Books Are Forever: Early Life Conditions, Education and Lifetime Earnings in Europe*

Giorgio Brunello, Guglielmo Weber and Christoph T. Weiss

We estimate the effect of education on lifetime earnings by distinguishing between individuals who lived in rural or urban areas during childhood and between individuals with access to many or few books at home at age ten. We instrument years of education using compulsory school reforms and find that, whereas individuals in rural areas were most affected by the reforms, those with many books enjoyed substantially higher returns to their additional education. We show that books retain explanatory power even when we select relatively homogeneous groups in terms of the economic position of the household, suggesting that their long-lasting beneficial effects are likely to be associated to the cultural environment in the household and the development of cognitive skills rather than to short-term liquidity constraints.

* Corresponding author: Guglielmo Weber, Department of Economics, University of Padova, Via del Santo 33, 35123 Padova, Italy. Email: guglielmo.weber@unipd.it

We thank the editor and three referees for particularly useful feedback. We are also grateful to Joshua Angrist, Erich Battistin, Tom Crossley, Margherita Fort, Richard Freeman, Lisa Kahn, Peter Kuhn, Michael Lechner, Sonia Orrefice, Mario Padula, Luigi Pistaferri, Alfonso Rosolia, Kjell Salvanes, Holger Sieg, Richard Spady and participants at seminars and conferences in Bologna, Cambridge (RES), Catanzaro (IWAEE), Essex (ISER), Geneva, Gothenburg (EEA), Groningen, Istanbul (Koç University), Kyoto (TPLS), Lugano (IIPF), Northwestern (NASM), Rome (Brucchi Luchino and Tor Vergata), St. Gallen and Surrey for comments and suggestions. We gratefully acknowledge financial support from the Antonveneta Centre for Economic Research (CSEA) and Fondazione Cariparo. The usual disclaimer applies.

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 *et al.* (2006). In this paper, we depart from standard practice in various ways. First, we estimate the effect of education on lifetime earnings, not just current earnings. Second, 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 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 and 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.[1] 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.[2]

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 when compared to those based on lifetime earnings, particularly if the sample includes many older workers (see Haider and Solon, 2006; and Bhuller *et al.* 2011). This evidence, however, relies on administrative data from 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, e.g., Acemoglu and Angrist, 2001; Oreopoulous, 2006; and Pischke and van Watcher, 2005), we recognise 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

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

[2] When we disaggregate by education level, we find converging profiles for individuals with and without a high school degree, and diverging profiles for individuals with college and a high school degree.

on current education, which vary with whether affected individuals lived in a rural or in an urban area at age ten.

Consistent with the evidence from Norway presented in Bhuller *et al.* (2011), we show that – on average – an additional year of education increases lifetime earnings by 9%. This is reassuring, given that the Norwegian earnings data are probably less affected by measurement error than survey data. We find that most 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. Some compliers also lived in urban areas, but were endowed with very few books in the household at age ten.

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 very 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, e.g., Cunha and Heckman, 2007; and Heckman, 2000).

Compulsory education has increased in most European countries after the Second World War. Our findings suggest that among the individuals induced by these reforms to attain higher education only those with enough books at home were able to reap significant private economic returns. A specific group of individuals, who lived in rural areas with very few books at home, attained higher education but much lower 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.

The paper is organised as follows. The next section presents the data, describes how we obtain individual measures of lifetime earnings and computes the bias associated to estimating returns to education on current rather than on lifetime earnings. Section 2 introduces the empirical model. In Section 3 we discuss the effects of compulsory school reforms on educational attainment in the European countries for which we have data. Section 4 examines our estimates of the returns to education using lifetime earnings. Section 5 considers how differences in early life conditions affect these returns and Section 6 discusses reasons why the number of books in the household at age ten matters. The last section concludes.

1. The Data

We use the Survey of Health, Ageing and Retirement in Europe (SHARE)[3], a

[3] 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.

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 draw our data from 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 issues 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 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). Monthly figures are multiplied by 12 to obtain annual earnings, and converted into 2006 Euro using PPP exchange rates and CPI indices so that wages are comparable across time and country.

As described in detail in Appendix A and in Weiss (2012), we use current and retrospective information on earnings, jobs and labour market experience to construct a measure of lifetime earnings (or permanent income), which we define as the income flowing from the asset value of working at age ten. The asset value of working at age ten is the discounted sum of wages earned from age ten up to retirement, using a discount rate of 2%.[5] In SHARELIFE, we observe the first wage in each job as well as the current or last wage. For those who have had only one job in their working life (more than 20% of the sample), we interpolate linearly between the first and the last (or current) wages. For those who have had more than one job, we regress current wages on labour market experience, a rich set of controls, which include

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

[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%. Our estimates are largely unaffected if we use an interest rate of 2.3% instead of 2% (results available from the authors upon request).

education, occupation, sector of activity, cohort and country effects and economic conditions at age ten, and the interactions of these controls with experience. We then 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.[6] For the individuals who are currently working, earnings are predicted until reported expected retirement age or until the statutory retirement age when expected retirement age is missing. A validation study which uses the German Socio-economic Panel suggests that our procedure to recover within-job earnings growth is quite accurate (see Appendix B).

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, which are subject to measurement error. Importantly, Bingley and Martinello (2014), using Danish administrative register information drawn from tax reports and civil registries to validate SHARE data, show that measurement error for annual income in these data is classical. Since we use lifetime earnings, this error is partly averaged out. Additional 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. It is also reassuring for us that our estimates are in line with those obtained using administrative data, as discussed below.

Our sample consists of 5,820 men born between 1920 and 1956 and residing in Austria,

[6] Our estimates in Section 1 and Sections 3 to 6 are broadly unaffected if we replace labour market experience with age and exclude education in the wage regressions used to generate both the final wage in each job and within-job earnings growth for individuals who have had more than one job (results available from the authors upon request).

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 for us to identify a pre-treatment and a post-treatment sample of cohorts.[7]

Table 1 provides some descriptive statistics on our sample. Average lifetime earnings flowing from the estimated asset value of working at age ten are equal to 7,759 real Euros. 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 66.94 and 36.55. 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.

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. Of the few studies focusing on lifetime earnings, most use administrative data. Haider and Solon (2006) draw their data from the Social Security earnings histories of participants in the US 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-

[7] In Belgium, we only keep 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. For Germany, we only include individuals from West Germany. 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 1995 devaluation and misreported earnings during the high inflation of the 80s and 90s.

cycle bias'.[8] 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), Brenner (2010) and Bhuller *et al.* (2011) use the Swedish Level of Living Survey and LINDA (Longitudinal Individual Database for Sweden), the German VVL longitudinal survey – which covers a sample of pension recipients born between 1939 and 1974 – and Norwegian administrative data, respectively. All these studies confirm that earnings profiles are not parallel with respect to education.

For each individual in the sample we compute both his lifetime earnings and the annual sequence of earnings from labour market entry until retirement. Using the longitudinal information provided by this sequence, we show in Appendix C that age-earnings profiles by educational attainment are not parallel but converge over time.[9] Following Haider and Solon (2006), we define the life cycle bias as the difference between the marginal effect of schooling on wages at age a and the marginal effect of schooling on lifetime earnings. When age-earnings profiles are parallel with respect to education, this bias is equal to zero for any value of a. When they are not parallel, the bias can be positive, negative or equal to zero at $a = a^*$.

Using the method described in Appendix C, we estimate the critical value a^* in our data, and find that, when schooling is at its mean level, this value is equal to 35.52 years, 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. Brenner (2010) reviews this literature and suggests that the critical age lies in the range 30 to 40. We consider the

[8] 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.

[9] Our evidence is based on net earnings. Because of the effects of progressive taxation, we cannot exclude that gross age-earnings profiles are parallel or even diverge.

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.

2. The Empirical Strategy

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

$$Y_i = \beta_1 S_i + X_i^{\,T} \beta_2 + U_i \tag{1}$$

$$S_i = Z_i^{\,T} \gamma_1 + X_i^{\,T} \gamma_2 + V_i \tag{2}$$

where Y denotes the logarithm of lifetime earnings, S 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.[10] 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.[11] Pooling data from different countries is unlikely to be useful when returns to schooling vary significantly across countries. To dispel this concern, in Section 5 we report test results that fail to reject the null hypothesis that they are the same.

[10] The linear trend is defined as year of birth – 1919. It ranges from 1 for individuals born in 1920 to 37 for those born in 1956.

[11] See also the discussion in Goldin and Katz (2003) and Stephens and Yang (2014).

Since education is affected by individual unobserved ability, which can also influence earnings, the covariance between the disturbances in equations (1) and (2) is unlikely to be zero. In addition, the correlation between *S* and *U* in (1) is nonzero if years of schooling are measured with error, or if lifetime earnings are misreported and misreporting is systematically related to educational attainment.[12] 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 literature[13], 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 literature. We apply this strategy to a multi-country setup, as in Brunello *et al.* (2009) and Brunello *et al.* (2013), and exploit the fact that school reforms occurred at different points in time and with varying intensity in several European countries.

Table 2 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

---

[12] Milligan *et al.* (2004) discuss the case when voting participation is misreported and the error is correlated with schooling. They also use instrumental variables to deal with the associated bias.

[13] See, e.g., Oreopoulos (2006), Pischke and von Wachter (2005) and Devereux and Hart (2010) for a review.

[14] We compute years of compulsory education using the country where the individual was living when the reform could have affected him (at age 11, 13 or 14, depending on the country) and not the country where he is residing now. Following Pischke and von Watcher (2005), we allow compulsory school reforms to occur at different moments of time across German states. Pischke and von Watcher (2005) discuss the implications of measurement error in the exact timing of reforms, due for instance to early school entry, grade repetition,

sample period were normally 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 were temporarily reduced, but increased overall by two and one year respectively.[15] Figure 3 plots average years of schooling against distance, measured as year of birth minus year of birth of the cohort first affected by each reform.[16] There is a clear discontinuity at the cut-off, suggesting that school reforms in Europe had a significant impact on educational attainment, as reported by Banks and Mazzonna (2012) for the UK and Brunello *et al.* (2013) for several countries of Continental Europe.

3. The Effect of Compulsory Education Reforms on Years of Schooling

The Becker-Card model suggests that privately optimal education is attained when marginal benefits equal marginal costs. Figure 4 illustrates school choice for four hypothetical individuals (A, B, C and D). The heterogeneity of outcomes shown in Figure 4 depends on heterogeneous marginal costs and revenues (see Ashenfelter and Rouse, 1998, and Appendix D for a derivation that highlights the identifying assumption we make). Consider first costs. Individuals in our sample were born between 1920 and 1956 – a period when the proportion

measurement error in the year of birth, and mobility between states. They argue that 'the various measurement problems [...] will tend to attenuate both the first stage and reduced form coefficients. Since the relative biases in both the first stage and the reduced form coefficients will be the same, instrumental variables will be consistent despite the measurement error for the standard reasons'. Pischke and von Watcher (2005, pp.14-15)

[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 *et al.* (2011).

[16] For countries with more than one compulsory school reform (the Czech Republic, France, the Netherlands and Sweden), we only show the effect of the last reform in Figure 3.

of households living and working in rural areas in Europe was substantially higher than today. The SHARELIFE survey asks respondents where they lived at age ten. They can choose the area of residence among the following options: a) big city; b) suburbs or outskirts of big city; c) large town; d) small town; e) rural area or village. It turns out that 42.8% of the sample lived in a rural area or a village. This percentage was 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 (marginal) 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. In Figure 4, the marginal costs of schooling for individuals living in rural areas at age ten (A and B) lie above the costs faced by individuals living in urban areas (C and D).

Turning to expected marginal revenues, SHARELIFE has a question on the number of books (excluding magazines, newspapers or school books) in the place where the individual was living at age ten. The answer to this question falls 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). Books can be considered as a proxy of two important drivers of returns, parental resources and skill formation early in life. 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 3).[17] Table 4 shows the proportion of individuals

[17] 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 a more equal proportion of Catholics and Protestants (Germany and the Netherlands) are somewhat in the middle between the two extremes.

living in a rural area or a village during their childhood, classified by country and number of books at age ten. For instance, 69.4% of the Austrians with very few books grew up in a rural area, compared to only 35.7% in the Netherlands. There is a higher proportion of people from rural areas 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 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 0.20 overall, ranging from 0.08 in the Netherlands to 0.27 in Austria.

We expect that for individuals with several books at age ten (B and D in the figure) the marginal benefits of education are higher than for individuals with very few books (A and C). In the absence of compulsory schooling, Figure 4 suggests that individual D would attain the highest level of education and individual A the lowest. Suppose that years of compulsory education are initially equal to YC_0, which corresponds to the dashed vertical line in Figure 4. 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 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 the compliers, who have increased their education because of the reform and face relatively high marginal costs of education but markedly different expected returns.

Their relatively high marginal costs suggest that, *ceteris paribus*, children living in rural areas (A and B in the figure) attained lower education than children living in cities, and were therefore exposed to a higher extent to reforms increasing years of compulsory education. To verify this conjecture, we estimate equation (2) by including in the set of instruments the interaction of years of compulsory education with the dummy variable 'rural area', an indicator equal to one if the individual lived in a rural area or a village at age ten and to zero otherwise. If compliers are drawn to a larger extent among those living in rural areas during childhood, we expect this interaction term to have 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 (1).

We find that reforms to compulsory education have had a statistically significant effect on educational attainment, especially in rural areas, with the exception of the relatively small group of individuals with more than two bookcases at home (more than 200 books). There are reasons to believe that this group is highly heterogeneous. For instance, it contains large fractions of individuals who report being discriminated in the labour market, particularly in countries such as the Czech Republic (related to its communist past) and Belgium (probably linked to the linguistic divide – recall that our Belgian sample is entirely Flemish). The percentage of 'always takers' in this group is close to 95%, much higher than in the rest of the sample (63%).[18] Machin *et al.* (2012) estimate the impact of education on regional mobility in Norway. In an effort to improve the strength of their selected compulsory school

[18] The estimated effects of compulsory education for this group are -0.009 (standard error: 0.374) for those living in urban areas and -0.328 (standard error: 0.425) for those living in rural areas during childhood. Including this group in our data leads to rejecting the null hypothesis that the interactions of years of compulsory schooling with country dummies are jointly equal to zero in a regression of years of schooling on compulsory years of schooling. In this case, pooling data from different countries – as we do in this paper – would not be advisable. High heterogeneity implies that the signal to noise ratio is too low for this group to provide useful information in our estimates. It is therefore not surprising that results based on a sample that includes these individuals are qualitatively similar to those presented in the main text, as we report below. Table A1 shows the descriptive statistics for the sample that excludes individuals with more than 200 books.

reform as an instrumental variable, they focus on the lower end of the educational distribution in their sample, where most of the changes took place. We follow their approach and exclude the group with more than 200 books from the analysis.

Our estimates of the first stage equation are shown in Table 5, which is organised in three columns, one for this new sample and the remaining two for the sub-samples of individuals with very few books (less than a shelf or at most 10 books) and with 11 to 200 books. Focusing on the first column, we find that one additional year of compulsory education has increased the years of schooling of individuals living in rural areas or in villages during childhood by 0.561 (the sum of 0.196 and 0.365, standard error: 0.106). In comparison, the effect on those who were living in urban areas is much smaller (0.196, standard error: 0.106).[19]

The effect of compulsory education on years of schooling is larger for those with very few books in the household during childhood than for those with more books, independently of whether they were living in a rural or in an urban area. We estimate that the effects on education of an additional year of compulsory education range from 0.375 (standard error: 0.116) to 0.629 (standard error: 0.128) in urban and rural areas for those with very few books, and from 0.004 (standard error: 0.143) to 0.316 (standard error: 0.147) in urban and rural areas for those with 11 to 200 books in the household.

As a test of the identification strategy, we regress years of schooling on years of compulsory education in a sample where each reform is artificially anticipated by five years (as also done by Machin *et al.*, 2012). If our strategy is correct, we should find that the estimated effect of compulsory education is much lower in this sample than in the true

[19] Standard errors of the sums of coefficients are computed using the command 'lincom' in Stata 12. We also find that living in a rural area during childhood reduces educational attainment by 3.72 years, which corresponds to 32.4% of average schooling, a sizeable effect.

sample, because of the addition of five cohorts of individuals who are too old to have been affected by reforms. We find that the effect of compulsory education on years of schooling is very small and not statistically significant (estimate: 0.036; standard error: 0.108), and conclude that this 'placebo' test with each reform anticipated by five years is supportive of our identification strategy.

4. The Causal Effect of Education on Lifetime Earnings

We initially estimate equation (1) by ordinary least squares (OLS) using as dependent variable either the logarithm of lifetime earnings or the logarithm of the current wage (or wage at the end of the main job in the career if the individual is retired).[20]. Table 6 reports our OLS estimates and shows that returns to education range from 3.1% to 4.2%, depending on the definition of the dependent variable (lifetime earnings or current wage). There is also evidence that living in a rural area at age ten has a negative and statistically significant effect on outcomes. We contrast these estimates with two stage least squares (2SLS) estimates, which we obtain by instrumenting years of schooling with years of compulsory education and their interaction with the dummy variable 'rural area'. 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 7 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

[20] Since the individuals in our sample are age 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 current (or main) job as the dependent variable, we also include as covariates a set of dummies for the age when the job started or ended. Our results are unaffected when we add to our regressions a country–specific business cycle indicator, measured as the residuals of country-specific regressions of real GDP on a linear and a quadratic trend.

above the rule of thumb value of 10, indicating that our instruments are not weak.[21] We estimate that an additional year of schooling in our sample increases lifetime earnings by 9%. It is re-assuring that a very similar return (8.7%) is found by Bhuller *et al.* (2011), who use Norwegian administrative data, which are less likely than our data to be affected by recall bias and measurement error.[22] In line with the evidence that age-earnings profiles are converging, the estimated effect on current earnings is lower than on lifetime earnings, and equal to 5.6%.[23]

Pooling data from different countries, as we do in this paper, is unlikely to be useful when returns to schooling vary significantly across countries. To dispel this concern, we run two auxiliary regressions on pooled data: first, we regress years of schooling both on compulsory years of schooling and on the interactions of this variable with country dummies. If the effect of compulsory education does not vary across countries, we should find that these interactions are not jointly statistically significant. We test this hypothesis and find that we cannot reject it (p-value of the test: 0.207). Second, we regress lifetime earnings both on years of schooling and the interactions of this variable with country dummies, using compulsory education as instrument. If returns to schooling do not vary by country, we

---

[21] Table 7 also reports the p-values of the Hansen J test, showing that we do not reject the over-identifying restrictions at the conventional 5% level of confidence.

[22] These results are also in line with those found by Gensowski (2014), who focuses on the lifetime returns of a sample of high-IQ individuals from California born around 1910. Our lifetime returns to education in a just-identified model where education is instrumented only with years of compulsory education (but not with their interaction with the indicator of rural area) are very similar to those reported in Table 7 with an estimated coefficient of 0.094 (standard error 0.047).

[23] Table A2 presents the OLS and 2SLS estimates of the effect of schooling on lifetime earnings when the few individuals who had more than 200 books at age ten are included. We show that including these individuals does not affect our estimates qualitatively.

should not reject the hypothesis that the interactions are not jointly significant. Again, we cannot reject the null (p-value: 0.825). In both regressions (as in all other estimation based on pooled data), we also allow for country – specific quadratic trends in birth cohorts.

Table A3 shows the OLS and 2SLS estimates when we run completely separate country-level regressions. We notice that sample sizes are small – they range from 217 for Austria to 772 for the Netherlands. We find that the OLS estimates of the returns to education (shown in the first column) are positive and significantly different from zero in all countries but Austria, and that they are not significantly different from each other (the p-value of a formal pooling test is 0.075). The 2SLS estimates (second column) are instead very imprecise in most countries – however, in the four cases where their t-ratios exceed unity they are all positive. We gain some precision when we drop the never significant control 'living in a rural area at age ten': in this specification (reported in the fourth column of Table A3), we find positive and statistically significant 2SLS coefficients in three countries, and positive coefficients in another three countries where the parameter estimate has a t-ratio higher than or close to unity.

This evidence clearly discourages working on a single-country basis, particularly if one intends to investigate possible heterogeneity in returns (as we do). Of comfort to us is that no single country plays a key role in our estimates of the causal effect of education on lifetime earnings. Table A4 shows 2SLS estimates when we drop from the sample one country at a time: point estimates of the key parameter of interest are fairly stable (ranging from 0.062 to 0.125) and always retain significance at the 10% level.

In our regressions, 2SLS 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 have been excluded from higher education because of even higher costs.[24] Yet Carneiro and Heckman (2002) warn against such an interpretation, which is based on the questionable assumption that OLS estimates measure the average treatment effect on the treated. In the next section we show that while the compliers with at least a shelf of books at age ten enjoyed high returns to schooling, those who had very few books received much lower returns from their additional schooling. The explanation of liquidity constraints seems to be much less compelling for the latter group.

5. 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 across individuals and between 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 *et al.* 2010). We measure early life conditions with the number of books available in the household when the individual was ten years old. We estimate separate regressions for two sub-groups, one with 0 to 10 books and the other with 11 to 200 books.

Table 8 reports both OLS and 2SLS estimates of the returns to education for each sub-group, using lifetime earnings as the dependent variable. Since the F-test statistic of the additional instruments is below the critical value of 10 in one of the two sub-samples, we also estimate the model by limited information maximum likelihood (LIML). The LIML estimator

[24] See, e.g., Card (1999, 2001) and Oreopoulos (2006).

is median-unbiased in over-identified models.[25] When we treat education as endogenous, we find a sharp contrast between the two groups of individuals: whereas the returns to schooling are small (around 5%) for the group with very few books, we estimate large (21%) and statistically significant returns for the group with more than 10 books in the household at age ten. The difference in returns is estimated at 0.159, and is statistically significant at the 5% level of confidence (confidence interval ranging from 0.009 to 0.424).[26] These findings are confirmed by the LIML estimates, which are very close to the 2SLS estimates. As shown in Appendix E, this large difference in returns cannot be driven by recall errors in the number of books at home, under plausible assumptions.

Our 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 a shelf of books are four times as large as the relatively low returns earned by the group with very few books.[27] Going back to Figure 4, the former group corresponds to 'individual B' and the latter group to 'individual A'. Both groups have high marginal costs of schooling but 'individual B' has substantially higher returns. Our estimates underscore the importance of early life conditions for the

[25] See Angrist and Pischke (2009) for a discussion and comparison of the sampling properties of 2SLS and LIML estimators. We also report in the table the p-values of Hansen J test, indicating that we do not reject the over-identifying restrictions at the conventional 5% confidence level, albeit only marginally so (p-value: 0.082) in the case of very few books.

[26] The confidence interval is computed using seemingly unrelated estimation and bootstrap standard errors (500 replications).

[27] To check whether our findings depend on cultural difference across countries, we 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. However, the estimates (available from the authors upon request) are very similar to those reported in Table 8.

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 *et al.* (2013).

Given the large heterogeneity across countries in the number of books at home at age ten, one may wonder how much of the heterogeneity in the returns to education between few and many books is due to the fact that the country composition changes between the two groups. For instance, as reported in Table 3, Sweden has the lowest proportion of people with very few books (20.1%) and Italy the highest (77.3%). If the Swedish compulsory school reform was more successful than the Italian one, the heterogeneity along the number of books may actually reflect heterogeneity across school reforms. If our estimates in Table 8 were driven by differences in the efficacy of school reforms, they should change drastically when we drop Italy and Sweden from the sample. When we drop these two countries, however, our estimates are broadly unaffected. In addition, the pooling tests discussed above clearly indicate that the first stage effects of compulsory school reforms on years of schooling do not change significantly among countries, suggesting that the reforms considered in this paper have been equally successful.[28]

One potential concern with the estimates presented so far is that we include in our sample individuals born during the 1920s, who entered the labour market before the Second World War and could have been both temporarily and permanently affected by the War. Table 9

[28] An additional source of concern is that the word 'rural' could mean substantially different things in the heavily populated Netherlands and in sparsely inhabited Sweden. To check whether our results are driven by the presence of countries with very high or very low population density, we drop from our sample the Netherlands, Belgium (both densely populated) and Sweden (sparsely populated). When we drop these three countries, we find that the estimated return to schooling are 0.214 (statistically significant) for the group with many books and -0.032 (imprecisely estimated) for the group with few books.

replicates the estimates in Table 8 for the sub-sample of individuals born after 1929 and shows that results are very similar.

Another concern one can have is that books are not randomly assigned across the population. Parents whose children do either well or badly at school may be more likely to buy more (non-school) books for them to read. However, it is hard to tell if a child at age ten is particularly bright or simply more mature than his peers – and the same applies for children whose performance at school is relatively poor.[29] We conclude from this that the discretion parents exert in purchasing books is unlikely to generate strong correlation between the number of books and unobserved ability.

A more plausible source of non-randomness in the assignment of books across the population (and within the rural and urban populations) has to do with parental background and children's unobserved ability. Suppose that unobserved ability depends on books at home at age ten, on urban/rural residence during childhood and on parental background in a linear way, then the vector of covariates X in equation (1) should include not only books at home and urban versus rural residence but also parental background.

We explore the importance of this second source of non-randomness in the assignment of books at home by adding indicators of parental background to the equation we estimate. In particular, we observe in our data the occupation of the main breadwinner when the child was

[29] According to Judson (1998, p.340), 'innate ability is measured with difficulty and with increasing clarity as education proceeds. Any test given will be a noisy signal, and the less education the person has had, the noisier the signal will be. Before primary school it is very difficult to discern levels of talent, but identification of talent is easier after a few years of primary school, still easier after high school, and so on.' Allen and Barnsley (1993) show that relative age plays an important role in determining minor hockey team selection for young children, and that the oldest children within each cohort are more likely to have successful hockey careers. This happens because the substantial variation in maturity within young cohorts makes it more likely that older children are selected for more competitive teams.

ten and classify occupations in two categories: white and blue collars. When we add the indicator that the main breadwinner was in a white collar job to the specification reported in Table 8, we find that the parameters of interest are virtually unaffected.[30]

Another indicator of parental environment is poor housing, which we measure with a dummy variable taking value 1 if the accommodation occupied by the household when the individual was aged ten lacked running water or an inside toilet, and zero otherwise.[31] We repeat the exercise performed above using this indicator in place of the occupation of the main breadwinner, and find that the estimated returns to education are very similar to those reported in the table (estimates available from the authors upon request). We conclude from this that the effect of books at home on the returns to education is not driven by the omission of either parental occupation (as a proxy for parental ability) or poor housing (as a proxy for material deprivation) in our specification.

6. Why Are Books at Home at Age Ten Associated for Lifetime Earnings?

Our finding that books at home at age ten are beneficial to lifetime income calls for an explanation. Books at home could matter because they induce children to read more and reading can have positive effects on school performance – as we know from many sources, including Kim (2007) and Allington *et al.* (2010), but see Fryer (2011) for a cautionary note. But children can borrow books from their schools, from local libraries or from friends – as

[30] The 2SLS point estimates are 5.4% for the sample with few books (instead of 5.2% shown in Table 8) and 20.2% for the sample with many books (instead of 21.1%). Standard errors are slightly larger for the sample of people with few books and slightly smaller for the sample of individuals with many books. The first stage F-test statistic is lower in the former sample (11 instead of 12.16) and higher in the latter (9.48 instead of 7.91).

[31] 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.

Roald Dahl's 'Matilda' powerfully reminds us! Our data do not tell us whether books were read, and how many, even though we can suspect a positive relation between the number of books at home and a child's reading activities.

Another candidate reason why books at home matter for lifetime earnings is that they capture the effects of early economic conditions. Table 10 shows the distribution of books by area of residence and by parental occupation at age ten. Unsurprisingly, households where the main breadwinner was in a white collar occupation are more likely to have at least a shelf of books, both in rural and in urban areas. At the same time, however, 45% of individuals who grew up in a rural household with a blue collar parent, and 58% of those from urban areas with a blue collar parent, had at least a shelf of books. This suggests that books do not simply capture differences in parental occupation. To confirm this, we also estimate lifetime earnings equations for each occupation group (see Table 11) and find that returns to education do not vary significantly with the occupation of the main breadwinner (the difference is 0.013, with a confidence interval ranging from -0.237 to 0.388).

Few books at home could also proxy for poor health at age ten. Table 12 shows the correlations between the number of books 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, that is indicative of parental investment in children's health, we find any evidence of a significant correlation with the number of books.

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 household. As reported in Table 13, there is clear evidence that books are associated with cognitive development, even after conditioning for parental education (and employment), most likely because they capture parental ability not measured by formal degrees.

Through its effects on the development of cognitive and socio-emotional skills during childhood, 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 relatively little in education (9.86 years on average for those with very few books at home and 12.69 year on average for those with more books at home). When forced to invest more by compulsory school leaving laws, these individuals earn low returns either because their cognitive ability is crystallised at a lower level or because they comply with the law by attending low quality education.

We also find that the individuals living in rural areas and having at least a shelf of books at age ten who were forced to take up additional education by the compulsory school reforms enjoyed high returns (21%). One may wonder why these individuals did not try to collect these high returns by investing in further education even in the absence of changes in minimum school leaving ages. A candidate explanation is that individuals living in rural areas were much more isolated than nowadays and chose education based on perceived rather than actual returns. Jensen (2010, p.517) shows that perceived returns to schooling are often inaccurate and underestimated, also because 'individuals rely almost exclusively on the earnings of workers in their own communities in forming their expectations of earnings.' 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. Not all people born in rural areas around World War II 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-World War I period.

In our data, we find that rural boys with books were more likely to move to the cities (46% versus 33%) and to have had as their first job a white collar job (33% versus 15%). 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 4, both rural boys A and B are pushed to increase their schooling up 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 21%) 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 realised 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.[32] Boys who grew up in a less cultured home, instead, went back to the countryside and to the standard farming jobs available there and failed to reap the benefits of their extra education.

[32] 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%.

7. Conclusion

In this paper we have investigated how lifetime earnings relate to education and socioeconomic background during childhood in a number of European countries using a rich data set containing detailed retrospective information on earnings and many variables of potential interest (including childhood characteristics). Our estimates suggest that an additional year of education increases average lifetime earnings by 9%. 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 captures educational attainment of the parents, which is not recorded in our data, we notice that evidence from recent cognitive test scores shows that the number of books predicts these scores even after controlling for parental education and occupation.

This may reflect the fact that books are simply an alternative measure of parental background that contains ability not captured by formal degrees. Also because of family ability, access to books when young is associated with home skill formation in cognitive and socio-emotional skills, something that has been emphasised as an important factor of economic success in life.

Given that today people have more books at home than in the sample studied in this paper, we may wonder whether our results would still hold for more recent cohorts. This is a difficult question, and we only offer here two speculative remarks. First, our comparison can be framed as the comparison of individuals with less than median and more than median number of books. Even if the median number of books rises over time, because the requirements of modern societies increase, we could still have that those with less than the median number of books earn lower returns from their education than those with more books. Second, since compulsory school reforms have been particularly effective in rural areas, we expect that the effectiveness of these reforms is bound to decline with increasing urbanization.

Table 1

Descriptive Statistics

	Mean	Std. Dev.	Min	Max	Median
Lifetime earnings	7,758.74	5,389.60	105.82	36,619.20	6,607.17
Years of education	11.76	3.99	2	25	11
Years of compulsory education	7.50	1.59	4	10	8
Age	66.94	8.72	52	89	66
Years of work	36.55	8.15	5	63	38
Number of jobs during career	3.13	2.05	1	18	3
Very few books at age ten	0.40	0.49	0	1	
Rural area or village at age ten	0.43	0.49	0	1	
Very few books × 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	0.11	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	0.13	0.33	0	1	
Netherlands	0.14	0.35	0	1	
Sweden	0.11	0.32	0	1	
Sample size	5,820	5,820	5,820	5,820	5,820

Table 2

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 Watcher (2005), Garrouste (2010) and Brunello, Fabbri and Fort (2013).

Table 3

Number of Books at Age Ten, by Country (in Percentage)

	Sample size	None or very few books (0-10)	One shelf (11-25 books)	One Bookcase (26-100 books)	Two bookcases (101-200 books)	More than two bookcases (> 200 books)
Austria	228	48.7	22.8	18.4	5.3	4.8
Belgium	669	55.0	19.3	17.1	4.9	3.7
Czech Republic	672	22.3	31.7	32.3	7.3	6.4
Denmark	747	23.6	21.6	29.4	11.2	14.2
Germany	516	34.7	26.4	23.8	7.2	7.9
France	773	46.2	21.2	19.5	5.7	7.4
Italy	734	77.3	12.3	7.2	1.6	1.6
Netherlands	829	34.5	25.9	26.5	6.2	6.9
Sweden	652	20.1	22.1	34.0	11.8	12.0
Full sample	5,820	39.9	22.4	23.4	6.9	7.4

Table 4

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 books (0-10)	One shelf (11-25 books)	One bookcase (26-100 books)	Two bookcases (101-200 books)	More than two bookcases (> 200 books)
Austria	228	54.8	69.4	50.0	42.9	33.3	0.0
Belgium	669	47.7	52.4	49.6	37.7	27.3	40.0
Czech Rep.	672	50.9	68.7	57.7	41.9	28.6	25.6
Denmark	747	42.3	57.4	43.5	40.9	35.7	23.6
Germany	516	46.5	62.0	50.7	35.8	29.7	12.2
France	773	36.0	46.8	35.4	18.5	31.8	19.3
Italy	734	57.8	63.5	44.4	32.1	33.3	25.0
Netherlands	829	30.5	35.7	32.1	29.1	17.6	15.8
Sweden	652	29.9	46.6	39.6	26.6	15.6	7.7
Full sample	5,820	42.8	54.8	44.2	33.3	26.8	18.6

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, 69.4% were living in a rural area.

Table 5

Schooling Regressions, by Number of Books. Dependent Variable: Years of Education

	Up to 200 books	None or very few books (0-10)	One shelf or more (11-200 books)
Compulsory edu.	0.196*	0.375***	0.004
	(0.106)	(0.116)	(0.143)
Rural × Comp. edu.	0.365***	0.254**	0.311**
	(0.070)	(0.102)	(0.081)
Rural area at age ten	-3.715***	-2.312***	-3.164***
	(0.555)	(0.798)	(0.635)
Sample size	5,390	2,325	3,065
R-squared	0.187	0.254	0.097
F-test statistic	20.70	12.16	7.91

Note. All regressions include birth cohort dummies, country dummies and country-specific quadratic cohort trends. 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 variable ‘rural area' and years of compulsory education.

Table 6

OLS Regressions. Dependent Variable: Lifetime Earnings and Current or Main Wage

	Lifetime earnings	Current or main wage
Years of education	0.031***	0.042***
	(0.003)	(0.003)
Rural area at age ten	-0.048**	-0.039**
	(0.019)	(0.019)
Sample size	5,390	5,390
R-squared	0.229	0.152

Note. All regressions include birth cohort dummies, country dummies and country-specific quadratic cohort trends. Robust standard errors in parentheses. ***p<0.01, **p<0.05, *p<0.1.

Table 7

2SLS Regressions. Dependent Variable: Lifetime Earnings and Current or Main Wage

	Lifetime earnings	Current or main wage
Years of education	0.090***	0.056**
	(0.030)	(0.026)
Rural area at age ten	0.010	-0.026
	(0.032)	(0.031)
Sample size	5,390	5,390
First stage F-test statistic	20.70	20.68
p-value Hansen J statistic	0.905	0.524

Note. All regressions include birth cohort dummies, country dummies and country-specific quadratic cohort trends. Standard errors clustered by country and cohort in parentheses. ***p<0.01, **p<0.05, *p<0.1.

Table 8

Very Few Books vs. a Shelf or More. Dependent Variable: Lifetime Earnings

	Very few books (0-10)			One shelf or more (11-200 books)		
	OLS	2SLS	LIML	OLS	2SLS	LIML
Years of education	0.032***	0.052	0.055	0.028***	0.211***	0.219***
	(0.005)	(0.048)	(0.055)	(0.004)	(0.059)	(0.062)
Rural area at age ten	0.005	0.014	0.016	-0.079***	0.065	0.071
	(0.032)	(0.034)	(0.036)	(0.024)	(0.058)	(0.060)
Sample size	2,325	2,325	2,325	3,065	3,065	3,065
R-squared	0.201			0.242		
First stage F-test statistic		12.16			7.91	
p-value Hansen J statistic		0.082			0.423	

Note. All regressions include birth cohort dummies, country dummies and country-specific quadratic cohort trends. OLS with robust standard errors, 2SLS and LIML with standard errors clustered by country and cohort in parentheses. ***p<0.01, **p<0.05, *p<0.1.

Table 9

Very Few Books vs. a Shelf or More. Dependent Variable: Lifetime Earnings. Excluding Individuals Born from 1920 to 1929

	Very few books (0-10)			One shelf or more (11-200 books)		
	OLS	2SLS	LIML	OLS	2SLS	LIML
Years of education	0.034***	0.058	0.065	0.026***	0.199***	0.201***
	(0.005)	(0.062)	(0.079)	(0.004)	(0.060)	(0.060)
Rural area at age ten	0.014	0.022	0.024	-0.062***	0.057	0.058
	(0.033)	(0.034)	(0.037)	(0.024)	(0.054)	(0.055)
Sample size	2,026	2,026	2,026	2,805	2,805	2,805
R-squared	0.199			0.239		
First stage F-test statistic		10.51			6.24	
p-value Hansen J statistic		0.124			0.714	

Note. All regressions include birth cohort dummies, country dummies and country-specific quadratic cohort trends. OLS with robust standard errors, 2SLS and LIML with standard errors clustered by country and cohort in parentheses. ***p<0.01, **p<0.05, *p<0.1.

Table 10

Distribution of Books at Home, by Occupation of the Main Breadwinner and Area of Residence at Age Ten

	White collar			Blue collar		
	Very few books	One shelf	One or two bookcases	Very few books	One shelf	One or two bookcases
Rural at age ten	0.28	0.27	0.45	0.55	0.24	0.21
Urban at age ten	0.16	0.21	0.63	0.42	0.26	0.32

Table 11

Occupation of the Main Breadwinner at Age Ten. Dependent Variable: Lifetime Earnings

	White collar			Blue collar		
	OLS	2SLS	LIML	OLS	2SLS	LIML
Years of education	0.035****	0.076	0.082	0.029***	0.090***	0.090***
	(0.006)	(0.092)	(0.106)	(0.003)	(0.031)	(0.031)
Rural area at age ten	-0.021	0.017	0.022	-0.041*	-0.009	-0.009
	(0.052)	(0.086)	(0.096)	(0.022)	(0.025)	(0.025)
Sample size	998	998	998	4,392	4,392	4,392
R-squared	0.321			0.219		
First stage F-test statistic		1.62			9.69	
p-value Hansen J statistic		0.535			0.944	

Note. All regressions include birth cohort dummies, country dummies and country-specific quadratic cohort trends. OLS with robust standard errors, 2SLS and LIML with standard errors clustered by country and cohort in parentheses. ***p<0.01, **p<0.05, *p<0.1.

Table 12

Percentage with Poor Health Conditions at Age Ten, by Number of Books at Home at Age Ten

	Ever missed school	Serious illness	Regular dentist	Poor health	Any vaccines
One shelf or more	14.1	28.3	63.8	8.1	95.8
Very few books	11.7	20.5	30.2	7.4	93.0

Table 13

Effects of the Number of Books at Home on Log Standardised Test Scores

Number of books at home	Reading skills; PIRLS 2006; tests taken at age 9-10	Math skills; TIMSS 1995; tests taken at age 8-11	Reading skills; PISA 2006; tests taken at age 15	Math skills; PISA 2006; tests taken at age 15
0-10	baseline	baseline	baseline	baseline
11-25	0.019***	0.067***	0.060***	0.047***
	(0.001)	(0.003)	(0.002)	(0.002)
26-100	0.045***	0.127***	0.117***	0.109***
	(0.001)	(0.002)	(0.002)	(0.001)
101-200	0.063***	0.154***	0.157***	0.147***
	(0.002)	(0.003)	(0.002)	(0.002)
More than 200	0.077***	0.161***	0.191***	0.187***
	(0.002)	(0.003)	(0.002)	(0.002)
Sample size	105,670	79,221	197,751	197,751

Note. The dependent variables are the logarithm of: PIRLS 2006 reading test scores in the fourth grade, TIMSS 1995 math test scores in the third and fourth grade, PISA 2006 reading test scores in the ninth grade, 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.

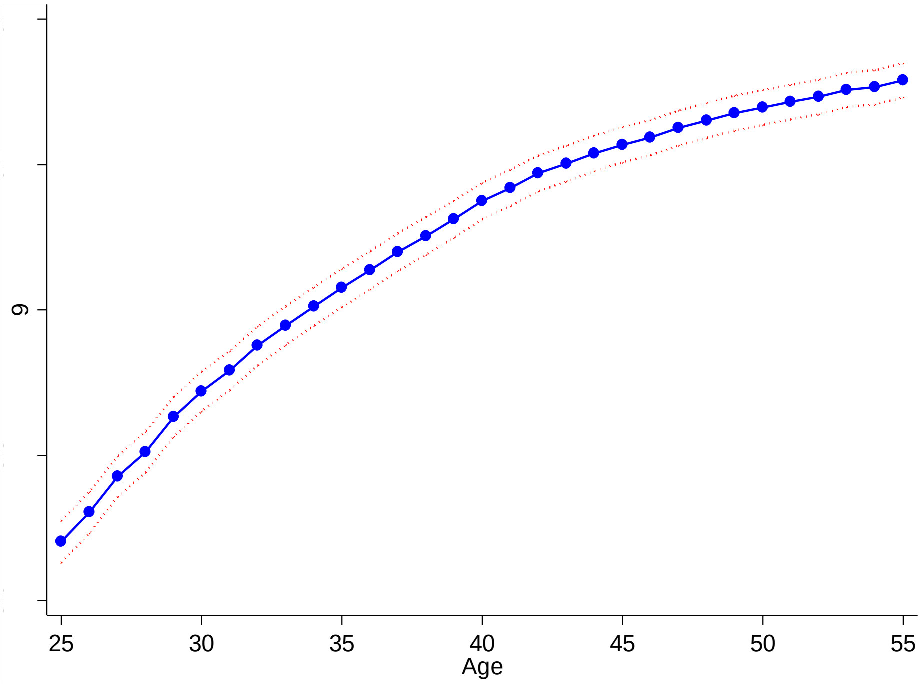

Fig. 1. *Age-earnings profiles net of country and cohort effects for individuals aged 25 to 55*

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 95% confidence interval for mean log earnings is also shown.

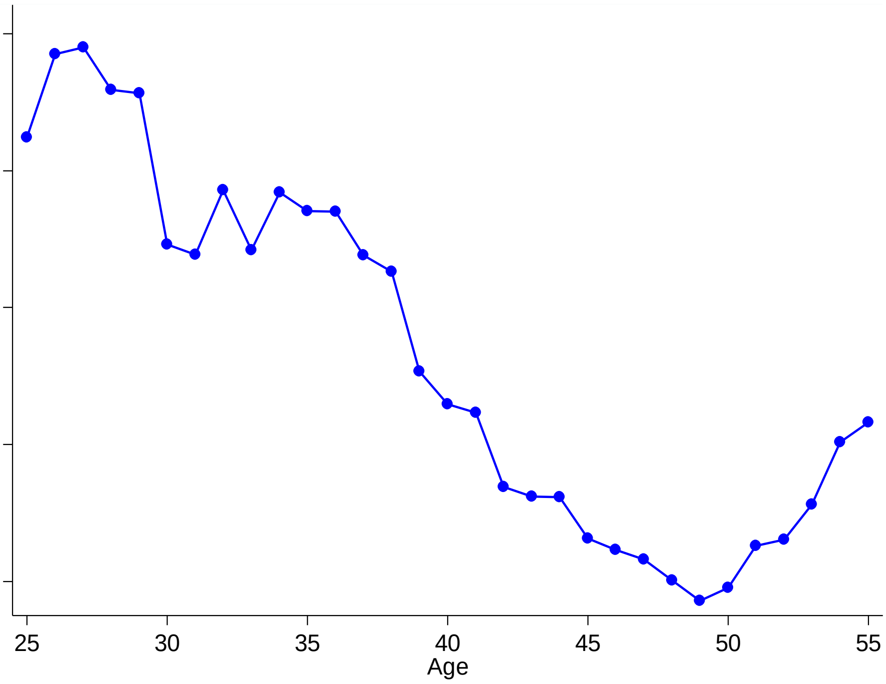

Fig 2. *Log earnings at or above country-specific median years of education (YS) minus log earnings below country-specific median years of education*

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.

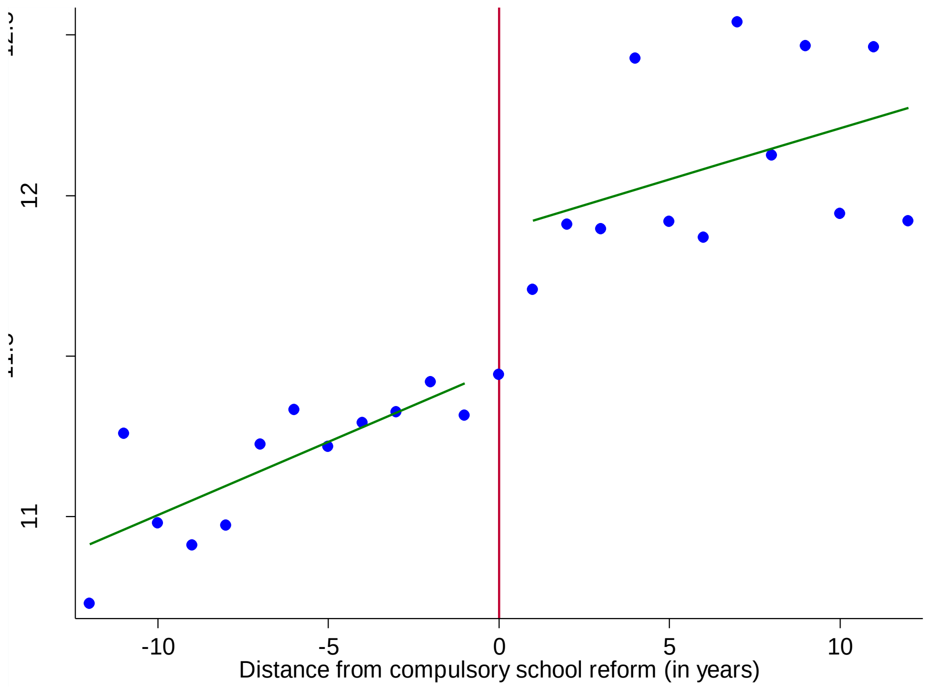

Fig 3. *Years of education, by distance from school reform*

Note. For countries with more than one compulsory school reform (the Czech Republic, France, the Netherlands and Sweden), we only use the last reform.

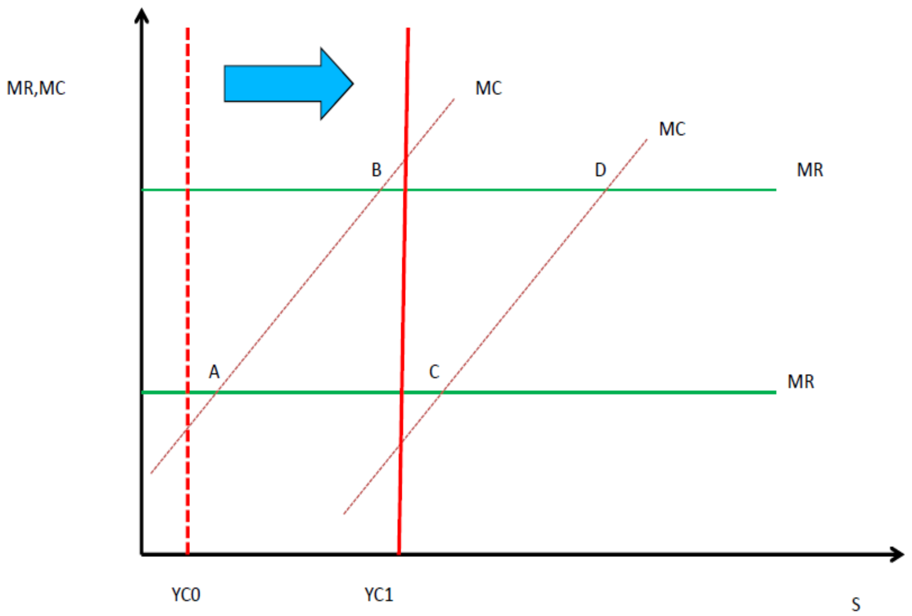

Fig 3. Compulsory school reforms and years of education

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 very few books at home, individuals B and D from families with many books at home.

University of Padova and IZA
University of Padova, CEPR and IFS
University of Padova

References

Acemoglu D. and Angrist, J. (2001). 'How large are human capital externalities? Evidence from compulsory schooling laws', in (B.S. Bernanke and K. Rogoff, eds.), *NBER Macroeconomics Annual 2000, Vol. 15*, pp. 9–59, Cambridge, MA: National Bureau of Economic Research.

Allen, J. and Barnsley, R. (1993). 'Streams and tiers: The interaction of ability, maturity and training in systems with age dependent recursive selection', *Journal of Human Resources*, vol. 28(3), pp. 649–659.

Allington, R.L., McGill-Franzen, A., Camilli, G., Williams, L., Graff, J., Zeig, J., Zmach, C. and Nowak, R. (2010). 'Addressing summer reading setback among economically disadvantaged elementary students', *Reading Psychology*, vol. 31(5), pp. 411–427.

Angrist, J.D. and Pischke, J.-S. (2009). '*Mostly Harmless Econometrics: An Empiricist's Companion*', Princeton, NJ: Princeton University Press.

Ashenfelter, O. and Rouse, C. (1998). 'Income, schooling, and ability: Evidence from a new sample of identical twins', *Quarterly Journal of Economics*, vol. 113(1), pp. 253–284.

Banks, J. and Mazzonna, F. (2012). ‘The effect of education on old age cognitive abilities: evidence from a regression discontinuity design’, *Economic Journal*, vol. 122, pp. 418–448.

Bhuller, M., Mogstad, M. and Salvanes, K.G. (2011). ‘Life-cycle bias and the returns to schooling in current and lifetime earnings’, IZA Discussion Paper No. 5788.

Bingley, P. and Martinello, A. (2014). ‘Measurement error in income and schooling, and its consequences for the estimation of linear models’, SFI Working Paper No. 01:2014.

Böhlmark, A. and Lindquist, M.J. (2006). ‘Life cycle variations in the association between current and lifetime income: replication and extension for Sweden’, *Journal of Labor Economics*, vol. 24(4), pp. 879–896.

Brenner, J. (2010). ‘Life-cycle variations in the association between current and lifetime earnings: evidence for German natives and guest workers’, *Labour Economics*, vol. 17(2), pp. 392–406.

Brunello, G., Fabbri, D. and Fort, M. (2013). ‘The causal effect of education on the body mass: Evidence from Europe’, *Journal of Labor Economics*, vol. 31(1), pp. 195-223.

Brunello, G., Fort, M. and Weber, G. (2009). ‘Changes in compulsory schooling, education and the distribution of wages in Europe’, *Economic Journal*, vol. 119, pp. 516–539.

Card, D. (1999). ‘The causal effect of education on earnings’, in (O.C. Ashenfelter and D.

Card, eds.), *Handbook of Labor Economics, Vol.3A*, pp. 1801–1863, Amsterdam: Elsevier.

Card, D. (2001). 'Estimating the return to schooling: Progress on some persistent econometric problems', *Econometrica*, vol. 69(5), pp. 1127–1160.

Carneiro, P. and Heckman, J.J. (2002). 'The evidence on credit constraints in post-secondary schooling', *Economic Journal*, vol. 112, pp. 705–734.

Cunha, F. and Heckman, J.J. (2007). 'The technology of skill formation', *American Economic Review*, vol. 97(2), pp. 31–47.

Cunha, F., Heckman, J.J. and Schennach, S. (2010). 'Estimating the technology of cognitive and noncognitive skill formation', *Econometrica*, vol. 78(3), pp. 883–931.

Devereux, P.J. and Hart, R.A. (2010). 'Forced to be rich? Returns to compulsory schooling in Britain', *Economic Journal*, vol. 120, pp. 1345–1364.

Fryer, R.G. (2011). 'Financial incentives and student achievement: Evidence from randomized trials', *Quarterly Journal of Economics*, vol. 126(4), pp. 1755–1798.

Garrouste, C. (2010). *100 Years of Educational Reforms in Europe: A Contextual Database,* EUR – Scientific and Technical Research series, Vol. 24487, Luxembourg: Publications Office of the European Union.

Garrouste, C. and Paccagnella, O. (2011). 'Data quality: Three examples of consistency across SHARE and SHARELIFE data', in (M. Schröder, ed.) *Retrospective Data Collection in the Survey of Health, Ageing and Retirement in Europe. SHARELIFE Methodology*, pp. 62–72, Mannheim: Mannheim Research Institute for the Economics of Ageing (MEA).

Gensowski, M. (2014). 'Personality, IQ and lifetime earnings', IZA Discussion Paper No. 8235.

Goldin, C. and Katz, L. (2003). 'Mass secondary education and the state: The role of state compulsion in the high school movement', NBER Working Paper No. 10075.

Gould, E.D., Lavy, V. and Paserman, M.D. (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*, vol. 78(3), pp. 938–973.

Haider, S. and Solon, G. (2006). 'Life cycle variation in the association between current and lifetime earnings', *American Economic Review*, vol. 96(4), pp. 1308–1320.

Havari, E. and Mazzonna, F. (2011). 'Can we trust older people's statements on their childhood circumstances? Evidence from SHARELIFE', SHARE Working Paper 05–2011.

Heckman, J.J. (2000). 'Policies to foster human capital', *Research in Economics*, vol. 54(1), pp. 3–56.

Heckman, J.J., Lochner, L.J and Todd, P.E. (2006). 'Earnings functions, rates of return and treatment effects: The Mincer equation and beyond' in (E. Hanushek and F. Welch, eds.), *Handbook of the Economics of Education, Vol.1*, pp. 307–458, Amsterdam: Elsevier.

Heckman, J.J., Pinto, R. and Savelyev, P. (2013). 'Understanding the mechanisms through which an influential early childhood program boosted adult outcomes', *American Economic Review*, vol. 103(6), pp. 2052–2086.

Jensen, R. (2010). 'The (perceived) returns to education and the demand for schooling', *Quarterly Journal of Economics*, vol. 125(2), pp. 515–548.

Judson, R. (1998). 'Economic growth and investment in education. How allocation matters', *Journal of Economic Growth*, vol. 3(4), pp. 337–359.

Kim, J.S. (2007). 'The effects of a voluntary summer reading intervention on reading activities and reading achievement', *Journal of Educational Psychology*, vol. 99(3), pp. 505–515.

van Kippersluis, H., O'Donnell, O. and van Doorslaer, E. (2011). 'Long-run returns to education: Does schooling lead to an extended old age?', *Journal of Human Resources*, vol. 46(5), pp. 695–721.

Lochner, L., and Monge-Naranjo, A. (2011). 'Credit constraints in education', NBER Working Paper No. 17435.

Machin, S., Salvanes, K.G. and Pelkonen, P. (2012). 'Education and mobility', *Journal of the*

European Economic Association, European Economic Association, vol. 10(2), pp. 417–450.

Milligan, K., Moretti, E. and Oreopoulos, P. (2004). ‘Does education improve citizenship? Evidence from the US and the UK’, *Journal of Public Economics*, vol. 88, pp. 1667–1695.

Mincer, J.A. (1974). *Schooling, Experience, and Earnings*, New York, NY: National Bureau of Economic Research. Distributed by Columbia University Press.

Moulton, B.R. (1990). ‘An illustration of a pitfall in estimating the effects of aggregate variables on micro units’, *Review of Economics and Statistics*, vol. 72(2), pp. 334–338.

Murphy, K.M. and Welch, F. (1990). ‘Empirical age-earnings profiles’, *Journal of Labor Economics*, vol. 8(2), pp. 202–229.

Oreopoulos, P. (2006). ‘Estimating average and local average treatment effects of education when compulsory schooling laws really matter’, *American Economic Review*, vol. 96(1), pp. 152–175.

Pischke, J.-S. and von Wachter, Till. (2005). ‘Zero returns to compulsory schooling in Germany: Evidence and interpretation’, NBER Working Paper No. 11414.

Stephens, M. and Yang, D.-Y. (2014). ‘Compulsory education and the benefits of schooling’, *American Economic Review*, vol. 104(6), pp. 1777-1792.

Trevisan, E., Pasini, G. and Rainato, R. (2011). ‘Cross-country comparison of monetary values from SHARELIFE’, SHARE Working Paper 02–2011.

Weiss, C.T. (2012). ‘Two measures of lifetime resources for Europe using SHARELIFE’, SHARE Working Paper 06–2012.

APPENDIX

(For Online Publication)

Appendix A. Computing Lifetime Income

A.1 *Initial and Final Earnings of Each Employment Spell*

We define lifetime earnings (or income) as the net present value at age ten of the stream of net wages earned over the life cycle from age ten using a discount rate of 2% (r = 0.02). We use all wages from labour market entry until retirement. If an individual started working before age ten, we only use the wages from age ten onwards. We mainly draw data on work histories from SHARELIFE but also some data from SHARE waves 1 and 2. Wages are converted using PPP exchange rates and CPI indices into 2006 Euro. PPP-adjusted exchange rates and CPI measures are taken from the OECD and national sources.[33]

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 1994 to March 1996. This implies that someone who reports to have started working in an employment spell in 1994 and stopped in 1994 will be treated equally to someone who started in 1994 but stopped in 1995.

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 income from employment in the previous year. If no value is reported in the imputation modules, we impute missing wage values using predictive mean matching. Predictive mean

[33] More details can be found in Trevisan *et al.* (2011).

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.[34] 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). Only the current and the main employment spells have wage measures both at the start and the end of the spell.

For those with more than one job in the career, we predict wages at the end of the spell 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. That is, potential experience starts at the end of full time education. 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

[34] 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.

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 ten.[35] We estimate the following linear model

$$y_{ic} = \beta_1 E_{ic} + \beta_2 E_{ic}^2 + \beta_3 E_{ic} S_i + \beta_4 E_{ic} X_{ic} + \beta_5 S_i + \beta_6 X_{ic} + \beta_7 Q_i + U_i \tag{A.1}$$

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 wages at the end of the job spell starting from earnings at the start of the spell, which we typically observe. We compute

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

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.[36] Armed with the wages at the start and the end of each

[35] 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.

[36] As a robustness check, we also used age instead of potential experience as a running variable and did not include education in the current wage regression. Despite the loss of precision, we find all results in the main analysis of the paper to be qualitatively the same (results available from the authors upon request).

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.[37] Table A5 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 and controls.[38] We obtain an empirical distribution of wages, and eliminate outliers in this distribution by censoring observations that are above the 99th and below the 1st percentile.

A.2 *Lifetime Earnings*

We multiply monthly earnings by 12 to obtain annual earnings. For each individual i, the discounted sum of the stream of annual earnings is

$$AW_i = r \cdot \sum_{j=1}^{J} Y_{ij} \sum_{k=1}^{K} (1 + gr_{ij})^{(k_i - 1)} / (1 + r)^{(ST_{ij} - (BY_i + 11) + k_i)} \tag{A.3}$$

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, ST_j the year when employment spell j started, BY the year of birth and r the interest rate.

[37] A validation study which uses the German Socio-economic Panel also suggests that our procedure is quite accurate (see Appendix B).

[38] 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.

To illustrate, if someone is born in 1950 and starts working in 1960, the first wage in 1960 is not discounted, but the wage in 1961 is discounted with $1+r$, the wage in 1962 with $(1+r)^2$ and so forth. While the first wage is reported by the individual, subsequent earnings are predicted using the within-job growth rate of earnings as discussed above.

We do not have information on actual retirement age for those who were still working at the time of the SHARELIFE interview, but we know from previous waves their expected retirement age.[39] 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 age at which they expect to collect pension benefits and their current age. When 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 being ever unemployed until retirement. 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

$$CWORK_i = r \cdot Y_{icurr} \cdot \sum_{s=1}^{V} (1+gr_{is})(1-qx_{is}) / (1+r)^{(INT_i - (BY_i + 11) + s_i)} \qquad \text{(A.4)}$$

[39] When individuals do not report at what age they will start collecting pension, we use information on statutory retirement age in their country.

where Y_{curr} denotes 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, $1 + gr_s$ the annual growth rate of income during the employment spell, qx_s the probability of death within the age interval $[BY+s, BY+s+1)$, INT the interview year, BY is the year of birth and r is the interest rate.

Thus, lifetime earnings are equal to AW if an individual is retired and to $AW + CWORK$ if he is still working.

Appendix B. 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 and 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 has equivalently defined variables for panel databases from the UK (BHPS), Australia (HILDA), South Korea (KLIPS), the US (PSID), and Canada (SLID).

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

$$y_{i2008} = \beta_1 E_{i2008} + \beta_2 E_{i2008}^2 + \beta_3 E_{i2008} S_i + \beta_4 E_{i2008} X_{i2008} + \beta_5 S_i + \beta_6 X_{i2008} + \beta_7 W_i + U_i \quad \text{(B.1)}$$

where y_{2008} denotes the logarithm of individual annual labour earnings in 2008, E is potential experience, S education, 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* a disturbance term. Our sample consists of all males born between 1945 and 1956 who report information on individual labour 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 B1 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 all 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.[40]

Appendix C. Age-Earnings Profiles by Education Levels

We investigate whether age-earnings profiles by education are parallel by estimating the following regression

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

[40] We repeated the same exercise using data from the Italian Survey of Household Income and Wealth (SHIW). The results (available from the authors upon request) are very similar to the ones using the German SOEP.

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, f is a time invariant individual effect and we exclude the individuals who have experienced unemployment during the sample period.[41]

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. We stress, however, 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, we find that convergence prevails.

We estimate equation (C.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 C1. There is evidence that earnings profiles are not parallel by education but mildly converge over time ($\beta_4 < 0$ and statistically significant).

Compared to Figure 2 in the text, Figure C.1 shows a more detailed break-down of earnings profiles (net of country and cohort effects) by education levels. The figure consists of two panels: in the upper panel we plot for each age the vertical distance between log earnings for

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

individuals with college and individuals with high school education; in the lower panel, we plot the vertical distance between log earnings for individuals with high school and individuals with 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.

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

$$W_{it} = \exp(\beta_0 + \beta_1 A_{it} + \beta_2 A_{it}^2 + \beta_3 S_i + \beta_4 S A_{it}) \tag{C.2}$$

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

$$W_{ia+x} = W_{ia} \exp(x_i \beta_1 + x_i^{\ 2} \beta_2 + 2 x_i \beta_2 A_{ia} + x_i \beta_4 S_i)\,. \tag{C.3}$$

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 ≥ 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

$$I_{ia} = W_{ia} + \frac{W_{ia+1}}{1+r} + \frac{W_{ia+2}}{(1+r)^2} + \ldots + \frac{W_{ia+q}}{(1+r)^q}\,. \tag{C.4}$$

Taking logarithms of equation (C.3) and using equation (C.4) we obtain

$$\ln I_{ia} = \ln W_{ia} + \ln\left\{1 + \Theta + \frac{\exp(\beta_4 S_i)}{1+r} + \frac{\exp(2\beta_4 S_i)}{(1+r)^2} + \ldots + \frac{\exp(q\beta_4 S_i)}{(1+r)^q}\right\} \tag{C.5}$$

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

$$\ln I_{ia} = \ln W_{ia+x} - x_i \beta_4 S_i + \ln\left\{1 + \frac{\exp(\beta_4 S_i)}{1+r} + \frac{\exp(2\beta_4 S_i)}{(1+r)^2} + \ldots + \frac{\exp(q\beta_4 S_i)}{(1+r)^q}\right\} . \tag{C.6}$$

By taking derivatives with respect to S on both sides of equation (C.5) we obtain

$$\frac{\partial \ln I_{ia}}{\partial S} = \frac{\partial \ln W_{ia+x}}{\partial S} - \beta_4 x_i + \frac{\beta_4 \left\{\frac{\exp(\beta_4 S_i)}{1+r} + \frac{2\exp(2\beta_4 S_i)}{(1+r)^2} + \ldots + \frac{q\exp(q\beta_4 S_i)}{(1+r)^q}\right\}}{1 + \frac{\exp(\beta_4 S_i)}{1+r} + \frac{\exp(2\beta_4 S_i)}{(1+r)^2} + \ldots + \frac{\exp(q\beta_4 S_i)}{(1+r)^q}} . \tag{C.7}$$

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

$$x_i^* = \frac{\left\{\frac{\exp(\beta_4 S_i)}{1+r} + \frac{2\exp(2\beta_4 S_i)}{(1+r)^2} + \ldots + \frac{q\exp(q\beta_4 S_i)}{(1+r)^q}\right\}}{1 + \frac{\exp(\beta_4 S_i)}{1+r} + \frac{\exp(2\beta_4 S_i)}{(1+r)^2} + \ldots + \frac{\exp(q\beta_4 S_i)}{(1+r)^q}} . \tag{C.8}$$

Appendix D. The Becker-Card Model of Schooling

Assume that individuals maximise

$$U(S_i, y_i) = \ln y_i - C(S_i) \tag{D.1}$$

(see Card, 1999), where y is lifetime earnings, C the cost of schooling, S years of schooling and i denotes the individual. Let lifetime earnings be described by the following log-linear specification

$$\ln y_i = b_0 + b_1 S_i + b_2 FB_i + b_3 R_i + b_4 FB_i R_i + b_5 FB_i S_i + a_i + e_i \tag{D.2}$$

where R is a dummy variable equal to 1 if the individual lived in a rural area at age ten and to 0 otherwise; FB a dummy variable equal to 1 if the individual had very few books (less than a shelf) in the household at age ten and to 0 otherwise; a denotes ability and e a disturbance term.

In the selected specification, the marginal returns to schooling are allowed to vary with the number of books. The identifying assumption we make is that the effect of books on marginal returns does not vary with the place of residence (rural or urban) at age ten. Conditional on this assumption, we cannot reject the hypothesis that in our data marginal returns to schooling MR are independent of R. Therefore we have

$$MR_i = b_1 + b_5 FB_i \,. \tag{D.3}$$

The costs of schooling C are given instead by

$$C_i = c_0 + \frac{c_1}{2} S_i^2 + c_2 FB_i + c_3 R_i + c_4 FB_i R_i + c_5 FB_i S_i + c_6 R_i S_i + c_7 FB_i R_i S_i + c_8 S_i a_i \tag{D.4}$$

which yields the following marginal cost function

$$MC_i = c_1 S_i + c_5 FB_i + c_6 R_i + c_7 FB_i R_i + c_8 a_i \,. \tag{D.5}$$

Privately optimal schooling S^* is obtained by equalizing marginal costs and marginal benefits ($MR = MC$), which gives

$$S_i^* = \frac{b_1}{c_1} + \frac{b_5 - c_5}{c_1} FB_i - \frac{c_6}{c_1} R_i - \frac{c_7}{c_1} FB_i R_i - \frac{c_8}{c_1} a_i \,. \tag{D.6}$$

Actual schooling S is higher than optimal schooling for the individuals induced to increase their education by compulsory schooling laws *Z*. We define actual schooling as

$$S_i = \frac{b_1}{c_1} + \frac{b_5 - c_5}{c_1} FB_i - \frac{c_6}{c_1} R_i - \frac{c_7}{c_1} FB_i R_i + \pi(FB_i, R_i) Z_i + v_i \qquad \text{(D.7)}$$

where v is the error term, $\pi \neq 0$ for compliers, $\pi = 0$ for the individuals not affected by school reforms and we assume that the effects of *Z* on *S* vary with *FB* and *R*.

When we estimate equation (E.7) on our data, we find that $c_1 > 0$, $b_5 - c_5 < 0$, $c_6 > 0$, and cannot reject the null hypothesis that $c_7 = 0$. These estimates are consistent with $b5 < 0$, $c5 = 0$, implying that: a) marginal benefits increase with the number of books at age ten; b) marginal costs are higher for individuals living in rural areas at age ten, as shown in Figure 4.

Appendix E. The Effect of Reporting Errors in the Books Dummy

Let *D* be a dummy variable for reported books at age ten (1: many books; 0: few books), D^* a dummy variable for the true number of books (1: many books; 0: few books), let *Y* denote the returns to education, *X* additional covariates and *f* the conditional density. Further assume that $E[Y|X, D^*, D] = E[Y|X, D^*]$, i.e. the report is uninformative on *Y* once the true value is taken into account.

Then the estimated returns to education for those with $D = 0$ and $D = 1$ are respectively

$$E[Y \mid X, D=0] = E[Y \mid X, D^*=0] f(D^*=0 \mid X, D=0)] + E[Y \mid X, D^*=1] f(D^*=1 \mid X, D=0) \qquad \text{(E.1)}$$

$$E[Y \mid X, D=1] = E[Y \mid X, D^*=1] f(D^*=1 \mid X, D=1)] + E[Y \mid X, D^*=0] f(D^*=0 \mid X, D=1)\ . \qquad \text{(E.2)}$$

Equations (E.1) and (E.2) can be rewritten as

$$E[Y \mid X, D=0] = E[Y \mid X, D^* = 0] + \{E[Y \mid X, D^* = 1] - E[Y \mid X, D^* = 0]\} f(D^* = 1 \mid X, D = 0) \quad \text{(E.3)}$$

$$E[Y \mid X, D=1] = E[Y \mid X, D^* = 1] - \{E[Y \mid X, D^* = 1] - E[Y \mid X, D^* = 0]\} f(D^* = 0 \mid X, D = 1)\,. \quad \text{(E.4)}$$

Assume the returns to education when books are correctly reported are higher with many books than with few books, so that $(E[Y \mid X, D^* = 1] - E[Y \mid X, D^* = 0] > 0)$ – books are good for earnings. Under this assumption, equation (E.3) implies

$$E[Y \mid X, D=0] > E[Y \mid X, D^* = 0] \quad \text{(E.5)}$$

That is, estimated returns for those with few books are upward biased: our very low estimated returns for those with few books cannot be due to measurement error in D. Under the same assumption, equation (E.4) implies

$$E[Y \mid X, D=1] < E[Y \mid X, D^* = 1] \quad \text{(E.6)}$$

Estimated returns for those with many books are downward biased: our high estimated returns for those with many books cannot be due to measurement error in D. Equations (E.5) and (E.6) together imply

$$E[Y \mid X, D=1] - E[Y \mid X, D=0] < E[Y \mid X, D^* = 1] - E[Y \mid X, D^* = 0]\,. \quad \text{(E.7)}$$

We conclude that the gap in the returns to education generated by differences in the reported number of books is lower than the gap induced by differences in the true number of books.

Table A1

Descriptive Statistics for the Sample Excluding Individuals with More Than 200 Books

	Mean	Std. Dev.	Min	Max	Median
Lifetime earnings	7,593.69	5,312.11	105.82	36,216.31	6,437.23
Years of education	11.47	3.87	2	25	11
Years of compulsory education	7.51	1.58	4	10	8
Age	67.13	8.74	52	89	66
Years of work	36.72	8.16	5	63	38
Number of jobs during career	3.09	2.01	1	18	3
Very few books at age ten	0.43	0.50	0	1	
Rural area or village at age ten	0.45	0.50	0	1	
Very few books × rural area	0.24	0.43	0	1	
Poor housing conditions at age ten	0.50	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	0.32	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	5,390	5,390	5,390	5,390	5,390

Table A2

OLS and 2SLS Regressions Including Individuals with More Than 200 Books. Dependent Variable: Lifetime Earnings

	Lifetime earnings OLS	Lifetime earnings 2SLS
Years of education	0.032***	0.103***
	(0.003)	(0.033)
Rural area at age ten	-0.057***	0.029
	(0.019)	(0.040)
Sample size	5,820	5,820
R-squared	0.235	
First stage F-test statistic		16.19

Note. All regressions include birth cohort dummies, country dummies and country-specific quadratic cohort trends. OLS with robust standard errors and 2SLS with standard errors clustered by country and cohort in parentheses. ***p<0.01, **p<0.05, *p<0.1.

Table A3

OLS and 2SLS Estimates by Country, with and without the Dummy 'Rural Area at Age Ten' as Additional Regressor.

Dependent Variable: Lifetime Earnings

	OLS including rural area at age ten	2SLS including rural area at age ten	OLS excluding rural area at age ten	2SLS excluding rural area at age ten	Sample size
All countries	0.031***	0.090***	0.032***	0.081***	5,390
	(0.003)	(0.030)	(0.003)	(0.021)	
Austria	0.013	0.185	0.012	-0.004	217
	(0.013)	(0.131)	(0.011)	(0.044)	
Belgium	0.022***	0.057	0.022***	0.070	644
	(0.007)	(0.047)	(0.007)	(0.048)	
Czech Republic	0.021***	0.204	0.022***	0.054	629
	(0.008)	(0.419)	(0.008)	(0.043)	
Denmark	0.046***	0.118***	0.049***	0.102***	641
	(0.009)	(0.041)	(0.009)	(0.027)	
France	0.029***	-0.091	0.031***	0.050	716
	(0.008)	(0.121)	(0.008)	(0.051)	
Germany	0.039***	0.072	0.041***	0.084*	475
	(0.009)	(0.057)	(0.009)	(0.047)	
Italy	0.034***	0.099	0.031***	-0.026	722
	(0.007)	(0.162)	(0.007)	(0.034)	
Netherlands	0.041***	-0.056	0.041***	-0.104	772
	(0.007)	(0.220)	(0.007)	(0.140)	
Sweden	0.019***	0.050	0.024***	0.185***	574
	(0.007)	(0.147)	(0.007)	(0.060)	

Note. All regressions include birth cohort dummies and quadratic cohort trends (these are country-specific in the first row). OLS with robust standard errors and 2SLS with standard errors clustered by cohort in parentheses. ***p<0.01, **p<0.05, *p<0.1.

Table A4

2SLS Estimates When We Drop From the Sample One Country at a Time.

Dependent Variable: Lifetime Earnings

	All countries	No AUT	No GER	No SWE	No NLD
Years of education	0.090***	0.094***	0.089***	0.072**	0.086***
	(0.030)	(0.032)	(0.030)	(0.029)	(0.032)
Rural area at age ten	0.010	0.021	0.012	0.014	-0.007
	(0.032)	(0.037)	(0.033)	(0.032)	(0.037)
Sample size	5,390	5,173	4,915	4,816	4,618
	No ITA	No FRA	No DNK	No BEL	No CZE
Years of education	0.125***	0.093***	0.062*	0.092***	0.094***
	(0.031)	(0.028)	(0.036)	(0.034)	(0.030)
Rural area at age ten	0.015	0.008	-0.007	0.015	0.014
	(0.034)	(0.030)	(0.034)	(0.038)	(0.035)
Sample size	4,668	4,674	4,749	4,746	4,761

Note. All regressions include birth cohort dummies, country dummies and country-specific quadratic cohort trends. Standard errors clustered by country and cohort in parentheses. ***p<0.01, **p<0.05, *p<0.1.

Table A5

Prediction Error for Current and Main Wages

	Sample size	Mean	Std. Dev.
Log current income	2,305	9.942	0.474
Predicted log current income	2,305	9.936	0.827
Prediction error of log current income	2,305	0.006	0.785
Log main income	4,687	9.824	0.684
Predicted log main income	4,687	9.756	1.028
Prediction error of log main income	4,687	0.068	1.089

Table B1

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	5.675	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	5.789	0.024	0.263
1989	784	5.811	5.782	0.029	0.241
1990	756	5.826	5.809	0.017	0.497
1991	729	5.836	5.820	0.016	0.534
1992	693	5.840	5.844	-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	5.904	0.011	0.724
1997	528	5.935	5.929	0.006	0.848
1998	480	5.935	5.948	-0.013	0.727
1999	447	5.957	5.954	0.002	0.951
2000	441	5.973	5.931	0.041	0.279
2001	401	5.963	5.934	0.028	0.467
2002	374	5.995	5.913	0.082	0.052
2003	344	6.027	5.899	0.128	0.004
2004	319	6.034	5.883	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	5.853	0.068	0.286

Table C1

First Difference Estimates of Age-Earnings Profiles

Variable	Age 21-55	Age 25-55	Age 21-50	Age 25-50
Age/1000	-0.536***	-0.449***	-0.653***	-0.558***
	(0.049)	(0.050)	(0.064)	(0.067)
Years of education/1000	-0.447***	-0.409***	-0.492***	-0.452***
	(0.057)	(0.056)	(0.063)	(0.064)
Constant	0.042***	0.038***	0.046***	0.042***
	(0.002)	(0.002)	(0.003)	(0.003)
Sample size	80,938	74,079	68,856	61,997

Note. Longitudinal panel of individuals always employed from age 21 to age 55. All regressions include country and country by age effects. Robust standard errors in parentheses. ***p<0.01, **p<0.05, *p<0.1.

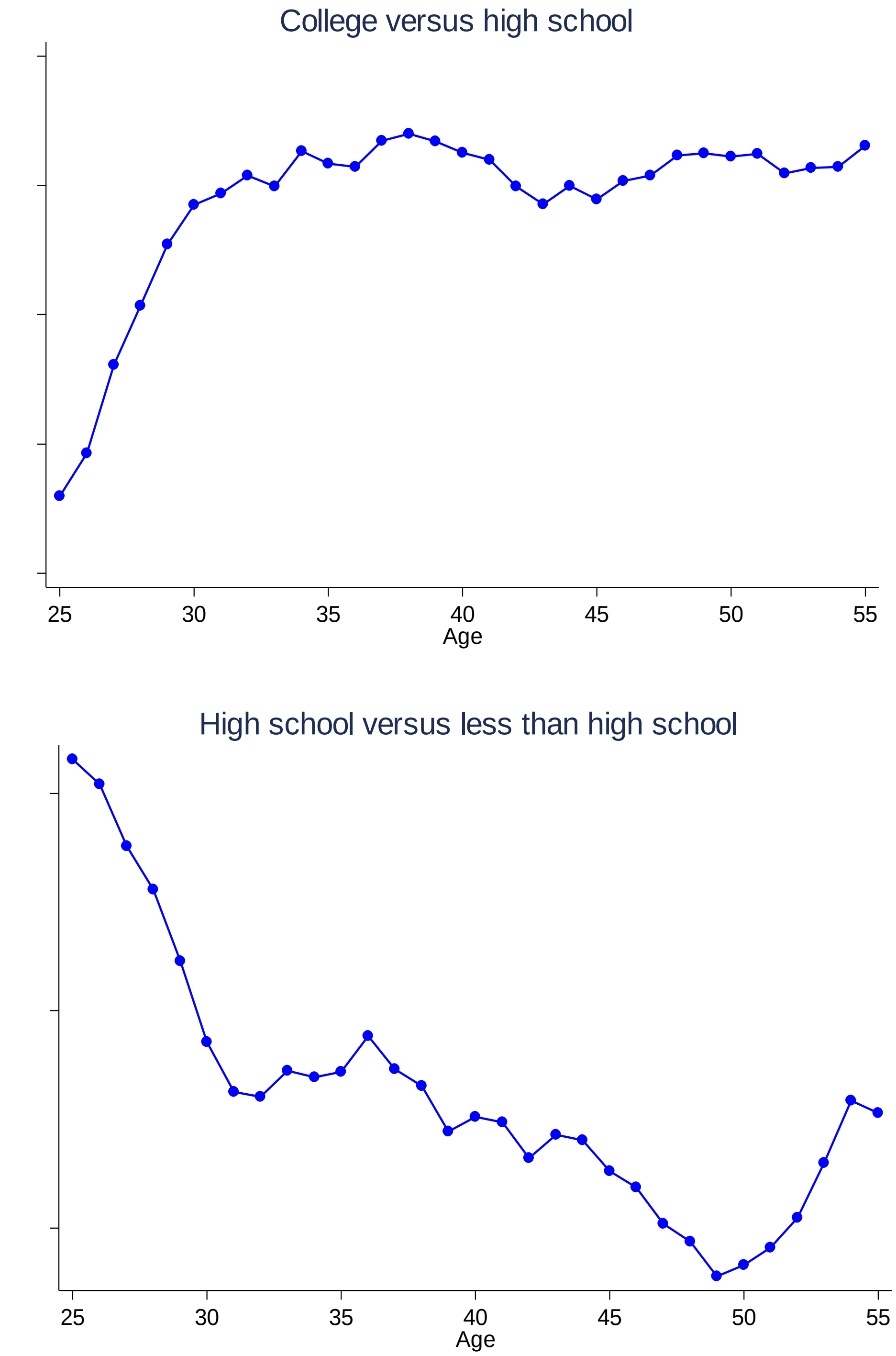

Fig. C1. Age-earnings profiles by education levels

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.