HCEO WORKING PAPER SERIES

The Causal Effects of Education on Adult Health, Mortality and Income: Evidence from Mendelian Randomization and the Raising of the School Leaving Age

Neil M. Davies
Matt Dickson
George Davey Smith
Frank Windmeijer
G.J. van den Berg

Working Paper

2019-029

04/2019

HUMAN CAPITAL AND
ECONOMIC OPPORTUNITY
GLOBAL WORKING GROUP

The University of Chicago
1126 E. 59th Street Box 107
Chicago IL 60637

www.hceconomics.org

ABSTRACT

The Causal Effects of Education on Adult Health, Mortality and Income: Evidence from Mendelian Randomization and the Raising of the School Leaving Age*

We compare estimates of the effects of education on health and health behaviour using two different instrumental variables in the UK Biobank data. One is based on a conventional natural experiment while the other, known as Mendelian randomization (MR), is based on genetic variants. The natural experiment exploits a compulsory schooling reform in the UK in 1972 which involved raising the minimum school leaving age (RoSLA). MR exploits perturbations of germline genetic variation associated with educational attainment, which occur at conception. It has been widely used in epidemiology and clinical sciences. Under monotonicity, each IV identifies a LATE, with potentially different sets of compliers. The RoSLA affected the amount of education for those at the lower end of the ability distribution whereas MR affects individuals across the entire distribution. We find that estimates using each approach are remarkably congruent for a wide range of health outcomes. Effect sizes of additional years of education thus seem to be similar across the education distribution. Our study corroborates the usefulness of MR as a source of instrumental variation in education.

JEL Classification: H52, I12, I21, I28

Keywords: returns to education, health, instrumental variables, RoSLA, genomic confounding, LATE

Corresponding author:
Neil M. Davies
Medical Research Council Integrative Epidemiology Unit
University of Bristol
BS8 2BN
United Kingdom
E-mail: neil.davies@bristol.ac.uk

* We are grateful to Ian Deary, William Hill and participants in seminars and conferences in Aberdeen, Brighton (Royal Economic Society), Bristol, Copenhagen, UC Dublin, Dundee, Edinburgh and Oxford (Nuffield College) for comments and suggestions. The Medical Research Council (MRC) and the University of Bristol support the MRC Integrative Epidemiology Unit [MC_UU_12013/1, MC_UU_12013/9, MC_UU_00011/1]. The Economics and Social Research Council (ESRC) support NMD via a Future Research Leaders grant [ES/N000757/1]. No funding body has influenced data collection, analysis or its interpretations. This publication is the work of the authors, who serve as the guarantors for the contents of this paper. This work was carried out using the computational facilities of the Advanced Computing Research Centre - http://www.bris.ac.uk/acrc/ and the Research Data Storage Facility of the University of Bristol - http://www.bris.ac.uk/acrc/storage/. This research was conducted using the UK Biobank Resource.

1. Introduction

The positive associations between years of education on the one hand and health, wealth and longevity on the other hand have been consistently recorded across countries and time periods (see e.g. overviews in Galama *et al.*, 2018, Hahn and Truman, 2015, Cutler and Lleras-Muney, 2014, Furnée *et al.*, 2008, and OECD, 2006). These associations may reflect causal effects of education, by affecting health-related behaviours (such as smoking or taking exercise) or by way of the impact of the increased income and labour market opportunities due to higher levels of education. Alternatively, the associations could be driven by unobserved confounders such as socioeconomic and family lifestyle factors and early-life health.

For public policy purposes it is important to understand the extent to which investments in education cause improvements in health. Natural experiments can potentially be used to identify such causal effects of education in the presence of unobserved confounders. Natural experiments exploit institutional features or reforms that influence when people leave school but are not dependent on unobserved confounders and only affect adult health through education (see Angrist and Krueger, 2001). A prime example concerns increases in the legal minimum school leaving age (RoSLA, after "Raising of School Leaving Age"). RoSLA forces some people to remain in school for longer than they would have otherwise chosen, and this tends to increase average education levels in the relevant cohorts. Exposure to RoSLA can then be used as an instrumental variable (IV) for a wide range of later-life outcomes, including health outcomes (see e.g. Clark and Royer, 2013, and Davies *et al.*, 2018).

A limitation of such approaches is that they identify a highly specific treatment parameter: the local average treatment effect (LATE) among the so-called compliers; that is, among those for whom the reform is binding. With RoSLA, the compliers are likely to be those at the lower end of the ability distribution. The estimates using the RoSLA do not necessarily generalize to the effects of education on health outcomes at other points in that distribution. This is particularly relevant for past reforms where the fraction of affected individuals was relatively small.

In this paper, we contrast estimates using a RoSLA design with estimates using a completely different IV approach, known as Mendelian randomization (MR) in epidemiology. This uses specific genetic variants (single nucleotide polymorphisms or SNPs; see below) as instrumental variables for an exposure on an outcome (Davey Smith and Ebrahim 2003, Davey Smith and

Hemani 2014, and Davies, Holmes and Davey Smith 2018). The natural experiment in this case occurs at conception when each child inherits half of each of its parents' genomes. In recent years, this has become widely used in epidemiology (Boef *et al.* 2015). As this approach is as yet rather unknown in economics, we discuss it and its pros and cons in detail (see Section 2 below). In our application, we use a polygenic education score based on the presence of variant forms of genes (alleles) that are known to associate with educational attainment. Movement from one standard deviation below the mean of this score to one standard deviation above the mean is associated with 0.30 additional years of education. Importantly, the differences in education induced by this genetic variation increase attainment across the entire population.

Both RoSLA and Mendelian randomization depend on a range of assumptions (see e.g. Glymour *et al.*, 2012, Davies, Holmes and Davey Smith 2018). As with any IV approach, the candidate instruments must be both relevant and valid. With Mendelian randomization, we are particularly concerned about the threat to validity from so-called dynastic effects, which occur if genotypes associated with education are systematically related to parental behaviours that affect child outcomes in ways that are not channelled through the child's education. It is difficult to rule out *a priori* that the presence of alleles with particularly strong behavioural effects do not simultaneously affect offspring education and also affect the style in which the offspring is raised in the family (Kong *et al.*, 2018). Controlling for early-life conditions, birth weight and breastfeeding exposure allows us to deal with dynastic effects to a certain extent. Towards the end of the paper we discuss the type of data that would need to be collected for a more comprehensive analysis controlling for dynastic effects in Mendelian randomization. Estimates using the RoSLA are unlikely to be affected by dynastic effects. However, candidate IVs based on natural experiments carry their own weaknesses.[1]

The core instrumental variable assumptions are insufficient to point identify a causal effect. A widely used assumption to achieve point identification is monotonicity – that the instrument affects everyone in the same direction. This assumption identifies the effect of the exposure on individuals whose treatment status was affected by the instrument, the local average treatment

[1] Consider for example the usage of distance to college or season of birth as IVs for effects of education. The residential location and the date of birth reflect parental choices which in turn reflect parental attitudes and preferences, and the latter may directly affect the offspring's economic and health outcomes (see e.g. Buckles and Hungerman, 2013). Also, parental awareness of the effect of distance to college or season of birth on education may influence how they teach their own children about health issues (Van den Berg, 2007).

effect (LATE). This assumption is likely to hold for the RoSLA natural experiments as well as for Mendelian randomization. The genetic score used as the instrument in the Mendelian randomization analysis uses SNPs and on an individual basis, there is little evidence that these SNPs have non-monotonic effects on education across the population (see section 3 below).

Epidemiologists have used Mendelian randomization to investigate a range of topics, for example to investigate the positive association between short-sightedness and education (Mountjoy *et al.*, 2018). A rare application of Mendelian randomization in the economics literature is in Von Hinke *et al.* (2016) which estimated effects of fat mass on academic performance and blood pressure. Böckerman *et al.* (2018) includes a brief overview of existing Mendelian randomization studies by economics researchers.

We use novel data from the UK Biobank which contains a wide range of health outcomes as well as genome-wide data. In particular, the Biobank includes questionnaire measures of health behaviours and clinical and self-reported measures of health. It is linked to national cancer registries and mortality records, providing precise administrative data on these outcomes. We use the 74 genetic variants detected in the Genome-Wide Association Study (GWAS) on educational attainment (Okbay *et al.*, 2016) to construct a weighted genetic score.[2] Many of the Biobank participants were in cohorts affected by the 1972 compulsory increase of the minimum school-leaving leaving age from 15 to 16. This reform affected individuals born from September 1957 onwards and increased average schooling by 0.25 years, which is similar to the increase in education association with a movement from one standard deviation below to one standard deviation above the mean genetic score.

Triangulating the results from both IV approaches is informative on the merits of either of the approaches. As it turns out, both approaches provide evidence that education causally reduces the risk of hypertension, diabetes, stroke and heart attack, as well as increased grip strength (a measure of health in ageing) and reduced BMI. We also find impacts on some health behaviours: reducing prevalence of smoking and hours watching television but increasing alcohol consumption. We demonstrate that our instrumental variables are not correlated with other allele scores for a vast range of other traits that may affect later health.

[2] We do not use more recent GWAS of educational attainment (Lee *et al.* 2018), because this used the UK Biobank and so could introduce weak instrument bias (Burgess *et al.* 2016).

The results based on the two approaches are congruent for most of the health outcomes. Indeed, the magnitudes of the causal effects are similar, despite each instrument identifying the effects in a different set of compliers. This suggests that the effects of education on health may be similar across the education distribution. This result is of importance for policy but also for the methodology of studying causal effects of education on health. In particular, our study suggests that Mendelian randomization is a potentially useful source of instrumental variation in education. In our setting, it allows for an extrapolation of effects among compliers in the RoSLA setting to other segments of the education distribution.

Our paper is organized as follows. Section 2 explains the concept of Mendelian randomization and its application as an IV for education. In this section we also discuss the RoSLA. We briefly discuss relevant existing empirical studies that exploit the 1972 RoSLA as an instrument, and we discuss the two studies that use Mendelian randomization as an instrument for education, both with data from Finland (Böckerman *et al.*, 2017, and Viinikainen *et al.*, 2018). Section 3 describes the data. Section 4 presents the results. Section 5 concludes. Supplementary tables and figures are in an extensive appendix.

2. Empirical Strategy

We use instrumental variable estimators to estimate the effects of additional schooling on each of the income, health and health-behaviour outcomes. As with the recent literature, our first approach is to use a change in the minimum school-leaving age to identify the causal effect of schooling. The Raising of the School Leaving Age Order passed in March 1972[3], required individuals in England and Wales to remain in school until the end of the academic year in which they turned 16, a one-year increase from the previous minimum age of 15. The change came into effect from 1st September 1972 and affected all individuals turning 15 from 1st September 1972 onwards i.e. those born from 1st September 1957 onwards. As children in the UK begin school in the academic year in which they turn 5, this change increased the minimum number of years of schooling from 10 to 11. The increased minimum school leaving age was supported by new buildings and the additional year for those extra students who previously were not staying on for it was part of a

[3] See: www.legislation.gov.uk/uksi/1972/444/pdfs/uksi_19720444_en.pdf

new five-year secondary education curriculum for all students, (i.e. it was not a remedial year for those who would otherwise drop out).[4]

Figure 1 shows the impact of this reform on the schooling distribution by quarter of birth. The passing of the new law (indicated by the vertical black line) is associated with a sharp drop, down to approximately zero, in the proportion of each cohort gaining 10 years of education or fewer. The appearance of some non-compliance with the new minimum in the figure is due to the structure of the academic year in the UK: individuals are assigned to a school cohort depending on birth date, with 1st September the key assignment date. As such, when some individuals born in the third quarter of the year (those born in July and August) officially leave school in the June of their final school year, it is before they have had their birthday that year. Our measure of years of schooling is derived from the age when the individual reports leaving school, therefore if individuals born in July/August report the age on the day that they left school rather than their age at the end of that school year (i.e. end of August) it will give the incorrect impression that they have not completed the compulsory years of schooling even though they have.

To more formally establish the impact of this reform on educational attainment we estimate the effect of the 1972 RoSLA reform on remaining in school beyond the age of 15. It has been shown in the literature (see Chevalier *et al.*, 2004, for example) that the reform induced some individuals to remain in school an additional year but there was very little increase beyond this. As such the increase in the proportion remaining beyond 15 is approximately equal to the increase in the average number of years of schooling. As ever with instrumental variables, the suitability of the instrument depends on satisfying the relevance and validity criteria. For the raising of the school-leaving age the relevance assumption will hold provided participants who attended school after the minimum school leaving age was increased stayed in school for longer on average.

We demonstrate that the instrument is relevant and strong via the first stage regressions reported in section 4. The RoSLA will be a valid instrument as long as it is independent of potential treatments and outcomes and is excludable from the structural equation i.e. as long as the reform did not directly affect the outcomes and is uncorrelated with any unobservables that affect the outcomes. We can demonstrate balancing tests to show that there are few detectable differences between the cohorts pre and post-RoSLA (see section 4), moreover because parents could not

[4] More of the historical context and detail of the 1972 RoSLA can be found in McCulloch, *et al.* (2012).

have anticipated the RoSLA reform (the exact implementation date was announced only months before coming into effect) it is highly unlikely to be associated with family background factors that could confound the association of education and later outcomes. Similarly, the decision to enact the reform was not taken in light of anything to do with the particular cohorts affected and as such assignment to the 'treatment' is as good as random. The validity of the RoSLA instrument could be questioned if the reform also affected the labour market around the time the participants entered the workforce, though other papers investigating this have shown that this is not the case (see Buscha and Dickson, 2015).

Since the seminal study of Lleras-Muney (2005), the use of compulsory schooling reforms to study causal effects of education on health has burgeoned (see Galama *et al.*, 2018, for an overview). A number of studies use the 1972 RoSLA for this purpose (Clark and Royer, 2013, Powdthavee, 2010, Jürges *et al.,* 2013, Janke *et al.*, 2018, Avendado *et al.*, 2018, and Davies *et al.*, 2018). These do not all focus on the same health outcomes, and they sometimes differ in terms of findings. To some extent the latter may be attributed to differences in the data sources, their sampling dates and sample sizes, the character of the outcome variables (self-reported or clinically determined) and the empirical strategy. For example, smaller studies tended to use larger regression discontinuity bandwidth thus including a greater number of cohorts before and after the reform. See Davies *et al.* (2018) for a detailed discussion. Additionally, our data is more recent than other data sources, therefore the participants are older and sicker thus for some of the binary outcomes (e.g. coronary heart disease) there are more events and thus greater statistical power.

Our second approach is to use genetic variants (alleles) associated with education to create an instrumental variable for education. This strategy exploits the natural experiment that occurs at conception – when each child inherits half of each of their parents' genomes. This process means that at each point (loci) at which DNA varies between people there is a 50% chance of inheriting one or other of each parents' alleles. These variants differ in terms of the chemical bases (**A**denine, **C**ytosine, **G**uanine, and **T**hymine) paired together. These loci where DNA varies between people are called polymorphisms and the most studied form of polymorphisms are Single Nucleotide Polymorphisms. This is a point in the genome in which just one of the chemical base pairs differ. The 2016 educational attainment Genome Wide Association Study (GWAS) of Okbay *et al.* reported the association of 8,259,394 genetic variants and the years of education in a meta-analysis of 64 studies, which did not include the UK Biobank. Okbay *et al.* identified 74 SNPs that associated with educational attainment ($p<5x10^{-08}$) after correcting for multiple testing. On

average, at each of these 74 SNPs, people with a particular allele are likely to have higher educational attainment than another individual with the alternative allele.

All of the eligible participants in our data provided a blood sample which was used to extract DNA.[5] Using the extracted genetic information allowed the genotyping of around 800,000 SNPs for each participant. Exploiting this information, we were able to construct an 'education allele score' for each individual determined by the presence or not of the 74 SNPs which were associated with years of education in the Okbay *et al.* discovery sample. The allele score is the sum of the number of education increasing alleles for each participant, with the contribution of each allele to the score weighted by the size of the coefficient reported by the GWAS. Combining many alleles into a single score mitigates the danger of many weak instruments bias (Bound *et al.* 1995). The allele score has a mean of 0.325 and standard deviation of 0.10, and represents the known effects of genetic variants on educational attainment.

Economists rarely use Mendelian randomisation and as such it is important to highlight the conditions under which Mendelian randomization can be used in an instrumental variables strategy seeking to identify the causal effect of the treatment (in our case education) on the outcomes of interest. Von Hinke *et al.* (2016) formally set out these conditions, relating them to the standard assumptions necessary for a suitable instrument in the econometrics literature. We refer the reader to von Hinke *et al.* (2016) for a more in-depth discussion of the use of genetic markers as instrumental variables along with more details on genetics. In brief, these assumptions are:

1. Relevance: there is a non-zero effect of the instrument on the treatment variable.
2. Validity, which comprises:
    a) Independence: the instrument is independent of all potential treatments and potential outcomes i.e. it is as good as randomly assigned.
    b) Exclusion: the exclusion restriction requires that there is no effect of the instrument on the outcome variable that is not mediated via the treatment variable (in our case education).
3. Monotonicity: the potential value of the treatment variable is at least as high when the instrument takes value z' as it is when the instrument has value z, or vice versa, that the potential value of the treatment variable when the instrument takes the value z' is equal to or lower than its potential value when the instrument takes the value z, for all individuals.

[5] These were extracted using the Axiom and BiLEVE genome-wide arrays. For more information on this see the Appendix.

In other words, whether positive or negative, the impact of the instrument on the treatment variable is *in the same direction* for all individuals.

It is worth highlighting the threats to identification relevant to each of these assumptions in our Mendelian randomization context of using alleles associated with educational attainment to construct an instrument for education, aiming to uncover the causal impact of education on later health behaviours and outcomes.

With respect to the first assumption, for any candidate instrument it is important to demonstrate its relevance. As detailed above, the allele score is constructed from 74 SNPs that associated with educational attainment at genome-wide significance levels ($p<5x10^{-08}$) in the discovery sample of the educational attainment GWAS (see Okbay *et al.*, 2016). The UK Biobank used in our study was not included in the Okbay study's discovery sample. These alleles used to construct the instrument have been shown to be robustly associated with educational attainment at the population level and as such there is a firm theoretical basis for their use as an instrument for education. We demonstrate empirically that the polygenic education score associates with educational attainment in the UK Biobank in section 4 where we report on the first stage regressions. Table 4 shows that the polygenic score strongly associates with educational attainment and exceeds the usual thresholds for weak instruments, with partial F-statistics ranging from 284 to 1101 depending on the health outcome of interest. Moreover, the change in years of education induced by the instrument is economically significant: moving from one standard deviation below the mean of the education allele score to one standard deviation above is associated with an additional 0.30 years of schooling. Another way to put it is that a unit increase in the allele score was associated with an additional 1.45 years of education, which is even greater than the treatment induced by the RoSLA, for those who were bound by the reform. Therefore, both instruments induce non-trivial increases in education.

In terms of the first aspect of validity i.e. independence, Mendelian randomization is analogous to a randomised experiment in which allocation to the treatment group is randomly assigned over all eligible individuals in the population. This randomization at conception means that at each SNP, alleles are likely to be independent of the environment conditional on parental genotype (Davey Smith *et al.* 2007). However, very few large studies have genetic data on parental genotype. In samples of unrelated individuals, it is possible for there to be differences in allele frequency across the population (e.g. between the North and South of the United Kingdom, see Haworth *et al.*

2018). If these differences also associate with differences in the outcomes, then the SNP-outcome association may be biased. In population genetics this is known as 'population stratification'. We deal with this issue of population stratification by controlling for the first 10 principal components of genetic variation. We assume that the alleles related to education are independent of confounders conditional on the principal components.

The allocation of genes is random at the family trio level (from two parents, to their off-spring). Therefore, the most robust identification for a Mendelian randomization study comes from comparison of mother-father-offspring trios or biological sibling pairs. Mendelian randomization assumes that social class, income and all other socioeconomic factors are balanced across genotypes such that the value of the instrument assigned is not related to these characteristics that may also affect the treatment and the outcome. The recent literature exploring the Mendelian randomization methodology in economics (see von Hinke *et al.*, 2016, and references therein) suggests that, at population level, genetic variants are largely unrelated to many socio-economic and behavioural characteristics that could confound estimates (once we have conditioned on the principal components of population stratification). As such, observational studies using genetic information from just one individual per family *should not* be confounded by relationships between genotype and socio-economic characteristics.

The presence of assortative mating based on characteristics that have particular genes associated with them could violate the independence assumption required for unbiased estimation, for example if more educated women partner with taller men. It is for this reason that within family identification is the cleanest way in which to operationalise the Mendelian randomization technique. However, even when observing only one individual per family, Mendelian randomization is valid if the assumption can be maintained that genotypes are unrelated to other characteristics that may affect the outcomes of interest, something that as noted above, the literature suggests *is* likely to hold.

We can partially explore this by testing the extent to which the distributions of observable characteristics are similar for different values of the instrument. Given that any observed characteristic could in theory be an outcome resulting from the treatment – i.e. in our case a result of the instrument working through its effect on education – we need to test for differences in pre-treatment characteristics. Figures 3 and 4, discussed in detail in section 4, report the balance of pre-treatment characteristics between those of different genotypes. There is some suggestion that

the genotypes relating to educational attainment are correlated with maternal (non)smoking during pregnancy, birthweight, and with factors relating to the geographic location of birth (closer to London) and the deprivation of birth area. However, in all cases, the correlation between these characteristics and the genotype is of similar magnitude to the correlation between these characteristics and an indicator for being affected by the 1972 RoSLA. Nevertheless, as a sensitivity analysis we control for these observable pre-treatment characteristics in our regression estimates.

The second part of the validity assumption, the exclusion restriction, requires that the instrument has no other effect on the outcomes except via their effect on education. This assumption could be violated in a number of ways. Firstly, if parents' behaviour is affected by the genotypes we are using as instruments and these behaviours affect their offspring's outcomes, then this can result in biased estimates of the effect of education. This is because if *as a result of their education*, individuals with higher levels of education parent in a different way to those with lower levels and this impacts their child's health, this will create a link between the child's genotype (inherited from their parents' genotypes) and the outcome, that does not work through the child's treatment (education). Similarly, there is a risk that parental behaviours that are not caused by their genotype but are correlated with it, affect child health by affecting child behaviours. For example, if people who choose more education are also the types who choose to exercise more (not because of education but because of other preferences), and they pass this behaviour on to their children, there will be a link between the child's genotype and their health outcomes that does not work through their education. Again, to some extent we can address this by examining the correlation between education variants and SNPs associated with other characteristics which may drive parenting behaviour. Figures 3 and 4 show that there is little evidence of correlations between the SNPs known to associate with educations and polygenic risk scores for many socio-economic and neuropsychiatric characteristics.

In general, since education is a treatment with heterogenous effects across the population, in most cases it is not possible to demonstrate that the exclusion restriction holds. Studies instead rely on economic theory and intuition to make the case for the validity of their instrument. In contrast, Mendelian randomization relies on biological and genetic rather than economic theory. Mendel's second law states that the inheritance of one genetic trait is independent of the inheritance of another trait. Therefore, the exclusion restriction is unlikely to be violated since the inheritance of

the alleles associated with educational attainment is unlikely to be related to inheritance of alleles related to other aspects of health (e.g. risk variants for coronary heart disease).[6]

In a model in which we expect there to be heterogeneous treatment effects, to be able to interpret the results as 'local average treatment effects' requires the final assumption: monotonicity. For the genetic instrument, this entails that for each individual in the population, putting an education increasing allele in place of one not associated with education would either increase or have no effect on their education. As we only ever observe an individual's realised genotype, we cannot prove monotonicity. However, there is no evidence to date that any of these SNPs have the opposite effects in any sub-groups (Okbay *et al.*, 2016) and as such we proceed on the assumption that monotonicity holds.

The literature using Mendelian Randomisation to derive IV estimates of causal effects of education on health outcomes is much sparser than is the case for instruments derived from compulsory schooling reforms. Two studies use Finnish data. Viinikainen *et al.* (2018) use the same 74 SNPs that we employ to derive a polygenic risk score for educational attainment and use this to estimate the causal effect of education on depression. Despite a strong association between the instrument and years of education, they found little evidence that education causally affects depressive symptoms: though consistent with the OLS estimates, their Mendelian randomization estimates were very imprecise. Similarly, Böckerman *et al.* (2017) use the 74 SNPs associated with educational attainment to study the causal effect of education on obesity. They find some evidence that education has a causal impact: the Mendelian randomization results suggest one additional year of schooling reduces BMI by 0.84 kg/m^2 though the 95% confidence interval is (-0.07, 1.77). In these Finnish studies, the small sample sizes may have hindered precise estimation, making it difficult to draw firm conclusions.

In operationalising the IV approach, we used two-stage least squares (see Angrist *et al.*, 1996). For the genetic instrument, in the first stage we estimate equation (1) in which the dependent variable, E_{ict}, is years of education and D_{ic} is the polygenic education score for individual *i* from cohort *c*. We include controls for the year and month of birth, gender, the year and month of birth dummies

[6] The phenomenon of 'linkage disequilibrium' could threaten Mendel's second law and see some traits co-inherited, however this is rare for SNPs on different chromosomes and the likelihood of linkage disequilibrium is largely determined by the distance between the loci of the alleles in question. In the covariate balance tests below only use SNPs for different traits that are a sufficient distance from the education SNPs, to investigate the independence assumption. For more information see the Appendix.

interacted with gender, and the first 10 principal components of population genetic stratification. In later analyses we include additional controls to capture background characteristics of the individuals: whether they were breastfed, whether their mother smoked during pregnancy, their birth weight, and the deprivation level of their birth location. All standard errors are robust and clustered by month of birth.

$$E_{ict} = \gamma_0 + \gamma_1 D_{ic} + \boldsymbol{X'_{ic}}\gamma_2 + u_{ict} \qquad (1)$$

In the second stage (equation (2)), we regress our measures of health behaviours, health outcomes and income, H_{ict}, on the fitted values of education derived from the first stage, again including the same control variables in the X_{ic} vector as for the first stage.

$$H_{ict} = \beta_0 + \beta_1 \hat{E}_{ict} + \boldsymbol{X'_{ic}}\beta_2 + \varepsilon_{ict} \qquad (2)$$

We use all of the cohorts in the data, dealing with the effects of age by controlling for year and month of birth.

For the RoSLA instrument, in the first stage we estimate equation (1) with the dependent variable, E_{ict}, being a dummy variable for remaining in school beyond age 15 for participant *i* of birth cohort *c,* at time *t*. The instrument, D_{ic}, is a dummy variable equal to one if the participant was a member of the cohort affected by the reform, and equal to zero if they were not affected.

The UK Biobank data is sufficiently large that for the RoSLA instrument, we can focus attention on just the two school cohorts born in the 12-months before and the 12-months after 1 September 1957, effectively calculating a mean difference estimate but correcting for gender and month of birth effects. In line with Davies *et al.* (2018) we then implement a difference-in-difference process to remove the average cohort-on-cohort differences in health outcomes that we observe between contiguous cohorts due to the effects of ageing. We estimated these average cohort-on-cohort differences using each of the five non-overlapping pairs of cohorts in the 10 years before and the 10 years after the reform. Within each of these cohort pairs the school leaving age is the same, and so any observed health differences between cohorts cannot be due to school leaving law differences but rather may be due to other factors such as the age difference between cohorts observed at a single time point in the UK Biobank. We calculated the average difference between cohorts pooling the estimates for these 10 cohort pairs and then remove this from the difference

between the cohorts immediately either side of the reform to leave the estimated treatment effect of the RoSLA. We estimated this difference and its standard error using the method described by Altman and Bland (2003).

While for the RoSLA we defined the treatment as remaining in school after the age of 15, for Mendelian randomization we used a continuous measure of years of education defined by the International Standard Classification of Education (ISCED), as used in Okbay *et al.* 2016 (more details below). Though the treatments differ slightly between the two instruments, the fact that the 1972 RoSLA induced those bound by the reform to remain in school for one additional year and very few remained any longer than that, means that the treatment effect of remaining in school beyond 15 translates to attaining one additional year of schooling. As such, the increase in the proportion of the cohort remaining in school beyond age 15 is equivalent to the increase in the average years of schooling for the affected cohorts, making the return to remaining in school beyond 15 approximately the return to one additional year of schooling i.e. the same scale as the Mendelian randomisation instrument estimates.

3. Data

Our data comes from the UK Biobank project. This project originally invited 9.2 million people aged between 40 and 69 to attend 23 centres across Great Britain. Of those invited, 503,325 (5.5%) were recruited to the study between 2006 and 2010. Figure 2 illustrates the inclusion/exclusion of individuals to our estimation sample of 315,436.

The participants were asked if they had a college or university degree and if they did not have a degree, they were asked what age they left full-time education. For RoSLA, we assumed that individuals who reported that they left school at age 15 or younger and who did not have a degree, left education prior to the age of 16, while everyone else was assumed to have left education after the age of 15. The participants also reported their highest qualification; we used this to derive a continuous measure of years of education based on the International Standard Classification of Education, which is the internationally harmonized measure used in the education GWAS for education. On average the UK Biobank participants were more educated that the British population: in the Biobank, 41.0% have a degree or equivalent, 64.0% have any post-16 education and 82.1% have at least one academic qualification. The corresponding figures from the 2011 UK

Census for individuals aged 40-70 are 27.9%, 61.8% and 76.5% respectively.[7] However, we use inverse probability weighting to make our sample match the population education distribution, and all of our results are robust to using this weighting or not (Hughes *et al.*, 2017).

With regard to health outcomes, participants were asked whether they had ever been diagnosed by a doctor with the following health conditions: high blood pressure, stroke, diabetes, or heart attack. They were asked if they had ever had a whole week where they felt depressed or down. The diagnoses of cancer and information about the death of the participants were taken from linkage to national cancer and mortality registries.[8] In addition to self-reports of health, a number of measurements were taken during participants' visits to a UK Biobank assessment centre, specifically: height and BMI, two measures of diastolic and systolic blood pressure (recorded via an electronic blood pressure monitor, with the measurements taken two minutes apart), arterial stiffness (measured using an electronic measure device), grip strength (measured in kilos using a hydraulic hand dynamometer). We residualized the measures of grip strength and arterial stiffness to control for potential between device heterogeneity. Verbal-numeric reasoning was measured via 13 logic puzzles that the participants had to answer in 2 minutes. Their score is the number of correct answers.

During their assessment centre visit the participants were asked to report their health behaviours. They were asked about how frequently they consumed alcohol. This is coded 6 if they drank every day, 5 for three or four times a week, 4 for once or twice a week, 3 for one to three times a week, 2 for special occasions only, and 1 for never. They were asked if they were a current, ex or never smoker. They were asked how often they vigorously and moderately exercised in a typical week. Finally, they were asked if their pre-tax income was below £18,000; between £18,000 and £30,999; between £31,000 and £50,999; between £52,000 and £100,000; or above £100,000. Participants who did not answer these questions were coded as missing. Table 1 summarises the characteristics of the individuals in the UK Biobank sample that we use.

4. Results

Covariate Balance Tests

[7] See https://www.ons.gov.uk/peoplepopulationandcommunity/culturalidentity/ethnicity/adhocs/006962ct06762011censusagebyhighestlevelofqualificationenglandandwales. Last accessed 09/08/17.

[8] Follow-up for the linked data ended with the last recorded death on 16th February 2014.

To address the concern that our instruments, particularly the polygenic education score, may be associated with other genetic and non-genetic characteristics that might also impact health and related outcomes, we estimated the relationship between each instrument, education and a number of genetic and non-genetic characteristics measured for our sample (Davies et al. 2018). These estimates are constructed from the ratio of the reduced form impact of the instrument on the characteristic of interest to the reduced form impact of the instrument on education (for the binary RoSLA instrument these are Wald estimates). The estimates adjust for month of birth, gender and the first 10 principal components of population stratification. These estimates allow any relationships to be considered on a comparable scale, indicating the 'impact' of a one-year increase in years of schooling induced by the instrument in question on the outcome. The OLS estimates of the impact of an additional year of education on these outcomes are included for comparison. In terms of the genetic, we look at polygenic risk scores for 45 traits that may also affect our outcomes of interest, constructing the scores from SNPs that were associated with each trait (for more information on the precise construction of these scores see the Appendix). In addition, we construct similar estimates for 16 non-genetic pre-determined characteristics relating to circumstances around the time of birth, childhood or family background.

Table 2 contains these estimates, which are also illustrated in Figures 3 and 4. The point estimates capturing the associations between educational attainment, the polygenic education score and the genetic risk scores for the 45 traits are almost all close to zero. For the most part they are similar in magnitude to the estimates derived using the RoSLA instrument, which is important given that we are very confident that being born in a cohort affected by the school leaving age reform is not related to an individual's genetic characteristics. The educational attainment genetic score is weakly associated with polygenic scores for other observed outcomes including bipolar disorder, childhood intelligence, inspection time, simple reaction time and infant head circumference. However, there was little evidence that these were any larger associations than those for the RoSLA. As we would expect, in all cases the estimates derived from the RoSLA instrument were less precise.

With regard to the 16 non-genetic pre-determined characteristics relating to circumstances around the time of birth, childhood or the family background of the individuals, Table 2 and Figure 4 show that the magnitude of any association between the polygenic education score and these outcomes is very small – in all but four cases smaller even than the association between RoSLA and these outcomes. There is evidence that the genetic variants for education are non-randomly

distributed around the UK – with greater prevalence in the south and east, and in general closer to London (for further details see Haworth *et al.,* 2018). There is also some evidence that the educational attainment genetic score associates with birthweight, having been breastfed, being taller than average at age 10 and whether the individual's mother smoked during pregnancy. These associations may have been induced by dynastic effects or assortative mating (see Hartwig *et al.,* 2018). For example, dynastic effects could arise because those with more education associated alleles will by definition have parents with more education associated alleles. If these more educated parents behave differently – smoking less in pregnancy – then this will create an association between the education associated SNPs and maternal smoking during pregnancy. Similarly, assortative mating could be a factor if more educated people partner with taller people, this will induce an association between education variants and height. Nevertheless, in all cases it is important to stress that these associations are weak and as noted there is little evidence that the polygenic education score associations are any stronger than those for RoSLA. In each case we can include these covariates as controls in the model to reduce any bias these associations could otherwise cause our estimates (see below). There is an association between RoSLA and the likelihood that the mother and father of the individual are alive. This is to be expected since at the discontinuity, those affected by RoSLA are on average one year younger than those unaffected and so it is therefore more likely that their parents are still alive.

First stage association of the instruments and education

As we have 25 outcome variables, there are 25 first stage regressions for each instrument, the sample size varying due to different degrees of missing values for the outcome variables. Figure 1 shows pictorially the first stage for the RoSLA instrument, and this drop of around 20 percentage points in the probability of leaving at age 15 is typical of our estimation results. People born from 1st September 1957 onwards and so potentially affected by the 1972 RoSLA were 23.0 (s.e. 0.69) percentage points more likely to remain in school beyond the age of 15. For the Mendelian randomization instrument, each unit increase in the polygenic education score was associated with 1.45 additional years of education (s.e. 0.05).[9] Neither instrument is likely to suffer from weak instrument biases, with the F-statistics for the exclusion of the instrument from the first stage ranging from 788 to 2,206 for the RoSLA instrument depending on the outcome variable, and from 284 to 1,101 for the polygenic education score. We display these F-statistics as a column in the main results table, Table 3. Both instruments induce a sizeable degree of variation in education

[9] These first stage estimates relate to the first stage of the mortality regressions where we have the largest sample size and thus most reliable estimate.

which is a particularly important finding for the polygenic score instrument. While each individual SNP only explains a very small proportion of the variation in education, in aggregate across all 74 SNPs they explain a substantial fraction the variation. This means the identification of the effect of education does not require extrapolation of very small induced variations. Figure 5 illustrates the ways in which each instrument impacts upon the education distribution (following Angrist and Imbens, 1995). The upper panel shows the difference in the cumulative distribution function of age left full-time education for the post-RoSLA cohorts compared with the pre-RoSLA cohorts. The negative difference for ages 15-17 which then flips to a small positive difference for ages 18-20 reflect that the post RoSLA distribution sees a shift in the density leaving at ages 16 and 17 up from age 15 but not much change further up the education distribution. The lower panel shows a series of differences in the cumulative distribution function of age left full-time education, but this time each curve represents the difference between the distribution for the highest (5^{th}) quintile of the polygenic education score and the other quintiles. The pattern is the same whichever quintile is being compared with the top quintile, in each case the lower quintile has greater density at lower leaving ages. This is more pronounced in the lowest versus highest quintile (solid line) and as we move up the quintiles of the polygenic education score, the difference with the top quintile reduces, but the pattern is still the same – more weight in the lower leaving ages the lower the quintile of the polygenic education score. This shows that increasing the polygenic education score shifts the whole distribution of leaving ages rightwards – that is: the increases in education associated with increasing values of the polygenic risk score are not concentrated in just one part of the education distribution, the whole distribution is shifted.

Results – health outcomes, health behaviours and income

Table 3 shows our estimation results capturing the impact of additional schooling on our outcomes of interest, using OLS and the two instrumental variables strategies: the 1972 RoSLA and Mendelian randomization. All of the estimates are also presented graphically in Figure 6.

The observational associations estimated by OLS generally follow the patterns that we would expect from the literature. An additional year of education is associated with lower risk of having had a diagnosis of hypertension, diabetes, having a stroke, having a heart attack, being a current or ever smoker or dying during the follow up of the sample. There is no impact on cancer risk but education is observationally associated with increased risk of a diagnosis of depression. Additional education is also associated with greater grip strength (indicating greater health), lower arterial stiffness, lower BMI, lower blood pressure, and greater fluid intelligence. Education is associated with a higher probability of having a higher income, though with a smaller absolute impact on the

probability in the higher parts of the income distribution (over £52,000 p.a. and over £100,000 p.a.). There was little association with happiness but the lifestyle relationships show an interesting pattern: more educated participants consumed more alcohol and did less moderate exercise, watched fewer hours television and did more vigorous exercise.

Moving to the IV columns, as we would anticipate, the estimates are less precise. For example, with respect to mortality, the Mendelian randomization estimate of a 0.40 percentage point lower risk has a standard error of 0.23 percentage points and so the 95% confidence interval just includes zero. The RoSLA estimate is notably larger at 1.4 pp reduction, though its standard error (0.5 pp) would not rule out the same size estimate as the Mendelian randomization instrument. For morbidity outcomes the two IVs give estimates of a similar magnitude in several cases: for risk of hypertension both IV estimates are a 1.0 percentage point reduction in risk for an additional year of education. For heart attack the estimates are a 1.2 percentage point reduction in risk (MR) and a 0.5 pp reduction (RoSLA). Risk of stroke again sees similar estimates: at 0.5 pp (MR) and 1.1pp (RoSLA). As with the OLS, there is no estimated impact of education on cancer risk using either IV, but unlike the observational associations, neither IV finds an impact of education on depression diagnosis. Diabetes is the one morbidity outcome for which the IV estimates differ more: the Mendelian randomization estimate is a reduction of 1.4 percentage points, whereas the RoSLA estimates a reduction of 3.1 pp for an additional year of education.

Both instrumental variable estimates suggest an additional year of schooling increases grip strength: by 0.42kg for the Mendelian randomization estimate, and by 1.00kg for the RoSLA estimate. However, the IV estimates for arterial stiffness, the other measure of healthy ageing, are very imprecisely estimated. For blood pressure, the OLS estimate is a reduction in both diastolic and systolic blood pressure, by 0.12 and 0.32 mmHg respectively per additional year of education. The Mendelian randomization analysis suggests much larger causal effects (-0.82mmHg diastolic and -1.20 mmHg systolic, both precisely estimated) whereas the RoSLA estimates are imprecise and both suggest an *increase* in blood pressure as a result of additional schooling.

Interestingly both IV estimates of the impact of education on BMI were very similar, -0.71 kg/m^2 for the Mendelian randomization estimate, -0.86 kg/m^2 for the RoSLA IV, both much larger in magnitude than the observational association that an additional year of education reduces BMI by 0.18 kg/m^2. As with the OLS estimate, neither instrument found an impact of additional education on happiness, however for intelligence both IV estimates are precisely estimated and suggest a

positive causal effect of education: the RoSLA IV is +0.46 and the Mendelian randomization +0.93. These are large impacts given the standard deviation of the intelligence measure is 2.10.

With regard to health behaviours, both instrumental variable estimates are very similar for the impact of a year of education on the propensity to be a current or ever smoker – reducing the risk of being a current smoker by approx. 4.5 percentage points and of ever being a smoker by 8.3 percentage points. Similarly, for alcohol consumption and hours watching TV, the IV estimates are in the same direction as the OLS but larger: an additional year of education is associated with 0.07 unit increase in alcohol consumption, whereas the corresponding IV estimates are 0.17 (RoSLA) and 0.19 (MR). For TV watching the IV estimates are -0.38 hours per day (RoSLA) and -0.48 (MR), both precisely estimated. The instrumental variable impacts on exercise are in general small and imprecise.

Estimated by OLS, the impact of an additional year of education on the absolute probability of earning over certain income thresholds follows a pattern whereby the greatest impact is on income over £31,000 p.a. with the impacts on income over £18,000 p.a. and income over £52,000 p.a. smaller and almost the same as each other. The impact on income over £100,000 p.a. is the smallest of all. Interestingly the IV point estimates follow exactly the same pattern, though in almost every case the IV estimates are larger than the OLS in magnitude but less precisely estimated. In each case the two instrumental variable estimates are similar to each other.

One of the outcome variables included in the UK Biobank is adult height. There is a positive association between height and educational attainment in the cross section, though this is likely to be driven by reverse causation from height to education as height is largely determined prior to completing full time education. Interestingly both of the IV estimates also find a positive effect of education on height. The Mendelian randomization results for height suggest that differences in education associated with genetic variants also correlate with differences in height which implies that there may be some influence of dynastic effects/assortative mating on the result. We investigated this in the covariate balance tests in Table 2 where we find a small positive association between educational attainment instrumented using Mendelian randomization and the probability of reporting having been taller than average at age 10. At the same time there is no impact of education (via MR) on comparative body size at age 10 which may suggest that any dynastic/assortative mating pathway is not straightforward.

Robustness

In light of the concerns raised in the empirical strategy section, we test the robustness of our Mendelian randomization estimates in a number of ways: firstly, we test the sensitivity to the inclusion of different sets of control variables. Our main analysis controls for year and month of birth, gender, the interaction of year and month of birth with gender, and the first 10 principal components of population genetic stratification. In later analyses we include additional controls to capture background characteristics of the individuals: whether they were breastfed, whether their mother smoked during pregnancy, their birth weight, and the deprivation level of their birth location. Estimates are displayed visually in Figures A1 and A2 with coefficient estimates reported in Appendix Table A1. Our estimates are extremely robust to the inclusion or exclusion of the different sets of controls, with point estimates and standard errors almost identical to each other in each specification. Only in the case of grip strength does the exclusion of all controls have a notable impact, the point estimate reducing from 0.42kg to 0.14kg and becoming less precise, the 95% confidence interval now including zero. Finally, the results were robust to whether or not we weight the sample to take account of the under-representation of the lower educated in the UK Biobank vis-à-vis the Census figures (see Figures A3 and appendix Table A2).

5. Conclusion

Our paper makes a substantial contribution to the literature on returns to education. We compared estimates of the effect of education on health outcomes using a commonly used policy reform "natural experiment" and a Mendelian randomization approach using genetic instrumental variables. Each method has distinct strengths and limitations, and under monotonicity they estimate the average effects of education in different subpopulations. As such they are complements rather than substitutes. As it turns out, the results using the two sources of identification are remarkably similar. Each approach provides significant evidence that education causally reduces the risk of hypertension, diabetes, stroke, heart attack and over-all mortality, and that it leads to an increased grip strength and lower BMI. Again, in each case, we find impacts on some health behaviours: education reduces prevalence of smoking and the number of hours watching television, but it increases alcohol consumption. Also, in each case, education increases the frequency of moderate exercise. Neither approach finds evidence of effects of education on depression, cancer or arterial stiffness. We find little evidence for correlations of the proposed instrumental variables with other allele scores for a range of other traits that may affect later health.

From a methodological point of view, perhaps the most fascinating implication is that the two approaches provide such remarkably congruent results for almost each measured health outcome. Indeed, the magnitudes of the estimated causal effects are similar, notwithstanding the fact that each instrument identifies a different set of LATE compliers. This can be interpreted as suggestive evidence for effects of education on health to be of a similar magnitude across the distribution of attained education levels. This result is of importance for policy but also for the methodology of studying causal effects of education on health. In particular, our study suggests that Mendelian randomization is a potentially useful source of instrumental variation in education. In our setting, it allows for an extrapolation of effects among compliers in the RoSLA setting to other segments of the education distribution.

Our study has identified some study limitations and hence gives rise to a number of topics for further research. Notably, the estimation results based on Mendelian randomization may be affected by dynastic effects which occur if there are effects of expression of the relevant SNPs in parents to their offspring's health outcomes. In the near future such limitations will be surmountable using large studies with genomic data from related individuals such as siblings and mother-father-offspring trios. These molecular genetically informed studies are likely to transform our understanding of the transmission of human capital from parents to their offspring.

References

Altman, D. and Bland, J. (2003) Interaction revisited: the difference between two estimates, *British Medical Journal*, 326:219.

Angrist, J. and Krueger, A. (2001) Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments, *Journal of Economic Perspectives,* 15(4):69-85.

Angrist, J., and Imbens, G. (1995) Two-stage least squares estimation of average causal effects in models with variable treatment intensity, *Journal of the American Statistical Association*, 90 (430): 431–442.

Angrist, J., Imbens, G., and Rubin, D. (1996) Identification of causal effects using instrumental variables, *Journal of the American Statistical Association*, 91(434):444–55.

Avendano, P., De Coulon, A. and Nafilyan, V. (2017) Does more education always improve mental health? Evidence from a British compulsory schooling reform. HEDG Working Paper series, no. 1710.

Böckerman, P., Viinikainen, J., Pulkki-Råback, L., Hakulinen, C., Lehtimäki, T., Pehkonen, J. and Raitakari, O. (2017) Does higher education protect against obesity? Evidence using Mendelian randomization, *Preventive Medicine*, 101: 195-198.

Böckerman, P., Cawley, J., Viinikainen, J., Lehtimäki, T., Rovio, S., Seppälä, I. Pehkonen, J., Raitakari, O. (2018) The effect of weight on labor market outcomes: An application of genetic instrumental variables, *Health Economics*, available at: https://doi.org/10.1002/hec.3828 (accessed 10 December 2018).

Boef, A., Dekkers, O., and le Cessie, S. (2015) Mendelian randomization studies: a review of the approaches used and the quality of reporting. *International Journal of Epidemiology*. 1;44(2):496–511.

Buckles, K. and Hungerman, D. (2013) Season of birth and later outcomes: olds questions, new answers, *The Review of Economics and Statistics*, 95(3):711-724.

Burgess, S., Timpson, N., Ebrahim, S. and Davey Smith, G. (2015) Mendelian randomization: where are we now and where are we going? *International Journal of Epidemiology*, 44(2): 379-388.

Burgess, S., Davies, N.M., Thompson, S.G., 2016. Bias due to participant overlap in two-sample Mendelian randomization: Burgess et al. Genetic Epidemiology 40, 597–608. https://doi.org/10.1002/gepi.21998

Buscha, F. and Dickson, M. (2015) The Wage Returns to Education over the Life-Cycle: Heterogeneity and the Role of Experience, *IZA* DP no. 9596.

Chevalier, A., Harmon, C., Walker, I. and Zhu, Y. (2004) Does education raise productivity or just reflect it? *The Economic Journal,* 114:F499-F517.

Clark, D. and Royer, H. (2013) The Effect of Education on Adult Mortality and Health: Evidence from Britain, *American Economic Review,* 103:2087-2120.

Clarke, P. and Windmeijer, F. (2012) Instrumental Variable Estimators for Binary Outcomes, *Journal of the American Statistical Association,* 107(500):1638–52.

Cutler, D. and Lleras-Muney, A. (2014) Education and Health: Insights from International Comparisons. In: *Encyclopedia of Health Economics*, Culyer, A. (ed.), Vol 1. San Diego: Elsevier; pp. 232-245.

Davey Smith, G. and Ebrahim, S. (2003) 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease?, International Journal of Epidemiology, 32(1):1-22.

Davey Smith, G. and Hemani, G. (2014) Mendelian randomization: genetic anchors for causal inference in epidemiological studies, *Human Molecular Genetics*, 23(R1):R89-98.

Davey Smith, G., Lawlor, D., Harbord, R., Timpson, N., Day, I., and Ebrahim, S. (2007) Clustered environments and randomized genes: a fundamental distinction between conventional and genetic epidemiology, *PLoS Med*, 4(12): e352.

Davies, N., Dickson, M., Davey Smith, G., van den Berg, G. and Windmeijer, F. (2018) The Causal Effects of Education on Health, Mortality, Cognition, Well Being, and Income: Evidence from the UK Biobank, *Nature Human Behaviour,* 2:117-125.

Davies, N., Holmes, M. and Davey Smith, G. (2018) Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians, *British Medical Journal*, 362:k601.

Delaney, J. and Devereux, P. (2018) More Education, Less Volatility? The Effect of Education on Earnings Volatility over the Life Cycle, *Journal of Labor Economics*, available at https://doi.org/10.1086/698897 (accessed 10 December 2018).

Furnée, C., Groot, W. and van den Brink, H. (2008) The health effects of education: a meta-analysis, *European Journal of Public Health*, 18(4):417-421.

Galama, T. Lleras-Muney, A., and van Kippersluis, H. (2018) The Effect of Education on Health and Mortality: A Review of Experimental and Quasi-Experimental Evidence. *Oxford Research Encyclopedia of Economics and Finance.* DOI: 10.1093/acrefore/9780190625979.013.7

Glymour, M., Tchetgen Tchetgen, E., Robins, J. (2012) Credible Mendelian randomization studies: approaches for evaluating the instrumental variable assumptions. *American Journal of Epidemiology*, 175(4):332-339.

Grenet, J. (2013) Is extending compulsory schooling alone enough to raise earnings? Evidence from French and British compulsory schooling laws, *The Scandinavian Journal of Economics,* 115(1): 176-210.

Hahn, A. and Truman, B. (2015) Education Improves Public Health and Promotes Health Equity, *International Journal of Health Services*, 45 (4):657-678.

Harmon, C. and Walker, I. (1995) Estimates of the economic return to schooling for the United Kingdom, *American Economic Review*, 85(5):1278-1286.

Hartwig F., Davies N., Davey Smith G. (2018) Bias in Mendelian randomization due to assortative mating. *Genet Epidemiol.* Jul 3; 42(7):608-620

Haworth, S., Mitchell, R., Corbin, L., Wade, K., Dudding, T., Budu-Aggrey, A., Carslake, D., Hemani, G., Paternoster, L., Davey Smith, G., Davies, N., Lawson, D., and Timpson, N. (2018) Common genetic variants and health outcomes appear geographically structured in the UK Biobank sample: Old concerns returning and their implications. DOI: 10.1101/294876

Hughes, R., Davies, N., Davey Smith, G., and Tilling, K. (2017) Selection bias in instrumental variable analyses. bioRxiv. DOI: 10.1101/192237

Janke, K., Johnston, D., Propper, C. and Shields, M. (2018) The Causal Effect of Education on Chronic Health Conditions, *IZA* DP no. 11353.

Jürges, H., Kruk, E. and Reinhold, S. (2013) The effect of compulsory schooling on health—evidence from biomarkers *Journal of Population Economics*, 26(2): 645-672.

Kong A., Thorleifsson, G., Frigge, M., Vilhjalmsson, B., Young, A., Thorgeirsson, T., Benonisdottir, S., Oddsson, A., Halldorsson, B., Masson, G., Gudbjartsson, D., Helgason, A., Bjornsdottir, G., Thorsteinsdottir, U., Stefansson, K. (2018) The nature of nurture: Effects of parental genotypes, *Science*, 359(6374):424-428.

Lee, J.J., 23andMe Research Team, COGENT (Cognitive Genomics Consortium), Social Science Genetic Association Consortium, Wedow, R., Okbay, A., Kong, E., Maghzian, O., Zacher, M., Nguyen-Viet, T.A., Bowers, P., Sidorenko, J., Karlsson Linnér, R., Fontana, M.A., Kundu, T., Lee, C., Li, H., Li, R., Royer, R., Timshel, P.N., Walters, R.K., Willoughby, E.A., Yengo, L., Alver, M., Bao, Y., Clark, D.W., Day, F.R., Furlotte, N.A., Joshi, P.K., Kemper, K.E., Kleinman, A., Langenberg, C., Mägi, R., Trampush, J.W., Verma, S.S., Wu, Y., Lam, M., Zhao, J.H., Zheng, Z., Boardman, J.D., Campbell, H., Freese, J., Harris, K.M., Hayward, C., Herd, P., Kumari, M., Lencz, T., Luan, J., Malhotra, A.K., Metspalu, A., Milani, L., Ong, K.K., Perry, J.R.B., Porteous, D.J., Ritchie, M.D., Smart, M.C., Smith, B.H., Tung, J.Y., Wareham, N.J., Wilson, J.F., Beauchamp, J.P., Conley, D.C., Esko, T., Lehrer, S.F., Magnusson, P.K.E., Oskarsson, S., Pers, T.H., Robinson, M.R., Thom, K., Watson, C., Chabris, C.F., Meyer, M.N., Laibson, D.I., Yang, J., Johannesson, M., Koellinger, P.D., Turley, P., Visscher, P.M., Benjamin, D.J., Cesarini, D., 2018. Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. Nature Genetics 50, 1112–1121. https://doi.org/10.1038/s41588-018-0147-3

Lleras-Muney, A. (2005) The relationship between education and adult mortality in the United States, *The Review of Economic Studies*, 72(1):189-221.

McCulloch, Gary, Steven Cowan, and Tom Woodin. 2012. The British conservative government and the raising of the school leaving age, 1959–1964. *Journal of Education Policy* 27:509-527.

Mountjoy, E., Davies, N., Plotnikov, D., Davey Smith, G., Rodriguez, S., Williams, C., Guggenheim, J., Atan, D. (2018) Education and myopia: assessing the direction of causality by mendelian randomisation, *British Medical Journal,* 361: k2022.

OECD (2006) Measuring the Effects of Education on Health and Civic Engagement Proceedings of the Copenhagen Symposium. Available at: http://www.oecd.org/education/innovation-education/measuringtheeffectsofeducationonhealthandcivicengagement.htm (accessed 10 December 2018)

Okbay, A. *et al.* (2016) Genome-wide association study identifies 74 loci associated with educational attainment. *Nature*, 533:539–542.

Powdthavee, N. (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.

Sniekers, S. et al. (2017) Genome-wide association meta-analysis of 78,308 individuals identifies new loci and genes influencing human intelligence, *Nature Genetics,* 49:1107–1112.

Trampush, J. *et al.* (2017) GWAS meta-analysis reveals novel loci and genetic correlates for general cognitive function: a report from the COGENT consortium, *Molecular Psychiatry*, 22:336–345

Van den Berg, G. (2007) An Economic Analysis of Exclusion Restrictions for Instrumental Variable Estimation, *IZA* DP no. 2585.

Viinikainen, J., Bryson, A., Böckerman, P., Elovainio, M., Pitkänen, N., Pulkki-Råback, L., Lehtimäki, T., Raitakari, O. and Pehkonen, J. (2018) Does education protect against depression? Evidence from the Young Finns Study using Mendelian randomization, *Preventive Medicine*, 115: 134-139

von Hinke, S., Davey Smith, G., Lawlor, D., Propper, C., Windmeijer, F. (2016) Genetic markers as instrumental variables, *Journal of Health Economics*, 45:131-148.

Zabaneh, D. *et al.* (2018) A genome-wide association study for extremely high intelligence, *Molecular Psychiatry*, 23:1226-1232

Table 1 – characteristics of UK BioBank participants

	N	Proportion	Count/ Std. Dev.
Male	315,436	0.46	146,571
Year of birth	315,436	1951	8
Age left education	315,436	18.19	3

Outcomes	N	Proportion	Count
Hypertension	307,496	0.25	76,638
Diabetes	313,766	0.04	13,877
Stroke	314,978	0.02	4,772
Heart attack	314,978	0.02	7,175
Depression	300,594	0.15	44,283
Cancer	314,152	0.13	40,014
Died	315,436	0.02	5,340
Ever smoked	314,422	0.10	31,259
Currently smoke	314,422	0.45	141,825
Income over £18k	274,617	0.78	215,423
Income over £31k	274,617	0.53	145,685
Income over £52k	274,617	0.27	72,867
Income over £100k	274,617	0.06	15,190
	N	Mean	Std. Dev.
Grip strength (kg)	314,788	0.33	10.89
Arterial stiffness	113,856	0.02	4.07
Height (cm)	314,760	168.92	9.26
BMI (kg/m2)	314,455	27.36	4.75
Diastolic (mmHg)	297,872	82.25	10.12
Systolic (mmHg)	297,871	138.11	18.62
Intelligence (0 to 13)	113,033	6.25	2.10
Happiness (0 to 5 Likert)	114,971	3.45	0.70
Alcohol consumption (0 low to 5 high)	315,239	3.16	1.48
Hours watching television per day	304,230	2.86	1.63
Moderate exercise (days/week)	301,195	3.61	2.33
Vigorous exercise (days/week)	301,440	1.82	1.94

Table 2: Estimates of the 'impact' of an additional year of education on polygeneic scores for 45 traits and 16 non-genetic characteristics

	Observed educational attainment		Education instrument by RoSLA		Education instrumented by Mendelian randomization	
Outcomes						
Polygeneic scores	Coefficient	s.e.	Coefficient	s.e.	Coefficient	s.e.
Height	0.002**	0.000	-0.017	0.020	0.000	0.003
Body mass index	-0.001**	0.000	-0.004	0.009	-0.009	0.003
Body fat percentage	-0.001	0.001	0.046	0.064	-0.025	0.009
Cigarettes smoked per day	-0.002	0.002	-0.139	0.111	-0.016	0.018
Ever vs never smoked	0.000	0.000	-0.009	0.018	-0.011**	0.003
Age of smoking initiation	0.000**	0.000	0.000	0.004	-0.002	0.001
Alcohol dependence	-0.001*	0.001	-0.110	0.067	-0.011	0.010
Birth weight	0.001**	0.000	-0.011	0.011	0.011	0.002
Birth length	0.000**	0.000	-0.008	0.014	0.008	0.002
Infant head circumference	0.000**	0.000	-0.002	0.005	0.020	0.001
Age at menarche	0.001**	0.000	-0.020	0.059	-0.012	0.011
Depressive symptoms	0.000	0.000	-0.012	0.006	-0.008**	0.002
Major depressive disorder	0.000	0.000	-0.023	0.021	-0.002	0.006
Autism	0.001	0.001	0.023	0.049	-0.022	0.010
Schizophrenia	0.000	0.001	0.032	0.046	0.013	0.011
Bipolar disorder	0.003**	0.001	0.030	0.057	0.041**	0.006
Migraine in bipolar disorder	0.000	0.001	0.015	0.063	0.017	0.015
PGC cross-disorder traits	0.001**	0.000	0.011	0.021	0.011*	0.004
Alzheimer's disease	-0.001**	0.000	-0.009	0.032	0.003	0.006
Father's age at death	0.000**	0.000	0.008	0.007	0.000	0.002
Mother's age at death	0.000**	0.000	-0.006	0.008	0.003*	0.002

Agreeableness	0.001	0.001	0.007	0.059	-0.003	0.016
Conscientiousness	0.000	0.001	0.010	0.049	-0.009	0.009
Extraversion	-0.001	0.001	0.140	0.094	-0.005	0.018
Openness to experience	0.000	0.001	0.002	0.104	-0.020	0.013
Neuroticism	0.000**	0.000	0.019	0.011	-0.004	0.002
Internalizing problems	-0.001	0.000	-0.036	0.035	-0.002	0.011
Subjective well being	0.000**	0.000	-0.003	0.004	0.000	0.001
Chronotype	0.000**	0.000	-0.014	0.011	0.008**	0.002
Sleep duration	0.000**	0.000	0.000	0.005	0.004**	0.002
G speed factor	0.000	0.000	-0.026	0.024	0.016	0.010
Symbol search	0.000	0.001	0.002	0.043	-0.114**	0.010
Digit symbol	0.001	0.001	0.024	0.021	0.004	0.007
Inspection time	-0.002*	0.001	-0.003	0.039	0.025	0.013
2-choice reaction time	-0.001*	0.001	-0.005	0.042	0.001	0.009
8-choice reaction time	-0.001*	0.001	0.030	0.049	-0.011	0.008
Simple reaction time	-0.001**	0.000	-0.009	0.014	0.020**	0.006
Childhood intelligence	0.000**	0.000	-0.004	0.009	0.028**	0.003
Cognition Sniekers et al.	0.001**	0.000	0.002	0.006	0.006**	0.001
Cognition Trampush p<5E-07	0.000**	0.000	0.003	0.004	0.002**	0.001
High IQ Zabaneh	0.000**	0.000	-0.007	0.008	0.001	0.002
Omega-3 fatty acids	0.001**	0.000	0.004	0.024	0.001	0.004
Omega-6 fatty acids	0.000	0.000	0.025	0.023	0.000	0.005
Omega-9 and sat. fatty acids	0.001**	0.000	0.012	0.028	-0.018**	0.006
Other PUFA	0.001**	0.000	-0.026	0.021	0.001	0.007
Linoleic acid (LA)	0.000	0.001	0.003	0.027	-0.007	0.007
Mono-unsaturated fatty acids	0.000	0.000	0.019	0.019	-0.023**	0.006
Zinc	-0.001*	0.001	-0.050	0.040	-0.011	0.006

	Observed educational attainment		Education instrument by RoSLA		Education instrumented by Mendelian randomization	
Outcomes						
Non-genetic characteristics	Coefficient	s.e.	Coefficient	s.e.	Coefficient	s.e.
Easting	0.012**	0.001	0.028	0.034	0.068**	0.009
Northing	-0.013**	0.001	-0.061	0.046	-0.068**	0.010
Index of Multiple Deprivation	0.025**	0.001	0.144	0.076	0.114**	0.010
Urban vs. rural	-0.002**	0.000	0.040**	0.013	-0.015**	0.005
Distance from London	-0.011**	0.001	-0.047	0.035	-0.069**	0.007
Birthweight	0.010**	0.002	0.077	0.071	0.061**	0.017
Breastfed	0.001**	0.000	0.034	0.034	0.012	0.007
Mother smoked in pregnancy	-0.008**	0.001	-0.109**	0.036	-0.069**	0.005
Comparative body size age 10	0.009**	0.001	0.062	0.086	0.017	0.012
Comparative height age 10	0.017**	0.002	0.043	0.066	0.081**	0.008
Father alive	0.029**	0.000	0.212**	0.030	0.003	0.006
Mother alive	0.037**	0.000	0.190**	0.017	0.016**	0.006
Number of brothers	-0.047**	0.001	-0.375**	0.099	-0.123**	0.018
Number of sisters	-0.043**	0.001	-0.178*	0.079	-0.114**	0.009

Notes: each coefficient is from a separate regression, in each regression controls are included for month of birth, gender and the first 10 principal components of population stratification.
*$p<0.05$, **$p<0.01$.

Table 3: The effect of one additional year of schooling on morbidity, mortality and socioeconomic outcomes

	OLS			IV: RoSLA				IV: Mendelian randomization			
Outcome	Coefficient	s.e.	N	Coefficient	s.e.	First stage F-statistic	N	Coefficient	s.e.	First stage F-statistic	N
Hypertension	-0.007**	0.000	307496	-0.009*	0.004	2126	21768	-0.010	0.006	1060	307496
Diabetes	-0.003**	0.000	313766	-0.031**	0.006	2187	22049	-0.014**	0.003	1096	313766
Stroke	-0.001**	0.000	314978	-0.011**	0.003	2202	22110	-0.005*	0.002	1101	314978
Heart attack	-0.003**	0.000	314978	-0.005*	0.002	2202	22110	-0.012**	0.003	1101	314978
Depression	0.006**	0.000	300594	-0.014	0.015	2028	21085	0.003	0.005	1040	300594
Cancer	0.000	0.000	314152	-0.001	0.012	2182	22011	0.000	0.005	1092	314152
Died	-0.001**	0.000	315436	-0.014**	0.005	2206	22138	-0.004	0.002	1099	315436
Ever smoked	-0.016**	0.000	314422	-0.084**	0.022	2202	22086	-0.083**	0.008	1094	314422
Currently smoke	-0.011**	0.001	314422	-0.047**	0.011	2202	22086	-0.044**	0.005	1094	314422
Income over £18k	0.039**	0.001	274617	0.058**	0.010	1866	19921	0.094**	0.008	935	274617
Income over £31k	0.046**	0.000	274617	0.129	0.014	1866	19921	0.113**	0.007	935	274617
Income over £52k	0.033**	0.001	274617	0.049*	0.022	1866	19921	0.092**	0.006	935	274617
Income over £100k	0.009**	0.000	274617	0.007	0.011	1866	19921	0.030**	0.003	935	274617
Gripstrength (kg)	0.247**	0.006	314788	1.002**	0.168	2161	21989	0.417**	0.101	1090	314788
Arterial stiffness	-0.057**	0.007	113856	-0.093	0.213	788	8537	-0.038**	0.093	304	113856
Height (cm)	0.277**	0.005	314760	0.334	0.190	2196	22077	0.986**	0.093	1096	314760
BMI (kg/m2)	-0.177**	0.004	314455	-0.858**	0.153	2197	22055	-0.714**	0.074	1094	314455
Diastolic (mmHg)	-0.119**	0.009	297872	0.215	0.424	2116	21494	-0.818**	0.133	1003	297872
Systolic (mmHg)	-0.319**	0.014	297871	1.273*	0.610	2114	21492	-1.203**	0.257	1003	297871
Intelligence (0 to 13)	0.251**	0.003	113033	0.455**	0.124	791	8540	0.925**	0.065	284	113033
Happiness (0 to 5 Likert)	0.000	0.001	114971	-0.010	0.047	807	8626	-0.002	0.018	302	114971
Alcohol consumption (0 low to 5 high)	0.075**	0.001	315239	0.173**	0.055	2204	22123	0.186**	0.021	1100	315239
Hours watching television per day	-0.157**	0.001	304230	-0.380**	0.071	2140	21206	-0.487**	0.026	1005	304230
Moderate exercise (days/week)	-0.020**	0.002	301195	0.160	0.087	2017	21330	-0.098**	0.032	1022	301195

Vigorous exercise (days/week)	0.004**	0.002	301440	0.025	0.058	2011	21379	-0.020	0.029	1022	301440

Notes: Standard errors clustered by month of birth. IV: MR includes controls for the year and month of birth, gender, the year and month of birth dummies interacted with gender, and the first 10 principal components of population genetic stratification. IV: RoSLA includes controls for month of birth and gender. *p<0.05, **p<0.01

Figure 1: The impact of the 1972 Raising of the School Leaving Age (RoSLA)

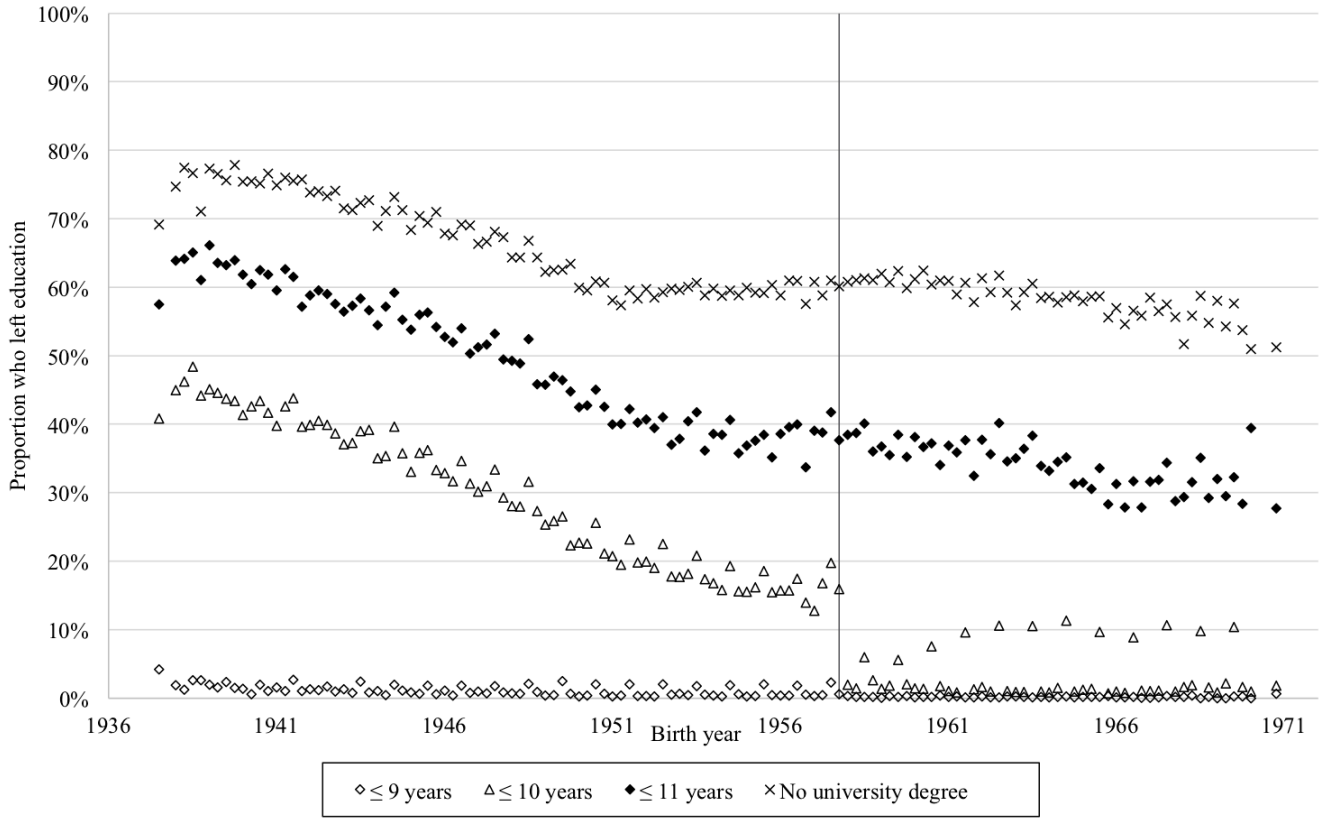


Figure 2: Selection of participants into the study

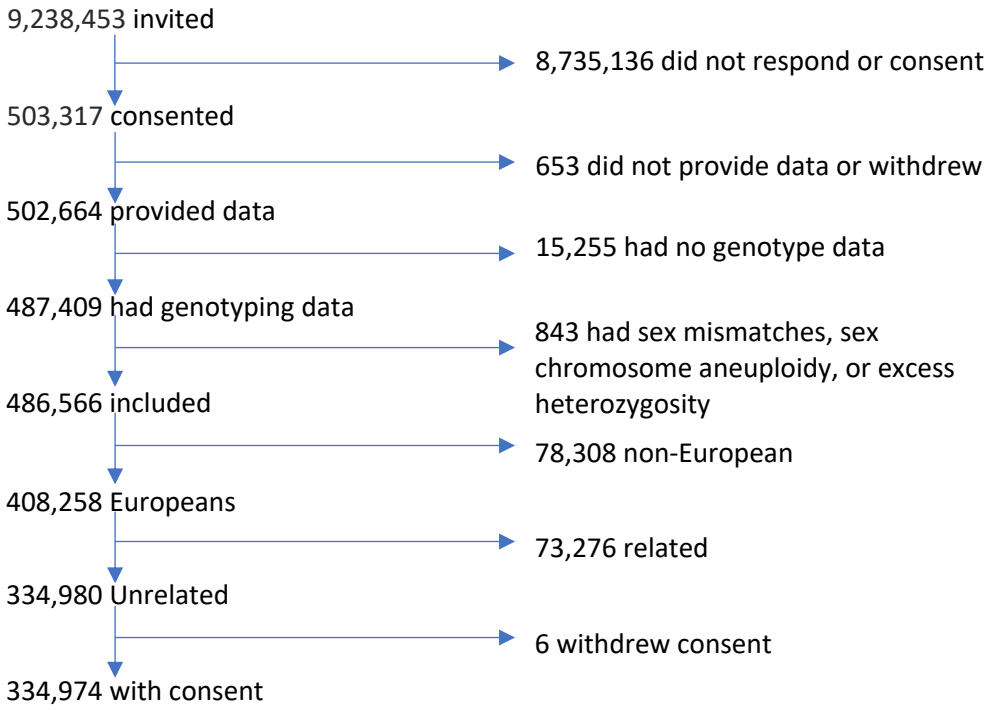
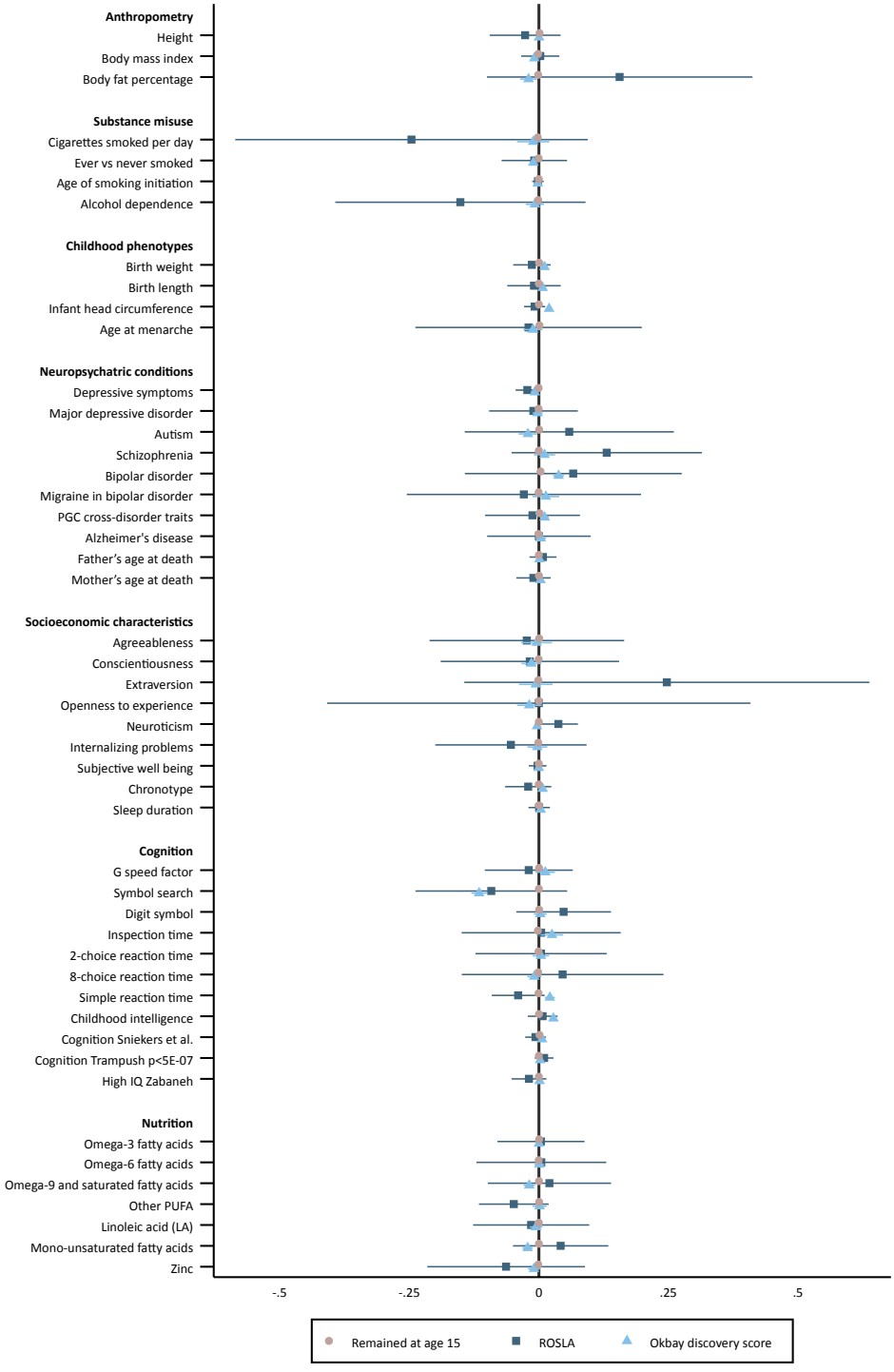


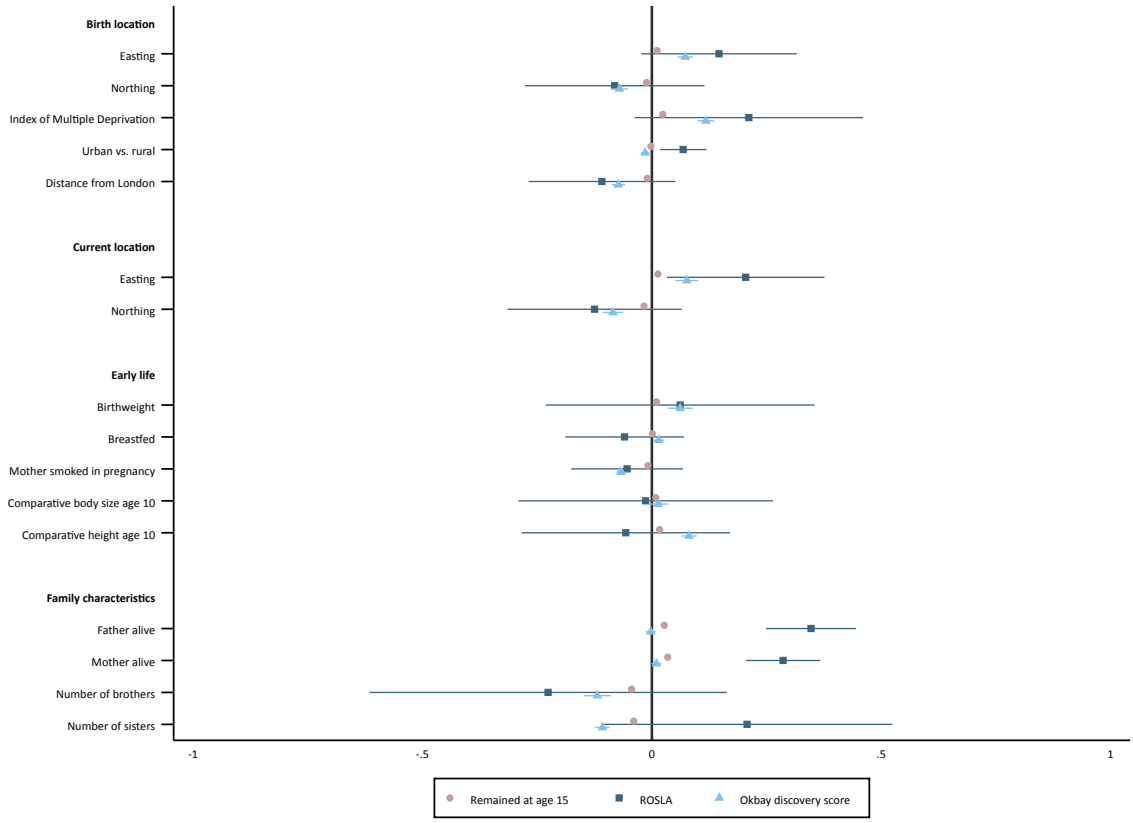
Figure 3: Estimates of the 'impact' of an additional year of education on polygeneic scores for 45 traits comparing observed educational attainment ●, education instrumented by RoSLA ■, and education instrumented by the polygenic education risk score ▲.



Notes: Adjusted for month of birth, sex, and the ten principal components of population stratification. Confidence intervals allowing for clustering by month of birth reported. Sample weighted to adjust for under sampling of less educated.

Figure 4: Estimates of the 'impact' of an additional year of education on 16 non-genetic characteristics comparing observed educational attainment ●, education instrumented by RoSLA ■, and education instrumented by the polygenic education risk score ▲.

.



Notes: Adjusted for month of birth, sex, and the ten principal components of population stratification. Confidence intervals allowing for clustering by month of birth reported. Sample weighted to adjust for under sampling of less educated.

Figure 5: Differences in age left school across the Raising of the School Leaving Age (top) and quintiles of the educational attainment polygenic risk score (bottom)

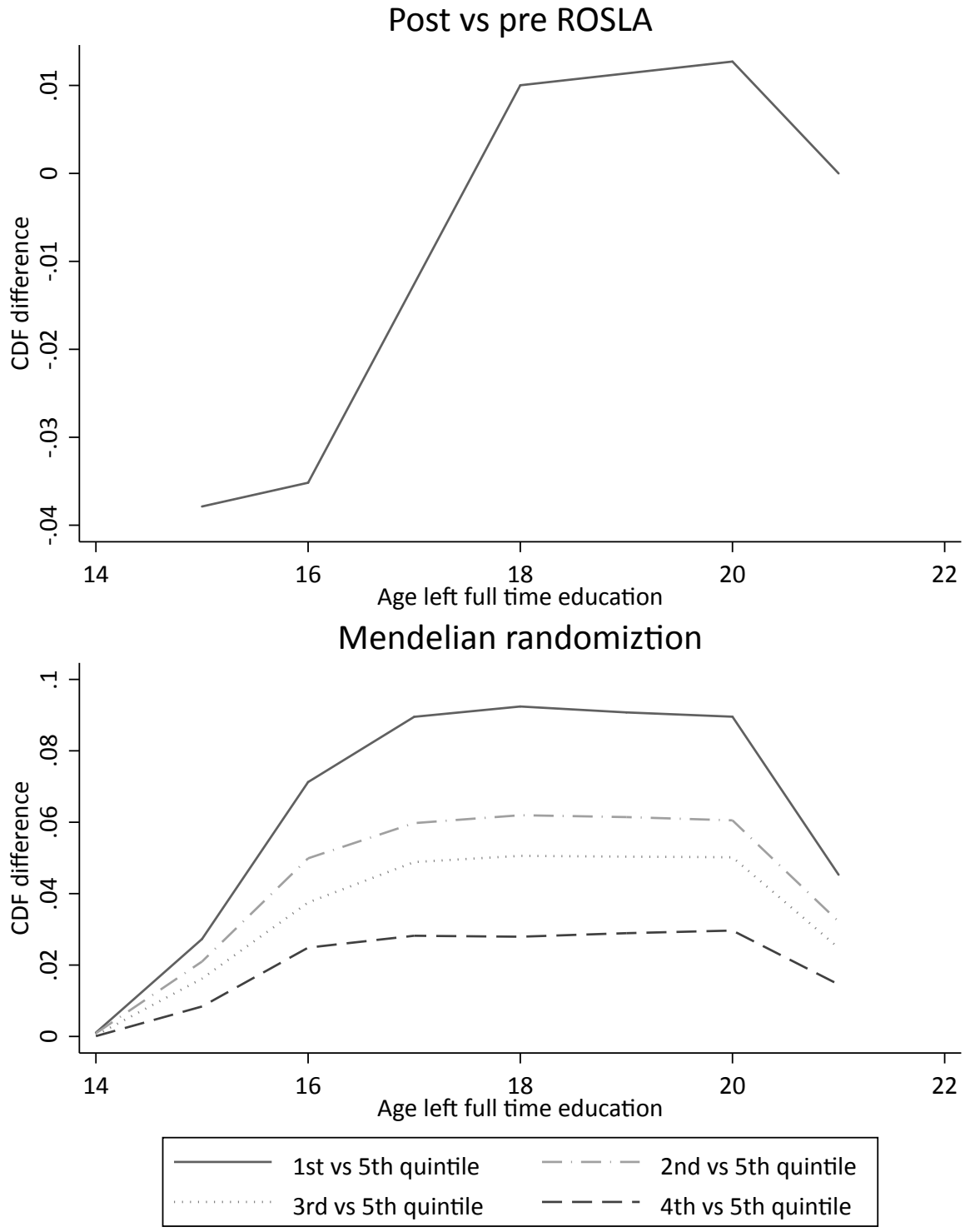
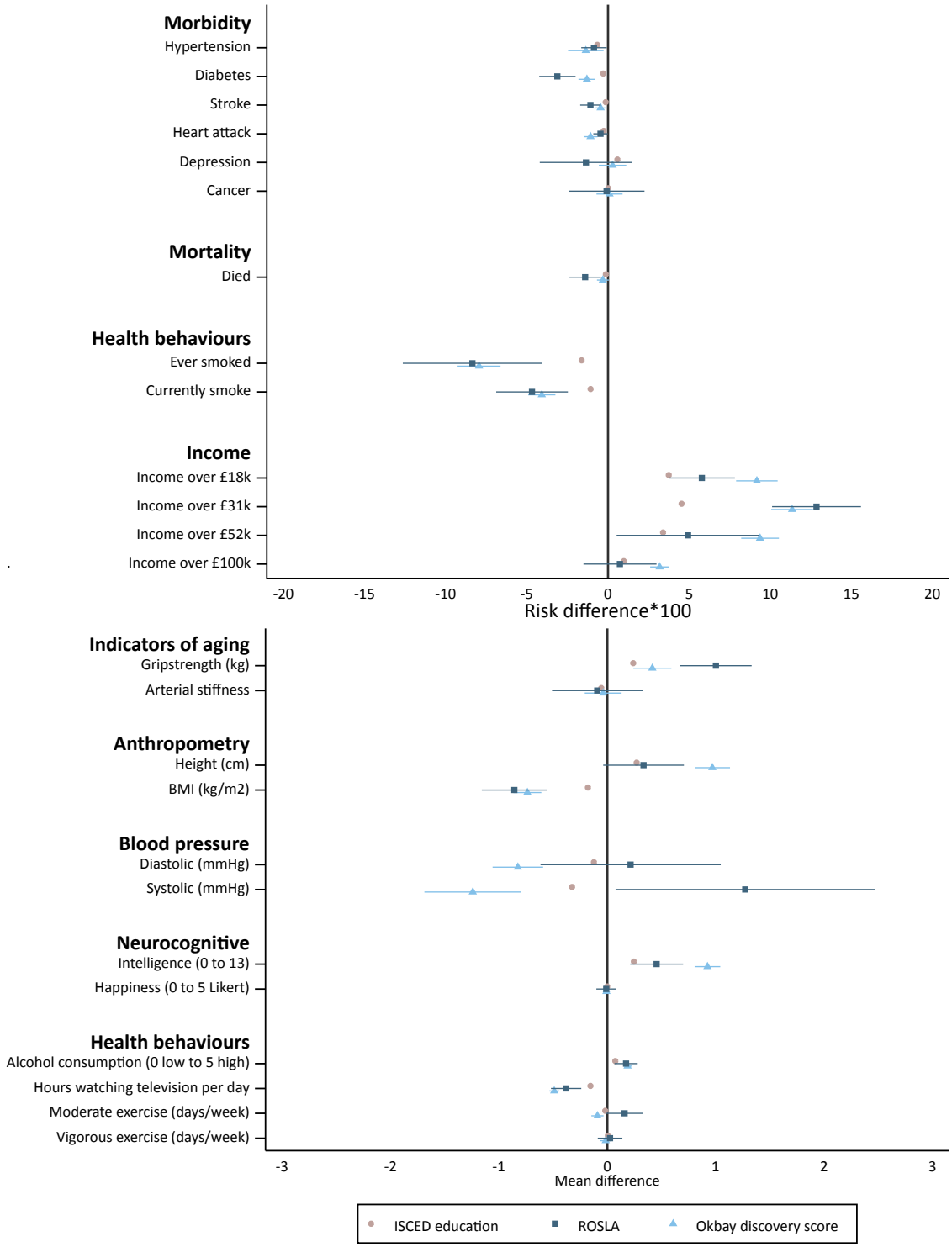


Figure 6: The effect of one additional year of schooling on morbidity, mortality and socioeconomic outcomes, estimated via OLS ●, and instrumenting education using the 1972 Raising of the School Leaving Age ■, and the polygenic educational attainment genetic risk score ▲.



Notes: Adjusted for month and year of birth, sex, and the ten principal components of population stratification. Confidence intervals allowing for clustering by month of birth reported. Sample weighted to adjust for under sampling of less educated.

Appendix

Construction of the polygenic scores for education and other traits

We constructed the scores from extracted SNPs that were associated with each trait at p<5e-5. We used a lower threshold than is usually used for genome-wide significance (p<5e-8) to define the scores because we wanted to maximise the explanatory power of the scores. We constructed the polygenic risk scores from extracted SNPs that were associated with each trait at p<5e-5. We used a lower threshold than is usually used for genome-wide significance (p<5e-8) to define the scores because we wanted to maximise the explanatory power of the scores. Furthermore, it is not possible for the educational attainment genetic score to have pleiotropic effects on the other polygenic scores. We LD pruned the SNPs for each trait using a threshold of r^2>0.001 across a distance of 10,000kb. We excluded SNPs from these scores that were in LD (r^2>0.001) with the 74 SNPs identified as associated with educational attainment at the genome-wide level (p<5e-08) the educational attainment GWAS (Okbay *et al.*, 2016). This resulted in a set of SNPs in independent points in the genome for each trait. We constructed allele scores equal to the sum of the effect alleles for each trait (as per Burgess and Thompson, 2014). The contribution of each SNP to the allele score was weighted by the coefficient reported in the GWAS for that trait. We harmonised the direction of SNP effects between UK Biobank and the GWAS. Finally, we checked for consistency of the allele frequency reported in the GWAS and the UK Biobank data. The allele frequencies were correlated 0.9913, and the maximum difference in allele frequency was 0.091.

Table A1: The effect of one additional year of schooling instrumented using Mendelian Randomisation on morbidity, mortality and socioeconomic outcomes, robustness analysis: controls

	Basic controls (Table 3)		No controls		Full set of controls	
Outcome	Coefficient	s.e.	Coefficient	s.e.	Coefficient	s.e.
Hypertension	-0.010	0.006	-0.012*	0.006	-0.012	0.011
Diabetes	-0.014**	0.003	-0.015**	0.003	-0.018**	0.005
Stroke	-0.005*	0.002	-0.005*	0.002	-0.002	0.003
Heart attack	-0.012**	0.003	-0.013**	0.003	-0.011**	0.004
Depression	0.003	0.005	0.002	0.005	-0.006	0.008
Cancer	0.000	0.005	0.000	0.005	-0.005	0.008
Died	-0.004	0.002	-0.004*	0.002	0.000	0.003
Ever smoked	-0.083**	0.007	-0.084**	0.008	-0.084**	0.012
Currently smoke	-0.044**	0.005	-0.043**	0.005	-0.046**	0.008
Income over £18k	0.094**	0.008	0.093**	0.008	0.078**	0.011
Income over £31k	0.113**	0.007	0.113**	0.007	0.110**	0.013
Income over £52k	0.092**	0.006	0.093**	0.006	0.097**	0.010
Income over £100k	0.030**	0.003	0.030**	0.003	0.038**	0.005
Gripstrength (kg)	0.417**	0.101	0.142	0.154	0.339*	0.166
Arterial stiffness	-0.038**	0.093	-0.085	0.090	-0.284	0.147
Height (cm)	0.986**	0.093	0.709**	0.128	0.961**	0.160
BMI (kg/m2)	-0.714**	0.074	-0.718**	0.074	-0.722**	0.127
Diastolic (mmHg)	-0.818**	0.132	-0.890**	0.136	-0.918**	0.240
Systolic (mmHg)	-1.203**	0.257	-1.324**	0.268	-1.621**	0.470
Intelligence (0 to 13)	0.925**	0.065	0.911**	0.065	0.894**	0.113
Happiness (0 to 5 Likert)	-0.002	0.018	-0.004	0.018	0.032	0.035
Alcohol consumption (0 low to 5 high)	0.186**	0.021	0.174**	0.021	0.142**	0.034
Hours watching television per day	-0.487**	0.026	-0.481**	0.027	-0.478**	0.044

Moderate exercise (days/week)	-0.098**	0.032	-0.103**	0.032	-0.119*	0.055
Vigorous exercise (days/week)	-0.020	0.029	-0.030	0.030	-0.039	0.044

Notes: Basic specification controls for the year and month of birth, gender, the year and month of birth dummies interacted with gender, and the first 10 principal components of population genetic stratification. Full set of controls contains the basic set plus whether the individual was breastfed, whether their mother smoked during pregnancy, their birth weight, and the deprivation level of their birth location. Standard errors clustered by month of birth. Sample weighted to adjust for under sampling of less educated. *p<0.05, **p<0.01

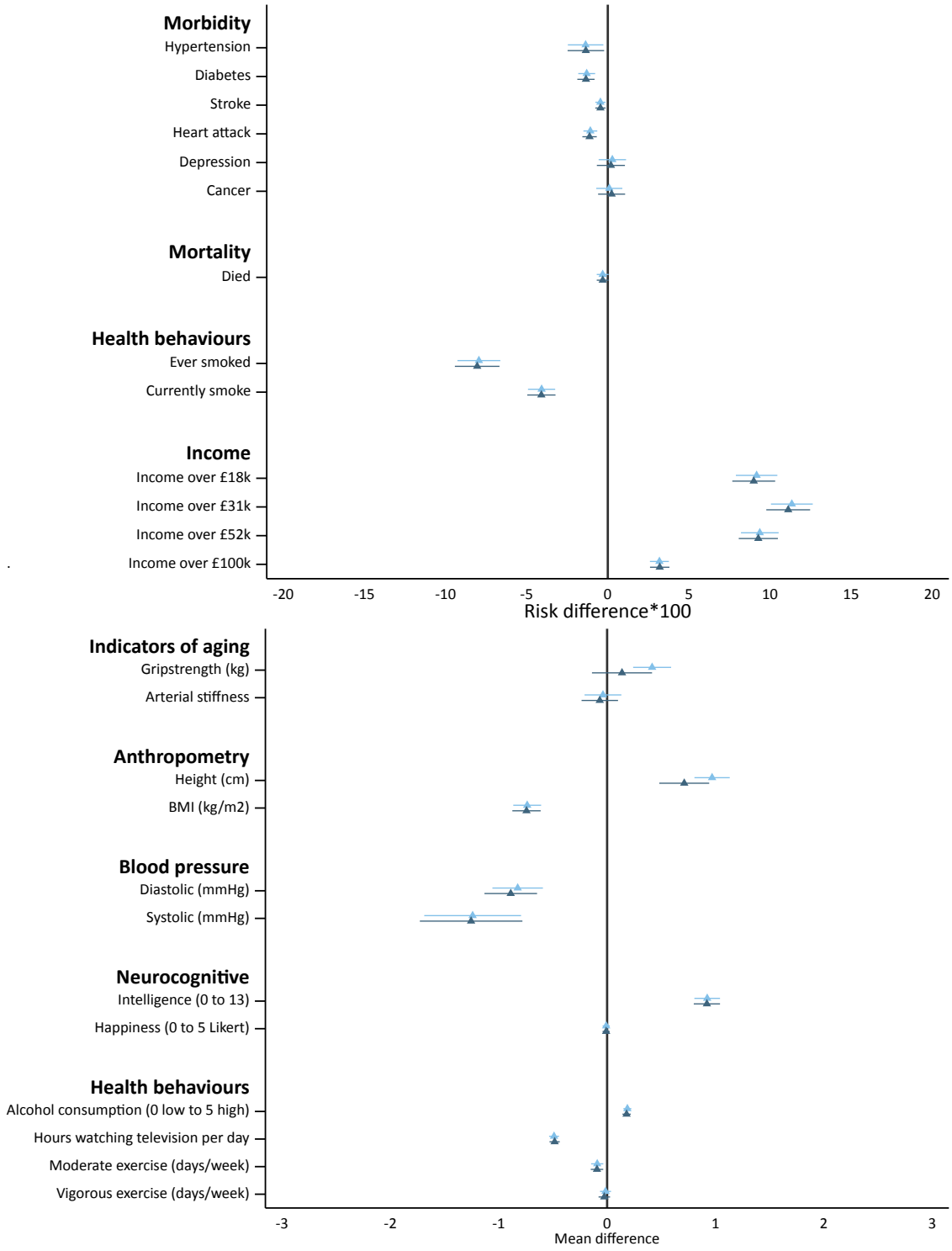
Table A2: The effect of one additional year of schooling instrumented using Mendelian Randomisation on morbidity, mortality and socioeconomic outcomes, robustness analysis: weighting

	Weighted results (Table 4)		Unweighted results	
Outcome	Coefficient	s.e.	Coefficient	s.e.
Hypertension	-0.010	0.006	-0.018**	0.005
Diabetes	-0.014**	0.003	-0.012**	0.003
Stroke	-0.005*	0.002	-0.004*	0.002
Heart attack	-0.012**	0.003	-0.009**	0.002
Depression	0.003	0.005	0.002	0.004
Cancer	0.000	0.005	0.002	0.004
Died	-0.004	0.002	-0.002	0.002
Ever smoked	-0.083**	0.007	-0.077**	0.007
Currently smoke	-0.044**	0.005	-0.038**	0.004
Income over £18k	0.094**	0.008	0.087**	0.006
Income over £31k	0.113**	0.007	0.114**	0.007
Income over £52k	0.092**	0.006	0.097**	0.007
Income over £100k	0.030**	0.003	0.035**	0.003
Gripstrength (kg)	0.417**	0.101	0.408**	0.085
Arterial stiffness	-0.038	0.093	-0.035	0.088
Height (cm)	0.986**	0.093	0.949**	0.080
BMI (kg/m2)	-0.714**	0.074	-0.781**	0.062
Diastolic (mmHg)	-0.818**	0.132	-0.871**	0.115
Systolic (mmHg)	-1.203**	0.257	-1.323**	0.216
Intelligence (0 to 13)	0.925**	0.065	0.929**	0.063
Happiness (0 to 5 Likert)	-0.002	0.018	-0.018	0.017
Alcohol consumption (0 low to 5 high)	0.186**	0.021	0.187**	0.018
Hours watching television per day	-0.487**	0.026	-0.500**	0.022

Moderate exercise (days/week)	-0.098**	0.032	-0.081**	0.028
Vigorous exercise (days/week)	-0.020	0.029	-0.005	0.025

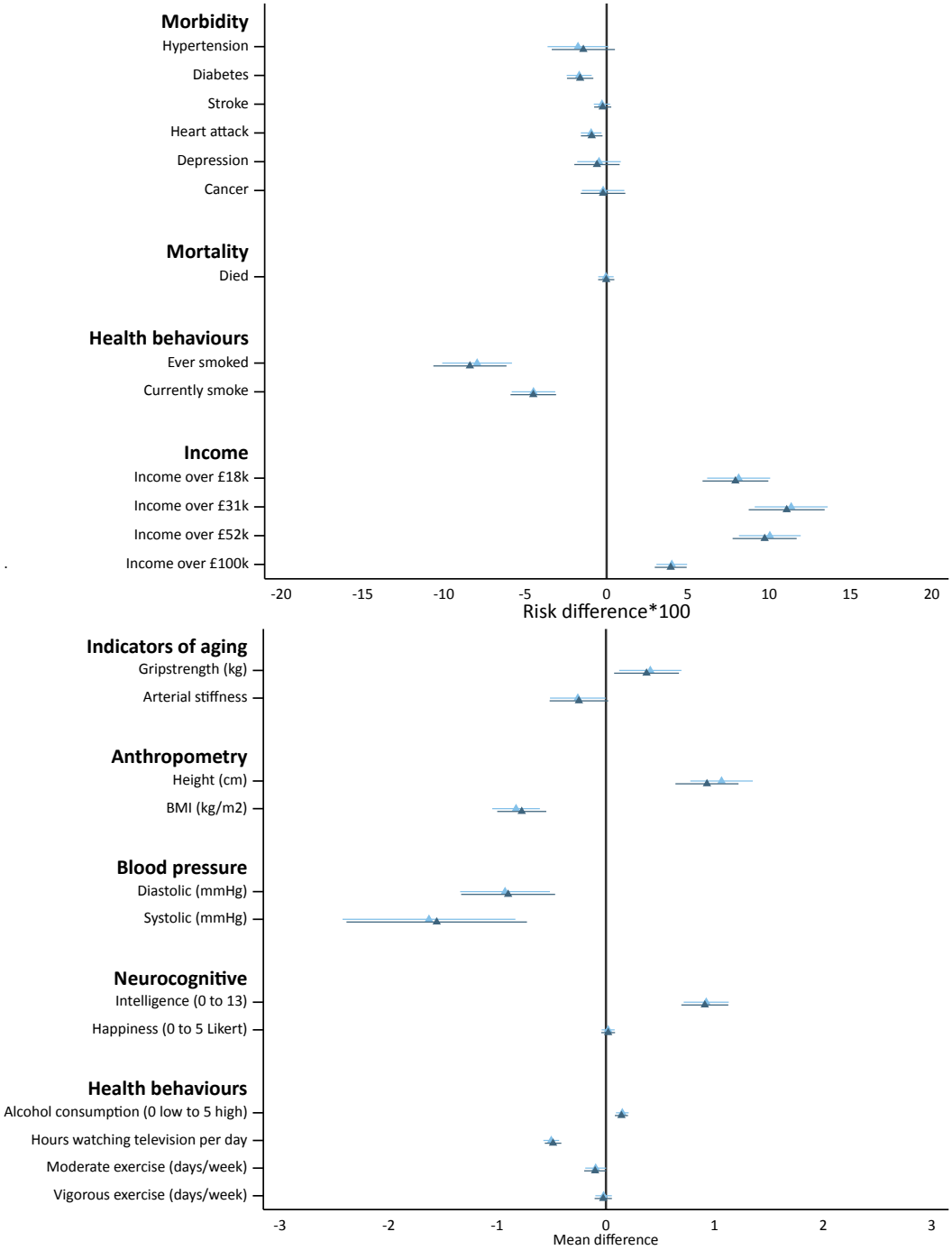
Notes: Adjusted for year, month of birth, sex, interaction of year of birth and sex, and the 10 principal components of population stratification. Standard errors clustered by month of birth. *p<0.05, **p<0.01

Figure A1: The effect of one additional year of schooling on morbidity, mortality and socioeconomic outcomes estimated using the educational attainment genetic score with and without adjusting for the sex, month and year of birth and principal components of population stratification ▲ and ▲ respectively.



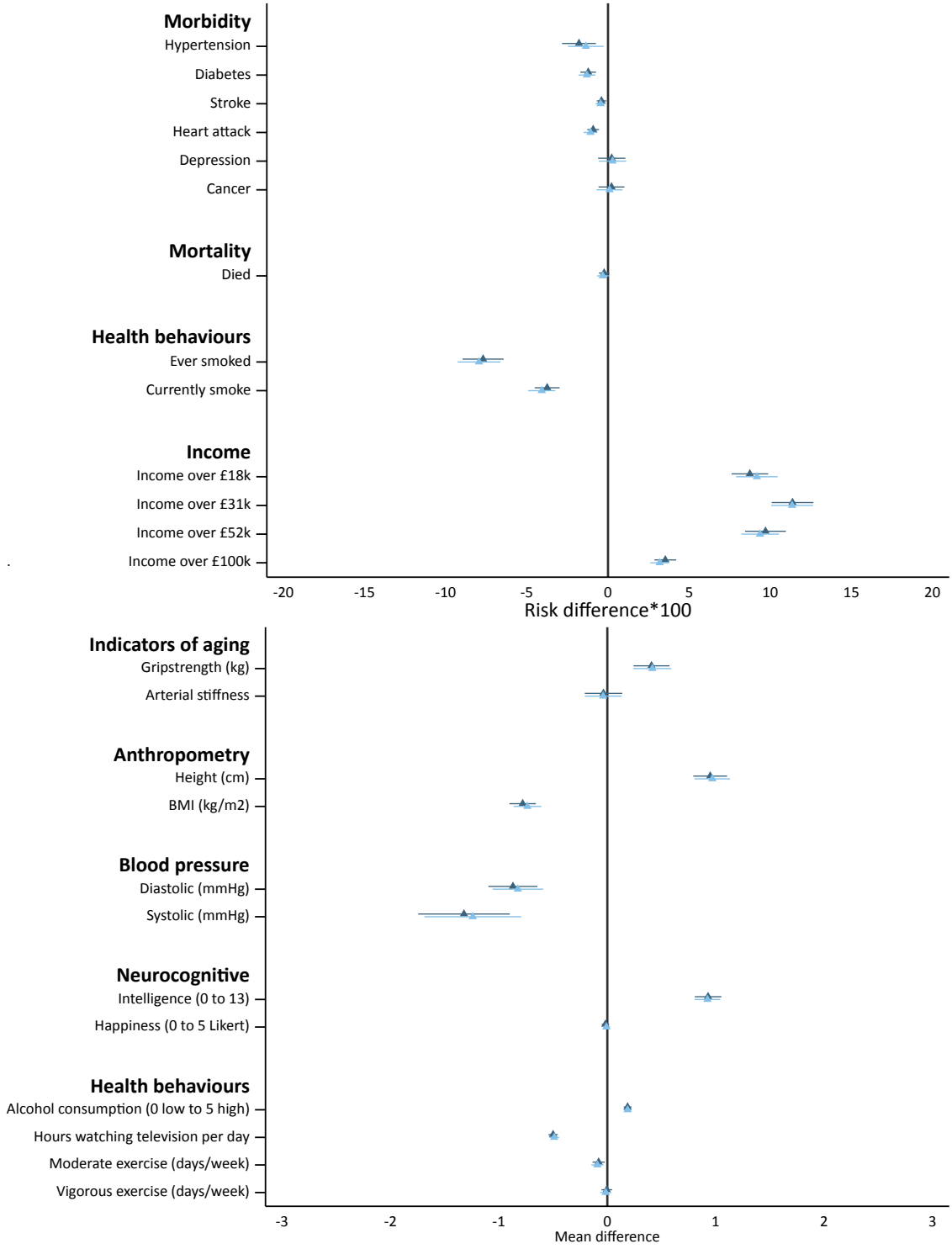
Notes: Confidence intervals allowing for clustering by month of birth reported. Sample weighted to adjust for under sampling of less educated.

Figure A2: The effect of one additional year of schooling on morbidity, mortality and socioeconomic outcomes estimated using the educational attainment genetic score with and without additionally adjusting breastfeeding, mother smoked during pregnancy, birth weight, birth location and deprivation (easting, northing, and distance to London) ▲ and ▲ respectively.



Notes: Confidence intervals clustered by month of birth reported. Sample weighted to adjust for under sampling of less educated. All results adjust for month and year of birth, sex, and the ten principal components of population stratification.

Figure A3: The effect of one additional year of schooling on morbidity, mortality and socioeconomic outcomes estimated using the educational attainment genetic score with and without weighting for under-sampling of less educated ▲ and ▲ respectively. The weighting did not affect the estimates.



Notes: Adjusted for month and year of birth, sex, and the ten principal components of population stratification. Confidence intervals allowing for clustering by month of birth reported.