Genetic Endowments and the Long-Run Health Returns to Compulsory Schooling

Johannes Hollenbach

August 25, 2026

Get the latest version

Abstract

Genetic differences tied to educational attainment predict later-life success and wellbeing. Using the English Longitudinal Study of Ageing and individuals aged 65 to 85, I document a sizable genetic gap in late-life health, measured by a frailty index that counts accumulated health deficits. The gradient weakens with educational attainment and disappears among the highest educated. I then estimate causal effects of education on old-age frailty by genetic propensity for education, using a fuzzy regression discontinuity design based on the 1947 UK reform that raised the school leaving age from 14 to 15. More education reduces frailty, but the effect is large and statistically significant only for those with high genetic predisposition for education. The improvement comes from the middle and upper part of the frailty distribution, indicating a delayed buildup of a larger number of deficits rather than a longer deficit-free period. Suggestive evidence on channels points to reduced smoking, higher physical activity, and a lower probability of receiving a disability pension, with no effects on income, employment history, or old-age pensions.

Johannes Hollenbach: RWI - Leibniz Institute for Economic Research, Hohenzollernstr. 1–3, 45128 Essen, Germany, and Paderborn University, Warburger Str. 100, 33098 Paderborn, Germany. I thank Hendrik Schmitz for helpful comments that substantially improved this manuscript.
I gratefully acknowledge financial support from the German Research Foundation (DFG, project number 437564156)

1 Introduction

Genetic differences are important drivers of success and wellbeing, influencing both the choices we make and moderating how environmental factors, that is, all non-genetic influences, affect us. (Benjamin et al. 2012; Houmark, Ronda, and Rosholm 2024; Biroli et al. 2026). Both are particularly apparent in education, where genetic endowments predict how much education individuals obtain as well as the differential returns to schooling (Papageorge and Thom 2020; Barcellos, Carvalho, and Turley 2021). Recent evidence shows that achievement gaps in school by genetic propensity for education are as large as socioeconomic gaps (Houmark et al. 2025). They also predicts success in outcomes outside school, including higher cognitive and noncognitive ability in childhood, earlier language acquisition, greater geographic mobility, higher income and occupational status, better financial planning, higher wealth, and partners of higher socioeconomic status (Belsky et al. 2016; Barth, Papageorge, and Thom 2020; Papageorge and Thom 2020). Part of these correlations reflect a genuine causal effect. Exploiting randomness in the genetic inheritance from parents to children, Carvalho (2025) shows that genetic propensity for education has a positive direct effect on educational attainment, occupation, income, and wealth.

Both children themselves and parents are responsible for these differences. Children with higher genetic endowments acquire and retain skills more effectively given the same inputs (Houmark, Ronda, and Rosholm 2024). Furthermore, genetic propensity for education likely influence educational choices by partially determining the marginal costs of schooling (Carvalho 2025). Children with higher genetic propensity therefore self-select into more schooling (Hollenbach, Schmitz, and Westphal 2026).

Parents reinforce genetic differences by investing more in children with higher genetic propensity for education, meaning genes and parental investment are complements (Houmark, Ronda, and Rosholm 2024). This explains how genetic differences moderate effects of environments, in this case, parental investments. Muslimova et al. (2026) show that being firstborn benefits siblings with higher genetic propensity for education the most in terms of educational attainment. Biroli et al. (2026) find that children with higher propensity benefit most from being old for their grade in terms of performance on a school entry test.

But how do schools respond to genetic gaps, given that they tend to compensate for differences between students, e.g., due to socioeconomic background? Biroli et al. (2026) also show that children with *lower* EA PGI benefit more from being old for their grade when looking at test performance after some years of schooling. Furthermore, higher-quality schools weaken the link between EA PGI and both educational attainment as well as student performance, narrowing preexisting genetic inequality (Arold, Hufe, and Stoeckli 2025; Cheesman et al. 2025). In other words, genes and school quality are substitutes.

For the investment I study here, additional instruction time induced by a higher compulsory schooling age, the gains in educational attainment are also largest among students with the lowest genetic propensity for educational attainment (Barcellos, Carvalho, and Turley 2018; Hollenbach, Schmitz, and Westphal 2026). This is because students with high propensity are more likely to be always-takers

to this type of reform, meaning they would have pursued more education anyway, regardless of any increase in the leaving age. Students with lower genetic propensity are more likely to be compliers, i.e., those who want to leave school at the earliest opportunity and are therefore directly affected by the reform, compelled to attend for another year (Hollenbach, Schmitz, and Westphal 2026).

Given the amplifying role of personal characteristics and family environments and the attenuating role of school for differences in skill development tied to the genetic propensity for education, it is unclear what this means for long-term returns to schooling and effects of policies aimed at increasing instruction time. It could be that individuals at lower genetic propensity for education benefit from reduced achievement gaps, or that individuals at the higher end do because they are more productive in building skills from the additional education. So far, Barcellos, Carvalho, and Turley (2021) show that the additional year of education from a 1972 UK compulsory schooling reform increased earnings only for those in the top tercile of genetic propensity for education. Using the same reform, Barcellos, Carvalho, and Turley (2018) find that the additional education reduced unhealthy body size most for those at higher genetic risk for obesity, while finding no differential effects by genetic propensity for education. Hollenbach, Schmitz, and Westphal (2026) show that additional schooling from the 1947 UK compulsory schooling reform led to higher old-age cognitive abilities only for those at the high end of genetic predisposition for educational attainment.

In this paper, I study the effect of education on health in old age and, importantly, whether the effect varies with genetic propensity for educational attainment. My sample usees data from the English Longitudinal Study of Ageing and includes individuals aged 65–85. The 1947 raising of the UK school leaving age from 14 to 15 provides exogenous variation in schooling at the lower end of the education distribution. I estimate the causal effect of education using a fuzzy regression discontinuity design, where being born in or after the first cohort affected by the reform instruments for years of schooling. My measure of health is the frailty index, the share of age-related health deficits present in an individual To quantify genetic propensity for educational attainment, I use an educational attainment polygenic index (EA PGI), a DNA-based measure that sums the effect sizes of the genetic variants an individual carries that correlate with schooling. It predicts a sizable share of the variation in educational attainment between individuals, but is not deterministic.

The results first show a sizable and relevant genetic health gap in late life. The difference in old-age frailty between the first and the fifth quintile of the EA PGI is about as large as the difference between the first and the fourth quintile of socioeconomic status. However, the PGI-health gradient weakens with educational attainment and disappears among the highest educated. This pattern suggests that increasing schooling at the lower margin could be beneficial for late-life health.

The pooled effect across the entire range of the EA PGI shows a small, statistically significant reduction in old-age frailty. When interacting with the EA PGI, however, only compliers in the highest quintile of the EA PGI experience a large reduction in frailty. The improvements come from the middle and upper part of the frailty index, suggesting a delayed buildup of a debilitating number of deficits rather than a longer period of remaining nearly deficit-free. Investigating potential channels, I find that the additional education from the reform improved smoking behavior and physical activity in old age

and reduced the probability of receiving a disability pension. I do not find any impacts on old-age pensions, income or employment history.

This paper contributes to the literature on the effects of the interaction between education and genetic predisposition on long-term outcomes beyond school. While the EA PGI has already been tied to small health differences during working life (Lombardi et al. 2026), I view it as an accelerant of educational policy, showing that this type of skill-building interaction in early life may have positive payoffs for old-age health – though only for those with high endowments. My study also contributes by investigating broad health consequences of this interaction, specifically late in life. The previous null results in Barcellos, Carvalho, and Turley (2018) come from individuals in mid-life and covers metabolic health.

Another contribution is to the literature on the causal effects of compulsory schooling on health. Specifically, is there a causal effect of education behind the usually large descriptive health differences, and if so, for whom and in which contexts? The existing evidence in the UK context, using compulsory schooling reforms as exogenous variation is mixed. Clark and Royer (2013) estimate very small effects of the 1947 and the 1972 UK compulsory schooling reforms on self-rated health. Silles (2009) finds larger effects. Jürges, Kruk, and Reinhold (2013) find no effects, including on biomarkers. More recent evidence from the 1972 reform, based on data from the UK Biobank finds that the additional education was responsible for reductions in body size at the upper end and increases in blood pressure at the lower end of the respective distributions (Barcellos, Carvalho, and Turley 2023), as well as for a reduced risk of diabetes and mortality (Davies et al. 2018). I contribute by studying older individuals. Furthermore, my measure of health, the frailty index, is a more robust and comprehensive measure of health in old age than, e.g., self-reported health, and it tracks the functional aspects of healthy aging particularly well. Finally, I investigate heterogeneity of this effect along a relevant source of differences in (economic) success over the life cycle.

The remainder of this paper is structured as follows. In Section 2, I introduce the institutional setting of the 1947 UK compulsory schooling reform and cover the measure for genetic propensity for educational attainment. Section 3 introduces the data and sample, variables, methods, and assumptions. In Section 4, I present the main results. First, I show descriptive patterns of the interaction between EA PGI and educational attainment. Second, I estimate the interaction causally. Third, I investigate patterns of selection on gains. Next, I present robustness checks and, finally, some suggestive evidence on non-health channels. Section 4.6 concludes.

2 Background and institutional setting

2.1 The 1947 raising of the school leaving age

To identify the causal effect of schooling on late-life health, I exploit the 1947 compulsory schooling reform in the UK.[1] The 1944 Education Act raised the minimum school-leaving age from 14 to 15, with

[1] A second reform raised the leaving age from 15 to 16 in September 1972. I do not use this reform because cohorts affected by it are still too young in the ELSA data to study late-life health outcomes.

implementation set by Ministerial order to April 1, 1947 (Clark and Royer 2013). Because children in the UK typically started school at age five, the reform extended the mandatory length of schooling from nine to ten years. Cohorts born from April 1933 onward, the pivotal cohort, were the first bound by the new rule. I later on exploit this discontinuity by comparing cohorts born before and after the cutoff.

The reform had a strong and broad effect on educational attainment, shifting a large share of the students from nine to ten years of schooling. Figure 1 illustrates this shift using cohort-level aggregates from ELSA. The figure plots the share of individuals still enrolled at ages 15 and 16 by birth cohort, with a vertical line marking the pivotal cohort. The top line (circles) tracks enrollment at age 15. For unaffected cohorts, this share sits around 40%, implying that roughly 60% of students left at the previous compulsory leaving age of 14. Both series trend upward in the cohorts just before the cutoff, but only enrollment at 15 jumps sharply at the pivotal cohort, rising from around 60% to nearly 100%. Enrollment at age 16 (diamonds) shows no such discontinuity. Compliance was high, but not complete. Some individuals report leaving school at age 14 even after the reform took effect. As Clark and Royer (2013) documents, this noncompliance is concentrated among individuals born in the summer months, who turned 15 between the end of one school year and the start of the next, and could therefore leave at age 14.

The reform raised the quantity of schooling without changing its quality. Nearly all compliers attended lower-track schools, which emphasized practical education and were of lower quality than grammar schools in terms of class size and teacher qualifications (Clark 2023). School resources adjusted to accommodate the increased enrollment, so these inputs did not change as a result of the reform (Clark and Royer 2013). The additional year did not raise the probability of obtaining formal academic or vocational qualifications on average (Clark 2023). Instead, schools used the extra year to extend the existing curriculum, introducing some students to more advanced material and helping others consolidate basic skills (Clark and Royer 2013). The effect of the reform can therefore be thought of as an increase in quantity of schooling.

The 1947 reform has been used to identify causal effects of education on wages (Harmon and Walker 1995; Oreopoulos 2006; Devereux and Hart 2010; Clark and Royer 2013), other labor market outcomes (Clark 2023), health (Silles 2009; Powdthavee 2010; Clark and Royer 2013; Jürges, Kruk, and Reinhold 2013), health knowledge (Johnston et al. 2015), mortality (Clark and Royer 2013; Gathmann, Jürges, and Reinhold 2015), and cognitive ability (Banks and Mazzonna 2012; Hollenbach, Schmitz, and Westphal 2026).

2.2 Measuring propensity for education

Frequently used measures of student ability include IQ, scores on academic or vocational tests, and college admission decisions (see, e.g., Brinch and Galloway 2011; Neal and Johnson 1996; Black and Smith 2006; Dale and Krueger 2002). Perhaps more cautiously interpreted as a propensity for education or learning, these measures have in common that they are malleable outputs of complex inputs (such as conditions and investments) early in life (Todd and Wolpin 2003; Cunha and Heckman 2007) and

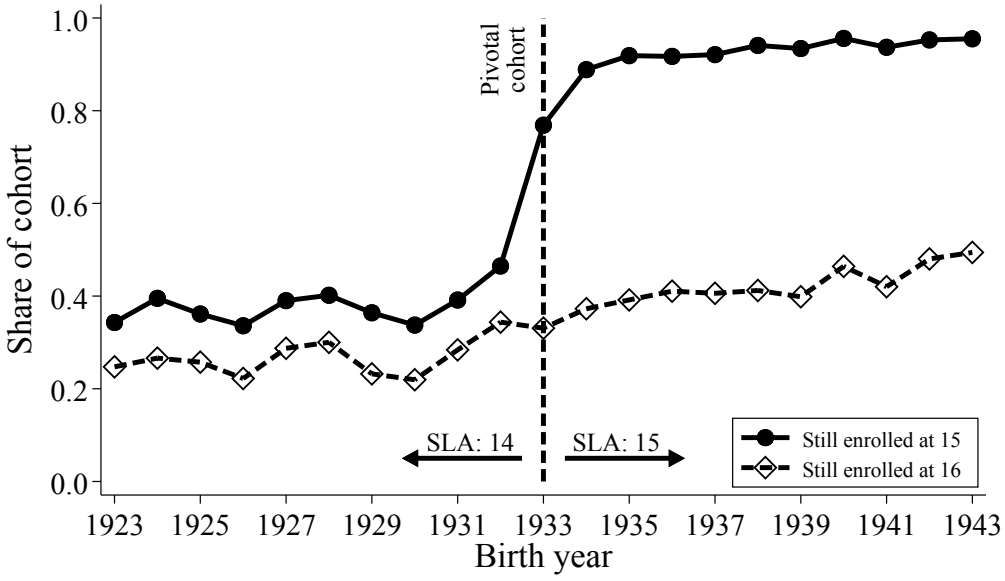


FIGURE 1 – School enrollment at ages 15 and 16 by birth cohort

Note: The 1944 Education Act raised the minimum school leaving age (SLA) from 14 to 15 for cohorts born from April 1933 onward. This figure shows yearly cohort means for enrollment at 15 and 16 since information on month of birth are unavailable. The mean of the pivotal cohort of 1933 therefore includes both individuals affected and unaffected by the schooling reform. The share still enrolled at age 16 is shown for comparison. Sample: 10,709 individuals across waves 1-10 of ELSA.

capture only part of innate ability. I rely on a DNA-based measure of the propensity for educational attainment. While these should also not be interpreted as innate ability, since they are probabilistic rather than deterministic, they have the advantage of being fixed at conception and not modifiable afterwards.

Two basic facts about human genetics motivate how it can be used to measure propensity (or risk) for complex traits like educational attainment:

i. *Most of the genome is shared between people, but small differences matter.* The human genome consists of roughly three billion DNA base pairs, and any two people share more than 99% of their DNA on average (Auton et al. 2015; Nurk et al. 2022). However, remaining variation accounts for a substantial part of differences in observable traits and behavior (Polderman et al. 2015).
ii. *There is no single gene for a complex trait.* Genetic influence on traits and behavior is mostly polygenic, meaning it is spread across many locations in the genome that each contribute a small effect (Chabris et al. 2015; Visscher et al. 2017).

Educational attainment fits this pattern, with thousands of genetic variants individually associated with years of schooling, each contributing only a small share (Lee et al. 2018; Okbay et al. 2022).

The standard tool for summarizing small additive contributions of many genetic variants into one measure is the polygenic index (PGI). Most genetic differences between individuals come down to single positions in the genome where the DNA base varies across the population. These are called

single-nucleotide polymorphisms, or SNPs. A PGI summarizes an individual's genetic predisposition for a given outcome by summing their SNPs, each weighted by its estimated association with that outcome. To quantify SNPs, one variant present in the population is chosen as the reference variant. Any person i can inherit up to two copies of each SNP j, one from each parent and therefore carries either zero, one, or two copies of the reference variant $r_{ij} \in \{0, 1, 2\}$. Weights w_j, i.e. the association of each SNP with the outcome, are obtained in large discovery studies, called genomic-wide association studies (GWAS). They estimate how strongly each SNP is correlated with the trait of interest by regressing the trait on each SNP, one at a time, in a large separate sample to avoid overfitting.[2] The PGI is then simply the weighted sum of counts of the reference variant across all SNPs:

$$PGI_i = \sum_j w_j r_{ij} \tag{1}$$

The PGI is a single individual-level index of additive genetic probability, in this paper for educational attainment. It is approximately normally distributed, with individuals at the upper end carrying more of the alleles that, on a population level, are positively associated with schooling and those at the lower end carrying fewer. The index is not deterministic. Individuals with a high educational attainment PGI can have low attainment, and vice versa. Nevertheless, at the individual level, there is a strong positive correlation with realized educational attainment.

3 Data and methods

3.1 Data and variables

I use data from the the English Longitudinal Study of Ageing (ELSA), a nationally representative longitudinal survey of adults aged 50 and older living in England. ELSA has been conducted biennially since 2002 and includes detailed information on physical and mental health, socioeconomic status, and, over time, has integrated biological and genetic data.

Health in old age. My outcome variable is the frailty index, the share of potential age-related health deficits that an individual exhibits in a given survey wave. Deficits are symptoms, signs, diagnosed diseases, disabilities, and functional impairments. The index therefore summarizes health across many dimensions in a single number bounded between zero and one. It is grounded in the view of aging as an accumulation of deficits and a gradual loss of the body's redundancies (Mitnitski, Mogilner, and Rockwood 2001; Dalgaard and Strulik 2014). A higher value indicates a more advanced state of aging rather than the presence of one particular condition. A key characteristic of the frailty index is its ability to capture health status and its variation with high granularity. Compared with binary or ordinal health measures common in applied work, it is continuous and finely graded, and it relies on many concrete items instead of a single overall self-assessment (Hosseini, Kopecky, and Zhao 2022).

[2] The educational attainment PGI that I use in this paper is based on the discovery study by Lee et al. (2018), covering roughly 1.1 million individuals. This PGI can explain about 12% of the variation in education between individuals.

The frailty index is an established measure of health. First introduced in gerontology (Mitnitski, Mogilner, and Rockwood 2001) and standardized by Searle et al. (2008), it predicts mortality, institutionalization, and further adverse health events, and it predicts these outcomes more accurately than self-reported health status (Rockwood and Mitnitski 2007; Hosseini, Kopecky, and Zhao 2022). Results also do not depend on the exact items chosen. As long as enough deficits from different areas of health enter, indices built from different item lists behave similarly (Searle et al. 2008; Hosseini, Kopecky, and Zhao 2022). Economists have taken up the measure both as the health state in models of aging (Dalgaard and Strulik 2014; Strulik 2018b) and as the health outcome in empirical work, in studies of aging patterns across Europe (Abeliansky and Strulik 2018), of retirement and occupational health gaps (Abeliansky and Strulik 2023), and of the role of health in life-cycle earnings inequality (Hosseini, Kopecky, and Zhao 2022, 2025).

I follow the standard procedure of Searle et al. (2008), implemented in ELSA, for example by Marshall et al. (2015) and Niederstrasser, Rogers, and Bandelow (2019) and use a broad range of 56 deficits available in ELSA. They include diagnoses of serious illnesses, difficulties with activities of daily living and mobility, sensory impairment, self-rated health, depressive symptoms, and results from cognitive tests. Every item takes a value between zero and one, where a binary item is either absent or present and an ordinal item can take intermediate values, and the index is the average across all 56 items. An individual reporting 14 of the 56 deficits, for instance, has a frailty index of 0.25. I list all components as well as how they enter the index in Appendix Table A.1.

Education. ELSA does not record years of schooling directly. Instead, it reports the age at which a person left full-time education. Most individuals in the final sample left school relatively early: 21 % at age 14 or below (the compulsory schooling age before the reform), and 35 % at age 15 (the compulsory schooling age afterwards). My measure of education is an indicator variable equal to 0 if an individual left school at age 14 or below and 1 if they stayed until at least age 15 (i.e., left at 15 or older). This is the margin targeted by the compulsory schooling reform.

Schooling reform. The compulsory schooling reform took effect in April 1947 and applied to everyone born on or after 1 April 1933. Accordingly, I define the instrument that exogenously shifts educational attainment as an indicator variable equal to 0 if an individual was born before 1933 and 1 if they were born in 1933 or later. ELSA does not provide birth months, so for individuals born in 1933 it is unclear whether they were subject to the new school leaving age. Because almost all individuals are interviewed multiple times and in different months, I can combine birth year, interview month, and age at interview to infer for part of the 1933 cohort whether they were born between January and March (pre-reform) or between April and December (reform). I drop the remaining individuals born in 1933, whom I cannot definitively assign to either side of the reform.

Genetic propensity for educational attainment. As introduced in Section 2.2, I use an educational attainment PGI to measure the genetic propensity for educational attainment. The PGI is readily available in ELSA and the weights used to calculate it are based on the discovery study by Lee et al.

(2018). Their score explains 11–13% of the variation in educational attainment in the discovery sample and about 10% of the variation in cognition. To investigate interactions of schooling and the PGI at different parts of its distribution, I divide the score into quintiles.

3.2 Sample and summary statistics

The analysis sample includes all ELSA respondents with genetic data who were born within ten years of the pivotal cohort of the compulsory schooling reform (1933), observed at ages 65–85 in waves 1–10. From the 1933 cohort itself, I drop individuals whom I cannot definitively assign to either side of the April 1 cutoff based on birth year, interview month and reported age at interview. These restrictions leave 18,705 observations from 3,415 individuals. Section 4.4 reports robustness checks that vary these sample restrictions.

Table 1 shows summary statistics for my sample, both in its entirety and separately by treatment and control group. The average frailty index is 0.16, with a substantial standard deviation. A majority of the sample consists of treated observations and individuals born after the reform cutoff. Treatment and control group differ significantly along several dimensions. Some of these differences are mechanical, most notably birth year and being born after the cutoff. Others concern individual characteristics: the treated group has a slightly higher share of women, is slightly younger, and is more concentrated in the upper part of the EA PGI distribution. The age gap is smaller than the gap in birth years, as repeated measurements at different ages seem to offset the latter to some degree. Crucially, individuals in the treated group are significantly less frail.

3.3 Method

To estimate the causal effect of schooling on old-age frailty, I implement the following fuzzy regression discontinuity design via two-stage least squares:

$$E_i \times \mathbb{1}[G_i = g] = \alpha_{g,0}\, \mathbb{1}[G_i = g] + \alpha_{g,1}\, Z_i \times \mathbb{1}[G_i = g] \qquad + X'\gamma + c_i + Z_i \times c_i + \varepsilon_{it} \tag{2}$$

$$Y_{it} = \sum_{g=1}^{5} \left[\beta_{g,0}\, \mathbb{1}[G_i = g] + \beta_{g,1}\, E_i \times \widehat{\mathbb{1}[G_i = g]} \right] + X'\delta + c_i + Z_i \times c_i + u_{it} \tag{3}$$

Equation 2 is the first stage. The schooling measure E_i indicates whether individual i stayed in school until at least age 15. G_i enters as five indicators for the quintile of the EA PGI: $g \in \{1, \ldots, 5\}$. The five interactions $Z_i \times \mathbb{1}[G_i = g]$, where Z_i indicates birth in 1933 or later, serve as instruments. They instrument the schooling variable separately within each quintile group, so there are five first stages. The coefficient $\alpha_{g,1}$ is the jump at the 1933 cutoff in the share of individuals in quintile g who stayed in school until at least age 15, and hence the complier share in that quintile.

Equation 3 is the second stage. The outcome Y_{it} is the frailty index of individual i in wave t. It is regressed on the five predicted interactions between schooling and quintile. Each $\beta_{g,1}$ is therefore the

TABLE 1 – Summary statistics

	Full sample		By E_i (Left school at 15 or later)		
	Mean	SD	0	1	Diff. (p-val.)
Outcome					
Frailty index	0.16	0.11	0.19	0.15	−0.04 (0.00)***
Treatment					
Left school at 15 or later	0.79		0.00	1.00	1.00
Instrument					
Born 1933 or later	0.68		0.14	0.83	0.69 (0.00)***
Educational attainment PGI					
1st quintile	0.19		0.23	0.17	−0.06 (0.00)***
2nd quintile	0.20		0.21	0.19	−0.02 (0.01)***
3rd quintile	0.20		0.19	0.20	0.01 (0.05)**
4th quintile	0.21		0.21	0.21	0.00 (0.58)
5th quintile	0.21		0.16	0.22	0.06 (0.00)***
Selected controls					
Female	0.54		0.53	0.55	0.02 (0.01)***
Age	74.24	5.34	77.05	73.49	−3.56 (0.00)***
Birth year	1,935.03	5.24	1,929.80	1,936.42	6.62 (0.00)***
Observations	18,705		3,927	14,778	

Note: Means and standard deviations (continuous variables) for the full estimation sample, as well as means for the treatment ($E_i = 1$) and control ($E_i = 0$) groups. The last column reports the raw difference in means and p-values of a t-test for equality of means. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ indicate significance levels. The full set of summary statistics, additionally including parental education, survey wave, and principal components of the genetic data, is reported in Appendix Table A.2.

full local average treatment effect (LATE) of schooling on frailty for compliers in quintile g. Differences across the $\beta_{g,1}$'s capture the gene-environment interaction, i.e., how the effect of schooling varies across PGI quintiles.

Both stages include the running variable c_i = birth year − 1933, a linear cohort trend centered at the 1933 cutoff, and its interaction with Z_i. The interaction allows the slope to differ on either side of the cutoff. X includes the remaining controls: indicators for gender, age, survey wave, and mother's and father's school-leaving age. I also include the first ten principal components of the genetic data, which capture broad genetic similarity among individuals in ELSA and therefore ensure that PGI quintiles reflect genetic propensity for education rather than shared ancestral origin. I discuss this in more detail in Section 3.4. ε_{it} and u_{it} are idiosyncratic error terms. Standard errors for all estimates are clustered at the individual level.

3.4 Assumptions and identification

Assumptions. Identification rests on three main assumptions. The first is that the 1947 reform is a valid instrument for education. Instrument relevance is not in doubt: the higher leaving age induced a large share of students to increase their attainment (see Figure 1), and my first stage is correspondingly strong (see Section 4.3). Instrument exogeneity is also likely to hold, given that eligibility for the compulsory schooling reform depends on date of birth and that the 1944 Education Act came years after these cohorts were born.

Second, I assume that the exclusion restriction holds. This requires that nothing but compulsory schooling shifts at the threshold. The leading concern is exposure to the Second World War. Both Clark and Royer (2013) and Banks and Mazzonna (2012), who also use this reform, conclude that wartime disruptions did not vary systematically across the cutoff. The cohort trends I include absorb gradual variation in such exposure, and my estimates hold for narrower bandwidths (see Section 4.4). The exclusion restriction could also be violated if the reform affected always-takers, for example through spillovers if it changed not only instruction time but also school quality. However, the compliers in this particular reform were almost exclusively students in lower-track schools that ended at the minimum leaving age (Clark 2023). This arguably left little scope for spillovers onto those who would have stayed on regardless. Furthermore, as enrolment expanded, so did school resources, leaving class sizes and teacher qualifications largely unaffected (Clark and Royer 2013). The additional year of schooling did not yield formal qualifications, which would have required remaining in school until at least age 17 (Banks and Mazzonna 2012; Clark 2023). The reform lengthened schooling without changing its character.

The third assumption, monotonicity, rules out defiers, that is, individuals who left school earlier only because they were subject to an additional compulsory year. Defiance of this kind seems unlikely.

Attrition. In longitudinal studies of older individuals, such as ELSA, nonrandom attrition may be problematic if it is related to the treatment. If differential attrition is present, those who remain in the sample skew the treatment (or control) group, potentially biasing the estimated treatment effect. For this study, a While the debate on a causal effect of compulsory schooling on mortality is not settled and depends, among other things, on institutional setting (Galama, Lleras-Muney, and van Kippersluis 2018), Clark and Royer (2013) find no effect on mortality from the 1947 compulsory schooling reform. Even if differential survival were present, the direction of the resulting bias is predictable in my application. Since frailty increases strongly with age, longer survival in my treated group would mean that more frail individuals stay alive, shifting the treatment group towards higher frailty. Given that I find a reduction in frailty due to education, this means I would likely underestimate the true effect. I nevertheless test for panel attrition and find no evidence of it, and I estimate my main specification, reweighted to account for attrition. See Section 4.4 for details.

Causality of genes. The association between an individual's EA PGI and their schooling does not reflect a clean causal effect of genetic endowment. Part of the association arises through what is

known as genetic nurture. Parents pass on both genes and a home environment, and those whose own genes are associated with more schooling also tend to provide a home environment that supports learning. The child's PGI is therefore correlated with their schooling partly because it proxies for family environment, not only because of the child's own DNA (Kong et al. 2018). However, conditional on the parents' genes, the child's genes are exogenous. Therefore, controlling for parental genotype yields a causal effect. A second strategy is a within-family design that compares siblings. Because they share the same parents, family fixed effects absorb parental genotype. Few datasets to date contain parent-child trios or enough siblings for a well-powered analysis, and in ELSA only between-family designs are feasible. In that case, Houmark, Ronda, and Rosholm (2024) show that controlling for parental education can absorb almost all of the family environment component.

Since I am only interested in the differential effect of education on frailty by EA PGI, it would be sufficient to use the PGI as a pre-determined (and therefore pre-treatment) interaction variable. Nevertheless, I include parental education as an additional control, in line with Barth, Papageorge, and Thom (2020), Papageorge and Thom (2020), and Barth et al. (2022), among others. In ELSA, parental education is provided as the age when each parent left full-time education, truncated at the low (14) and high (19) ends. Unfortunately, this information is not available for 17.5 % (mothers) and 18.1 % (fathers) of the observations in my final sample. Therefore, I include an additional category for missing parental school leaving age instead of dropping these observations.

A second threat is population stratification. Because of historical migration and non-random mating, subpopulations differ systematically in the frequency of genetic variants. If they also differ in schooling or health for cultural or environmental reasons, the discovery study underlying the PGI assigns weight to variants that signal ancestry rather than the trait itself. The PGI then partly captures the environments associated with ancestry, such as region, language, or local socioeconomic conditions, in a similar way as it partly captures the family environment. I therefore follow the convention in this literature (see e.g. Barth, Papageorge, and Thom 2020; Papageorge and Thom 2020; Barth et al. 2022) and additionally control for the first ten principal components of the genetic data. These are summary indices of general genetic similarity between individuals and account for population stratification (Price et al. 2006). ELSA provides them as pre-computed variables.

4 Results

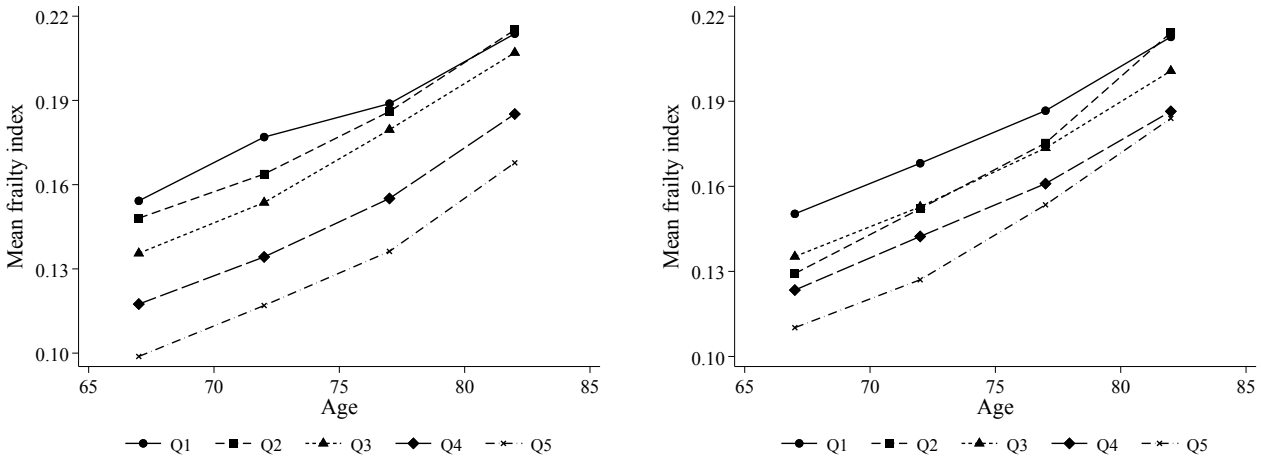
4.1 Descriptive patterns

Houmark et al. (2025) show that academic achievement gaps produced by the EA PGI have the same size and trajectory across time in school as socioeconomic achievement gaps. I compute old-age health gaps by socioeconomic status (SES) and EA PGI in Figure 2. It shows means of the frailty index for 5-year bins from 65 to 84 separately for quintiles of an SES index (Panel A) and quintiles of the EA PGI (Panel B).[3] There are clear health differences in both SES and EA PGI, with those in the highest quintile

[3] I construct the SES index using level of education, total household wealth and pension income. Household wealth is broken down by dividing by the square root of household size. For each of the three components, I rank respondents within

being healthiest (lowest frailty) and those in the lowest quintile being frailest. These differences exist initially for 65-to-69-year-olds and increase almost in parallel over time.

The most striking takeaway is the comparison between the two patterns. The total difference in frailty according to the EA PGI is roughly the size of the frailty difference between the first and the fourth SES quintile. The EA PGI seems to be not only an important predictor of early-life skill gaps, but also of late-life health gaps of a considerable size.



A. Socioeconomic gap in old-age frailty

B. EA PGI gap in old-age frailty

FIGURE 2 – Socioeconomic and genetic frailty gaps

Note: Progression of the average frailty index from age 65 to 84 in 5-year bins and separately by quintile of a socioeconomic score (Panel A) that includes education, household wealth and pension income and quintile of the EA PGI (Panel B). Averages are plotted at the midpoint of each 5-year bin. Sample: 26,850 observations from 6,083 individuals

Next, I add educational attainment (school leaving age) as an additional dimension to visualize broad health differences across both education and EA PGI (Figure 3). The figure shows predicted frailty for combinations of school leaving age and EA PGI, controlling for age, gender, and parental education. Differences in predicted frailty by educational attainment are pronounced. Higher attainment is generally associated with better health in later life (that is, lower frailty), and lower attainment with higher frailty. However, these differences are mainly present at low and medium values of the genetic predisposition to educational attainment. At the upper end of the EA PGI, there are barely any differences in predicted frailty. In addition, the group with the highest educational attainment has low predicted frailty regardless of its EA PGI.

This already hints at potential interactions between educational attainment and genetic predisposition in old-age frailty. To illustrate these descriptive patterns more explicitly, I use the same predictions as before but now construct frailty-PGI gradients from them, separately by school leaving age (Figure 4). The predictions again come from an OLS regression of frailty on the interaction of school leaving

each wave and convert the ranks to percentiles. Ranking within wave ensures that the index is unaffected by inflation or by respondents being observed in different waves. I then average each respondent's percentile rank over all waves and take the average of the three percentile ranks to get the SES index. Each component receives equal weight. By averaging across all waves, there is one observation per respondent, so that the cut-offs do not depend on how many times each person was observed. Finally, I assign each respondent to a quintile of the index.

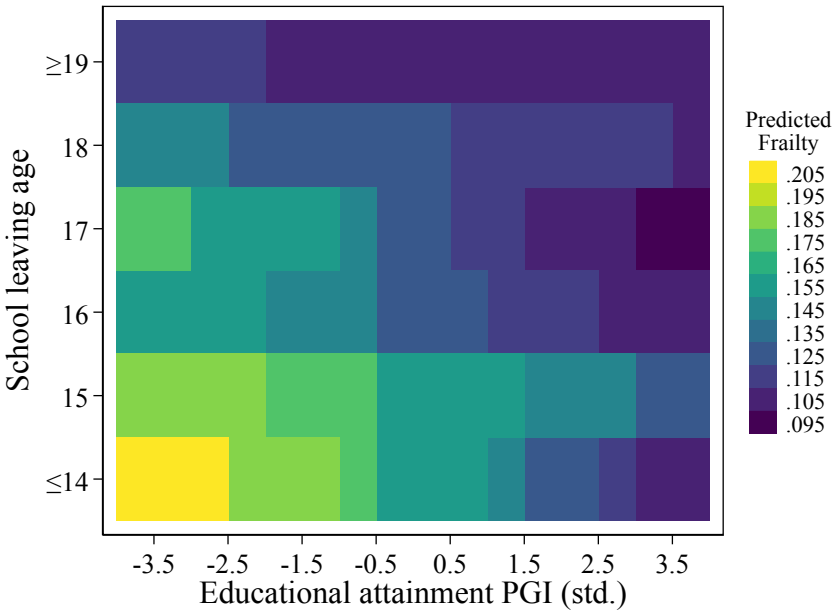


FIGURE 3 – The association of frailty with school leaving age by EA PGI

Note: Predicted frailty index for combinations of school leaving age and EA PGI, from a regression of frailty on the interaction of school leaving age and EA PGI, sex, age, genetic principal components and parental years of education. Predictions are computed at each school leaving age × EA PGI (discretized to 0.5 SD cells) combination and averaged over the sample distribution of the remaining controls. Sample: 45,538 observations from 7,033 individuals across waves 1-10 of ELSA.

age and the EA PGI, as well as on controls. Because the PGI enters the regression linearly but is fully interacted with school leaving age indicators, the predictions produce straight but not necessarily parallel lines. Each school leaving age has its own intercept and its own slope in the PGI. Any level or slope differences across school leaving ages therefore reflect the data rather than functional form restrictions. Within each school leaving age, the slope captures the total association between the PGI and frailty. Level differences across lines reflect the association between educational attainment and frailty at a given PGI, and differences in slopes across lines indicate $G \times E$ interactions. Horizontal lines would imply no PGI gradient in frailty, overlapping lines no education gradient, and parallel lines no interaction between the two.

Figure 4 shows variation along all three dimensions and supports three descriptive observations. First, predicted frailty decreases in the EA PGI within every school leaving age bin, and the slope is steeper for those who left school earlier. A relatively flat line, as for those leaving at 18 and at 19 or later, indicates almost no difference in predicted frailty between the lower and upper parts of the PGI distribution. Second, individuals who left school early have higher predicted frailty across the lower range of the PGI distribution, while differences across school leaving ages shrink at the upper end (as in Figure 3). Third, the resulting differences in slopes point to substantial $G \times E$ interactions: the PGI gradient in frailty varies by school leaving age, and the education gradient varies across the PGI distribution.

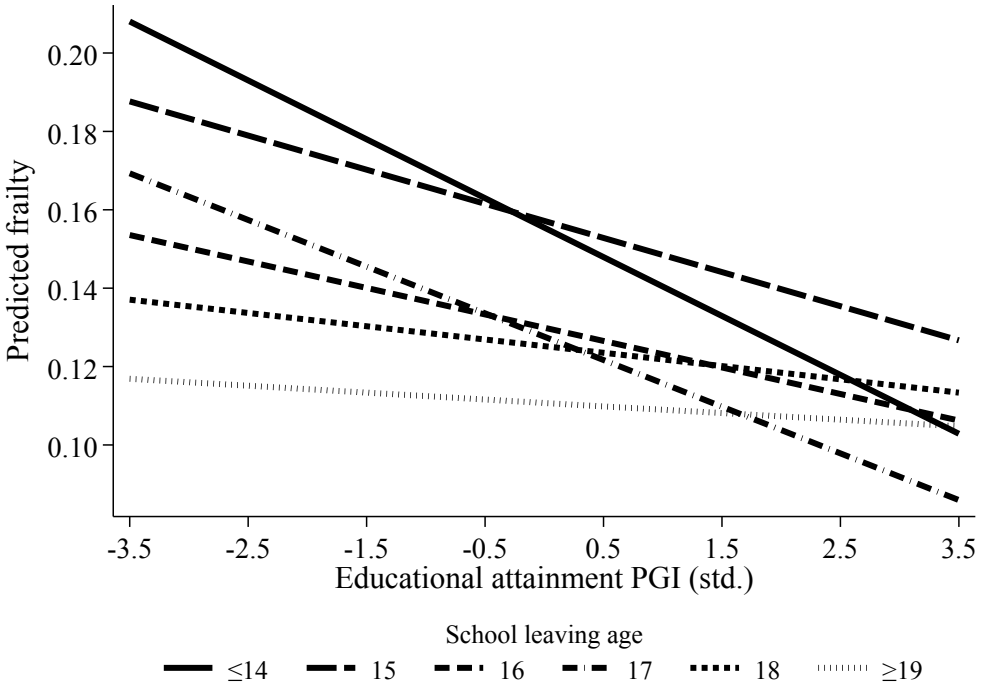


FIGURE 4 – Frailty-PGI gradient by school leaving age

Note: Predicted frailty index as a function of the standardized EA PGI, separately by school leaving age. Predictions come from a regression of frailty on the interaction of school leaving age and EA PGI, sex, age, genetic principal components and parental years of education. EA PGI enters the model linearly, so each line is straight and its slope is the frailty-PGI gradient for the corresponding school leaving age. Level differences across lines reflect the association between education and frailty at a given PGI level. Sample: 45,538 observations from 7,033 individuals across waves 1-10 of ELSA.

4.2 The causal effect of compulsory schooling on frailty by genetic propensity for education

It is unclear whether these descriptive patterns reflect a causal effect of education that varies with the EA PGI. Generally healthier individuals may instead acquire more education, and both are related to the EA PGI: a higher PGI is robustly associated with more schooling, and given the possibility of genetic nurture I discuss above in Section 3.4, parents with a higher EA PGI plausibly provide a healthier early-life environment (leading to better health later in life) while passing on their genetic variants. Therefore, in my main results in Table 2, I report both OLS and IV regressions of frailty on the gene-education interactions following Eq. 3. These analyses test both whether education has a causal effect on old-age frailty, identified at the low end of the educational distribution, and whether compulsory schooling policies matter for old-age frailty.

OLS regressions show a statistically significant negative correlation between staying in school until at least age 15 (my schooling treatment) and old-age frailty, both in the pooled regression (Column 1) and in each PGI quintile (Column 2). Across all quintiles of the EA PGI, staying in school for the additional year is associated with a frailty index that is about 0.02 to 0.04 points lower. All estimates are relatively small, but they consistently have a negative sign. The pooled IV estimate in Column (3)

TABLE 2 – Main results

DV: Frailty index	(1) OLS	(2) OLS	(3) IV	(4) IV	(5) IV
E_i	−0.028***		−0.029**		−0.010
	(0.005)		(0.014)		(0.016)
$E_i \times (G_i = 1)$		−0.021**		−0.010	Ref.
		(0.010)		(0.016)	
$E_i \times (G_i = 2)$		−0.022**		−0.029	−0.019
		(0.009)		(0.018)	(0.018)
$E_i \times (G_i = 3)$		−0.038***		−0.026	−0.016
		(0.010)		(0.020)	(0.020)
$E_i \times (G_i = 4)$		−0.020**		−0.026	−0.016
		(0.008)		(0.019)	(0.019)
$E_i \times (G_i = 5)$		−0.033***		−0.059**	−0.048**
		(0.010)		(0.026)	(0.024)

Note: The dependent variable is the frailty index. E_i indicates staying in school to at least age 15 and G_i quintiles of the educational attainment PGI. Columns (1) and (2) report OLS estimates, (3) to (5) IV estimates instrumenting E_i with exposure to the 1947 compulsory schooling reform (born 1933 or later). Columns (2) and (4) report the full effect of E_i within each PGI quintile as given by Equation 3. Column (5) re-parameterizes column (4) with an intercept and $G_i = 1$ as the reference category, so that the coefficient on E_i is the effect in the lowest quintile and the interactions are differences relative to it. All specifications control for gender, survey wave, age, and mother's and father's age when leaving full-time education. IV specifications add the running variable (re-centered birth year) and its interaction with the instrument. Interacted specifications also include the first ten genetic principal components. Estimation sample as described in Section 3.1 with 18,705 observations from 3,415 individuals. Standard errors clustered at the individual level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

is the same size as the pooled OLS estimate and is also statistically significant. The two, however, refer to different estimands: the OLS estimate to an average treatment effect and the IV estimate to a local average treatment effect for the compliers.

For the IV regression in Column (4), only those in the highest PGI quintile have a statistically significant local average treatment effect. For them, the additional year of education induced by the reform reduced frailty by 0.059 on average, about half the standard deviation of the frailty index. While the estimate for the lowest quintile is very small, the coefficients for quintiles 2 to 4 may still be relevant in size, but they are at most about half the size of the estimate for the highest quintile. Overall, this implies increasing differences in the effect of schooling on old-age frailty across the EA PGI and therefore a gene-environment interaction. In Column (5), I estimate the model of Column (4) but take the lowest quintile as the reference group. There, the estimate for E_i is the effect in the lowest quintile and the interaction terms are markups to this effect, so that the total coefficient is the sum of the coefficient on E_i and the respective interaction. This shifts the hypothesis test from whether there is a statistically significant effect of schooling on frailty within each PGI quintile to whether there are statistically significant differences between quintiles. Again, only the highest quintile shows a statistically significant difference compared to the lowest.

Appendix Table A.3 shows reduced-form and first-stage estimates, where frailty and the education treatment, respectively, are regressed directly on interactions between the instrument (being born in 1933 or later) and the PGI quintiles. These are the two components that define the LATE. Since

the LATE is the ratio of the two, a large effect could in principle arise from a small reduced-form effect divided by a small first stage. There are large first-stage differences between the PGI quintiles, which I discuss in more detail in the following section. The reduced-form regressions show that the main results in Column (4) of Table 2 are not merely driven by smaller complier shares, which would mechanically entail a larger LATE. The reduced form shows the same pattern observed in the main results: Only those in the highest PGI quintile have a statistically significant intention-to-treat effect. Overall, there is a strong negative correlation between education and old-age frailty across the entire EA PGI distribution. However, the causal effect of the reform itself, and of the additional education, is concentrated among those with the highest EA PGI.

An effect on the frailty index does not reveal where in its distribution it occurs. Knowing at which part of the distribution of the frailty index education improves it has implications for how to interpret the effect since the frailest individuals in particular have lower quality of life, higher care needs and lower chances of survival (Song, Mitnitski, and Rockwood 2010). A decrease in the probability of being frail due to more education would mean avoiding the worst consequences of frailty, whereas an increase in the probability of being regarded as robust would mean more individuals thriving in old age.

I therefore split the frailty index into the three commonly used categories: robust (frailty ≤ 0.08), pre-frail ($0.08 <$ frailty < 0.25), and frail (frailty ≥ 0.25), and use them as outcomes in my main IV model (see, e.g. Song, Mitnitski, and Rockwood 2010; Davies et al. 2021). Table 3 reports the results of IV regressions of each category on instrumented education. I do not investigate the effect within each PGI quintile as this may be demanding on the data, that is, some strata might have only a small number of observations. Nevertheless, the takeaways from the pooled regressions are interesting. There is no effect of an additional year of education on being regarded as robust. However, I see an increase in the probability of being pre-frail, the middle category. The estimate is only statistically significant at the 10 percent level, but has an economically meaningful magnitude of almost 9 percentage points. Furthermore, there is a strong, 10-percentage-point decrease in the likelihood of being frail. This indicates that the additional year of schooling has led to improvements at the upper end of the frailty index, consistent with individuals avoiding (or delaying) the buildup of a debilitating amount of deficits rather than staying nearly deficit-free in old age.

The analyses closest to mine come from Barcellos, Carvalho, and Turley (2018) and Hollenbach, Schmitz, and Westphal (2026). The former estimate the effect of an additional compulsory year of education from the 1972 UK schooling reform on body size, lung function, and blood pressure in middle age, and whether these effects differ by genetic predisposition for BMI and educational attainment. While they find that the effect of education on these midlife health outcomes depends on the BMI PGI, with the largest gains concentrated among those at highest genetic risk for overweight, they find no differences by the EA PGI. Hollenbach, Schmitz, and Westphal (2026) show that the additional year of education from the 1947 reform substantially increased old-age cognitive ability, with gains increasing in the EA PGI and largest for those in the highest quintile.

In terms of magnitude, my pooled estimate corresponds in size to the effect of retirement on the frailty index that Abeliansky and Strulik (2023) find for women. My estimate for the fifth EA PGI quintile

TABLE 3 – IV results for frailty categories

Dependent variable	Robust	Pre-frail	Frail
E_i	0.013	0.087*	−0.101**
	(0.046)	(0.051)	(0.047)
Observations	18,705	18,705	18,705
Individuals	3,415	3,415	3,415

Note: Dependent variables are indicators for being robust (frailty index ≤ 0.10), pre-frail (0.10 < frailty index ≤ 0.21), and frail (frailty index > 0.21). E_i is staying in school to at least age 15. IV estimates instrumenting E_i with exposure to the compulsory schooling reform (born 1933 or later). All specifications control for gender, survey wave, age, the running variable (re-centered birth year) and its interaction with the instrument. Estimation sample as described in Section 3.1. Standard errors clustered at the individual level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

exceeds their retirement effect for men with low education, the group with the largest physical health benefit from retiring. Huang et al. (2024) construct a Wald estimate of education on frailty from genetic discovery studies. For a set of genetic variants that predict schooling, they divide the effect of each variant on frailty (the "reduced form") by its effect on schooling (the "first stage") and combine the resulting ratios. Their results suggest that a one SD increase in schooling corresponds to a drop of roughly half an SD in frailty.

4.3 First-stage dynamics and selection on gains

My analysis aims to identify how the effect of additional schooling on old-age frailty differs by EA PGI, i.e. the gene-environment interaction effect between schooling and endowments for schooling. However, standard IV and reduced-form estimations may not always yield a true interaction effect, because they cannot cleanly separate the interaction effect from latent shifts in complier groups along the interaction variable (Hollenbach, Schmitz, and Westphal 2026). If this is the case, the computation of interaction effects involves comparing potential outcomes from different complier groups, instead of well-defined (i.e. fixed) ones that would yield the true LATE. The estimate can then numerically differ from the true interaction effect if there is also treatment effect heterogeneity between these complier groups. Hollenbach, Schmitz, and Westphal (2026) formalize two conditions that must be met for 2SLS or reduced-form estimates to differ in size from the true interaction effects: First, compliers, who identify the local average treatment effect of the interaction, must have different unobserved characteristics across strata of the interaction variable. In my setting, this would entail that compliance behavior with the schooling reform depends on the EA PGI. Second, there must be essential heterogeneity in the treatment effects, meaning the propensity for treatment is correlated with unobserved effect heterogeneity (see, e.g., Heckman, Urzua, and Vytlacil 2006). In other words, there is selection into treatment based on expected gains (or losses).

Both conditions are likely in my setting. First, that the reaction to a schooling reform differs by genetic endowment for educational attainment has been shown, e.g., in Barcellos, Carvalho, and Turley (2018) for the 1972 UK compulsory schooling reform and in Hollenbach, Schmitz, and Westphal

(2026) for the 1947 reform I investigate. In the latter, those with high EA PGI are more likely to be always-takers (choose more education regardless of the reform) and those with low EA PGI are more likely to be compliers (choose more education because of the reform). Second, selection on gains has been established as a frequent feature of educational choices (Carneiro, Heckman, and Vytlacil 2011; Kamhöfer, Schmitz, and Westphal 2019; Westphal, Kamhöfer, and Schmitz 2022; Hollenbach, Schmitz, and Westphal 2026). Nevertheless, as Kowalski (2023) points out, even within the same (natural) experiment, patterns of selection into treatment can differ substantially depending on the outcome. In fact, Barcellos, Carvalho, and Turley (2021) do not find evidence of selection on gains in wage returns to the 1972 UK schooling reform.

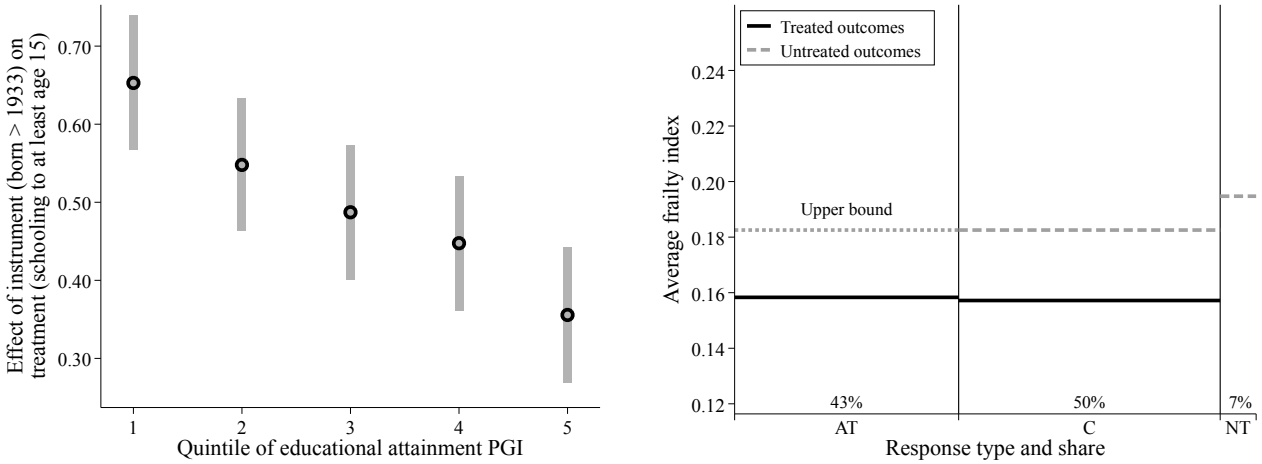
Figure 5A shows the differences in the first-stage estimates by quintiles of the EA PGI. These correspond to the share of compliers in each quintile. There are sizable differences and a monotonic decline in the share of compliers along the EA PGI, from about 65 % in the lowest PGI quintile down to 35 % in the highest. The latter is still a sizable share, showing that the 1947 compulsory schooling reform had a broad impact. The differences by EA PGI are expected: children with low EA PGI are more likely to want to leave school early and are therefore more likely to be prompted by the reform to stay for another year, while the reverse holds for children with high EA PGI.

To investigate potential patterns of selection related to treatment propensity, I start by calculating means of the frailty index for different groups that are informative about these patterns. Always-takers have the highest propensity: they take more education despite not being affected by the schooling reform. Never-takers have the least: they do not take more education, despite being affected by the reform. Compliers are in between. I plot the means alongside the respective group sizes in Panel 5B.[4] The three groups are ordered on the horizontal axis according to their propensity to stay in school until at least age 15 (the treatment). Of my sample, 43% are always-takers, 50% are compliers, and 7% are never-takers. Treated potential outcomes are shown in black color with solid lines, untreated in grey with dashed lines. The treated outcome for always-takers and the untreated outcome for never-takers can be calculated directly from the data since we can identify them in the sample from their behavior. However, the counterfactual outcomes for these groups can never be observed since one group always takes the treatment and the other never takes it. For compliers, both treated and untreated potential outcomes can be calculated, and I use the method from Imbens and Rubin (1997) to do so.[5] The difference between them is the local average treatment effect of one more year of schooling due to the reform on old-age frailty, without PGI interactions.

Comparing treated potential outcomes of always-takers and compliers involves comparing the average frailty of those most eager to take on more education and do so regardless of the reform to the average of those who were prompted by the reform to take an additional year of education. This comparison comprises both differences in selection and treatment effect heterogeneity (Kowalski 2023). Panel 5B

[4] I do this without interacting by PGI quintile since selection on gains, which I ultimately test for, is a pattern that affects, and can be shown in, the entire population.

[5] Appendix Section B describes in detail how all numbers in this section are calculated.



A. First-stage differences by PGI quintile

B. No differences in treated and untreated potential outcomes

FIGURE 5 – Evidence on differences in the first-stage and second-stage potential outcomes

Note: Panel A: Strength of the first stage (i.e., complier shares) by PGI quintile with 95% confidence intervals. Appendix Table A.3 reports the exact estimates, standard errors, and first-stage F-statistics. Panel B: Treated potential outcome means for always-takers (AT) and compliers (C), and untreated means for compliers and never-takers (NT), without gene interactions. Black solid lines denotes treated potential outcomes, grey dashed lines untreated. The horizontal axis shows each group and its share of the sample, with the width of each group corresponding to that share. Groups are ordered from most to least willing to take the treatment (more education). Additionally assuming weak monotonicity of untreated potential outcomes yields an upper bound on the local average treatment effect for always-takers. Appendix Section B describes in detail how all numbers are calculated. Both analyses are calculated on the main sample described in Section 3.1.

shows no difference in these treated outcomes. But since we cannot separate the two channels, this could indicate either no selection and no heterogeneity, or that both work in opposite directions.

Therefore, I use two additional strategies following Kowalski (2023). First, I consider differences in untreated potential outcomes between compliers and never-takers. By definition, the untreated potential outcomes are unaffected by the treatment, and therefore by heterogeneous treatment effects, but are informative about selection on levels in the absence of treatment. Panel 5B shows possible small differences in untreated frailty of compliers and never-takers. However, never-takers are a potentially special group of individuals who were able to leave school early despite being in the reform cohorts, as we establish in Hollenbach, Schmitz, and Westphal (2026).

Second, I go one step further and make an additional assumption in the spirit of Kowalski (2023). I assume weak monotonicity in untreated outcomes related to willingness to take the treatment. This means that, in the absence of the treatment, the frailty of always-takers is at least as low as that of compliers, whose frailty is at least as low as that of never-takers. This assumption should be unproblematic, especially when compared to assuming the same for treated outcomes, where effect heterogeneity would complicate the ordering. At worst, it implies that – absent the treatment – there should be no differences in frailty between always-takers, compliers and never-takers.[6] What this

[6]Comparing pre-treatment covariates related to the outcome across the three groups could be used to further support the credibility of this ordering (Kowalski 2023). In my case, this is not easy to achieve, since treatment occurred when ELSA participants were young. In Appendix Figure A.1 I plot shares of respondents recalling being in excellent or very good health as a child. A larger share of AT report having been very healthy than C and NT, confirming the ordering. However, this

TABLE 4 – Frailty by response type: potential outcomes and selection on gains

Type	Share	Potential outcome	Mean frailty	Difference of means	Slope in potential outcomes	Selection on gains
Always-taker	0.430	Treated AT	0.158	-0.001	-0.002	
				(0.013)	(0.028)	
Complier	0.502	Treated C	0.157			-0.045
		Untreated C	0.183	0.012	0.043	(0.049)
				(0.014)	(0.050)	
Never-taker	0.069	Untreated NT	0.195			

Note: Shares of always-takers (AT), compliers (C), and never-takers (NT) alongside treated potential outcome means for always-takers and compliers, untreated potential outcome means for compliers and never-takers, as well as treated and untreated potential-outcome tests and a test for selection on gains. All estimates use the main sample described in Section 3.1. The shares come from a first-stage regression of the treatment (in school until at least 15) on the instrument (born 1933 or later), without gene interactions. The means are based on a regression of the frailty index on interactions of the treatment with the instrument. Appendix Section B includes a detailed description of how each number is calculated. Standard errors are computed using the Bayesian bootstrap (Rubin 1981) with 1,000 replications and are shown in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ indicate significance levels.

assumption allows for is the use of the untreated complier mean as an upper bound for the unobserved untreated potential outcome for always-takers, which also implies an upper bound on the local average treatment effect for always-takers. This is informative about treatment effect heterogeneity because I can now compare two LATEs. This bounding exercise shows that the maximum possible LATE for always-takers has the same size as the point-identified LATE for compliers. However, if there were selection on gains, we would expect always-takers (who are more eager to take the treatment) to have a larger LATE.

Table 4 shows the means and shares from Figure 5B, as well as tests for the difference between treated and untreated means, for slopes in treated and untreated outcomes (which additionally account for cohort size), and for selection on gains. The latter tests are based on Brinch, Mogstad, and Wiswall (2017). The difference between each pair of outcomes is small and not statistically significant. So are the slopes in potential outcomes as well as their difference, which yields a test for selection on gains across all groups. A negative coefficient on this test implies selection on gains (whereas a positive would imply selection on losses), meaning that those at the left of the horizontal axis in Figure 5B, the always-takers with the strongest taste for education, also benefit more from it (i.e. have higher treatment effects), and returns fall as willingness to take more education declines, that is, to the right. However, as Table 4 shows, the estimate for selection on gains is close to zero and not statistically significant.

Therefore, in my setting, due to the pronounced first-stage differences by EA PGI, the complier groups that 2SLS uses to calculate the interaction effects still differ along the interaction variable EA PGI. Therefore, strictly speaking, 2SLS makes "wrong comparisons", meaning it computes interaction effects on potential outcomes from different complier groups (the first condition from Hollenbach,

measure was asked retrospectively and refers only to "childhood", and thus may also reflect treated periods. Therefore, I also include means of the standardized subjective wellbeing PGI. This is set at conception, but does not refer to realized outcomes due to its probabilistic nature. Nevertheless, a higher PGI predicts better subjective wellbeing. There are few differences between the groups, but in general, the order fits, with AT having at least as many variants predicting wellbeing as C, who have slightly more variants than NT.

Schmitz, and Westphal 2026). However, there is no treatment effect heterogeneity along the unobserved willingness to take the treatment (the second condition). This entails that the comparisons 2SLS makes to compute the interaction effect do not differ numerically from a computation using fixed complier groups that could be achieved by estimating marginal treatment effects. As a result, my main 2SLS estimates reflect the true interaction effects.

4.4 Robustness

I investigate the robustness of my main results with respect to the bandwidth, age composition, linearity of the cohort trends, and the choice of quintiles, the split for the EA PGI, account for attrition probability by reweighting, and investigate an alternative outcome related to functional aging without the mortality aspect. Table 5 reports the results.

Lowering the bandwidth to 8 and 6 years does not change the nature of the results. Only the coefficient for the highest quintile shows a statistically significant effect of a year of education on frailty. In Column (3), I lower the age threshold from 65 to 50 and include some younger individuals as well as many mid-life observations of the older individuals that are the focus of this paper. Doing so does not change the results.

Next, I use quartiles of the EA PGI instead of quintiles to estimate the interaction terms. The results show a coefficient close to zero for the lowest quartile of the EA PGI distribution and, again, a sizable reduction in frailty for the highest quartile. As expected, this effect is smaller than the one for the top 20 %. In this specification, there is also a statistically significant reduction in frailty of 0.04 from the additional year of education for the second quartile.

As discussed above in Section 3.4, attrition in old-age samples may interfere with treatment effects. To test for differential attrition, I first fill in missing waves for all individuals (after the first wave in which they are observed) up to wave 10, keeping only the observations in which the individual would have been within my age range of 65 to 85, so that aging out of the sample is not counted as attrition. I then create an attrition indicator equal to one if the individual did not respond in a given wave and use it as the outcome in two specifications: the pooled 2SLS estimation (without PGI interactions) and a reduced-form regression of attrition directly on the instrument and controls. Appendix Table A.4 reports the results. Both coefficients are close to zero and no statistically significant. So, the reform, and schooling induced by the reform, are unrelated to panel attrition in my sample.

Nevertheless, I also estimate the main regression with inverse probability weights that reweight for panel response (Table 5, Column 7). I obtain the individual-level weights from a logit regression of a response indicator (1 – *attrition*, as defined above) on observables, including the reform dummy, the education dummy, the EA PGI, birth cohort, a wave indicator, as well as age and frailty measured in the individual's first observed wave. The sample for this regression includes all waves after an individual's first appearance until wave 10, as long as their projected age still falls into the 65 to 85 window. Everyone's first wave is excluded, since the inclusion probability is one there. The estimate in

TABLE 5 – Robustness

DV: Frailty index	(1) Main	(2) BW 8	(3) BW 6	(4) 50–85	(5) Quartiles	(6) IPW	(7) First obs.	(8) Functional
$E_i \times (G_i = 1)$	−0.010	−0.008	−0.017	−0.012	−0.009	−0.009	−0.007	−0.011
	(0.016)	(0.018)	(0.021)	(0.016)	(0.015)	(0.019)	(0.020)	(0.019)
$E_i \times (G_i = 2)$	−0.029	−0.027	−0.035	−0.030*	−0.041**	−0.030	−0.030	−0.027
	(0.018)	(0.020)	(0.022)	(0.018)	(0.018)	(0.021)	(0.023)	(0.022)
$E_i \times (G_i = 3)$	−0.026	−0.029	−0.029	−0.024	−0.019	−0.034	−0.032	−0.029
	(0.020)	(0.021)	(0.024)	(0.020)	(0.017)	(0.022)	(0.025)	(0.024)
$E_i \times (G_i = 4)$	−0.026	−0.028	−0.035	−0.025	−0.052**	−0.030	−0.018	−0.024
	(0.019)	(0.021)	(0.024)	(0.019)	(0.024)	(0.021)	(0.028)	(0.023)
$E_i \times (G_i = 5)$	−0.059**	−0.064**	−0.076**	−0.056**		−0.070**	−0.063*	−0.074**
	(0.026)	(0.030)	(0.035)	(0.026)		(0.029)	(0.033)	(0.032)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	18,705	16,057	12,962	21,083	18,705	18,705	3,415	18,705
Individuals	3,415	2,844	2,230	3,504	3,415	3,415	3,415	3,415

Note: The dependent variable is the frailty index. E_i is the education indicator (stayed in school until at least age 15) and G_i denotes quintiles of the educational attainment PGI. All columns report IV estimates of the interaction terms of the education indicator with each PGI group, instrumenting education with exposure to the 1947 compulsory schooling reform (born in 1933 or later). Column (1) reproduces the main IV estimates reported in Table 2. Columns (2) and (3) narrow the bandwidth around the reform cutoff to eight and six birth cohorts on either side. Column (4) widens the age range of the estimation sample to ages 50 to 85. Column (5) uses only the first observation per individual. Column (6) weights observations by the inverse of the estimated probability of being observed in a given wave, obtained from a logit model of panel response on baseline characteristics. Column (7) uses an alternative dependent variable, a frailty index constructed from the functional items only. Column (8) splits the PGI into quartiles instead of quintiles, so that only four interaction terms are estimated. All regressions include indicator variables for gender, survey wave, age, and mother's and father's age when they left full-time education, the running variable (re-centered birth year) and its interaction with the instrument, as well as the first ten principal components of the genetic data. Standard errors clustered at the individual level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ indicate significance levels.

Column (7) is my main regression, but weighted by the inverse of the predicted response probability from this logit regression. Again, the results are very similar to my main set of results.

Furthermore, I estimate a specification using only the first observation of each individual. By definition, there can be no panel attrition, since it only uses cross-sectional information. The results do not differ substantially from the main results.

Finally, I estimate the main regression for an alternative outcome, a sub-index of the frailty index that omits severe conditions. These omitted items are cardiovascular, metabolic, cognitive, and other severe diagnoses. The remaining 40 items include impairments in activities of daily living, self-assessments of physical and mental health, and cognitive tests, which cover functional and sensory aspects of the frailty index, but do not relate as much to the mortality dimension. The results are qualitatively and quantitatively (as a share of the standard deviation) comparable to the main results. This shows that the results hold for an alternative outcome that tracks mortality less closely, but instead measures quality of life and autonomy in old age.

4.5 Suggestive evidence on channels

Next, I turn to non-health outcomes that relate to potential channels of an effect of education on old-age health are affected by the reform. I investigate outcomes in three categories: income and

pensions (Panel A of Table 6), employment history (Panel B), and health behavior (Panel C). The table also includes mean values of these outcomes and the number of observations and individuals. Each line is a separate IV regression of the respective outcome on my education indicator instrumented by the reform dummy without PGI interactions using the same sample limitations as the main analysis (10 birth cohorts to either side of the 1933 cutoff and observations from individuals aged 65–85).

The youngest respondents in my sample are 65 years old, so right at the official retirement age. Many have been retired since the first time they were observed in ELSA. Therefore, data on current jobs or earnings reflects jobs respondents have taken on in retirement, rather than the career these individuals have had. I deal with this issue in three ways. First, I use information from ELSA's life history module. There, respondents are asked about their careers and the earnings of their final job in their main career. Second, I look at subjective financial situation, that is, whether a person has ever experienced financial hardship. Third, I use information about the size of public and employer pensions as well as information about receiving a disability pension. The more a person has paid into public or private pension schemes, the more they should be able to take out, giving me a rough idea on whether income might have been affected.

Panel A suggests that the additional education from the 1947 schooling reform did not affect income or pensions. Coefficients on pay in final job, financial hardship, public and employer pensions are close to zero and not statistically significant. Overall, this seems consistent with the small to zero returns from this reform found by Devereux and Hart (2010) or Clark (2023). However, there is a sizable decrease in the probability of receiving a disability pension. Abeliansky and Strulik (2023) show an occupational gradient in the frailty index. Workers in physically more demanding jobs accumulate health deficits faster. A decrease in the probability of relying on disability pension may hint at less burden and injuries through the job. Unfortunately, ELSA does not provide detailed information about the nature of the jobs in the life history module, like occupational classifiers. Therefore, I am unable to investigate patterns of physical or mental strain of the former jobs.

In terms of employment history, I am able to identify instances of unemployment, self-employment, part-time emplyment, and the number of jobs a person has held. I also investigate the impact of education on taking early retirement, identified from the main sample. The results suggest that there is no effect of the additional year of compulsory education on whether a respondent was ever unemployed, self-employed, ever in part time or on the number of jobs they have had during their working life or has retired early. The literature broadly finds the same: Clark (2023) investigates the effect of the 1947 compulsory schooling reform on several labor market outcomes, including labor force participation, unemployment, selfemployment and finds no effects.

Panel C investigates different aspects of (risky) health behaviors. I observe a large decrease in the probability of ever having smoked and a statistically insignificant, but still sizable estimate suggesting a reduction in the probability of currently smoking. However, there does not seem to be an effect on current alcohol consumption or heavy (risky) drinking, which, according to official UK advisory, is more than 14 units per week (UK Department of Health and Social Care 2016). In terms of physical activity, there is no effect of education on doing any light activity, but sizable positive effects on doing

TABLE 6 – Effect of education on non-health outcomes

	IV estimate	Mean	Observations	Individuals
A. Income and pensions				
Log final pay in last job	−0.115	16,760	2,122	2,122
	(0.236)			
Ever severe financial hardship	−0.024	0.177	3,141	3,141
	(0.055)			
Log public old-age pension income	0.042	5,704	30,943	6,712
	(0.150)			
Log private/employer pension income	0.745	4,342	30,942	6,698
	(0.720)			
Receives disability pension	−0.066**	0.137	31,253	6,717
	(0.032)			
B. Employment history				
Ever unemployed	−0.014	0.342	3,612	3,612
	(0.064)			
Ever self-employed	0.016	0.207	3,588	3,588
	(0.054)			
Ever part-time	0.022	0.475	3,588	3,588
	(0.057)			
Number of jobs	−0.008	5.294	3,612	3,612
	(0.343)			
Early retirement	0.017	0.533	5,928	5,928
	(0.056)			
C. Health behavior				
Ever smoked	−0.137**	0.606	3,610	3,610
	(0.064)			
Smokes currently	−0.047	0.098	31,287	6,679
	(0.031)			
Drank alcohol last 12 months	−0.038	0.845	27,206	6,488
	(0.038)			
Risky drinking (14+/week)	−0.009	0.047	37,693	8,649
	(0.014)			
Any light activity	0.023	0.731	37,693	8,649
	(0.021)			
Any moderate activity	0.117***	0.631	37,693	8,649
	(0.029)			
Any vigorous activity	0.068**	0.254	37,693	8,649
	(0.029)			
BMI	−0.538	27.976	11,426	4,951
	(0.616)			

Note: Standard errors clustered at the individual level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ indicate significance levels.

moderate and vigorous activity. Nevertheless, this does not seem to translate to weight reductions, as the estimate on BMI is not statistically significant and not of relevant size given the high average.

Better health behavior may be able to explain reductions in old-age frailty from increased education. Strulik (2018a) uses a health deficit model centered around frailty and calibrated to US data to show that more educated inviduals delay the accumulation of deficits through lower unhealthy consumption

(e.g. smoking) and increased health investments (e.g. exercise). In Strulik (2018b), he shows that smoking drives long-run health trajectories and finds longevity gains from not smoking. Lack of exercise has long been identified as a key risk factor for old-age morbidity (World Health Organization 2009). Futhermore, evidence from RCTs randomizing incentives for exercise and exercise itself show that increased physical activity reduces blood sugar in middle-aged diabetics and helps manage the disease (Aggarwal, Dizon-Ross, and Zucker 2026) and increases functional health in old age (Pahor et al. 2014). Furthermore, both increased exercise and reduced smoking are frequently reported as key mechanisms for the effect of retirement on old-age health (see, e.g. Insler 2014 or Eibich 2015).

4.6 Conclusion

In this study, I provide evidence on the relationship between genetic propensity for education, educational attainment and old-age frailty using data from the English Longitudinal Study of Ageing and individuals aged 65–85. First, I document a sizable gap in old-age frailty by genetic predisposition for education that is almost as big as the socioeconomic gap at this age. These health differences decline with educational attainment and disappear amont the highest educated. Second, I estimate a causal effect of education on old-age frailty by genetic propensity for education using a fuzzy regression discontinuity design where exposure to the 1947 UK compulsory schooling reform serve as exogenous variation in schooling. The results show that more education decreases old-age frailty. This decrease is highest and only statistically significant for those with high genetic propensity for education. This shows that only those most genetically inclined to succeed in school have positive long-term returns to increased instruction time in terms of health. Furthermore, I show that the improvements in frailty originate in a delayed buildup of a larger number of deficits rather than from staying (nearly) deficit-free for longer. Suggestive evidence on channels points to more advantageous health behavior due to education and a reduced likelihood of relying on disability benefits, while I find no indication for income or labor market effects.

This is in line with the notion that those with a higher genetic predisposition for education are able to extract larger benefits from their education (Houmark, Ronda, and Rosholm 2024). My results show that this type of skill-building interaction in early life may extend to the additional instruction time generated by a compulsory schooling reform, where the treatment effect is concentrated among compliers in the upper part of the PGI distribution, and that this has positive health payoffs many years later for those with a high genetic predisposition, in addition to the already established positive income and cognitive returns to schooling for this group (Barcellos, Carvalho, and Turley 2021; Hollenbach, Schmitz, and Westphal 2026).

References

Abeliansky, Ana Lucia and Holger Strulik. 2018. "How We Fall Apart: Similarities of Human Aging in 10 European Countries." *Demography* 55 (1): 341–359.

Abeliansky, Ana Lucia and Holger Strulik. 2023. "Health and aging before and after retirement." *Journal of Population Economics* 36 (4): 2825–2855.

Aggarwal, Shilpa, Rebecca Dizon-Ross, and Ariel D. Zucker. 2026. "Designing Incentives for Impatient People: An RCT Promoting Exercise to Manage Diabetes." *American Economic Journal: Applied Economics*. Forthcoming.

Arold, Benjamin W., Paul Hufe, and Marc Stoeckli. 2025. "Genetic endowments, educational outcomes and the mediating influence of school investments." *Journal of Political Economy Microeconomics*. Forthcoming.

Auton, Adam, Gonçalo R. Abecasis, David M. Altshuler, Richard M. Durbin, Gonçalo R. Abecasis, David R. Bentley, Aravinda Chakravarti, Andrew G. Clark, Peter Donnelly, Evan E. Eichler, et al. 2015. "A global reference for human genetic variation." *Nature* 526 (7571): 68–74.

Banks, James and Fabrizio Mazzonna. 2012. "The effect of education on old age cognitive abilities: evidence from a regression discontinuity design." *The Economic Journal* 122 (560): 418–448.

Barcellos, Silvia H., Leandro Carvalho, and Patrick Turley. 2021. "The Effect of Education on the Relationship between Genetics, Early-Life Disadvantages, and Later-Life SES." NBER Working paper 28750.

Barcellos, Silvia H., Leandro S. Carvalho, and Patrick Turley. 2018. "Education can reduce health differences related to genetic risk of obesity." *Proceedings of the National Academy of Sciences* 115 (42): E9765–E9772.

Barcellos, Silvia H., Leandro S. Carvalho, and Patrick Turley. 2023. "Distributional Effects of Education on Health." *Journal of Human Resources* 58 (4): 1273–1306.

Barth, Daniel, Nicholas W. Papageorge, and Kevin Thom. 2020. "Genetic Endowments and Wealth Inequality." *Journal of Political Economy* 128 (4): 1474–1522.

Barth, Daniel, Nicholas W. Papageorge, Kevin Thom, and Mateo Velásquez-Giraldo. 2022. "Genetic endowments, income dynamics, and wealth accumulation over the lifecycle." Working paper 30350, National Bureau of Economic Research.

Belsky, Daniel W., Terrie E. Moffitt, David L. Corcoran, Benjamin Domingue, HonaLee Harrington, Sean Hogan, Renate Houts, Sandhya Ramrakha, Karen Sugden, Benjamin S. Williams, et al. 2016. "The Genetics of Success: How Single-Nucleotide Polymorphisms Associated With Educational Attainment Relate to Life-Course Development." *Psychological Science* 27 (7): 957–972.

Benjamin, Daniel J., David Cesarini, Christopher F. Chabris, Edward L. Glaeser, David I. Laibson, Vilmundur Guðnason, Tamara B. Harris, Lenore J. Launer, Shaun Purcell, Albert Vernon Smith, et al. 2012. "The Promises and Pitfalls of Genoeconomics." *Annual Review of Economics* 4 (1): 627–662.

Biroli, Pietro, Titus Galama, Stephanie von Hinke, Hans van Kippersluis, Cornelius A. Rietveld, and Kevin Thom. 2026. "The Economics and Econometrics of Gene–Environment Interplay." *Review of Economic Studies* 93 (1): 144–180.

Black, Dan A. and Jeffrey A. Smith. 2006. "Estimating the Returns to College Quality with Multiple Proxies for Quality." *Journal of Labor Economics* 24 (3): 701–728.

Brinch, Christian N. and Taryn Ann Galloway. 2011. "Schooling in adolescence raises IQ scores." *Proceedings of the National Academy of Sciences* 109 (2): 425–430.

Brinch, Christian N., Magne Mogstad, and Matthew Wiswall. 2017. "Beyond LATE with a discrete instrument." *Journal of Political Economy* 125 (4): 985–1039.

Carneiro, Pedro, James J. Heckman, and Edward J. Vytlacil. 2011. "Estimating marginal returns to education." *American Economic Review* 101 (6): 2754–2781.

Carvalho, Leandro S. 2025. "Genetics and socioeconomic status: some preliminary evidence on mechanisms." *Journal of Political Economy Microeconomics* 3 (3): 429–476.

Chabris, Christopher F., James J. Lee, David Cesarini, Daniel J. Benjamin, and David I. Laibson. 2015. "The Fourth Law of Behavior Genetics." *Current Directions in Psychological Science* 24 (4): 304–312.

Cheesman, Rosa, Nicolai Borgen, Astrid M. J. Sandsør, and Paul Hufe. 2025. "The genetic lottery goes to school: Better schools compensate for the effects of students' genetic differences." *Proceedings of the National Academy of Sciences* 122 (43).

Clark, Damon. 2023. "School quality and the return to schooling in Britain: new evidence from a large-scale compulsory schooling reform." *Journal of Public Economics* 223: 104902.

Clark, Damon and Heather Royer. 2013. "The effect of education on adult mortality and health: Evidence from Britain." *American Economic Review* 103 (6): 2087–2120.

Cunha, Flavio and James J. Heckman. 2007. "The Technology of Skill Formation." *American Economic Review* 97 (2): 31–47.

Dale, S. B. and A. B. Krueger. 2002. "Estimating the Payoff to Attending a More Selective College: An Application of Selection on Observables and Unobservables." *The Quarterly Journal of Economics* 117 (4): 1491–1527.

Dalgaard, Carl-Johan and Holger Strulik. 2014. "Optimal Aging and Death: Understanding the Preston Curve." *Journal of the European Economic Association* 12 (3): 672–701.

Davies, Katie, Asri Maharani, Tarani Chandola, Chris Todd, and Neil Pendleton. 2021. "The longitudinal relationship between loneliness, social isolation, and frailty in older adults in England: a prospective analysis." *The Lancet Healthy Longevity* 2 (2): e70–e77.

Davies, Neil M., Matt Dickson, George Davey Smith, Gerard J. van den Berg, and Frank Windmeijer. 2018. "The causal effects of education on health outcomes in the UK Biobank." *Nature Human Behaviour* 2 (2): 117–125.

Devereux, Paul J. and Robert A. Hart. 2010. "Forced to be rich? Returns to compulsory schooling in Britain." *The Economic Journal* 120 (549): 1345–1364.

Eibich, Peter. 2015. "Understanding the effect of retirement on health: Mechanisms and heterogeneity." *Journal of Health Economics* 43: 1–12.

Galama, Titus, Adriana Lleras-Muney, and Hans van Kippersluis. 2018. "The Effect of Education on Health and Mortality: A Review of Experimental and Quasi-Experimental Evidence." *Oxford Research Encyclopedia of Economics and Finance*.

Gathmann, Christina, Hendrik Jürges, and Steffen Reinhold. 2015. "Compulsory schooling reforms, education and mortality in twentieth century Europe." *Social Science & Medicine* 127: 74–82.

Harmon, Colm and Ian Walker. 1995. "Estimates of the economic return to schooling for the United Kingdom." *American Economic Review* 85 (5): 1278–1286.

Heckman, James J., Sergio Urzua, and Edward Vytlacil. 2006. "Understanding instrumental variables in models with essential heterogeneity." *The Review of Economics and Statistics* 88 (3): 389–432.

Hollenbach, Johannes, Hendrik Schmitz, and Matthias Westphal. 2026. "Gene-environment interactions with essential heterogeneity." *The Economic Journal*. Forthcoming.

Hosseini, Roozbeh, Karen Kopecky, and Kai Zhao. 2025. "How Important Is Health Inequality for Lifetime Earnings Inequality?" *Review of Economic Studies* 93 (1): 556–599.

Hosseini, Roozbeh, Karen A. Kopecky, and Kai Zhao. 2022. "The evolution of health over the life cycle." *Review of Economic Dynamics* 45: 237–263.

Houmark, Mikkel Aagaard, Victor Ronda, Esben Agerbo, Preben Bo Mortensen, and Michael Rosholm. 2025. "Genetic and Socioeconomic Achievement Gaps in Elementary School." *The Economic Journal* 135 (672): 2426–2454.

Houmark, Mikkel Aagaard, Victor Ronda, and Michael Rosholm. 2024. "The nurture of nature and the nature of nurture: how genes and investments interact in the formation of skills." *American Economic Review* 114 (2): 385–425.

Huang, Jian, Ying Gui, Jing Wu, and Yubo Xie. 2024. "Causal effects of socioeconomic traits on frailty: a Mendelian randomization study." *Frontiers in Medicine* 11.

Imbens, G. W. and D. B. Rubin. 1997. "Estimating outcome distributions for compliers in instrumental variables models." *The Review of Economic Studies* 64 (4): 555–574.

Insler, Michael. 2014. "The Health Consequences of Retirement." *Journal of Human Resources* 49 (1): 195–233.

Johnston, David W., Grace Lordan, Michael A. Shields, and Agne Suziedelyte. 2015. "Education and health knowledge: evidence from UK compulsory schooling reform." *Social Science & Medicine* 127: 92–100.

Jürges, Hendrik, Eberhard Kruk, and Steffen Reinhold. 2013. "The effect of compulsory schooling on health—evidence from biomarkers." *Journal of Population Economics* 26 (2): 645–672.

Kamhöfer, Daniel A., Hendrik Schmitz, and Matthias Westphal. 2019. "Heterogeneity in Marginal Non-Monetary Returns to Higher Education." *Journal of the European Economic Association* 17 (1): 205–244.

Kong, Augustine, Gudmar Thorleifsson, Michael L. Frigge, Bjarni J. Vilhjalmsson, Alexander I. Young, Thorgeir E. Thorgeirsson, Stefania Benonisdottir, Asmundur Oddsson, Bjarni V. Halldorsson, Gisli Masson, et al. 2018. "The nature of nurture: Effects of parental genotypes." *Science* 359 (6374): 424–428.

Kowalski, Amanda E. 2023. "How to examine external validity within an experiment." *Journal of Economics & Management Strategy* 32 (3): 491–509.

Lee, James J., Robbee Wedow, Aysu Okbay, Edward Kong, Omeed Maghzian, Meghan Zacher, Tuan Anh Nguyen-Viet, Peter Bowers, Julia Sidorenko, Richard Karlsson Linnér, et al. 2018. "Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals." *Nature Genetics* 50 (8): 1112–1121.

Lombardi, Stefano, Nurfatima Jandarova, Kristina Zguro, Jarkko Harju, Aldo Rustichini, and Andrea Ganna. 2026. "Effects of Genetic Propensity for Education on Labor Market and Health Trajectories across the Working Life." Working Paper. arXiv:2604.24336.

Marshall, Alan, James Nazroo, Gindo Tampubolon, and Bram Vanhoutte. 2015. "Cohort differences in the levels and trajectories of frailty among older people in England." *Journal of Epidemiology & Community Health* 69 (4): 316–321.

Mitnitski, Arnold B., Alexander J. Mogilner, and Kenneth Rockwood. 2001. "Accumulation of Deficits as a Proxy Measure of Aging." *The Scientific World Journal* 1: 323–336.

Muslimova, Dilnoza, Hans van Kippersluis, Cornelius A. Rietveld, Stephanie von Hinke, and S. Fleur W. Meddens. 2026. "Gene-Environment Complementarity in Educational Attainment." *Journal of Labor Economics*. Forthcoming.

Neal, Derek A. and William R. Johnson. 1996. "The Role of Premarket Factors in Black-White Wage Differences." *Journal of Political Economy* 104 (5): 869–895.

Niederstrasser, Nils Georg, Nina Trivedy Rogers, and Stephan Bandelow. 2019. "Determinants of frailty development and progression using a multidimensional frailty index: Evidence from the English Longitudinal Study of Ageing." *PLOS ONE* 14 (10): e0223799.

Nurk, Sergey, Sergey Koren, Arang Rhie, Mikko Rautiainen, Andrey V. Bzikadze, Alla Mikheenko, Mitchell R. Vollger, Nicolas Altemose, Lev Uralsky, Ariel Gershman, et al. 2022. "The complete sequence of a human genome." *Science* 376 (6588): 44–53.

Okbay, Aysu, Yeda Wu, Nancy Wang, Hariharan Jayashankar, Michael Bennett, Seyed Moeen Nehzati, Julia Sidorenko, Hyeokmoon Kweon, Grant Goldman, Tamara Gjorgjieva, et al. 2022. "Polygenic prediction of educational attainment within and between families from genome-wide association analyses in 3 million individuals." *Nature Genetics* 54 (4): 437–449.

Oreopoulos, Philip. 2006. "Estimating Average and Local Average Treatment Effects of Education when Compulsory Schooling Laws Really Matter." *American Economic Review* 96 (1): 152–175.

Pahor, Marco, Jack M. Guralnik, Walter T. Ambrosius, Steven Blair, Denise E. Bonds, Timothy S. Church, Mark A. Espeland, Roger A. Fielding, Thomas M. Gill, Erik J. Groessl, et al. 2014. "Effect of Structured Physical Activity on Prevention of Major Mobility Disability in Older Adults: The LIFE Study Randomized Clinical Trial." *JAMA* 311 (23): 2387.

Papageorge, Nicholas W. and Kevin Thom. 2020. "Genes, education, and labor market outcomes: evidence from the Health and Retirement Study." *Journal of the European Economic Association* 18 (3): 1351–1399.

Polderman, Tinca J. C., Beben Benyamin, Christiaan A. de Leeuw, Patrick F. Sullivan, Arjen van Bochoven, Peter M. Visscher, and Danielle Posthuma. 2015. "Meta-analysis of the heritability of human traits based on fifty years of twin studies." *Nature Genetics* 47 (7): 702–709.

Powdthavee, Nattavudh. 2010. "Does education reduce the risk of hypertension? Estimating the biomarker effect of compulsory schooling in England." *Journal of Human Capital* 4 (2): 173–202.

Price, Alkes L., Nick J. Patterson, Robert M. Plenge, Michael E. Weinblatt, Nancy A. Shadick, and David Reich. 2006. "Principal components analysis corrects for stratification in genome-wide association studies." *Nature Genetics* 38 (8): 904–909.

Rockwood, K. and A. Mitnitski. 2007. "Frailty in Relation to the Accumulation of Deficits." *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences* 62 (7): 722–727.

Rubin, Donald B. 1981. "The Bayesian Bootstrap." *The Annals of Statistics* 9 (1): 130–134.

Searle, Samuel D., Arnold Mitnitski, Evelyne A. Gahbauer, Thomas M. Gill, and Kenneth Rockwood. 2008. "A standard procedure for creating a frailty index." *BMC Geriatrics* 8 (1): 24.

Silles, Mary A. 2009. "The causal effect of education on health: Evidence from the United Kingdom." *Economics of Education Review* 28 (1): 122–128.

Song, Xiaowei, Arnold Mitnitski, and Kenneth Rockwood. 2010. "Prevalence and 10-Year Outcomes of Frailty in Older Adults in Relation to Deficit Accumulation." *Journal of the American Geriatrics Society* 58 (4): 681–687.

Strulik, Holger. 2018a. "The return to education in terms of wealth and health." *The Journal of the Economics of Ageing* 12: 1–14.

Strulik, Holger. 2018b. "Smoking kills: An economic theory of addiction, health deficit accumulation, and longevity." *Journal of Health Economics* 62: 1–12.

Todd, Petra E. and Kenneth I. Wolpin. 2003. "On the Specification and Estimation of the Production Function for Cognitive Achievement." *The Economic Journal* 113 (485): F3–F33.

UK Department of Health and Social Care. 2016. "UK Chief Medical Officers' Low Risk Drinking Guidelines." https://www.gov.uk/government/publications/alcohol-consumption-advice-on-low-risk-drinking. Accessed 19 August 2026.

Visscher, Peter M., Naomi R. Wray, Qian Zhang, Pamela Sklar, Mark I. McCarthy, Matthew A. Brown, and Jian Yang. 2017. "10 Years of GWAS Discovery: Biology, Function, and Translation." *The American Journal of*

Human Genetics 101 (1): 5–22.

Westphal, Matthias, Daniel A. Kamhöfer, and Hendrik Schmitz. 2022. "Marginal college wage premiums under selection into employment." *The Economic Journal* 132 (646): 2231–2272.

World Health Organization. 2009. "Global health risks - Mortality and burden of disease attributable toselected major risks."

Appendix A. Additional figures and tables

TABLE A.1 – Frailty index items and coding

Description	Coding
Difficulty with walking 100 yards	No=0, Yes=1
Difficulty sitting for about two hours	No=0, Yes=1
Difficulty getting up from a chair after sitting for long periods	No=0, Yes=1
Difficulty climbing several flights of stairs without resting	No=0, Yes=1
Difficulty climbing one flight of stairs without resting	No=0, Yes=1
Difficulty stooping, kneeling, or crouching	No=0, Yes=1
Difficulty reaching or extending arms above shoulder level	No=0, Yes=1
Difficulty pulling or pushing large objects like a living room chair	No=0, Yes=1
Difficulty lifting or carrying weights over 10 pounds, like a heavy bag	No=0, Yes=1
Difficulty picking up a 5p coin from a table	No=0, Yes=1
Difficulty dressing, including putting on shoes and socks	No=0, Yes=1
Difficulty walking across a room	No=0, Yes=1
Difficulty bathing or showering	No=0, Yes=1
Difficulty eating, such as cutting up your food	No=0, Yes=1
Difficulty getting in or out of bed	No=0, Yes=1
Difficulty using the toilet, including getting up or down	No=0, Yes=1
Difficulty using a map to figure out how to get around in a strange place	No=0, Yes=1
Difficulty preparing a hot meal	No=0, Yes=1
Difficulty shopping for groceries	No=0, Yes=1
Difficulty making telephone calls	No=0, Yes=1
Difficulty taking medications	No=0, Yes=1
Difficulty managing money (e.g. paying bills, keeping track of expenses)	No=0, Yes=1
Difficulty doing work around the house or garden	No=0, Yes=1
Self-reported general health	Excellent=0, v.good=0.25, Good=0.5, Fair=0.75, Poor=1
Felt depressed much of the time during the past week	No=0, Yes=1
Felt everything done during the past week was an effort	No=0, Yes=1
Felt sleep was restless much of the time during the past week	No=0, Yes=1
Was happy much of the time during the past week	Yes=0, No=1
Felt lonely much of the time during the past week	No=0, Yes=1
Enjoyed life much of the time during the past week	Yes=0, No=1
Felt sad much of the time during the past week	No=0, Yes=1
Could not get going much of the time during the past week	No=0, Yes=1
High blood pressure or hypertension	No=0, Yes=1
Angina	No=0, Yes=1
Heart attack (including MI or coronary thrombosis)	No=0, Yes=1
Congestive heart failure	No=0, Yes=1
Abnormal heart rhythm	No=0, Yes=1
Diabetes or high blood sugar	No=0, Yes=1
Stroke (cerebral vascular disease)	No=0, Yes=1
Chronic lung disease (e.g. chronic bronchitis or emphysema)	No=0, Yes=1
Asthma	No=0, Yes=1
Arthritis (including osteoarthritis or rheumatism)	No=0, Yes=1
Osteoporosis	No=0, Yes=1
Cancer or malignant tumor (excluding minor skin cancers)	No=0, Yes=1

Continued on next page

Table A.1 continued from previous page

Description	Coding
Parkinson's disease	No=0, Yes=1
Any emotional, nervous or psychiatric problems	No=0, Yes=1
Alzheimer's disease	No=0, Yes=1
Dementia, organic brain syndrome, senility or other serious memory impairment	No=0, Yes=1
Self-reported eyesight function (with lenses, if appropriate)	Excellent=0, v.good=0.2, Good=0.4, Fair=0.6, Poor=0.8, Blind=1
Self-reported hearing function (with hearing aid, if appropriate)	Excellent=0, v.good=0.25, Good=0.5, Fair=0.75, Poor=1
Identify today's date: day of month	No=0, Yes=1
Identify today's date: month	No=0, Yes=1
Identify today's date: year	No=0, Yes=1
Identify the day of the week	No=0, Yes=1
Immediate word recall (quartiles)	Q1=1, Q2=0.67, Q3=0.33, Q4=0
Delayed word recall (quintiles)	Q1=1, Q2=0.75, Q3=0.5, Q4=0.25, Q5=0

TABLE A.2 – Summary statistics (full table)

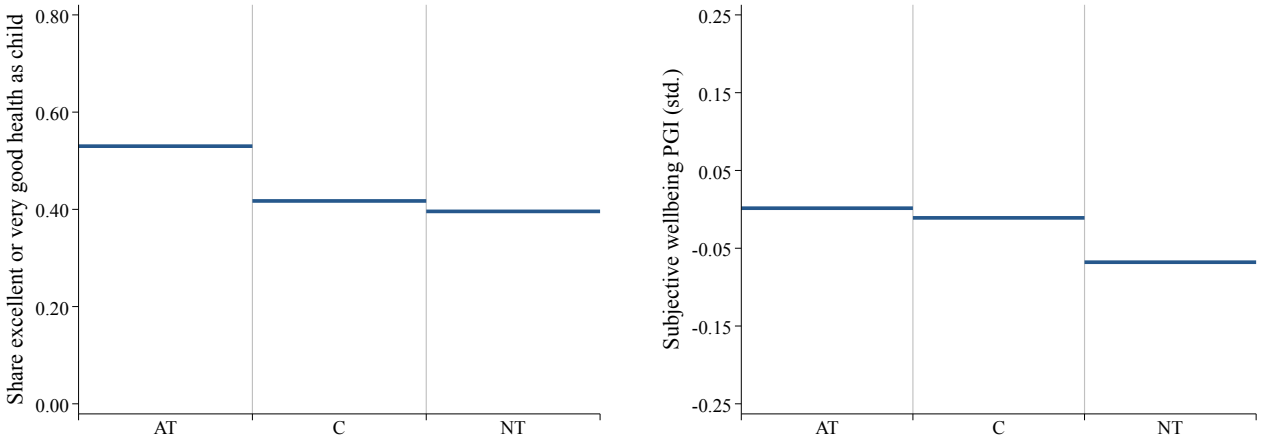
	Full sample				By E_i (Left school at 15 or later)		
	Mean	SD	Min	Max	0	1	Diff. (p-val.)
Frailty index	0.162	0.110	0.000	0.772	0.194	0.153	−0.041 (0.000)***
Left school at 15 or later	0.790		0	1	0.000	1.000	1.000
Born 1933 or later	0.685		0	1	0.143	0.829	0.686 (0.000)***
Educational attainment PGI							
1st quintile	0.186				0.234	0.174	−0.061 (0.000)***
2nd quintile	0.199				0.214	0.195	−0.019 (0.008)***
3rd quintile	0.200				0.189	0.203	0.014 (0.048)**
4th quintile	0.209				0.206	0.210	0.004 (0.582)
5th quintile	0.206				0.158	0.219	0.061 (0.000)***
Female	0.544		0	1	0.525	0.549	0.024 (0.009)***
Age	74.238	5.344	65.000	85.000	77.049	73.491	−3.559 (0.000)***
Birth year	1,935.027	5.241	1,923.000	1,943.000	1,929.798	1,936.416	6.619 (0.000)***
Age mother left education							
≤ 14	0.678				0.656	0.683	0.027 (0.001)***
15	0.060				0.018	0.071	0.053 (0.000)***
16	0.039				0.007	0.048	0.040 (0.000)***
17	0.013				0.002	0.016	0.015 (0.000)***
18	0.017				0.004	0.020	0.016 (0.000)***
≥ 19	0.019				0.005	0.022	0.017 (0.000)***
Missing	0.175				0.308	0.140	−0.168 (0.000)***
Age father left education							
≤ 14	0.663				0.634	0.671	0.037 (0.000)***
15	0.054				0.022	0.062	0.039 (0.000)***
16	0.044				0.005	0.054	0.049 (0.000)***
17	0.010				0.004	0.011	0.008 (0.000)***
18	0.019				0.005	0.022	0.018 (0.000)***
≥ 19	0.030				0.005	0.037	0.031 (0.000)***
Missing	0.181				0.325	0.143	−0.182 (0.000)***
Survey wave							
1	0.104				0.188	0.081	−0.107 (0.000)***
2	0.125				0.189	0.108	−0.082 (0.000)***
3	0.124				0.164	0.114	−0.051 (0.000)***
4	0.146				0.152	0.145	−0.008 (0.226)
5	0.131				0.118	0.135	0.017 (0.006)***
6	0.114				0.087	0.122	0.035 (0.000)***
7	0.093				0.055	0.103	0.048 (0.000)***
8	0.075				0.030	0.087	0.057 (0.000)***
9	0.058				0.012	0.071	0.059 (0.000)***
10	0.029				0.004	0.036	0.032 (0.000)***
PC1	−0.000	0.012	−0.027	0.041	−0.001	−0.000	0.000 (0.120)
PC2	0.000	0.012	−0.027	0.030	−0.000	0.000	0.000 (0.039)**
PC3	0.000	0.011	−0.057	0.047	0.000	0.000	0.000 (0.261)
PC4	0.000	0.012	−0.036	0.054	0.000	0.000	−0.000 (0.543)
PC5	0.000	0.012	−0.039	0.040	0.000	0.000	−0.000 (0.088)*
PC6	0.000	0.012	−0.034	0.044	−0.001	0.000	0.001 (0.000)***
PC7	0.000	0.012	−0.034	0.042	−0.001	0.000	0.001 (0.000)***
PC8	0.000	0.012	−0.039	0.042	0.000	−0.000	−0.000 (0.186)
PC9	0.000	0.012	−0.039	0.040	−0.000	0.000	0.000 (0.068)*
PC10	−0.000	0.012	−0.048	0.048	−0.000	−0.000	0.000 (0.603)
Observations			18,705		3,927	14,778	

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ indicate significance levels.

TABLE A.3 – Reduced-form and first-stage estimates

	Reduced form	First stage	
	DV: Frailty (Y_i)	DV: Left school at 15 or later (E_i)	F
$Z_i \times (G_i = 1)$	−0.006	0.653***	221.6
	(0.011)	(0.044)	
$Z_i \times (G_i = 2)$	−0.016	0.548***	159.1
	(0.010)	(0.043)	
$Z_i \times (G_i = 3)$	−0.013	0.487***	123.1
	(0.010)	(0.044)	
$Z_i \times (G_i = 4)$	−0.011	0.448***	104.2
	(0.009)	(0.044)	
$Z_i \times (G_i = 5)$	−0.022**	0.356***	64.9
	(0.010)	(0.044)	
Controls		Yes	
Observations		18,705	
Individuals		3,415	

Note: Reduced-form and first-stage (Eq. 2) estimates. Etimated on the main sample described in Section 3.1. The first-stage estimates are plotted in Figure 5A. Each interaction of the instrument Z_i with a PGI quintile G_i is a separate first stage (though estimated in one model). Therefore, the last column also reports F-statistics for the excluded instrument separately for each interaction. Both models include the same controls: a linear cohort trend, its interaction with the instrument, indicator variables for gender, survey wave, age, and mother's and father's age when they left full-time education, and the first ten principal components of the genetic data. Standard errors clustered at the individual level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ indicate significance levels.



A. Shares of AT, C, and NT who recall very good childhood health

B. Means of wellbeing PGI for AT, C, and NT

FIGURE A.1 – Possible support for weak monotonicity in untreated outcomes

Note: Panel A: Strength of the first stage (i.e., complier shares) by PGI quintile with 95% confidence intervals. Appendix Table A.3 reports the exact estimates, standard errors, and first-stage F-statistics. Panel B: Treated potential outcome means for always-takers (AT) and compliers (C), and untreated means for compliers and never-takers (NT), without gene interactions. Blue denotes treated potential outcomes, orange untreated. The horizontal axis shows each group and its share of the sample, with the width of each group corresponding to that share. Groups are ordered from most to least willing to take the treatment (more education). Additionally assuming weak monotonicity of untreated potential outcomes yields an upper bound on the local average treatment effect for always-takers. Appendix Section B describes in detail how all numbers are calculated. Both analyses are calculated on the main sample described in Section 3.1.

TABLE A.4 – Panel attrition

DV: Attrition	(1) 2SLS	(2) RF
E_i	−0.007	
	(0.037)	
Z_i		−0.003
		(0.020)
Controls	Yes	Yes
Observations	25,805	25,805
Individuals	3,415	3,415

Note: Notes: The dependent variable is an indicator equal to one if an individual eligible for interview in a given wave does not respond. E_i is the education indicator (stayed in school until at least age 15), instrumented by Z_i in column (1). Z_i is an indicator for exposure to the 1947 compulsory schooling reform (born in 1933 or later). Column (1) reports the 2SLS estimate, column (2) the reduced-form estimate. All specifications control for the running variable, its interaction with Z_i, gender, and wave fixed effects. Standard errors clustered at the individual level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ indicate significance levels.

Appendix B. Potential outcomes and a test for selection on gains

In Figure 5B, I show the shares of always-takers (AT), compliers (C), and never-takers (NT) in my setting, alongside four potential outcomes: the treated potential outcome for always-takers, $\mathbb{E}[Y_i^1|AT]$, the treated and untreated potential outcomes for compliers, $\mathbb{E}[Y_i^1|C]$ and $\mathbb{E}[Y_i^0|C]$, and the untreated potential outcome for never-takers, $\mathbb{E}[Y_i^0|NT]$. Imbens and Rubin (1997) show how to calculate these seven numbers under the standard LATE assumptions and given a binary treatment E_i (for education) and instrument Z_i.

Based on the data, i.e., observed E_i and Z_i, it is possible to identify only the following four objects:

i. $\mathbb{E}[Y_i|E_i = 1, Z_i = 0] = \mathbb{E}[Y_i^1|AT]$. This is the potential outcome for those who take the treatment even when the instrument is switched off. Under the monotonicity assumption (no defiers), they must be always-takers. If the instrument is exogenous, taking the mean of Y_i for individuals with $E_i = 1, Z_i = 0$ cleanly identifies the treated potential outcome for always-takers.
ii. $\mathbb{E}[Y_i|E_i = 0, Z_i = 1] = \mathbb{E}[Y_i^0|NT]$. Similarly, we can obtain the untreated potential outcome for never-takers from those who do not take the treatment even though the instrument is turned on.
iii. $\mathbb{E}[Y_i|E_i = 1, Z_i = 1]$. This is the potential outcome for a mix of compliers and always-takers. At the individual level, we cannot distinguish, among those who take the treatment when the instrument is turned on, who is an always-taker and who is a complier.
iv. $\mathbb{E}[Y_i|E_i = 0, Z_i = 0]$. This is the potential outcome for a mix of compliers and never-takers. Again, we are unable to determine who is who.

However, ideally, we are looking for the treated and untreated potential outcomes for compliers, $\mathbb{E}[Y_i^1|C]$ and $\mathbb{E}[Y_i^0|C]$ since the difference between them is the local average treatment effect. The fact that we can identify some individuals as always-takers and some as never-takers (and compute their outcome means) helps us lern about compliers. To obtain $\mathbb{E}[Y_i^1|C]$, we start by expressing object (iii.) as a weighted average of the treated potential outcomes of compliers and always-takers as per the law of total probability:

$$\begin{aligned}\mathbb{E}[Y_i|E_i = 1, Z_i = 1] \quad = \quad & \mathbb{E}[Y_i|E_i = 1, Z_i = 1, AT] \cdot Pr(AT|E_i = 1, Z_i = 1) \\ & + \mathbb{E}[Y_i|E_i = 1, Z_i = 1, C] \cdot Pr(C|E_i = 1, Z_i = 1)\end{aligned} \tag{4}$$

Due to instrument exogeneity, the potential outcome for always-takers is the same regardless of whether the instrument is turned on. Therefore, $\mathbb{E}[Y_i|E_i = 1, Z_i = 1, AT]$ must be equal to object (i.). $Pr(AT|E_i = 1, Z_i = 1)$ can be calculated as $\phi_{NT}/(\phi_{AT} + \phi_C)$, where ϕ_{AT} is the share of always-takers and ϕ_C the share of compliers in the sample. Similarly, $Pr(C|E_i = 1, Z_i = 1)$ can be calculated as $\phi_C/(\phi_{AT} + \phi_C)$. Plugging in and solving for the object of interest, the treated potential outcome for compliers $\mathbb{E}[Y_i|E_i = 1, Z_i = 1, C] = \mathbb{E}[Y_i^1|C]$, we get

$$\mathbb{E}[Y_i^1|C] = \mathbb{E}[Y_i|E_i = 1, Z_i = 1]\frac{\phi_{AT} + \phi_C}{\phi_C} - \mathbb{E}[Y_i^1|AT]\frac{\phi_{AT}}{\phi_C} \tag{5}$$

Crucially, everything on the right-hand side is directly obtainable from the data. Next, we obtain the untreated potential outcome for compliers by starting with object (iv.), the untreated potential outcome for a mix of compliers and never-takers, and performing analogous steps on it:

$$\begin{aligned}
\mathbb{E}[Y_i|E_i=0,Z_i=0] &= \mathbb{E}[Y_i|E_i=0,Z_i=0,NT]\cdot Pr(NT|E_i=0,Z_i=0) \\
&\quad + \mathbb{E}[Y_i|E_i=0,Z_i=0,C]\cdot Pr(C|E_i=0,Z_i=0) \\
&= \mathbb{E}[Y_i|E_i=0,Z_i=1]\cdot\frac{\phi_{NT}}{\phi_{NT}+\phi_C} + \mathbb{E}[Y_i^0|C]\cdot\frac{\phi_C}{\phi_{NT}+\phi_C} \\
\iff \quad \mathbb{E}[Y_i^0|C] &= \mathbb{E}[Y_i|E_i=0,Z_i=0]\,\frac{\phi_{NT}+\phi_C}{\phi_C} - \mathbb{E}[Y_i^0|NT]\,\frac{\phi_{NT}}{\phi_C}
\end{aligned} \tag{6}$$

If the instrument is random, the shares of always-takers are the same for $Z_i = 1$ and $Z_i = 0$, and therefore, in the whole population. Thus, we can calculate $Pr(AT|E_i = 1, Z_i = 1)$ from the known always-takers and use this as the sample share of AT: $Pr(E_i = 1, Z_i = 0) = \phi_{AT}$. The same holds for never-takers. We can obtain $Pr(NT|E_i = 0, Z_i = 0)$ as $Pr(E_i = 0, Z_i = 1) = \phi_{NT}$. Because we assume monotonicity (no defiers), there are only three groups in the data and their shares must add up to one. The complier share is therefore calculated as $\phi_C = 1 - \phi_{AT} - \phi_{NT}$.

I obtain all shares and potential outcomes via regression. For the type shares, I run the following first-stage regression: $E_i = \zeta_0 + \zeta_1 Z_i + c_i + Z_i \times c_i + X'\eta + \kappa$, which corresponds to Eq. 2 but without interacting with G. Here, ζ_1 gives the complier share ϕ_C. I evaluate at the running variable $c_i = 0$ and demean the other control variables. Then ζ_0 gives ϕ_{AT}. Consequently, $\phi_{NT} = 1 - \zeta_0 - \zeta_1$.

To get the sample means (i.) – (iv.) introduced above, I regress the frailty index on a full set of interactions between E_i and Z_i without an intercept:

$$\begin{aligned}
Y_{it} = {} & \lambda_1 \mathbb{1}\{E_i=1\}\,\mathbb{1}\{Z_i=0\} \\
& + \lambda_2 \mathbb{1}\{E_i=1\}\,\mathbb{1}\{Z_i=1\} \\
& + \lambda_3 \mathbb{1}\{E_i=0\}\,\mathbb{1}\{Z_i=0\} \\
& + \lambda_4 \mathbb{1}\{E_i=0\}\,\mathbb{1}\{Z_i=1\} + c_i + Z_i \times c_i + X'\mu + \nu
\end{aligned} \tag{7}$$

Again, evaluated at $c_i = 0$ and demeaning the other control variables means that λ_1 provides the treated mean for always-takers, λ_2 the treated mean for a mix of always-takers and compliers, λ_3 the untreated mean for a mix of never-takers and compliers, and λ_4 the untreated mean for never-takers.

todo: discuss the existence of NT