cycle bias induced by using current rather than lifetime earnings is minimized at around 35 years, an age very similar to the one estimated by Bhuller et al (2011) and in line with the international evidence reviewed by Brenner (2010).

We find that compliers – defined as the individuals induced to increase their educational attainment because of the exogenous change in minimum school leaving age – lived in the rural areas of Europe during their childhood. As suggested by Lochner and Monge-Naranjo (2011), these individuals had relatively low education either because of liquidity constraints or because they shared a high distaste for schooling and a high opportunity value of time. Using information on the number of books in the household at age ten, we show that compliers with few books at home have enjoyed markedly lower returns to education than compliers with many books. This result suggests that early life conditions have long-lasting effects on individual welfare, and adds to the growing literature on the importance of early life interventions, which finds, for instance, lower returns to college for individuals who grew up in disadvantaged households (see among others Cunha and Heckman, 2007 and Heckman, 2000).

The paper is organized as follows. The next section presents the data and describes how we

compute individual measures of lifetime earnings. This section also contains an explicit test of the hypothesis that age-earnings profiles are parallel by educational attainment. Section 3 introduces the empirical model. In section 4 we discuss the effects of compulsory school reforms on educational attainment in the European countries for which we have data. Section 5 present our estimates of the returns to education using lifetime earnings. Section 6 considers how differences in early life conditions affect these returns and Section 7 presents a discussion of reasons why the number of books in the household at age ten matters. The last section concludes.

2. The data

One of the main reasons why researchers estimate returns to schooling using current rather than lifetime income is that longitudinal data on earnings which span entire working lives are seldom available. The few studies focusing on lifetime earnings use either administrative or survey data. Haider and Solon (2006) use Social Security earnings histories of participants in the U.S. Health Retirement Study (HRS) for the period 1951-1991 and find that using current rather than lifetime earnings to estimate returns to education generates a life-cycle bias.³ Heckman et al. (2006) use US Census data from 1940 to 1990 and reject the hypothesis of parallel experience-log earnings profiles for whites during all years except 1940 and 1950. Böhlmark and Lindquist (2006) use longitudinal data from the Swedish Level of Living Survey and data from LINDA (Longitudinal Individual Database for Sweden); Brenner (2010) use the German VVL longitudinal survey, which covers a sample of pension recipients born between 1939 and 1974; finally Bhuller et al. (2011) use Norwegian administrative data. All these studies confirm that earnings profiles are not parallel with respect to education.

In this paper, we use the Survey of Health, Ageing and Retirement in Europe (SHARE), a multidisciplinary and cross-national European data set containing current and retrospective information on labour market activity, retirement, health and socioeconomic status of more than 25,000 individuals aged 50 or older. We use all three waves of the survey, and in particular the third wave, SHARELIFE, which contains detailed retrospective life and labour market histories. We focus on males because of the problems associated with female labour force participation and exclude the self-employed and people who have worked less than 5 years.⁴ In SHARELIFE, survey participants are asked to report the amount they were paid monthly after taxes each time they started an employment spell. They are also

³ In their data, earnings are only available for jobs covered by U.S. Social Security. In some years, a large proportion of the sample is right-censored because of the Social Security taxable limit for that year.

⁴ Murphy and Welch (1990) also exclude the self-employed in their analysis of age-earnings profiles.

asked the monthly net wage in their current job (if they are still working) and the monthly net wage at the end of the main job in their career (if they have already retired). For wages and other benefits to be comparable across time and country, we transform them into 2006 Euro using PPP exchange rates and CPI indices.

As described in detail in Appendix B and in Weiss (2012), we use current and retrospective information on earnings, jobs and labour market experience to construct a measure of lifetime (or permanent) earnings, which we define as the income flowing from the asset value of working at age ten. In short, we regress current wages on labour market experience, a rich set of controls, which include education, occupation, sector of activity, cohort and country effects and economic conditions at age ten, and the interactions of these controls with experience. We use the estimated coefficients and the first wage in each job to generate both the final wage in the job and within-job earnings growth. A validation study which uses the German Socio-economic Panel suggests that our procedure is quite accurate (see Appendix C).

The asset value of working at age ten is the discounted sum of wages and other benefits received from age ten up to retirement, using a discount rate of 2%.⁵ We also construct a measure of lifetime income which includes actual and/or expected pension income until death, using cohort and country specific mortality tables. This measure seems appropriate as pension benefits are typically affected by either earnings or contributions, and can be seen as part of the return to the investment in education.

Our dataset has the advantage that it covers nine European countries, which gives a broader perspective on European earnings than previous studies in this area, and the potential drawback that it uses long recall data. These data are subject to measurement error, possibly not of the classical type.⁶ However, the validation studies by Garrouste and Paccagnella (2011) and Havari and Mazzonna (2011) find that recall bias is not severe in SHARELIFE data, arguably because of the state-of-the-art elicitation methods used: respondents are helped to locate events along the time line, starting from domains that are more easily remembered, and then asked progressively more details about them.⁷ It is also reassuring for us that our estimates are in line with estimates obtained using administrative data, as

⁵ Haider and Solon (2006), Böhlmark and Lindquist (2006) and Brenner (2010) also assume a constant real interest rate of 2% to construct a measure of lifetime income. Bhuller et al. (2011) use instead an interest rate of 2.3%.

⁶ Several studies find a negative correlation between the true value of earnings and measurement error (e.g. low earners tend to over-report their earnings in survey data). Pischke (1995) uses the Panel Study of Income Dynamics Validation Study and finds a negative correlation for hourly earnings but no correlation for monthly earnings.

⁷ In their study of the effect of childhood environment on economic and social outcomes for Yemenites who immigrated to Israel in 1949, Gould, Lavy and Paserman (2011) find evidence that retrospective data on childhood environment from more than 50 years ago is of high quality.

discussed below.

Our final sample consists of 5804 individuals born between 1920 and 1956 and residing in Austria, Belgium, the Czech Republic, Denmark, France, Germany, Italy, the Netherlands and Sweden. We are forced to exclude data for Greece, Spain and Switzerland because of the selected estimation strategy, which uses the exogenous within and cross-country variations in minimum schooling laws to identify the causal relationship between education and lifetime income. In the excluded countries, the existing variations in compulsory schooling occur too late to allow us to identify a pre-treatment and a post-treatment sample of cohorts.⁸

Table 1 provides some descriptive statistics on our sample. The average lifetime income flowing from the estimated asset value of working at ten is equal to 8452 real euro when this value is net of pension benefits and to 10824 real euro when pension benefits are included. Median years of schooling and of compulsory schooling are equal to 11 and 8 respectively. Average age at the time of the interview and years of work are equal to 66.97 and 36.53. The table shows that almost 30% of the individuals in the sample are still working and that they have had on average three different jobs during the career. More than 40% of the individuals lived in a rural area or a village during their childhood and 40% lived at age ten in a household with less than 10 books. Only 22% lived in a rural area and had less than ten books at age ten (the correlation between these two indicators is relatively low, at only 0.20).

2.1 Are earnings profiles parallel?

For each individual in the sample we estimate both his lifetime earnings and the annual sequence of earnings from labour market entry until retirement. Using this retrospective panel, we investigate whether earnings profiles are parallel with respect to educational attainment by estimating the following regression

$$\ln W_{it} = \beta_0 + \beta_c + \beta_d + \beta_1 A_{it} + \beta_2 A_{it}^2 + \beta_3 S_i + \beta_4 S_i A_{it} + \varepsilon_{it} \quad (1)$$

⁸ In Belgium, we only keep in the sample the individuals who went to school in Flanders, because the school reform of 1953 took place in this region and not in the rest of the country. We do not consider individuals from Poland because of unreliable income data. Trevisan et al (2011) argue that Poles answering the SHARE questionnaire got confused between new and old Zloty around the devaluation in 1995 and misreported earnings during the high inflation of the 80s and 90s. For Germany, we only include individuals from West Germany. Our empirical results are qualitatively unchanged if we exclude from the sample the Czech Republic, which belonged to the Soviet bloc before 1989.

where A denotes age, S years of schooling, W annual earnings, i is for the individual, t for time in the labour market, β_c is a vector of country dummies, β_d a vector of cohort dummies and we exclude the individuals who have experienced unemployment during the sample period.⁹

Earnings profiles are parallel with respect to education if $\beta_4 = 0$. Figure 2 plots the vertical distance between log earnings when education is equal to or above the country-specific median (S_a) and log earnings when education is below this median (S_b). This distance is equal to $[S_a - S_b](\beta_3 + \beta_4 A)$, and is constant if $\beta_4 = 0$. Inspection of the figure suggests that this vertical distance declines from age 26 until age 50 and mildly increases thereafter. It is worth stressing that a more detailed break-down of education categories confirms convergence for lower education levels, but shows divergence for higher education levels, particularly between college and high school graduates. Given the low proportion of college graduates in our sample, convergence prevails.¹⁰

We estimate equation (1) in first differences. By so doing, we difference out time invariant unobserved individual heterogeneity, which is correlated with education. Our results for the samples of individuals aged 21 to 55 and 25 to 55 are reported in the first two columns of Table 2. There is evidence of converging profiles ($\beta_4 < 0$ and statistically significant). According to these estimates, the returns to schooling accruing to individuals aged 55 are 1.78 percentage points lower than the returns accruing to individuals aged 21. When we restrict the sample to the age range 21 to 50 or 25 to 50, we find that the magnitude of the estimated coefficient on age is higher (in absolute terms), in line with what we observe in Figure 2. Overall, we interpret this evidence as suggesting that earnings profiles are not parallel by education but mildly converge over time.

Following Haider and Solon (2006), we define the life cycle bias (LCB) at age $a+x$ as

$$LCB_{a+x} = \frac{\partial \ln W_{a+x}}{\partial S} - \frac{\partial \ln I_a}{\partial S} \quad (2)$$

where W_{a+x} is real wage at age $a+x$, I_a is lifetime income evaluated at age a and S denotes educational attainment. When age-earnings profiles are parallel with respect to education, this bias is equal to zero for any value of x . When they are not parallel, the bias can be positive, negative or equal to zero at

⁹ See Bhuller et al, (2011), for a similar choice. This exclusion is dictated by our method to compute the life cycle bias. Estimates including those who have experienced unemployment are very similar to the ones presented in this paper (results available from the authors upon request).

¹⁰ See Appendix D for further details.

$x = x^*$.

Following the method suggested in Appendix A, we estimate the critical value x^* in our data, and find that, when schooling is at its mean level, this value is equal to 14.53 years. Since a is equal to 21, $a + x^*$ is equal to 35.53, a number very similar to the one estimated by Bhuller et al. (2011) using Norwegian administrative data and in line with the evidence for the US, Sweden and Germany.¹¹ We consider the similarity of our results with those found in longitudinal and administrative data reassuring for the quality of our data, which rely heavily on retrospective information.

3. The empirical strategy

We estimate the effects of education on lifetime earnings using the following empirical model

$$Y_i = \beta_1 S_i + X_i^T \beta_2 + U_i \quad (3)$$

$$S_i = Z_i^T \gamma_1 + X_i^T \gamma_2 + V_i \quad (4)$$

where Y denotes the logarithm of lifetime earnings, S is years of education, $X \equiv \{X_k\}_{k=1}^K$ is a vector of covariates, $Z \equiv \{Z_l\}_{l=1}^L$ a vector of instruments and U and V are disturbance terms. We pool data for the selected nine European countries and include in the vector X country fixed effects, cohort fixed effects and country-specific quadratic trends in birth cohorts. Country fixed effects control for national differences, both in reporting styles and institutions affecting lifetime income. As pointed out by Lochner and Monge-Naranjo (2011), country specific (quadratic) trends in the year of birth are required to avoid that we incorrectly attribute trends in earnings to school reforms.¹²

Since education is affected by individual unobserved ability, which also influences earnings, the covariance between the disturbances in equations (3) and (4) is unlikely to be zero. We address the endogeneity of education by instrumental variables. Instrument validity requires that the selected instrument affects individual earnings only indirectly by influencing years of schooling. Following an established literature¹³, we use the exogenous variation provided by changes of minimum school leaving age within and between countries to identify the causal relationship of education on earnings. This identification strategy is widely considered to be credible and has been extensively used in the

¹¹ Brenner (2010) reviews this literature and suggests that the critical age lies in the range 30 to 40.

¹² See also the discussion in Goldin and Katz (2003).

¹³ See, e.g., Oreopoulos (2006), Pischke and von Wachter (2008), Devereux and Hart (2010) and Lochner and Monge-Naranjo (2011) for a review.

literature. We apply this strategy to a multi-country setup, as in Brunello, Fort and Weber (2009) and Brunello, Fort, Schneeweis and Winter-Ebmer (2011), and exploit the fact that school reforms have occurred at different points in time and with varying intensity in several European countries.

Table 3 documents the reforms of minimum school leaving age which occurred in the European countries included in our sample from the 1930s until the late 1960s. For each reform, the table presents the year of the reform, the first birth cohort affected by the reform (or pivotal cohort), the change in the minimum school leaving age, the years of compulsory education, and the age at school entry.¹⁴ Compulsory years of schooling during the relevant sample period have broadly increased, from one year in Austria, Belgium and Germany to three years in France, Sweden, Denmark and Italy. In the Netherlands and the Czech Republic, compulsory years have temporarily declined but increased overall by two and one year respectively.¹⁵

4. The effect of compulsory education reforms on years of schooling

Suppose that years of compulsory education for cohort j are initially equal to YC_0 , which corresponds to the dashed vertical line in Figure 3. When a reform increases compulsory education to YC_1 – the continuous vertical line in the figure – compliance with the reform implies that all four hypothetical individuals in the figure (A, B, C and D) attain at least YC_1 . However, while individuals C and D would have attained higher education even in the absence of the reform, individuals A and B are forced by the reform to attain a level of education where their marginal costs (MC) are higher than their expected marginal revenues from the investment (MR). This second pair of individuals consists of compliers who have increased their education because of the reform and face relatively high marginal costs of education but markedly different expected returns.

Since individuals select their education by comparing costs to returns, the heterogeneity of outcomes described in Figure 3 depends on heterogeneous marginal costs and revenues. Consider first costs. The individuals in our sample were born between 1920 and 1956 - a period when the proportion of households living and working in rural areas was substantially higher than in Europe today. The

¹⁴ Years of compulsory education are computed using the country where the individual was living when the reform could have affected him and not the country where he is residing now. This is a potentially important issue for Germany where we use information at the state level and where people have moved across states during their life.

¹⁵ In the Netherlands, the 1942 reform (from 7 to 8 years) was enacted by German occupiers who wanted the Dutch youth to learn German. After the war, a law extended the increase to 9 years, but set its start in 1950 - so the legal limit was back to 7 from 1947 to 1949. See van Kippersluis, O'Donnell and van Doorslaer (2009).

SHARELIFE survey asks these individuals where they lived at age ten. It turns out that 42.8 percent of the sample was living in a rural area, this percentage being lowest in Sweden and the Netherlands and highest in Italy and Austria. We conjecture that for the children born in rural households during the selected period the direct and indirect costs of attending school were substantially higher than for children living in cities: child labour was fairly common in rural areas in Europe for those cohorts, and travelling to the nearest school was much more expensive for children living in remote villages.

Turning to expected marginal revenues, SHARELIFE asks the number of books in the place where the individual was living at age ten (excluding magazines, newspapers or school books), which can be considered as a proxy of two important drivers of returns, parental resources and skill formation early in life. Answers to this question are classified in five categories: none or very few (less than 10 books), a shelf of books (11 to 25 books), a bookcase (26 to 100 books), two bookcases (101 to 200 books), and more than two bookcases (more than 200 books). While more than 75% of the Italians report to have had none or very few books at age ten, this is the case for less than 25% of the people living in Czech Republic, Denmark or Sweden (see Table 4).¹⁶ Table 5 shows the proportion of individuals living in a rural area or a village during their childhood, classified by country and the number of books at age ten. For instance, 69.6 percent of the Austrians with very few books grew up in a rural area, compared to only 35.8 percent in the Netherlands. Unsurprisingly, there is a higher proportion of people who grew up in a rural area among those with very few books at home at age ten. At the same time, however, more than a quarter of the individuals with 101 to 200 books in the house grew up in a rural area. There is thus some association between living in a rural area and having very few books, but it is not very strong: the correlation coefficient is 0.19 overall, ranging from 0.08 in the Netherlands to 0.29 in Austria.

The heterogeneity of returns and the relatively high marginal costs suggest that, *ceteris paribus*, children living in rural areas had attained lower education than children living in cities, and were therefore exposed to a higher extent to reforms increasing the years of compulsory education. We identify these individuals with A and B in Figure 3. To verify this conjecture, we estimate equation (4) by including in the set of instruments the interaction of years of compulsory education with the dummy variable RURAL, an indicator equal to one if the individual lived in a rural area at age ten and to zero

¹⁶ Perhaps unsurprisingly, the figures are correlated with the predominant religion in these countries: individuals living in predominantly Catholic countries tend to have fewer books (with the notable exception of the Czech Republic). People from countries where there is a majority of Protestants (Denmark and Sweden) have more books. Finally, countries where there is an equal proportion of Catholics and Protestants (Germany and the Netherlands) are somewhat in the middle between the two extremes.

otherwise. If compliers are drawn to a larger extent among those living in rural areas during childhood, we expect this interaction term to attract a positive and statistically significant coefficient. Since growing up in a rural area could have affected lifetime earnings independently of education, we also add this dummy variable to the set of regressors in equation (3).

The results of our estimates are shown in Table 6, which is organized in six columns, one for the full sample and the remaining for different sub-samples based on the number of books in the household at age ten. Focusing on the first column, we find that living in a rural area or a village during childhood reduces educational attainment by 31.7%, a sizeable effect. In all remaining columns but the last, we find that years of compulsory education have a statistically significant effect on educational attainment only when interacted with the dummy *RURAL*, in support to our view that compliers are drawn *exclusively* among those who were living in rural areas at age ten. For these individuals, one additional year of compulsory schooling raises actual education by 0.2 to 0.7 years, depending on the number of books in the house. An exception to this pattern is the small group of individuals with more than two bookcases (more than 200 books) in the household at age ten. For these individuals, a reform of compulsory education has a negative effect on years of schooling. As argued by Angrist and Pischke (2009), the interpretation of IV estimates as local average treatment effects requires monotonicity. To preserve this property in our data, we focus from now on only on individuals with less than 200 books in the household. Table A.1 provides the same descriptive statistics as in Table 1 for the individuals with less than 200 books in the household at age ten.

5. The causal effect of education on lifetime earnings

We initially estimate equation (3) by ordinary least squares using as dependent variable the following four alternatives: the log of lifetime earnings net of pension benefits, the log of lifetime income with pension benefits, the log of the current wage (or the wage at the end of the main job in the career if the individual is retired), and the log of the wage in the first job. When pension benefits are funded exclusively from social security contributions, their addition to lifetime income does not generate any double counting if wages are net of these contributions, as it is in our case. We cannot exclude, however, that some pension benefits are funded by savings on net income. If this is the case, our measure of lifetime income which adds pension benefits may suffer from double counting. For this reason, we prefer to use lifetime earnings net of pensions as the dependent variable in our baseline estimates.

Table 7 reports our estimates and shows that returns to education range from 2.8% to 5.6%, depending on the definition of the dependent variable. There is also evidence that living in a rural area has a negative and statistically significant effect on all outcomes except for the first wage. We contrast these estimates with 2SLS (two stages least squares) estimates, which we obtain by instrumenting years of schooling with years of compulsory education and their interaction with the dummy RURAL. In all specifications we cluster standard errors by country and cohort, the dimensions of relevant variation for years of compulsory education (see Moulton, 1990). Table 8 presents our results: independently of the selected dependent variable, the F-test statistic for the inclusion of additional instruments in the first stage regressions is always above the rule of thumb value of 10, which allows us to reject the hypothesis that our instruments are weak. We estimate that an additional year of schooling in our sample increases lifetime earnings net of pension benefits by 8.4%. It is re-assuring that a very similar return (8.7%) was found by Bhuller et al. (2011), who use high quality Norwegian administrative data, which are less likely than our data to be affected by recall bias and measurement error.¹⁷

When we include pension benefits in the measure of permanent income, estimated returns decline slightly to 7.1%. If age-earnings profiles were parallel with respect to education, we should obtain similar estimates when we consider the wage in the first job and the current wage (or the wage at the end of the main job in the career if the individual is retired). This is not the case, however, as we find that returns are higher – at 10.7% – in the first job and lower – at 3.6% – in the current (or main) job.¹⁸ Our estimates confirm the evidence in Section 3 that age-earnings profiles are not parallel with respect to education but converge as individuals age in the labour market, a result in line with recent findings for the US by Heckman, Lochner and Todd (2008) but in contrast with the results for Norway by Bhuller et al. (2011).

In our regressions, IV estimates are larger in absolute value than the OLS estimates. This finding is fairly common in this literature, and has often been interpreted as evidence of the presence of liquidity constraints: compliers with high returns to schooling must have been excluded from higher education because of even higher costs of schooling.¹⁹ Yet Carneiro and Heckman (2002) warn against

¹⁷ When we cluster standard errors by country rather than by country and cohort, the second row in the table is very similar to Table 8 and is equal to (0.027), (0.025) and (0.038). Therefore, changes in the clustering variable do not affect qualitatively our results.

¹⁸ Since our individuals are aged 50+ at the time of the survey, their current job is towards the end of their working history. In the regressions using the wage in the first job and the wage in the current (or main) job as the dependent variable, we also include as covariates a set of dummies for the age when the job was started or ended.

¹⁹ Card (2001), Heckman (2000), Carneiro and Heckman (2002) and Brenner and Rubinstein (2011) tackle this issue. Heckman (2000) argues that skills beget skills and that cognitive ability develops early in life. Therefore, growing up in a family who is interested in the intellectual development of their children is crucially important. Poor access to opportunities

such an interpretation, which is based on the questionable assumption that ordinary least squares estimates measure the average treatment effect for the treated. In line with Carneiro and Heckman, we show in the next section that liquidity constraints cannot be the explanation for an important share of compliers, who earn very low returns from their additional education.

6. Early life conditions and returns to schooling

Early life conditions matter for individual development and labour market success. Cunha and Heckman (2007) show that ability gaps between individuals and across socioeconomic groups open up at early ages, for both cognitive and socio-emotional skills. Cognitive abilities become stable around the age of ten, suggesting that environmental conditions below this age are important and that early policy interventions pay off more than later interventions (Cunha, Heckman and Schennach, 2010). We measure early life conditions with the number of books available in the household when the individual was ten years old. As already reported in Table 4, Sweden has the lowest proportion of people with very few books (20.2%) and Italy the highest (77.3%). We consider individuals with at most 200 books in the household and estimate separate regressions for two sub-groups, one with 0 to 10 books and the other with 11 to 200 books.

Table 9 reports both OLS and 2SLS estimates of the returns to education for each sub-group, using lifetime income net of pension benefits as the dependent variable. Since the F-test statistic of the additional instruments in each sub-sample is below the critical value of 10, we also estimate the model by limited information maximum likelihood (LIML). The LIML estimator is median-unbiased in over-identified models.²⁰ When we treat education as endogenous, we find a sharp contrast between the two groups of individuals: whereas the returns to schooling are close to zero for the group with few books, there are large (18.1%) and positive returns for the group with more than 10 books in the household at age ten. These findings are confirmed by the LIML estimates, which are very close to the 2SLS

to learn early in life has negative long-term consequences that later policy interventions cannot easily undo. Carneiro and Heckman (2002) critically review the argument that high IV returns to education indicate the presence of binding credit constraints. They point out that heterogeneous opportunities can be considered as a type of long-term credit constraints, or “long term factors that promote cognitive and noncognitive abilities”, and that these are the major determinant of the relationship between parental income and education. Brenner and Rubinstein (2011) recognize the strength of Carneiro and Heckman’s argument but argue that the group of poor children is heterogeneous in terms of cognitive and non-cognitive abilities.

²⁰ See Angrist and Pischke (2009) for a discussion and comparison of the sampling properties of 2SLS and LIML estimators.

estimates.²¹

These estimates highlight the presence of substantial heterogeneity in the group of compliers, who have increased their education because of the reforms to compulsory education: the estimated returns to education for the group with more than 10 books in the household at age ten are more than five times as large as the relatively low returns earned by the group with few books. Going back to Figure 3, the former group corresponds to “individual B” and the latter group to “individual A”. Our estimates underscore the importance of early life conditions for the returns to education, and suggest that early interventions which improve learning in the first years of life may have large payoffs for less privileged individuals, as pointed out by Heckman, Malofeeva, Pinto and Savelyev (2010).

Table 10 reports the estimates of a more restrictive specification which uses the whole sample and adds to the regressors the interaction of years of education with the indicator of very few books at age ten.²² We also add to the list of instruments Z the interactions between the indicator of very few books and years of compulsory education and between rural area, the indicator of very few books and years of compulsory education. The estimates are in line with the results in Table 9: as indicated by the negative sign of the estimated coefficient associated to the interaction between education and the very few books indicator, an additional year of education yields a significantly lower return for those with very few books in the house at age ten than for the rest of the sample.

We estimate the same specification using as dependent variable log current earnings rather than lifetime earnings. Compared to the latter, the former does not rely on long recalls and on imputation, and is therefore less prone to be affected by measurement error. Since in all specifications the indicator of rural area is never statistically significant, we report in Table 11 the estimates when this variable is omitted from the list of regressors (but not of the instruments). Confirming the findings in Table 8, returns to education measured using current earnings are lower than returns measured using lifetime earnings. Reassuringly, we find that these returns are significantly higher for those with more books in the household at age ten.

²¹ To check whether these results do not depend on cultural difference across countries, we also define an alternative indicator of few books, which includes households with less than 25 books in the Czech Republic, Denmark and Sweden. As reported in Table 4, individuals in these countries have a higher than average number of books in the house at age ten. The estimates (not reported here) are very similar to those reported in Table 9.

²² This specification also includes the indicator of few books in the household and its interaction with the dummy variable `RURAL`.

7. Why do books at home at age ten matter for lifetime earnings?

One candidate reason why books at home matter for lifetime earnings is that they capture the effects of poor early economic conditions. If this was the case, including these conditions in the regression would affect in a significant way the impact of the number of books on lifetime earnings. An indicator of poor economic conditions is poor housing, which we measure with a dummy variable taking value 1 (and 0 otherwise) if the accommodation occupied by the household when the individual was aged ten lacked running water or an inside toilet.²³ Mazzonna (2011) argues that this variable is an asset indicator which proxies household long-run wealth.

In Table 12, we repeat the exercise performed in Table 10 but augment the specification with the interaction of years of education with the indicator of poor housing conditions. We also add to the list of regressors the indicator of poor housing conditions and its interaction with the dummy variable *RURAL* and augment the list of instruments with the interactions of poor housing conditions with years of compulsory education and of the dummy variable *RURAL* with poor housing conditions and years of compulsory education. In the model estimated by 2SLS, we find that, while the variables containing the indicator of very few books remain jointly significant and with coefficients similar to those in Table 10, the variables containing poor housing are not jointly significant. We interpret this as evidence that the number of books in the house is not a proxy of the economic conditions of the household when the individuals was aged ten.

Few books at home could proxy for poor health at age ten. Table 13 shows the correlations between the former variable and alternative measures of poor health when young, which include: ever missed school for more than a month because of health problems, serious illness, poor health, any vaccination, and regular visits to the dentist when young. Only for this last variable is there any evidence of a significant correlation with the number of books in the house. This correlation could reflect parental education: better educated parents buy more books and are more willing to screen and take preventive health care.

To further check whether the number of books simply capture the economic conditions of the household, we use the information on the occupation of the main breadwinner at age ten and classify occupations in 3 categories: white collar (legislator, senior official, manager, professional, technician or clerk); service worker (service worker, market sales worker, skilled agricultural or fishery worker, craft

²³ A similar indicator of housing conditions is used by Gould et al. (2011). Their variable is based on whether the individual lived in a house with running water, bathroom and electricity.

worker); elementary occupation (plant operator or assembler, elementary occupation). We then estimate lifetime earnings equations for each occupation group (see Table 14). Albeit a bit imprecise, the estimates do not vary significantly with the occupation of the father.

The evidence presented so far suggests that books at home capture the cultural background in the household and the development of cognitive skills rather than the presence of short-term liquidity constraints due to scarce financial resources. In the parlance of Carneiro and Heckman (2002), books at home may indicate the presence of long-term constraints. To further support this view, we look at international data on cognitive test scores, which typically include information on the number of books at home.

We draw our data from three different surveys: PIRLS 2006 for the reading test scores of primary school children aged 9 to 10, TIMSS 1995 for the math test scores of students aged 8 to 11, and PISA 2006 for the math and reading test scores of 15 years old pupils. We regress individual test scores on country dummies, measures of parental education and occupation, immigrant status, language spoken at home, gender, age and the number of books in the house. As reported in Table 15, there is clear evidence that books matter for cognitive development, even after conditioning for parental education (and employment).

Through its effects on the development of cognitive and socio-emotional skills, parental investment has been shown to be a key determinant of the economic and social success of children at an adult age (Cunha et al., 2010). We find that individuals with disadvantaged cultural background invest little in education (9.83 years on average for those with very few books at home and 12.66 year on average for those with more books at home). Figure 4 shows kernel density estimates of years of education for individuals with books and without. When forced to invest more by compulsory school leaving laws, these individuals earn low returns either because their cognitive ability is crystallized at a lower level or because they comply with the law by attending low quality education.

For our cohorts of interest, individuals with very few books and living in rural areas were much more isolated than nowadays. An alternative explanation for their lower returns could be that individuals chose education based on perceived rather than actual returns.²⁴ High paying jobs were located in cities but rural boys lived in communities where few adults had information on urban jobs. Perceived returns may have been much lower than actual returns, especially if productivity growth in urban jobs was under-estimated. In our data, we find that rural boys with books are more likely to move

²⁴ Using data from the Dominican Republic, Jensen (2010, p.517) suggests that perceived returns to schooling are inaccurate and underestimated, also because “individuals rely almost exclusively on the earnings of workers in their own communities in forming their expectations of earnings.”

to cities (47% versus 36%) and to have as their first job a white collar job (42% versus 18%). Not all people born in rural areas around WW2 could foresee the important, long-lasting positive macroeconomic shock of the post-war period as well as the increasing skill premia, especially if they were basing their expectations on the experience of the post-WW1 period.

The different migration pattern of rural boys into cities may very well be the key to understanding the massive difference in returns to education that we document in this paper. In Figure 3 both rural boys A and B are pushed to increasing their schooling to the point where the expected marginal return is lower than the marginal cost. However, it is most likely that type B actual marginal return (that we estimated at 18%) proved much higher than originally expected. A possible channel for this was migration into the fast growing cities. Rural youngsters who grew up in homes with books may have not realized how high the marginal return to education could be ex-ante, but ex-post they reaped the full benefits of their education by moving into fast-growing urban areas. We conjecture that the more cultured environment they grew up in made them more willing to take a chance and move to the city.²⁵ Less cultured boys, instead, went back to the countryside and to the standard farming jobs available there and failed to reap the benefits of their extra education.

8. Conclusions

In this paper we have investigated how lifetime income relates to education and socioeconomic background during childhood in a number of European countries using a rich data set containing detailed retrospective information on earnings, pensions and many variables of potential interest (including childhood characteristics). Our estimates suggest that an additional year of education increases on average lifetime earnings by almost 9 percent. These returns vary markedly with socioeconomic background early in life, and are significantly lower for those with very few books at home at age ten. Even though we cannot rule out that the presence of books at home capture educational attainment of the parents, which is not recorded in our data, we notice that evidence from recent cognitive test scores shows that number of books predict these scores even after controlling for parental education and occupation (see Table 15 for further details). Access to books when young seems to reflect home skill formation in cognitive and socio-emotional skills, something that has been emphasized as an important factor of economic success in life.

²⁵ For people who grew up in rural areas, the correlation coefficient between the indicator of very few books and whether they are still living in a rural area is positive and statistically significant at 1%.

Compulsory education has increased in most European countries after the Second World War (Murtin and Viarengo, 2008). Our findings suggest that among the individuals induced by these reforms to attain higher education only those with enough books at home have been able to reap significant private economic returns. A significant group of individuals, who lived in rural areas with very few books at home, attained higher education but low private returns. This might suggest that alternative education policies, targeted at reducing the marginal cost of education - such as education vouchers – could have been a more efficient way of increasing the education of individuals with potentially high returns. Yet this view fails to consider that education has both private and social returns. If additional education has substantial positive externalities – either because it reduces crime rates or because of productivity spillovers – these social returns may more than compensate the low private returns obtained by compliers with few books at home. In any case, our results speak clearly about the importance of early economic conditions, and of policies affecting these conditions, in line with the important findings of recent economic research in this field (Cunha et al., 2006).

Table 1: Descriptive statistics

Variable	Mean	Std.Dev.	Min	Max	Median
Lifetime income	10823.70	6091.20	485.39	39123.13	9712.618
Lifetime earnings net of pension	8352.35	5422.70	265.93	38253.88	7230.454
Years of education	11.75	3.98	2	25	11
Years of compulsory education	7.51	1.58	4	10	8
Age	66.97	8.75	52	89	66
Years of work	36.53	8.20	5	63	38
Number of jobs during career	3.12	2.04	1	18	3
Very few books at age ten	0.40	0.49	0	1	
Rural area or village during childhood	0.43	0.49	0	1	
Very few books x rural area	0.22	0.41	0	1	
Poor housing conditions at age ten	0.47	0.50	0	1	
Ever unemployed	0.09	0.29	0	1	
Retired	0.73	0.44	0	1	
Austria	0.04	0.19	0	1	
Belgium	0.12	0.32	0	1	
Czech Republic	0.12	0.32	0	1	
Denmark	0.13	0.33	0	1	
France	0.13	0.34	0	1	
Germany	0.09	0.28	0	1	
Italy	0.12	0.33	0	1	
Netherlands	0.14	0.35	0	1	
Sweden	0.11	0.32	0	1	
Sample size	5804	5804	5804	5804	5804

Table 2: First difference estimates of age earnings profiles. Longitudinal panel of individuals always employed from age 21 to age 55.

Variable	Age 21-55	Age 25-55	Age 21-50	Age 25-50
Age/1000	-0.524*** (0.048)	-0.436*** (0.048)	-0.642*** (0.063)	-0.546*** (0.064)
Years of education/1000	-0.435*** (0.055)	-0.389*** (0.055)	-0.482*** (0.062)	-0.433*** (0.062)
Constant	0.042*** (0.002)	0.037*** (0.002)	0.046*** (0.003)	0.042*** (0.003)
Sample size	80751	73905	68699	61853

Note: All regressions include country and country by age effects. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 3: Compulsory school reforms, by country

	Reform year	Pivotal cohort	Change in min. school leaving age	Years of compulsory education	Age at school entry
Austria	1962	1951	14 to 15	8 to 9	6
Belgium (Flanders)	1953	1939	14 to 15	8 to 9	6
Czech Republic	1948	1934	14 to 15	8 to 9	6
-	1953	1939	15 to 14	9 to 8	6
-	1960	1947	14 to 15	8 to 9	6
Denmark	1958	1947	11 to 14	4 to 7	7
France	1936	1923	13 to 14	7 to 8	6
-	1959	1953	14 to 16	8 to 10	6
Germany (Baden-Württemberg)	1967	1953	14 to 15	8 to 9	6
Germany (Bayern)	1969	1955	14 to 15	8 to 9	6
Germany (Bremen)	1958	1943	14 to 15	8 to 9	6
Germany (Hamburg)	1949	1934	14 to 15	8 to 9	6
Germany (Hessen)	1967	1953	14 to 15	8 to 9	6
Germany (Niedersachsen)	1962	1947	14 to 15	8 to 9	6
Germany (Nordrhein-Westfalen)	1967	1953	14 to 15	8 to 9	6
Germany (Rheinland-Pfalz)	1967	1953	14 to 15	8 to 9	6
Germany (Saarland)	1964	1949	14 to 15	8 to 9	6
Germany (Schleswig-Holstein)	1956	1941	14 to 15	8 to 9	6
Italy	1963	1949	11 to 14	5 to 8	6
Netherlands	1942	1929	13 to 14	7 to 8	6
-	1947	1933	14 to 13	8 to 7	6
-	1950	1936	13 to 15	7 to 9	6
Sweden	1949	1936	13 to 14	6 to 7	7
-	1962	1950	14 to 16	7 to 9	7

Note: Data on school reforms are taken from Pischke and van Wachter (2005), Garrouste (2009) and Brunello et al. (2011).

Table 4: Number of books at age ten, by country (in percentage)

	Sample size	None or very few (0-10 books)	One shelf (11-25 books)	One bookcase (26-100 books)	Two bookcases (101-200 books)	More than two bookcases (> 200 books)
Austria	229	48.9	22.7	18.4	5.2	4.8
Belgium	688	54.4	19.8	17.3	4.6	3.9
Czech Republic	674	22.2	31.9	32.2	7.3	6.4
Denmark	743	23.3	21.7	29.6	11.3	14.1
Germany	500	35.2	26.6	24.0	6.4	7.8
France	769	46.4	21.2	19.2	5.9	7.3
Italy	723	77.3	12.2	7.2	1.7	1.6
Netherlands	828	34.3	26.0	26.6	6.2	6.9
Sweden	650	20.2	22.2	34.0	11.8	11.8
Full sample	5804	39.9	22.5	23.4	6.8	7.4

Table 5: Proportion of individuals living in a rural area, by country (in percentage)

	Sample size	Lived in rural area at age ten	None or very few (0-10 books)	One shelf (11-25 books)	One bookcase (26-100 books)	Two bookcases (101-200 books)	More than two bookcases (> 200 books)
Austria	229	55.0	69.6	50.0	42.9	33.3	0.0
Belgium	688	47.4	52.7	50.0	36.1	25.0	37.0
Czech Republic	674	50.9	68.7	57.7	41.9	28.6	25.6
Denmark	743	42.1	56.6	43.5	40.9	35.7	23.8
Germany	500	47.4	61.4	51.9	35.8	34.4	15.4
France	769	35.9	46.8	35.6	18.2	31.1	17.9
Italy	723	57.5	63.1	43.2	34.6	33.3	25.0
Netherlands	828	30.6	35.8	32.1	29.1	17.6	15.8
Sweden	650	29.8	46.6	39.6	26.2	15.6	7.8
Full sample	5804	42.8	54.7	44.3	33.3	26.9	18.7

Note: The figures do not add up to 100%. They refer to the proportion of people living in a rural area within each cell (country and number of books). For instance, among the Austrians who had very few books at age ten, 69.6% were living in a rural area.

Table 6: Schooling regressions, by number of books. Dependent variable: years of education

Variable	Full sample	None or very few books	One shelf	One bookcase	Two bookcases	More than two bookcases
Compulsory edu.	0.030 (0.089)	0.212 (0.131)	-0.007 (0.186)	0.047 (0.200)	-0.091 (0.347)	-0.561** (0.272)
Rural x Comp. edu.	0.338*** (0.068)	0.256** (0.102)	0.234* (0.132)	0.298** (0.146)	0.727** (0.289)	-0.176 (0.281)
Rural area	-3.726*** (0.541)	-2.355*** (0.802)	-2.361** (1.030)	-2.970** (1.178)	-6.046*** (2.189)	0.190 (2.135)
Sample size	5804	2317	1307	1359	394	427
R-squared	0.184	0.250	0.131	0.109	0.233	0.200
F-test statistic	13.38	6.29	1.65	2.32	3.32	2.97

Note: All regressions include controls for birth cohort, country and country-specific quadratic cohort trends (interactions of age and its square with country dummies). Standard errors clustered by country and cohort in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The F-test statistic refers to the joint significance of years of compulsory education and the interaction between the dummy RURAL and years of compulsory education.

Table 7: OLS regressions. Dependent variable: different measures of earnings

Variable	Lifetime earnings net of pensions	Lifetime income gross of pensions	Current or main wage	First wage
Years of education	0.028*** (0.002)	0.031*** (0.002)	0.042*** (0.003)	0.056*** (0.004)
Rural area	-0.056*** (0.017)	-0.052*** (0.014)	-0.043** (0.019)	-0.009 (0.028)
Sample size	5377	5377	5377	5377
R-squared	0.193	0.276	0.151	0.294

Note: All regressions include controls for birth cohort, country and country-specific quadratic cohort trends (interactions of age and its square with country dummies). Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 8: IV regressions. Dependent variable: different measures of earnings

Variable	Lifetime earnings net of pensions	Lifetime income gross of pensions	Current or main wage	First wage
Years of education	0.084*** (0.028)	0.071*** (0.024)	0.036 (0.028)	0.107** (0.042)
Rural area during childhood	-0.000 (0.031)	-0.012 (0.026)	-0.048 (0.033)	0.033 (0.041)
Sample size	5377	5377	5377	5377
First stage F-test statistic	16.63	16.63	17.22	20.68

Note: All regressions include controls for birth cohort, country and country-specific quadratic cohort trends (interactions of age and its square with country dummies). Cluster standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 9: Lifetime earnings regressions: very few books vs. a shelf or more. Dependent variable: lifetime earnings net of pensions

Variable	Very few books			A shelf or more		
	OLS	2SLS	LIML	OLS	2SLS	LIML
Years of education	0.028*** (0.004)	0.033 (0.047)	0.034 (0.053)	0.024*** (0.003)	0.181*** (0.050)	0.181*** (0.050)
Rural area during childhood	-0.020 (0.027)	-0.018 (0.033)	-0.017 (0.034)	-0.078*** (0.022)	0.046 (0.050)	0.047 (0.051)
Sample size	2317	2317	2317	3060	3060	3060
R-squared	0.163			0.227		
First stage F-test statistic		6.29				8.20

Note: All regressions include controls for birth cohort, country and country-specific quadratic cohort trends (interactions of age and its square with country dummies). OLS with robust standard errors, 2SLS and LIML with cluster standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 10: Interaction between years of education and very few books. Dependent variable: lifetime earnings net of pensions

Variable	OLS	2SLS	LIML
Years of education	0.023*** (0.003)	0.179*** (0.051)	0.190*** (0.056)
Very few books x Years of education	0.007 (0.005)	-0.124** (0.052)	-0.132** (0.055)
Rural area during childhood	-0.079*** (0.021)	0.042 (0.049)	0.050 (0.053)
Very few books at age ten (VFB)	-0.166*** (0.062)	1.591** (0.656)	1.702** (0.705)
Rural x Very few books	0.065* (0.034)	-0.042 (0.053)	-0.049 (0.056)
Sample size	5377	5377	5377
R-squared	0.195		
Angrist-Pischke first stage F-test statistics		4.89 (Edu) 19.60 (VFBxEdu)	

Note: All regressions include controls for birth cohort, country and country-specific quadratic cohort trends (interactions of age and its square with country dummies). OLS with robust standard errors, 2SLS and LIML with cluster standard errors in parentheses. The list of instruments includes: years of compulsory education, the interaction between rural area and years of compulsory education, the interaction term between VFB and years of compulsory education, and the interaction term of rural area, VFB and years of compulsory education. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 11: Interaction between years of education and very few books. Dependent variable: current earnings

Variable	OLS	2SLS	LIML
Years of education	0.038*** (0.004)	0.075*** (0.026)	0.075*** (0.026)
Very few books x Years of education	0.004 (0.006)	-0.070** (0.034)	-0.070** (0.034)
Very few books at age ten (VFB)	-0.103 (0.074)	0.755* (0.402)	0.756* (0.403)
Rural x Very few books	-0.042 (0.032)	-0.060* (0.032)	-0.060* (0.032)
Sample size	5377	5377	5377
R-squared	0.153		
Angrist-Pischke first stage F-test statistics		11.68 (Edu) 14.86 (VFBxEdu)	

Note: All regressions include controls for birth cohort, country and country-specific quadratic cohort trends (interactions of age and its square with country dummies). OLS with robust standard errors, 2SLS and LIML with cluster standard errors in parentheses. The list of instruments includes: years of compulsory education, the interaction between rural area and years of compulsory education, the interaction term between VFB and years of compulsory education, and the interaction term of rural area, VFB and years of compulsory education. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 12: Lifetime earnings regressions, with indicators of very few books and poor housing conditions

Variable	OLS	2SLS	LIML
Years of education	0.019*** (0.003)	0.201*** (0.065)	0.236*** (0.085)
Very few books x Years of education	0.004 (0.005)	-0.109** (0.046)	-0.128** (0.055)
Poor housing conditions x Years of education	0.008 (0.005)	-0.060 (0.037)	-0.072 (0.044)
Rural area during childhood	-0.039 (0.025)	0.085 (0.056)	0.109 (0.069)
Very few books at age ten (VFB)	-0.130** (0.065)	1.414** (0.598)	1.679** (0.722)
Rural x Very few books	0.092*** (0.035)	-0.049 (0.066)	-0.074 (0.079)
Poor housing conditions at age ten (PHC)	-0.121* (0.062)	0.751* (0.456)	0.908* (0.546)
Rural x Poor housing conditions	-0.075** (0.035)	-0.100** (0.044)	-0.103** (0.047)
Sample size	5377	5377	5377
R-squared	0.198		
F-test statistic (p-value): VFB variables	0.008	0.047	
F-test statistic (p-value): PHC variables	0.000	0.097	
Angrist-Pischke first stage F-test statistics		2.19 (Edu) 16.84 (VFBxEdu) 9.76 (PHCxEdu)	

Note: All regressions include controls for birth cohort, country and country-specific quadratic cohort trends (interactions of age and its square with country dummies). OLS with robust standard errors, 2SLS and LIML with clustered standard errors in parentheses. The list of instruments includes: years of compulsory education, the interaction term of rural area and years of compulsory education, the interaction of VFB and years of compulsory education, the interaction of rural area, VFB and years of compulsory education, the interaction term of PHC and years of compulsory education, and the interaction of rural area, PHC and years of compulsory education. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 13: Very few books in the house at age ten and poor health conditions at the same age

		Ever missed school		Serious illness		Regular dentist		Sample size
		0	1	0	1	0	1	
Very few books	0	85.8	14.2	71.8	28.2	35.9	64.1	3060
at age ten	1	88.4	11.7	79.5	20.5	69.8	30.2	2317
Sample size		4692	709	4056	1345	2723	2678	

		Poor health		Any vaccines		Sample size
		0	1	0	1	
Very few books	0	92.0	8.0	4.3	95.7	3060
at age ten	1	92.6	7.4	7.1	92.9	2317
Sample size		4983	418	298	5103	

Table 14: Lifetime earnings regressions, by the occupation of the father

Variable	White collar			Service, agricultural, craft			Elementary occupation		
	OLS	2SLS	LIML	OLS	2SLS	LIML	OLS	2SLS	LIML
Years of edu.	0.028*** (0.005)	0.105 (0.096)	0.113 (0.108)	0.030*** (0.004)	0.072* (0.038)	0.072* (0.038)	0.021*** (0.004)	0.106** (0.052)	0.110** (0.055)
Rural area	0.014 (0.046)	0.072 (0.074)	0.077 (0.081)	-0.057** (0.024)	-0.028 (0.032)	-0.028 (0.032)	-0.027 (0.030)	0.017 (0.042)	0.019 (0.043)
Sample size	887	887	887	2851	2851	2851	1639	1639	1639
R-squared	0.262			0.189			0.210		
FS F-test stat.	1.38			14.03			5.97		

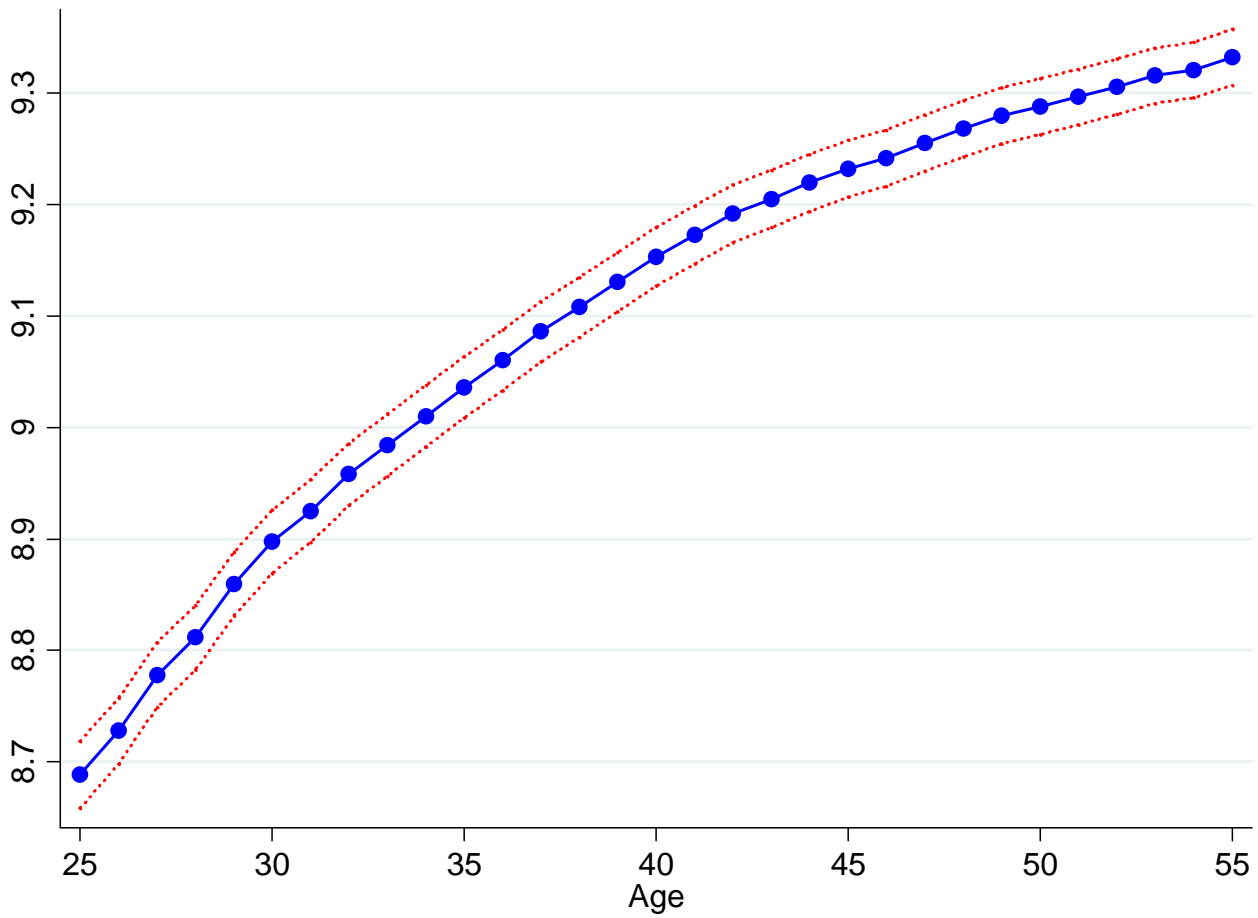
Note: All regressions include controls for birth cohort, country and country-specific quadratic cohort trends (interactions of age and its square with country dummies). OLS with robust standard errors, 2SLS and LIML with cluster standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 15: Effects of the number of books at home on log standardized test scores.

Number of books at home	(1) Reading skills; PIRLS 2006; tests taken at age 9-10	(2) Math skills; TIMSS 1995; tests taken at age 8-11	(3) Reading skills; PISA 2006; tests taken at age 15	(4) Math skills; PISA 2006; tests taken at age 15
0-10	baseline	baseline	baseline	baseline
11-25	0.019*** (0.001)	0.067*** (0.003)	0.060*** (0.002)	0.047*** (0.002)
26-100	0.045*** (0.001)	0.127*** (0.002)	0.117*** (0.002)	0.109*** (0.001)
101-200	0.063*** (0.002)	0.154*** (0.003)	0.157*** (0.002)	0.147*** (0.002)
More than 200	0.077*** (0.002)	0.161*** (0.003)	0.191*** (0.002)	0.187*** (0.002)
Sample size	105670	79221	197751	197751

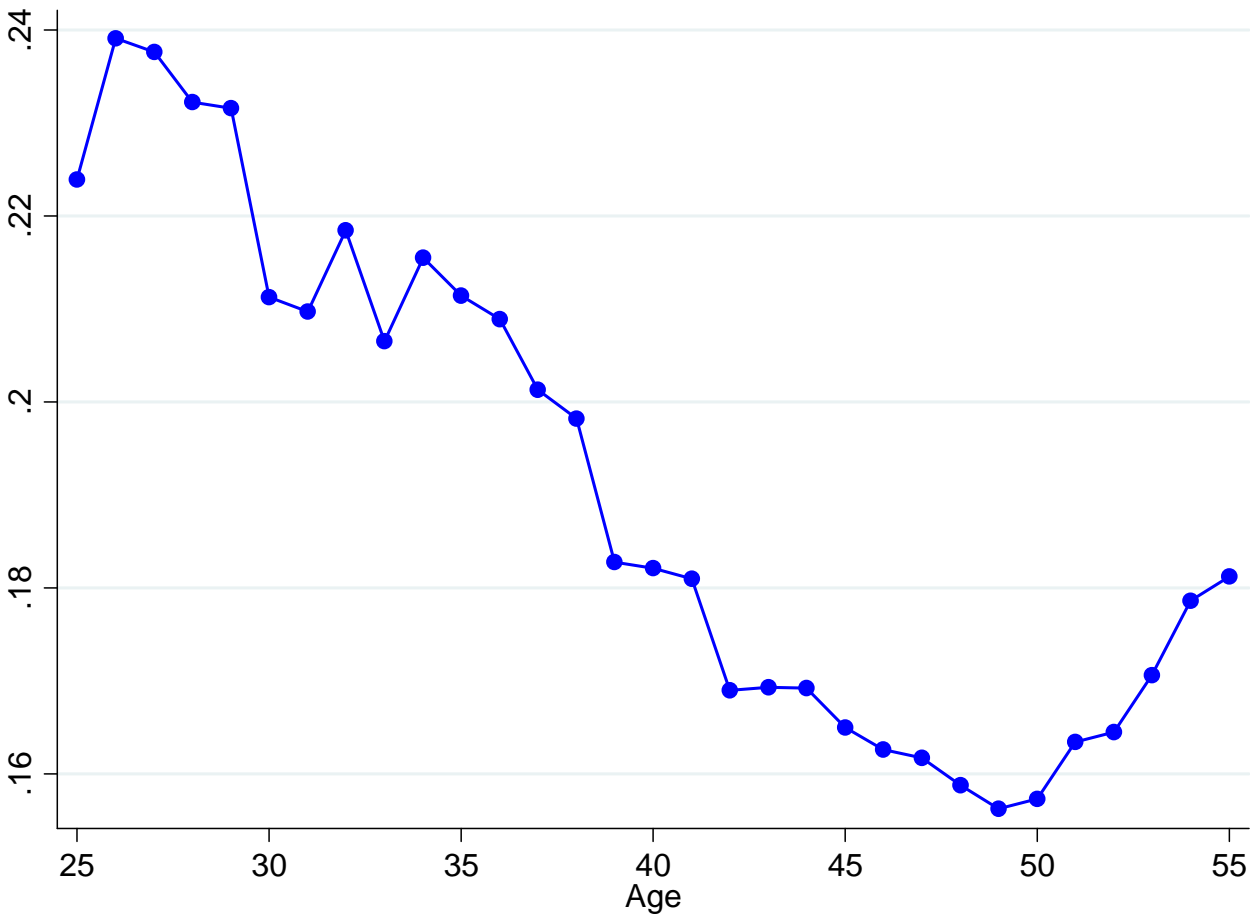
Note: The dependent variables are the logarithm of: (1) PIRLS 2006 reading test scores in the fourth grade, (2) TIMSS 1995 math test scores in the third and fourth grade, (3) PISA 2006 reading test scores in the ninth grade, (4) PISA 2006 math test scores in the ninth grade. PIRLS: the regression includes country dummies, parental education and employment, immigrant status, language spoken at home, gender, age. TIMSS: the regression includes country dummies, age, gender, household conditions, language spoken at home and immigrant status. PISA: each regression includes country dummies, parental education, language spoken at home, immigrant status, age and gender. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Figure 1: Age-earnings profiles net of country and cohort effects for individuals aged 25 to 55



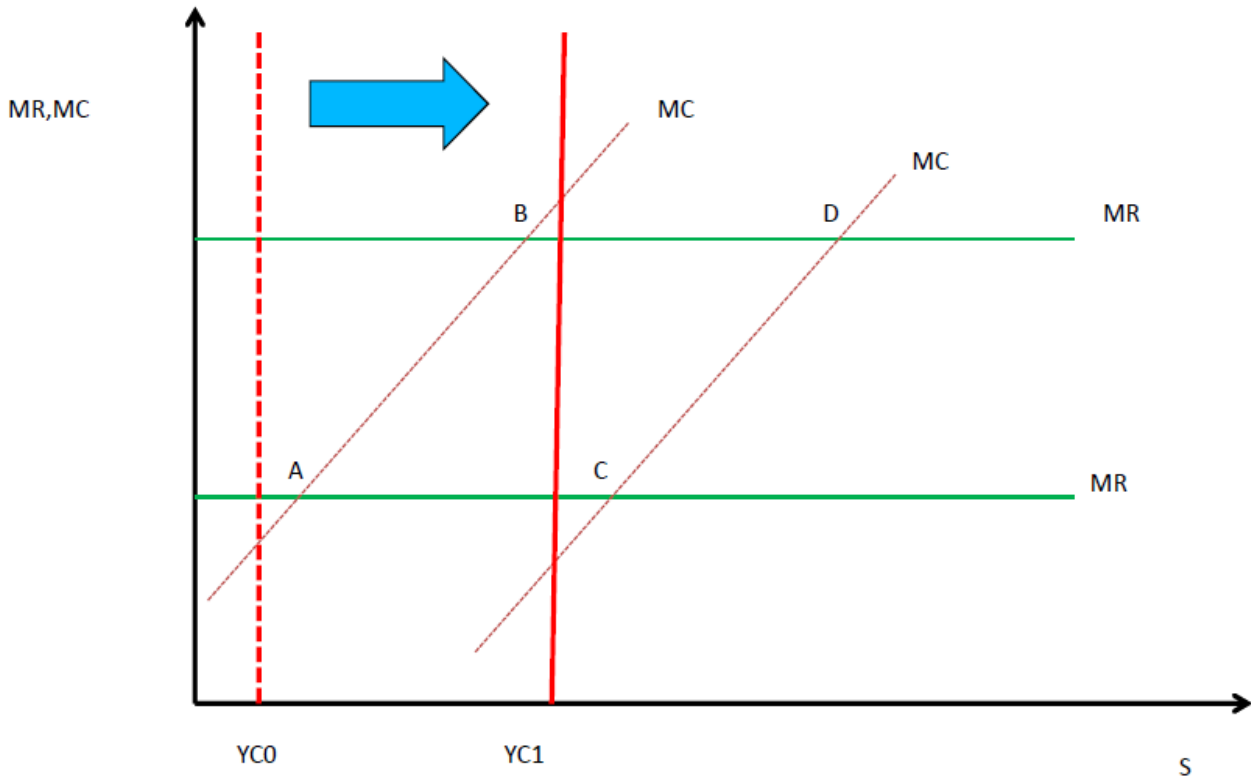
Note: This graph shows the estimated age profile for log earnings (net of country and cohort effects) for men who were never self-employed, but could have experienced unemployment spells. The earnings data are partly reported by respondents, partly imputed as explained in Appendix B. The 95% confidence interval for mean log earnings is also shown.

Figure 2: Log earnings at or above country-specific median years of schooling (YS) minus log earnings below country-specific median YS



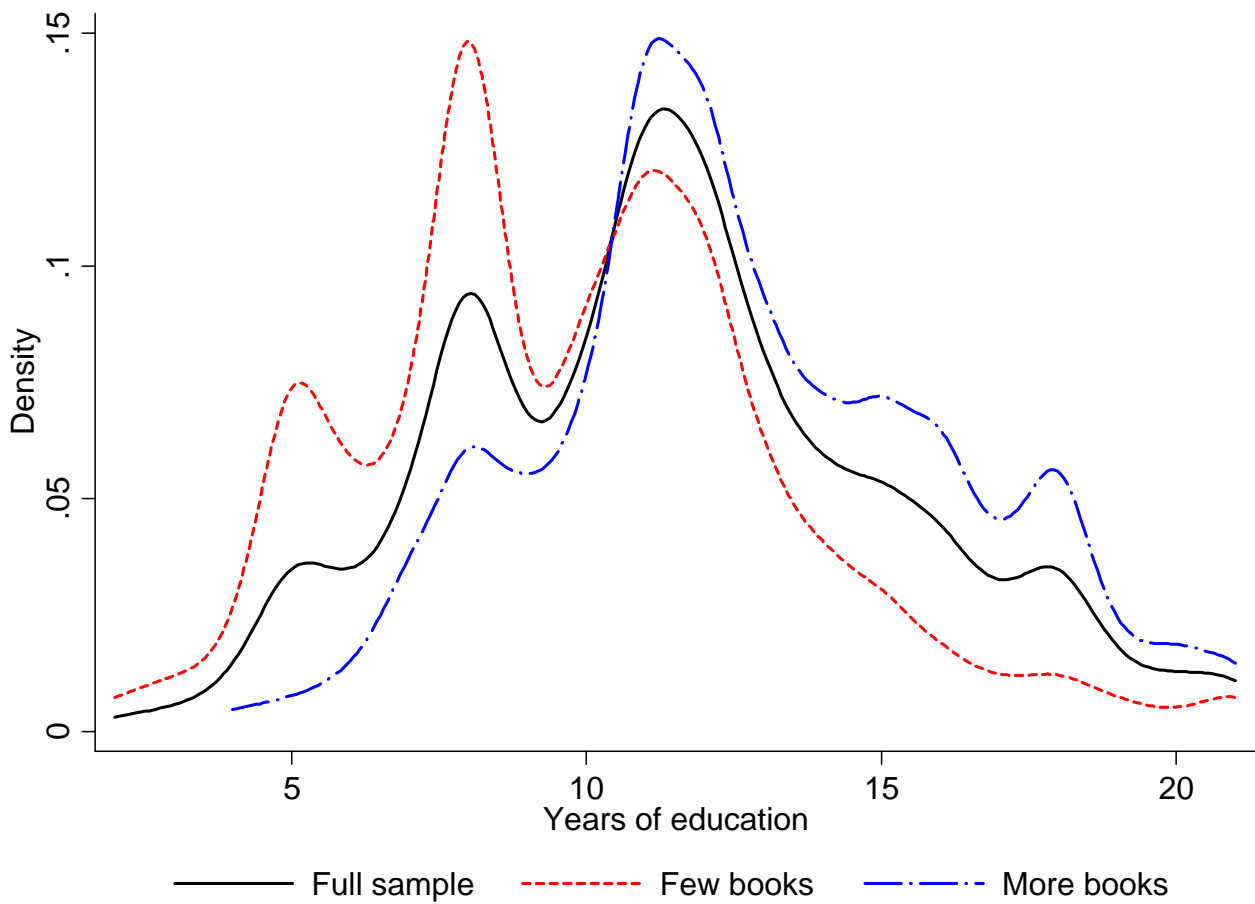
Note: This graph plots the vertical distance between log earnings (net of country and cohort effects) when education is equal to or above the country-specific median and log earnings when education is below this median. The earnings data are partly reported by respondents, partly imputed as explained in Appendix B.

Figure 3: Compulsory school reforms and schooling



Note: This graph shows hypothetical marginal cost (MC) and expected marginal revenue (MR) curves for individuals who choose their optimal level of schooling (S). YC0 and YC1 refer to years of compulsory schooling before and after a school reform, respectively. Individuals A and B live in rural areas, individuals C and D live in urban areas. Individuals A and C come from families with few books at home, individuals B and D from families with many books at home.

Figure 4: Kernel density estimates of years of education, by number of books



Note: This graph plots kernel density estimates of years of education for individuals with books and individuals without books in the household at age ten. It uses a Gaussian kernel and the bandwidths are chosen following Silverman's rule-of-thumb.

References

- Acemoglu Daron and Joshua Angrist (2001). "How Large are Human Capital Externalities? Evidence from Compulsory Schooling Laws." In *NBER Macroeconomics Annual 2000*, Vol. 15. Bernanke, B. and K. Rogoff (eds.), National Bureau of Economic Research, Cambridge, MA, 9-59.
- Angrist, Joshua D. and Jörn-Steffen Pischke (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press, Princeton, NJ.
- Böhlmark, Anders and Matthew J. Lindquist (2006). "Life Cycle Variations in the Association between Current and Lifetime Income: Replication and Extension for Sweden." *Journal of Labor Economics*, 24(4), 879-896.
- Bhuller, Manudeep, Magne Mogstad and Kjell G. Salvanes (2011). "Life-cycle Bias and the Returns to Schooling in Current and Lifetime Earnings." IZA Discussion Paper No. 5788.
- Brenner, Dror and Yona Rubinstein (2011). "The Returns to Education and Family Income." Department of Economics, Brown University, unpublished manuscript.
- Brenner, Jan (2010). "Life-cycle Variations in the Association between Current and Lifetime Earnings: Evidence for German natives and Guest Workers." *Labour Economics*, 17(2), 392-406.
- Brunello, Giorgio, Margherita Fort, Nicole Schneeweis and Rudolf Winter-Ebmer (2011). "The Causal Effect of Health on Education: What is the Role of Health Behaviors?" IZA Discussion Paper No. 5944.
- Brunello, Giorgio, Margherita Fort and Guglielmo Weber (2009). "Changes in Compulsory Schooling, Education and the Distribution of Wages in Europe." *Economic Journal*, 119, 516-539.
- Card, David (2001). "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems." *Econometrica*, 69(5), 1127-1160.
- Carneiro, Pedro and James J. Heckman (2002). "The Evidence on Credit Constraints in Post-secondary Schooling." *Economic Journal*, 112, 705-734.
- Cunha, Flavio and James J. Heckman (2007). "The Technology of Skill Formation." *American Economic Review*, 97(2), 31-47.
- Cunha, Flavio, James J. Heckman, Lance Lochner and Dimitriy V. Masterov (2006). "Interpreting the Evidence on Life Cycle Skill Formation." In *Handbook of the Economics of Education*, Vol.1. Hanushek, E. and F. Welch (eds.), North-Holland, Amsterdam, 697-812.
- Cunha, Flavio, James J. Heckman and Susanne Schennach (2010). "Estimating the Technology of Cognitive and Noncognitive Skill Formation." *Econometrica*, 78(3), 883-931.
- Devereux, Paul J. and Robert A. Hart (2010). "Forced to be Rich? Returns to Compulsory Schooling in Britain." *Economic Journal*, 120, 1345-1364.

- Garrouste, Christelle (2010). *100 Years of Educational Reforms in Europe: A Contextual Database*. EUR - Scientific and Technical Research series, Vol. 24487, Publications Office of the European Union, Luxembourg.
- Garrouste, Christelle and Omar Paccagnella (2011). "Data Quality: Three Examples of Consistency Across SHARE and SHARELIFE Data." In *Retrospective Data Collection in the Survey of Health, Ageing and Retirement in Europe. SHARELIFE Methodology*. Schröder, M. (ed.), Mannheim Research Institute for the Economics of Ageing (MEA), Mannheim, 62-72.
- Goldin, Claudia and Lawrence Katz (2003). "Mass Secondary Education and the State: The Role of State Compulsion in the High School Movement." NBER Working Paper No. 10075.
- Gould, Eric D., Victor Lavy and M. Daniele Paserman (2011). "Sixty Years after the Magic Carpet Ride: The Long-Run Effect of the Early Childhood Environment on Social and Economic Outcomes." *Review of Economic Studies*, 78(3), 938-973.
- Haider, Steven and Gary Solon (2006). "Life Cycle Variation in the Association between Current and Lifetime Earnings." *American Economic Review*, 96(4), 1308-1320.
- Havari, Enkelejda and Fabrizio Mazzonna (2011). "Can We Trust Older People's Statements on Their Childhood Circumstances? Evidence from SHARELIFE." SHARE Working Paper 05-2011.
- Heckman, James J. (2000). "Policies to Foster Human Capital." *Research in Economics*, 54(1), 3-56.
- Heckman, James J., Lance J. Lochner and Petra E. Todd (2006). "Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond." In *Handbook of the Economics of Education*, Vol.1. Hanushek, E. and F. Welch (eds.), North-Holland, Amsterdam, 307-458.
- Heckman, James J., Lance J. Lochner and Petra E. Todd (2008). "Earnings Functions and Rates of Return." *Journal of Human Capital*, 2(1), 1-31.
- Heckman, James J., Lena Malofeeva, Rodrigo Pinto and Peter Savelyev (2010). "Understanding the Mechanisms Through Which an Influential Early Childhood Program Boosted Adult Outcomes." Department of Economics, University of Chicago, unpublished manuscript.
- Jensen, Robert (2010). "The (Perceived) Returns to Education and the Demand for Schooling." *Quarterly Journal of Economics*, 125(2), 515-548.
- Lochner, Lance and Alexander Monge-Naranjo (2011). "Credit Constraints in Education." NBER Working Paper No. 17435.
- Mazzonna, Fabrizio (2011). "The Long-Lasting Effects of Family Background: A European Cross-Country Comparison." MEA Discussion Paper No. 245-11.
- Mincer, Jacob A. (1974). *Schooling, Experience, and Earnings*. National Bureau of Economic Research, New York, NY. Distributed by Columbia University Press.

- Moulton, Brent R. (1990). "An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Units." *Review of Economics and Statistics*, 72(2), 334-338.
- Murphy, Kevin M. and Finis Welch (1990). "Empirical Age-Earnings Profiles." *Journal of Labor Economics*, 8(2), 202-229.
- Murtin, Fabrice and Martina Viarengo (2008). "The Convergence of Compulsory Schooling in Western Europe: 1950-2000." CEE Discussion Paper No. 95.
- Oreopoulos, Philip (2006). "Estimating Average and Local Average Treatment Effects of Education when Compulsory Schooling Laws Really Matter." *American Economic Review*, 96(1), 152-175.
- Pischke, Jörn-Steffen (1995). "Measurement Error and Earnings Dynamics: Some Estimates From the PSID Validation Study." *Journal of Business & Economics Statistics*, 13(3), 305-314.
- Pischke, Jörn-Steffen and Till von Wachter (2005). "Zero Returns to Compulsory Schooling in Germany: Evidence and Interpretation." NBER Working Paper No. 11414.
- Trevisan, Elisabetta, Giacomo Pasini and Roberta Rainato (2011). "Cross-Country Comparison of Monetary Values from SHARELIFE." SHARE Working Paper 02-2011.
- Van Kippersluis, Hans, Owen O'Donnell and Eddy van Doorslaer (2011). "Long-Run Returns to Education: Does Schooling Lead to an Extended Old Age?" *Journal of Human Resources*, 46(5), 695-721.
- Weiss, Christoph T. (2012). "Two Measures of Lifetime Resources for Europe Using SHARELIFE." SHARE Working Paper 06-2012.

Appendix A: The life cycle bias

In this appendix we propose a method to identify the value at which the life cycle bias is minimized and apply this method to our data. Assume that wage profiles are not parallel. In particular, let

$$W_{it} = \exp(\beta_0 + \beta_1 A_{it} + \beta_2 A_{it}^2 + \beta_3 S_i + \beta_4 S A_{it}) \quad (\text{A.1})$$

where $\beta_4 \neq 0$. Wages at $t=a+x$ are given by

$$W_{ia+x} = W_{ia} \exp(x_i \beta_1 + x_i^2 \beta_2 + 2x_i \beta_2 A_{ia} + x_i \beta_4 S_i). \quad (\text{A.2})$$

As in Bhuller et al (2011), we focus on males from age 21 to 55 who have never been unemployed. In this case, $a = 21$ and $q = 34 \geq x$ is the length of the age span between 21 and the terminal year, which is typically before retirement. For these individuals, lifetime income is defined as

$$I_{ia} = W_{ia} + \frac{W_{ia+1}}{1+r} + \frac{W_{ia+2}}{(1+r)^2} + \dots + \frac{W_{ia+q}}{(1+r)^q} \quad (\text{A.3})$$

Taking logarithms of equation (A.2) and using equation (A.3) we obtain

$$\ln I_{ia} = \ln W_{ia} + \ln \left\{ 1 + \Theta + \frac{\exp(\beta_4 S_i)}{1+r} + \frac{\exp(2\beta_4 S_i)}{(1+r)^2} + \dots + \frac{\exp(q\beta_4 S_i)}{(1+r)^q} \right\} \quad (\text{A.4})$$

where Θ includes all terms which do not depend schooling S . The above equation can be rewritten as

$$\ln I_{ia} = \ln W_{ia+x} - x_i \beta_4 S_i + \ln \left\{ 1 + \frac{\exp(\beta_4 S_i)}{1+r} + \frac{\exp(2\beta_4 S_i)}{(1+r)^2} + \dots + \frac{\exp(q\beta_4 S_i)}{(1+r)^q} \right\} \quad (\text{A.5})$$

By taking derivatives with respect to S on both sides of equation (A.4) we obtain

$$\frac{\partial \ln I_{ia}}{\partial S} = \frac{\partial \ln W_{ia+x}}{\partial S} - \beta_4 x_i + \frac{\beta_4 \left\{ \frac{\exp(\beta_4 S_i)}{1+r} + \frac{2 \exp(2\beta_4 S_i)}{(1+r)^2} + \dots + \frac{q \exp(q\beta_4 S_i)}{(1+r)^q} \right\}}{1 + \frac{\exp(\beta_4 S_i)}{1+r} + \frac{\exp(2\beta_4 S_i)}{(1+r)^2} + \dots + \frac{\exp(q\beta_4 S_i)}{(1+r)^q}} \quad (\text{A.6})$$

The life cycle bias LCB_{ia+x} is equal to zero when

$$x_i^* = \frac{\left\{ \frac{\exp(\beta_4 S_i)}{1+r} + \frac{2 \exp(2\beta_4 S_i)}{(1+r)^2} + \dots + \frac{q \exp(q\beta_4 S_i)}{(1+r)^q} \right\}}{1 + \frac{\exp(\beta_4 S_i)}{1+r} + \frac{\exp(2\beta_4 S_i)}{(1+r)^2} + \dots + \frac{\exp(q\beta_4 S_i)}{(1+r)^q}} \quad (\text{A.7})$$

Appendix B: Computing lifetime income

A. Initial and final earnings of each employment spell

We define lifetime income as the net present value at age ten of the stream of net wages and pension benefits earned over the life cycle from age ten using a discount rate of 2% ($r = 0.02$). We mainly use data on work histories from SHARELIFE but also some data from SHARE waves 1 and 2. Wages and pension benefits are transformed using PPP exchange rates and CPI measures into 2006 Euro. PPP-adjusted exchange rates and CPI measures are taken from the OECD and national sources.²⁶

We start by computing the length of each employment spell. When the years at the beginning and at the end of the spell are identical, we assume that the individual spent an entire year in the job, i.e. working from January 1 to December 31. When the years are different, we assume that he started and stopped working in the same month, e.g. working from March 1974 to March 1986. This implies that someone who reports to have started working in an employment spell in 1974 and stopped in 1974 will be treated equally to someone who started in 1974 but stopped in 1975.

Whenever the current income in SHARELIFE is missing but an income measure was reported at the beginning of the current employment spell, we use the income measure from the imputation module in wave 2 (if the current employment spell started before the interview year of wave 2) or from wave 1 (if the current employment spell started before the interview year of wave 1). The imputation modules in waves 1 and 2 contain a measure of annual net income from employment in the previous

²⁶ More details can be found in Trevisan et al (2011).

year.²⁷ We impute missing wage values using predictive mean matching. Predictive mean matching is an imputation method used for continuous variables and is similar to a regression method. It finds the observation whose predicted value is closest to the predicted value of the missing observation but uses the observed value for imputation.²⁸ Predicted values are obtained by regressing annual wages on ISCED education levels (3 different levels), birth cohort (3 cohorts), decade when the employment spell started (4 different decades), whether the worker is a white collar during the spell, whether he worked part-time during the spell, and country. Imputation is used for approximately 25% of wages at the start of job spells. Unsurprisingly, there are more missing values for jobs that started in earlier decades.

In SHARELIFE, individuals are asked to report their monthly net pay at the start of each job. They are not asked to report how much they were paid at the end of each spell, except for the main spell in their career (if they have retired) or their current employment spell (if they are still working). Therefore, only the current and the main employment spells have wage measures both at the beginning and the end of the spell.

We predict wages at the end of the spell by using potential labour market experience as the running variable. Potential experience is defined as $A_t - S - IS$ where A_t denotes age in year t , S years of education and IS age at school entry. We regress log current earnings on potential experience, potential experience squared, education, occupation and industry dummies, and interactions of these variables with experience. We also control for characteristics that are constant over the life cycle: country, 3 birth cohorts, whether the individual was better (or much better) than others in mathematics at age ten (as opposed to about the same, worse or much worse), whether the individual was better (or much better) to others in the country's language at age ten (as opposed to about the same, worse or much worse), accommodation conditions at age ten (5 indicators for whether or not the accommodation had a fixed bath, cold running water supply, hot running water supply, inside toilet and central heating), and an indicator of the number of rooms occupied by the household divided by the number of people living in the household at age ten²⁹. We estimate the following linear model

²⁷ See Christelis (2011) for more details on the imputed variables in SHARE.

²⁸ One can also draw at random from a set of observed values whose predicted values are close to the one of the observation with missing value. See, e.g., Heitjan and Little (1991), Schenker and Taylor (1996) and Horton and Kleinmann (2007).

²⁹ The indicator takes value one if the number of rooms occupied by the household at age ten (including bedrooms but excluding kitchen, bathrooms, and hallways) divided by the number of people living in the household is equal or higher to one, and zero otherwise. That is, we compute whether there are more rooms than people in the household at age ten.

$$y_{ic} = \beta_1 E_{ic} + \beta_2 E_{ic}^2 + \beta_3 E_{ic} S_i + \beta_4 E_{ic} X_{ic} + \beta_5 S_i + \beta_6 X_{ic} + \beta_7 Q_i + U_i$$

where y_c is the log current wage, E experience, S education, X the characteristics that are specific to the employment spell (i.e. occupation and industry), Q the characteristics that are constant over the life cycle, and U a disturbance term. We then use the estimated parameters to predict end of spell wages starting from earnings at the start of a job spell, which we typically observe. We compute

$$\hat{y}_{1ij} = y_{0ij} + b_1(E_{1ij} - E_{0ij}) + b_2(E_{1ij}^2 - E_{0ij}^2) + b_3(E_{1ij}S - E_{0ij}S) + b_4(E_{1ij}X_j - E_{0ij}X_j)$$

where \hat{y}_{1ij} is predicted end of spell log earnings, y_{0ij} is the log observed (or imputed) wage at the beginning of spell, and E_{1ij} and E_{0ij} are potential experience at the end and the beginning of the spell respectively. Armed with the wages at the beginning and the end of each spell, we compute spell – specific annual earnings growth rate and use this growth rate to generate annual earnings in each employment spell.

To check the accuracy of our procedure, we apply it to current and main employment spells, for which we have information on wages at the end of the spell, and compare predicted with actual values. Table B.1 shows that predicted final wages are close to reported values, an unsurprising result given that the estimated coefficients used for predictions are obtained from regressions on current income on controls.³⁰ We obtain an empirical distribution of wages, and eliminate outliers in this distribution by excluding observations that are above the 99th and below the 1st percentiles. We proceed in a similar fashion with pension benefits.

B. Lifetime earnings

We multiply monthly earnings by 12 to obtain annual earnings. Earnings are annualized because employment spells are in years. For each individual i , the discounted sum of the stream of annual earnings is

³⁰ While the means of the current and predicted income are not statistically different at the 1% confidence level, this is not case for the means of the main income and the predicted main income.

$$AW_i = r \cdot \sum_{j=1}^J Y_{ij} \sum_{k=1}^K (1 + gr_j)^{(k-1)} / (1 + r)^{(S_j - (BY_i + 11) + k)}$$

where j is the job, J the total number of jobs, Y_j annual earnings at the beginning of each spell j , k the year in the employment spell, K is the total length of each employment spell (in years), $1 + gr_j$ the annual growth rate of earnings in the employment spell j , S_j the year when employment spell j started, BY the year of birth and r the interest rate.

To illustrate, if someone is born in 1940 and starts working in 1950, the first wage in 1950 is not discounted, but the wage in 1951 is discounted with $1+r$, in 1952 with $(1+r)^2$ and so forth. While the first wage is reported by the individual, subsequent earnings are predicted as discussed above.

C. Pension benefits

We also add to discounted sum AW the discounted flow of expected incomes and pension benefits from the time of the interview until age 110. These benefits are multiplied by the survival probability between age_t and age_{t+1} , which we take from the Human Mortality Database (Department of Demography at the University of California, Berkeley and Max Planck Institute for Demographic Research).

While some individuals are still working at the time of their interview, others have already retired. For those who have retired before the interview, their future income flow consists of pension benefits and expected pension benefits. This income flow is given by

$$BRET_i = r \cdot pension_i \cdot \sum_{l=1}^L 1 / (1 + r)^{(RET_i - (BY_i + 11) + l)}$$

where $pension$ refers to the annual pension benefit currently received, l is the year in the pension spell, L is the length of the pension spell (i.e. the difference between the interview year and the retirement year), RET is the retirement year, BY is the year of birth and r is the interest rate. We assume that real pension benefits do not increase or decrease during retirement.

Expected pension benefits from time of the interview year to age 110 is computed as

$$DRET_i = r \cdot pension_i \cdot \sum_{m=1}^M (1 - qx_m) / (1 + r)^{(INT_i - (BY_i + 1) + m)}$$

where m is the year spent in the expected pension spell, M is the expected length of the pension spell (i.e. the difference between age 110 and the age at the time of the interview), qx_m is the probability of death within the age interval $[BY+m, BY+m+1)$, INT is the interview year, BY is the year of birth and r is the interest rate.

We use SHARELIFE data on net monthly pension benefits. We multiply these monthly benefits by 12 and obtain annual measures. When the sum of pension benefits is missing or below/above the censoring thresholds (1st and 99th percentiles), we use information from waves 1 and 2 of SHARE. We compute the sum of all annual pension benefits reported in wave 2, which includes public old age pensions, public early or pre-retirement pensions, public disability insurance, public unemployment benefit or insurance, public survivor pension from partner, war pensions, private (occupational) old age pensions, private (occupational) early retirement pensions, private (occupational) disability insurance, private (occupational) survivor pensions from the partner's job, public old age supplementary pensions, secondary public disability insurance pensions, secondary public survivor pensions from spouse/partner, occupational old age pensions from a second job, occupational old age pensions from a third job, and private (occupational) disability insurance. We use this measure to replace the missing values of pension benefits that could not be recovered using SHARELIFE. When the sum of pension benefits is still missing, we use information from wave 1. The remaining missing values for pension benefits (approx. 5% of the sample) are imputed using predictive mean matching. We regress pensions on ISCED education level (3 different levels), birth cohort (3 cohorts), decade of the retirement year (4 different decades), and country and predict using the estimated coefficients.

We do not have information on pension benefits for those who are still working, but we know from previous waves their expected retirement age and replacement rate.³¹ To compute a measure of lifetime income which includes all working episodes over the life cycle for all individuals, we create a new artificial employment spell that should correspond to the last employment spell until retirement. Obviously, for those who have already retired, the length of this artificial employment spell is equal to zero. For those who are still working, the length of the employment spell is the difference between the

³¹ When individuals do not report at what age they will start collecting pension, we use information on statutory retirement age in their country. Information on statutory retirement age is mainly taken from Angelini, Brugiavini and Weber (2009).

age at which they expect to collect pension benefits and their current age. If these two ages are equal, we assume that they retire immediately and start collecting pension benefits.

In this artificial employment spell, we assume that individuals who are still working at the time of the interview in SHARELIFE will continue working until their expected retirement age, without any unemployment spell. We also assume that, upon reaching retirement age, they immediately stop working and retire. We predict their wage at the end of this artificial spell in a similar fashion as done for the final wage of each real employment spell. We compute the discounted sum of expected incomes up to expected pension age for each individual as

$$CW_i = r \cdot Y_{i,curr} \cdot \sum_{s=1}^V (1 + gr_s)(1 - qx_s)/(1 + r)^{(INT_i - (BY_i + 1) + s_i)}$$

where Y_{curr} is current earnings, s is the year spent in the current employment spell until expected pension age, V is the expected length of the artificial employment spell, $1 + gr$ the annual growth rate of income during the employment spell, qx_s the probability of death, INT the interview year, BY is the year of birth and r is the interest rate. Expected pension benefits are given by

$$DWORK_i = r \cdot rebrate \cdot Y_{i,curr} \cdot \sum_{t=1}^T (1 - qx_t)/(1 + r)^{(PY_i - (BY_i + 1) + t_i)}$$

where $rebrate$ is the replacement rate (or percentage of salary received as pension), Y_{curr} is current income, t is the year spent in the expected pension spell, T is the expected length of the retirement spell (i.e. the difference between age 110 and the expected pension age), qx_t is the probability of death, PY the expected retirement year (the year in which the individual will start receiving pension benefits), BY is the year of birth and r is the interest rate. We use current wage and not the predicted wage at expected retirement age because individuals are asked what is the percentage of current wage they will received as pension benefits.

The expected pension replacement rate is reported by individuals who are working in wave 2 of SHARE. We aggregate over all types of pensions for each individuals and consider only those in the range between 50 and 100 percent. When values drawn from wave 2 are missing, we use values from wave 1. We fill the remaining missing values by using the median replacement rate within each country

and 3 birth cohorts.

D. Total lifetime earnings

For individuals who have retired at the time of the interview, their lifetime income at age 10 is $NPV_{10} = AW + BRET + DRET$. For individuals who are still working at the time of the interview, this income is $NPV_{10} = AW + CWORK + DWORK$. We only retain in our final dataset those individuals with an estimated lifetime income between the 1st and the 99th percentile of the distribution of lifetime income.

Appendix C: Validation of the wage prediction procedure using German panel data

We attempt to validate the procedure used to predict wages at the end of an employment spell using data from the German Socio-Economic Panel Study (SOEP). The SOEP is a longitudinal panel dataset of the population in Germany which started in 1984, which contains information on household composition, occupation, employment, earnings, health and life satisfaction. We use annual data from 1984 to 2008. SOEP data are integrated into the Cross National Equivalent File (CNEF), which contains equivalently defined variables for panel databases from the UK (BHPS), Australia (HILDA), South Korea (KLIPS), the U.S. (PSID), and Canada (SLID).

We perform the validation study using the variables in the CNEF file. We define potential experience as: age - schooling - 6 (the age at school entry in Germany) and estimate the following model

$$y_{i2008} = \beta_1 E_{i2008} + \beta_2 E_{i2008}^2 + \beta_3 E_{i2008} S_i + \beta_4 E_{i2008} X_{i2008} + \beta_5 S_i + \beta_6 X_{i2008} + \beta_7 W_i + U_i$$

where y_{2008} denotes the logarithm of individual annual labour earnings in 2008, E is potential experience, S education, and X the characteristics that are specific to the employment spell (i.e. occupation and industry), W the characteristics that are constant over the life cycle (i.e. the birth cohort), and U the disturbance term. Our sample consists of all males born between 1945 and 1956 who report information on individual labor earnings, age, schooling, occupation and industry in 2008. We

then use the wage in 1984 (the first year in SOEP) and the estimated coefficients of the above regression to predict wages from 1985 to 2008, in the same fashion as done for our retrospective panel. While the variables used in this regression are very similar to those used for predicting wages at the end of a job in SHARE, the main difference is that we focus on a single country – Germany – and that we do not include covariates describing early life conditions as they are not available in SOEP.

Table C.1 reports the mean observed and predicted wage, the prediction error, and the p-value of the hypothesis that the mean predicted wage and the mean observed wage are equal using SOEP data. For each year, the sample consists of the individuals who report information in 1984 (the starting year) and year t . The table shows that we reject the hypothesis that the means of the predicted and observed wages are equal in only 4 years (2003 to 2006) out of 25. This suggests that our procedure to estimate earnings at the end of a job is rather accurate.³²

Appendix D: Age-earnings profiles by education levels

Compared to Figure 2 in the text, Figure D.1 shows a more detailed break-down of earnings profiles (net of country and cohort effects) by education levels. The figure consists of two parts: in the upper part we plot for each age the vertical distance between log earnings for individuals with college and high school education; in the lower part, we plot the vertical distance between log earnings for individuals with high school and less than high school education. There is convergence among lower education levels, but divergence between college and high school graduates. Given that there are relatively few college graduates in our sample, convergence prevails overall.

³² We repeated the same exercise using data from the Italian Survey of Household Income and Wealth (SHIW). The results (not reported here) are very similar to the ones using the German SOEP.

Table A.1: Descriptive statistics for the sample excluding individuals with more than 200 books

Variable	Mean	Std.Dev.	Min	Max	Median
Lifetime income	10621.79	6001.71	529.66	39123.13	9521.26
Lifetime earnings net of pension	8218.19	5361.27	265.93	38253.88	7096.58
Years of education	11.46	3.86	2	25	11
Years of compulsory education	7.51	1.58	4	10	8
Age	67.16	8.78	52	89	66
Years of work	36.70	8.22	5	63	38
Number of jobs during career	3.08	2.00	1	18	3
Very few books at age ten	0.43	0.50	0	1	
Rural area or village during childhood	0.45	0.50	0	1	
Very few books x rural area	0.24	0.42	0	1	
Poor housing conditions at age ten	0.49	0.50	0	1	
Ever unemployed	0.09	0.28	0	1	
Retired	0.74	0.44	0	1	
Austria	0.04	0.20	0	1	
Belgium	0.12	0.33	0	1	
Czech Republic	0.12	0.32	0	1	
Denmark	0.12	0.32	0	1	
France	0.13	0.34	0	1	
Germany	0.09	0.28	0	1	
Italy	0.13	0.34	0	1	
Netherlands	0.14	0.35	0	1	
Sweden	0.11	0.31	0	1	
Sample size	5377	5377	5377	5377	5377

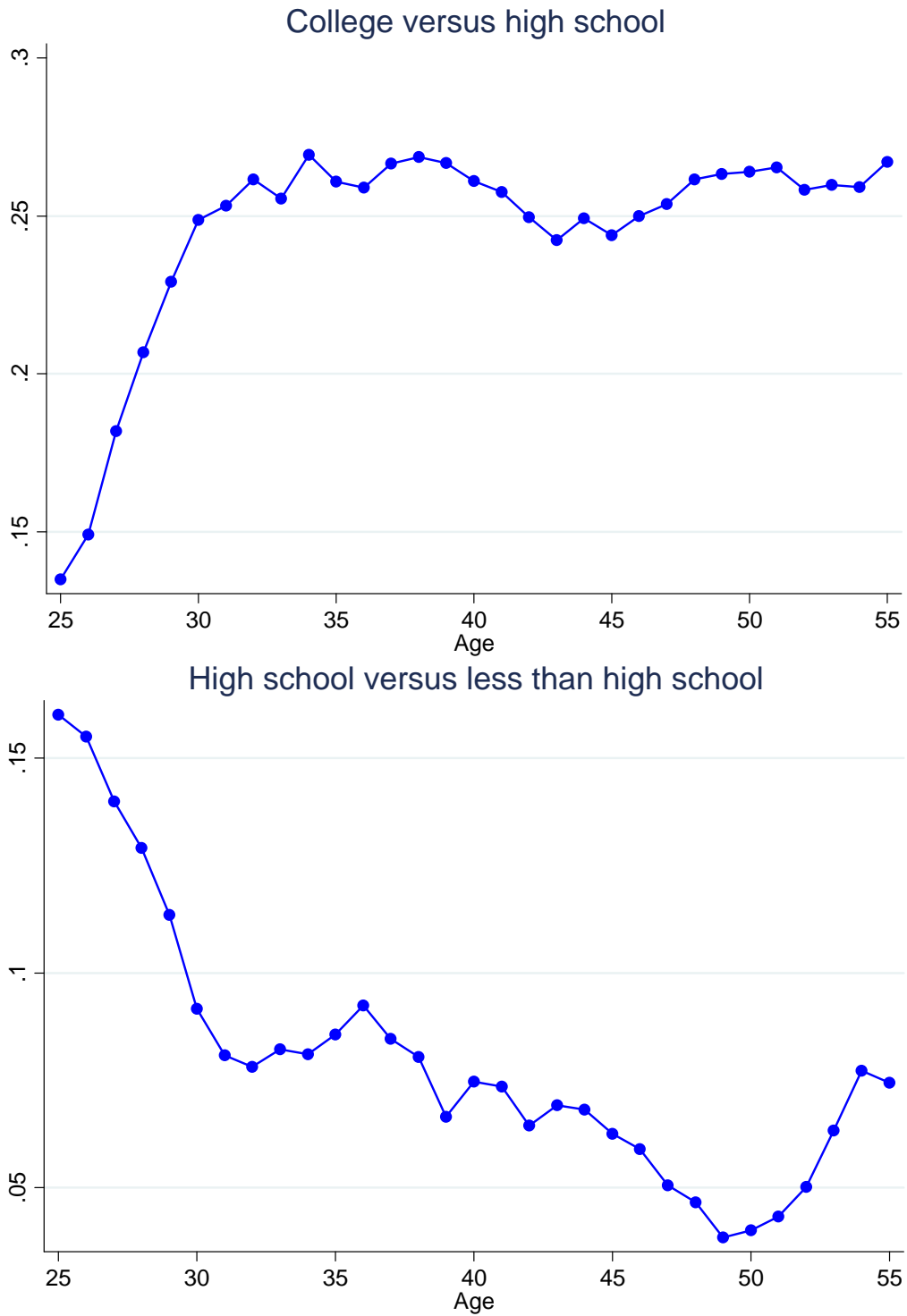
Table B.1: Prediction error for current and main wages

Variable	Sample size	Mean	Std. Dev.
Log current income	2305	9.9422	0.4739
Predicted log current income	2305	9.9359	0.8274
Prediction error of log current income	2305	0.0062	0.7851
Log main income	4687	9.8242	0.6843
Predicted log main income	4687	9.7560	1.0285
Prediction error of log main income	4687	0.0682	1.0888

Table C.1: Means of the observed wage, predicted wage and prediction error: German SOEP

Year	Sample Size	Observed wage	Predicted wage	Prediction error	p-value of the error
1985	984	5.700	5.673	0.024	0.270
1986	877	5.736	5.713	0.023	0.321
1987	838	5.772	5.760	0.011	0.630
1988	772	5.814	5.785	0.024	0.263
1989	784	5.811	5.779	0.029	0.241
1990	756	5.826	5.809	0.017	0.497
1991	729	5.836	5.819	0.016	0.534
1992	693	5.840	5.843	-0.004	0.866
1993	14	5.769	5.885	-0.116	0.366
1994	17	5.875	5.976	-0.101	0.401
1995	585	5.904	5.879	0.025	0.409
1996	556	5.915	5.905	0.011	0.724
1997	528	5.935	5.929	0.006	0.848
1998	480	5.935	5.947	-0.013	0.727
1999	447	5.957	5.952	0.002	0.951
2000	441	5.973	5.931	0.041	0.279
2001	401	5.963	5.927	0.028	0.467
2002	374	5.995	5.912	0.082	0.052
2003	344	6.027	5.897	0.128	0.004
2004	319	6.034	5.884	0.151	0.001
2005	295	6.016	5.895	0.121	0.009
2006	260	5.980	5.848	0.132	0.009
2007	249	5.925	5.855	0.070	0.229
2008	217	5.921	5.850	0.068	0.286

Figure D.1: Age-earnings profiles by education levels



Note: Vertical distance of log earnings between individuals with college education and individuals with high school education in the upper panel, and vertical distance of log earnings between individuals with high school education and individuals who did not finish high school in the bottom panel. The earnings data are partly reported by respondents, partly imputed as explained in Appendix B.

Additional references (Appendix)

- Angelini, Viola, Agar Brugiavini and Guglielmo Weber (2009). "Ageing and Unused Capacity in Europe: Is There an Early Retirement Trap?" *Economic Policy*, 24, 463-508.
- Christelis, Dimitris (2011). "Imputation of Missing Data in Waves 1 and 2 of SHARE." SHARE Working Paper 01-2011.
- Heitjan, Daniel F. and Roderick J.A. Little (1991). "Multiple Imputation for the Fatal Accident Reporting System." *Journal of the Royal Statistical Society C*, 40(1), 13-39.
- Horton, Nicholas J. and Ken P. Kleinman (2007). "Much Ado about Nothing: A Comparison of Missing Data Methods and Software to Fit Incomplete Data Regression Models." *American Statistician*, 61(1), 79-90.
- Schenker, Nathaniel and Jeremy M.G. Taylor (1996). "Partially Parametric Techniques for Multiple Imputation." *Computational Statistics & Data Analysis*, 22(4), 425-446.