USC Dornsife
Center for Economic and Social Research

USC Schaeffer
Leonard D. Schaeffer Center for Health Policy & Economics

DISTRIBUTIONAL EFFECTS OF EDUCATION ON HEALTH

*Silvia H. Barcellos, Leandro S. Carvalho, and Patrick Turley**

Paper No: 2018-002

CESR-SCHAEFFER WORKING PAPER SERIES

DISTRIBUTIONAL EFFECTS OF EDUCATION ON HEALTH

*Silvia H. Barcellos, Leandro S. Carvalho, and Patrick Turley**

This paper studies distributional effects of education on health. In 1972, England, Scotland, and Wales raised their minimum school-leaving age from 15 to 16 for students born after 9/1/1957. Using a regression discontinuity design and objective health measures for 0.27 million individuals, we find that education reduced body size and increased blood pressure in middle age. The reduction in body size was concentrated at the upper tail of the distribution with a 7.5 percentage point reduction in obesity. The increase in blood pressure was concentrated at the lower tail of the distribution with no effect on stage 2 hypertension. JEL codes: I10, I20.

Studies of the effects of education on health have focused on the effects on *average* health.[1] However, the morbidity and mortality consequences of a change in some health outcome may depend on which parts of the distribution of this health outcome are affected. There is, for example, evidence of U-shaped relationships between body mass index (BMI) and disability (Månsson et al. 1996; Al Snih et al. 2007) and between BMI and mortality (Fogel 1994; Aune et al. 2016). The benefit of a given reduction in BMI will be therefore larger for an obese individual than an individual with normal weight. In the presence of

[1] See, for example, Clark & Royer 2013; Lleras-Muney 2005; Albouy and Lequien 2009; Silles 2009; Powdthavee 2010; Kemptner et al. 2011; Jürges et al. 2013; Davies et al. 2017.

* Barcellos: University of Southern California, Center for Economic and Social Research, 635 Downey Way, Los Angeles, CA 90089-3332 (silvia.barcellos@usc.edu); Carvalho: University of Southern California, Center for Economic and Social Research, 635 Downey Way, Los Angeles, CA 90089-3332 (leandro.carvalho@usc.edu); Turley: Massachusetts General Hospital, Richard B. Simches Research Center, 185 Cambridge Street, CPZN-6818, Boston MA 02114 (paturley@broadinstitute.org). This paper benefited from discussions with James Banks, Dan Benjamin, Damon Clark, Maria Fitzpatrick, Dana Goldman, Mireille Jacobson, Arie Kapteyn, Adriana Lleras-Muney, Michael Mechine, Heather Royer, Jon Skinner, Jim P. Smith, and from the feedback of seminar participants at a number of universities, institutes and conferences. Joao Vilela and Sean Lee provided excellent research assistance. Research reported in this publication was supported through the Roybal Center for Health Decision Making and Financial Independence in Old Age (P30AG024962-13S1 and P30AG024962-13S2), NIA grant K01AG050811-01 (Barcellos), RF1AG055654 (Carvalho), and by the USC Population Research Center. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. This research has been conducted using the UK Biobank Resource under Application Number 15666. The authors declare that they have no relevant or material financial interests that relate to the research described in this paper.

such nonlinearities, distributional analyses are key to understanding the welfare effects of education.

This paper uses a regression discontinuity design to study the *distributional* effects of education on health. In 1972, England, Scotland, and Wales raised their minimum school-leaving age from 15 to 16 for students born on or after September 1, 1957 (students born before this date could drop out at age 15), generating a discontinuity in the relationship between education and date of birth at the September 1, 1957 "cutoff." Previous studies exploiting this reform (or a similar reform introduced in 1947) estimated the effect of education on average health from the discontinuity in average health at the cutoff. We estimate the distributional effects from discontinuities in the cumulative distribution function (CDF) of health at the cutoff.

We use data from the UK Biobank, a study that collected multiple *objective* and *continuous* measures of health between 2006 and 2010; 34-38 years after the policy change. Using standardized protocols, healthcare technicians and nurses measured the BMI, body fat percentage, waist and hip circumferences, lung function, and blood pressure of more than a quarter of a million individuals born in England, Scotland, and Wales between September 1, 1947 and August 31, 1967. The unprecedented availability of objective health measures for such a large sample permits well-powered estimation of distributional treatment effects. To ease concerns regarding multiple hypotheses testing, we focus our analyses on three health indices constructed from the multiple objective measures available: body size, lung function, and blood pressure.

We begin our analysis by estimating the effect of education on *average* health, finding only suggestive evidence of an impact. Specifically, staying in school until age 16 may improve one dimension of health—body size decreases on average by 0.15 of a standard deviation—while *worsening* another dimension of health—blood pressure *increases* on average by 0.15 of a standard deviation.[2] These point estimate are, however, only significant at the 10% level. We also find a marginally significant improvement in lung function that loses statistical significance once controls are added. Our estimates lie within the 95% confidence intervals of Clark and Royer (2010, 2013), who use the 1972 school leaving-age reform to study the effects of education on average health.

If the effects of education on health are heterogeneous, then a distributional test may be better powered than a test of difference in means. This would be the case, for example, if the effects of education on health are concentrated at particular parts of the health distribution (e.g., we may not expect health improvements for individuals who would have been healthy even in the absence of the additional year of schooling). We illustrate this point in Appendix F.

There are three main takeaways from the distributional analysis. First, it confirms the suggestive treatment effect estimates on the averages: staying in school until age 16 reduces body size, while also increasing blood pressure. Second, it reveals that these effects vary considerably along the health distribution. To give a sense of how effects are concentrated, staying in school until age 16 reduces the 90th percentile of the body size distribution by 0.38 of a standard deviation—that is equal to 2.5 times the average treatment effect of –0.15 of a standard deviation. Third, the effects on body size and on blood pressure occur in different parts of their respective distributions: while the effect on body size is concentrated at the upper tail (i.e., among the least healthy), the effect on blood pressure is concentrated at the lower tail (i.e., among the most healthy).[3]

We conduct a distributional test based on Shen and Zhang (2016) to formally investigate whether these changes are statistically significant, testing

---

[2] In our sample, there are very few participants with low blood pressure or who are underweight.

[3] We do not have the power to test whether some people experienced both a reduction in body size and an increase in blood pressure. However, given that there is a moderate positive correlation between these two outcomes ($\rho = 0.3$), it is unlikely (though possible) that those who experienced reductions in body size were the same individuals experiencing increases in blood pressure. See Appendix Figure D3.

differences in the bottom and top halves of the CDFs of our three health indices. The test rejects at the 5% significance level the null of equality for the top half of the body size distribution (p-value of 0.013) and for the bottom half of the blood pressure distribution (p-value of 0.010). Our test also rejects the null of equality of the top half of the lung function distribution, but only at the 10% confidence level (p-value of 0.062). For this reason, we take our lung function results as suggestive only.

These results illustrate the policy relevance of studying distributional effects. Even though the average treatment effects show a reduction in body size and an increase in blood pressure of identical magnitudes (0.15 of a standard deviation), a policy maker may wish to trade-off these effects based on which parts of the respective distributions are affected. Because the reductions in BMI occur at the right tail of the BMI distribution, staying in school until age 16 reduces obesity rates (i.e., BMI above 30) by 7.5 percentage points. In contrast, the increase in blood pressure is concentrated below the clinical threshold for stage 2 hypertension (i.e., diastolic blood pressure above 90 mmHg or systolic blood pressure above 140) with no statistically significant consequences for prevalence.[4]

We also conduct an exploratory analysis to investigate possible channels through which education may affect health. While improvements in SES and diet may explain the reduction in body size, we are not able to uncover in our data channels through which education increases blood pressure. We speculate that, by changing the types of occupations and careers individuals have, education might have an effect on job responsibilities, expectations, and work-related stress.

[4] There could still be harmful health consequences, as observational studies have linked increases in blood pressure in the range documented here to increased ischemic heart disease and stroke mortality risk (Lewington et al. 2002). In fact, the hypertension guidelines have changed in 2017, with what was previously defined as prehypertension now being called stage 1 hypertension. This reclassification aimed at earlier intervention and ultimately lowering cardiovascular event rates (Bakris and Sorrentino 2018)

There are a number of studies that have exploited changes in compulsory schooling laws to study the causal effect of education on average health (e.g., Lleras-Muney 2005; Albouy and Lequien 2009; Silles 2009; Powdthavee 2010; Kemptner et al. 2011; Clark & Royer 2013; Jürges et al. 2013; Davies et al. 2017; Janke et al. 2018; Meghir et al. forthcoming).[5] This paper adds to this literature in the following ways. First, to the best of our knowledge, this is the first paper to estimate the distributional effects of education on health using a quasi-experimental approach.[6] Second, we provide evidence about the importance of conceptualizing health as a multidimensional construct. We show that education improves body size while also increasing blood pressure. Finally, we consider only objective and continuous measures of health. Previous research shows that discrepancies between subjective (e.g. self-reported hypertension) and objective measures of health (e.g. objectively measured hypertension) vary with socioeconomic status (e.g. Johnston et al. 2009).

The paper is structured as follows. Section 1 discusses the 1972 raising of the school leaving-age reform and the data. In Section 2 we present the effects of the reform on education and the effects on average health. Section 3 discusses the methods used to estimate the distributional effects with results shown in Section 4. Section 5 presents suggestive evidence on mechanisms and Section 6 concludes.

1. Background and Data

A. The 1972 Raising of the School Leaving Age

The British compulsory schooling laws specify the maximum age by which children must start school and the minimum age at which they can leave school.

[5] This literature has produced mixed evidence, resulting in no consensus on whether there is a causal effect of education on health. Grossman (2015) reviews this literature and concludes that "there is enough conflicting evidence to warrant more research on the question of whether more schooling does in fact cause better health outcomes."
[6] Conti, Heckman and Urzua (2011) have used a structural approach to estimate the distributional effects of education on health.

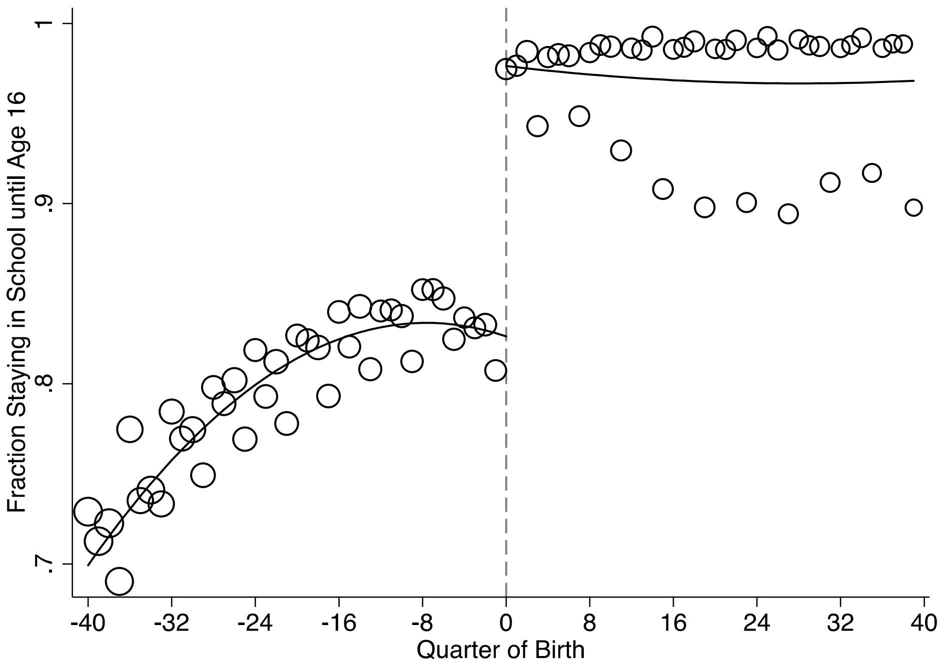
In this paper, we exploit the 1972 Raising of School Leaving Age (ROSLA) legislation, which increased the minimum school-leaving age from 15 to 16 years of age in England, Scotland, and Wales. These laws and their implementation have been extensively documented in other studies (see Clark and Royer 2010, 2013) so we only include a brief summary of its main features here.

The UK's 1944 Education Act raised the minimum school-leaving age from 14 to 15 years of age in England, Wales, and Scotland and gave the Minister of Education the power to further raise it to 16 years when conditions allowed. The Minister did so in January 1972 for Scotland (Statutory Instrument No. 59)[7] and in March 1972 for England and Wales (Statutory Instrument No. 444)[8]. Both changes took effect in September 1, 1972, implying that those who were 15 or younger before that date (born on September 1, 1957 or later) had to stay in school until at least age 16 in the three countries (hereafter, we use the term "stayed in school until age 16" to refer to those who stayed in school until *at least* age 16). Infrastructure investments, such as school building to absorb the additional students, preceded the 1972 ROSLA but key elements of the school system did not change with the policy.

Figure 1, which displays the fraction of study participants who stayed in school until age 16 (y-axis) by quarter of birth (x-axis), shows that the policy generated a discontinuous relationship between these two variables. There is a large jump at the September 1, 1957 cutoff marked by the vertical dashed line. Those born during the summer months could in practice drop out at age 15 even after the 1972 ROSLA, since the law required students to be 16 by the start of the next school year. We estimate that the policy increased the fraction of student participants who stayed in school until age 16 by 15 percentage points – see Table 1.

[7] http://www.legislation.gov.uk/uksi/1972/59/pdfs/uksi_19720059_en.pdf
[8] http://www.legislation.gov.uk/uksi/1972/444/pdfs/uksi_19720444_en.pdf

Figure 1: Fraction Staying in School until Age 16 by Quarter of Birth



Notes: The figure shows the fraction of study participants who stayed in school until age 16 by quarter of birth. The dashed vertical line marks the first birth cohort affected by the 1972 school-leaving age reform. Cohorts born to the right of the line had to stay in school until age 16 while cohorts born before could leave at age 15. The curves show quadratic polynomials in quarter of birth that capture birth cohort trends. The circumference of each circle reflects the number of participants born in that quarter. N = 271,082.

Notice there is a cyclical drop in the fraction of students staying in school until age 16, corresponding to those born between June and August. This phenomenon is not specific to our data and has been noted by others. According to Clark and Royer (2013), "when the minimum leaving age became age 16, students had to stay until part way through grade 10. Grade 10 finishes with the "O level" exam period and, technically, students finish when they complete their last exam. Since the exam period starts in late May and finishes in mid-June, starting in 1972, students born in late June, July, and August could leave at 15, technically younger than the minimum leaving age (16)" (pg. 2 of Online Appendix). We include calendar month of birth dummies in our regressions to control for this seasonality.

B. Data

We use data from the UK Biobank, a large, population-based prospective study initiated by the UK National Health Service (NHS) (Sudlow et al. 2015). Between 2006 and 2010, invitations were mailed to 9.2 million people between the ages of 40 and 69 who were registered with the NHS and lived up to about 25 miles from one of 22 study assessment centers distributed throughout the UK (Allen et al. 2012) – see Appendix Figure D1.[9] The sample is formed by 503,325 individuals who agreed to participate (i.e. a response rate of 5.47%). Although the sample is not nationally representative, our estimates have internal validity because there is no differential selection on the two sides of the September 1, 1957 cutoff – see Appendix A.

Physical measures, such as anthropometrics, spirometry, and blood pressure, were collected of survey participants. The collection was standardized across centers and was conducted by trained nurses or healthcare practitioners. Every participant was genotyped.

In this paper, we focus on objective and continuous measures of health. Continuous measures because we are interested in studying how education affects the distribution of health. Objective measures because research shows that discrepancies between subjective (e.g. self-reported hypertension) and objective measures of health (e.g. objectively measured hypertension) vary with socioeconomic status (e.g. Johnston et al. 2009).

We restrict ourselves to three dimensions of health (that satisfy the two criteria above and) that can be arguably affected by education: *body size*, *lung function*, and *blood pressure*.[10] Weight and body size can be affected by education through diet and physical exercise. Lung function can be affected by education through smoking and pollution. Blood pressure may be affected by

[9] The NHS has contact details for an estimated 98% of the UK population.

[10] The other objective and continuous measures of health currently available in the UK Biobank are either available for just a subsample (e.g., arterial stiffness, bone densitometry of heel, ECG) or there is no clear hypothesis on how they could be affected by education (e.g., hand grip strength).

education through diet, physical exercise, medication, and stress. The UK Biobank has multiple measures of each of these health dimensions.

Next, we describe how each of these health dimensions is measured in the data.

Body Size

We use three measures of body size: BMI, body fat percentage, and waist-hip ratio.[11] A bioimpedance analyzer was used to calculate body fat percentage. This device passes a low electrical current through the body. Water conducts electricity. While fat contains very little water, muscle contains 70% water. The bioimpedance analyzer calculates body fat from the speed of the current: The slower the signal travels, the greater the fat content.

Lung Function

A spirometry test was conducted to measure participants' lung function. The spirometer is a small machine attached to a mouthpiece by a cable that measures the volume and speed of air after a forced exhale. Participants were asked to fill their lungs as much as possible and to blow air out as hard and as fast as possible in the mouthpiece.[12] Three parameters were measured: 1) *forced expiratory volume in the first second* is the amount of air exhaled during the first second; 2) *forced vital capacity* is the total amount of air exhaled during the forced breath; and 3) *peak expiratory flow* is the fastest rate of exhalation. These parameters are used to assess pulmonary conditions, such as chronic obstructive pulmonary disease and asthma. We follow DeMateis et al. (2016)'s

[11] The UK Biobank provides two measures of BMI: one calculated as weight in kilograms divided by height squared (in meters) and one using height and electrical impedance to quantify mass. We take the average of these two measures. We can get very similar results if we use exclusively the first measure of BMI (i.e., the weight in kilograms divided by height in meters squared). The waist-hip ratio is equal to the waist circumference divided by the hip circumference.

[12] They were instructed to continue blowing until no more air came out of their lungs. Up to three attempts were allowed. The participant was allowed a third attempt if the first two blows did not satisfy the reproducibility criteria of the spirometry protocol.

criteria to identify acceptable expiratory maneuvers in the UK Biobank data. Valid spirometry measures are available for 79% of our sample.[13]

Blood Pressure

Two measurements were taken of the diastolic and systolic blood pressures of each study participant. We use the average of these two measurements.

Summary Indices

In order to reduce the number of outcomes and partly address concerns about multiple hypothesis testing, we construct for each health dimension a summary index that is a weighted average of the different outcomes measuring that dimension:

1. *Body size*: body mass index, waist-to-hip ratio, and body fat percentage;
2. *Lung function*: forced expiratory volume in the first second, forced vital capacity, and peak expiratory flow;
3. *Blood pressure*: diastolic and systolic blood pressures.

First, each measure is standardized separately by gender, using as a reference those born in the 12 months before September 1, 1957. We then follow the procedure proposed by Anderson (2008), weighting the measures by their variance-covariance matrix. The weights are calculated to maximize the amount of information captured in the index. Finally, we construct a fourth "summary index" that is a summary of the body size, the lung function, and the blood pressure indices, using the same weighting procedure. We construct all four indices so that a higher number corresponds to worse health.

The correlation between the body size and lung function indices is 0.20. The correlation between the body size and the blood pressure indices is 0.30. The correlation between the lung function and the blood pressure indices is 0.10. The correlations between the summary index and the body size, lung function, and blood pressure indices are respectively 0.69, 0.67, and 0.68.

[13] Appendix Figure C5 and Appendix Table C1 show that participants born before and after September 1957 are equally likely to have valid spirometry measures.

2. Mean Effects

A. Effects of the Compulsory Schooling Change on Education

We use a regression discontinuity design (RDD) to estimate the "first stage", i.e., the effect of the 1972 ROSLA on education. In particular, we estimate the following regression:

$$Educ_i = a_0 + a_1 Post_i + f(DoB_i) + \boldsymbol{x}_i' \boldsymbol{a}_2 + \varepsilon_i, \tag{1}$$

where $Educ_i$ is a measure of the educational attainment of individual i; $Post_i$ is 1 if individual i was born on or after September 1, 1957 (and 0 otherwise); DoB_i is individual i's date of birth; and the vector $\boldsymbol{x}_i$ contains predetermined characteristics. Date of birth is measured in days relative to the cutoff, such that $DoB = 0$ for someone born on September 1, 1957. The function $f(\cdot)$ captures birth cohort trends in educational attainment, which are allowed to differ on either side of the September 1, 1957 cutoff. The coefficient a_1 gives the effect of the 1972 ROSLA on educational attainment.[14]

We restrict the data to study participants born within 10 years of September 1957 – that is, born between September 1, 1947 and August 31, 1967 – and use a quadratic polynomial in date of birth to capture cohort trends (i.e., function $f(\cdot)$ in equation (1)). In Appendix B we show our main results are robust to the choice of bandwidth and to the use of linear trends.[15] We use triangular kernel weights that give greater weight to study participants born closer to the cutoff. The set of predetermined characteristics include gender, age in days (at the time of the baseline assessment) and age squared, dummies for

[14] The inclusion of predetermined controls in equation (1) is not needed for identification but can improve the precision of estimates.

[15] Gelman and Imbens (2016) caution against the use of higher order polynomials (higher than 2) in RDD.

ethnicity, dummies for country of birth, and dummies for calendar month of birth (to control for seasonality).[16]

Notice that even though previous work studying the 1972 ROSLA clustered standard errors by month-year of birth (e.g., Clark and Royer 2013; Davies et al. 2017), we do not need to cluster our standard errors because our data include exact date of birth. As Card and Lee (2008) discuss, in applications where the running variable is only reported in coarse intervals (e.g., month-year of birth), researchers have to choose a particular functional form for the model relating the outcomes of interest to the running variable. The deviation between the expected value of the outcome and the predicted value from a given functional form is modeled as a random specification error, which is incorporated in inference by clustering the standard errors for different values of the running variable. This specification error should be negligible in our context because our data include day-month-year of birth. Appendix Table D4 shows that we get virtually identical standard errors estimates irrespective of whether we cluster by date of birth or not.

Table 1 shows estimates of effects of the 1972 ROSLA on education. Each cell reports results from a separate ordinary least squares estimation of (1), where we vary the dependent variable (listed in the column) and whether the predetermined characteristics are included as controls. The table shows the coefficient on the indicator variable for being born on or after September 1, 1957, a_1, and the mean of the dependent variable among those born in the 12 months before September 1, 1957. Robust standard errors are reported between brackets.

We estimate that the 1972 ROSLA increased the fraction of study participants staying in school until age 16 by 14-15 percentage points, an estimate significant at the 1% significance level.[17] Studies using nationally

[16] Because participants were surveyed for the baseline assessment between 2006 and 2010, date of birth and age are not perfectly collinear.

[17] Estimates of the effect of the 1972 school-leaving age reform on staying in school until age 17 or later are an order of magnitude smaller than the effect on staying in school till age 16

representative data, such as Clark and Royer (2013), estimate this figure to be closer to 25 percentage points. This difference is likely due to the composition of the UK Biobank sample, which is more educated than the overall population (despite the selectivity of the UK Biobank sample, our estimates have internal validity because there is no differential selection on the two sides of the September 1, 1957 cutoff – see Appendix A). One consequence is that the standard errors of our two stages least squares (2SLS) estimates will be *ceteris paribus* larger than of studies with nationally representative data, something that is compensated by the larger sample size of the UK Biobank.

Table 1: Effects on Education

	Left school at age ≥ 16		*No qualification*		*CSE*	
Post	0.150	0.139	-0.048	-0.050	0.059	0.070
	[0.004]***	[0.004]***	[0.004]***	[0.004]***	[0.005]***	[0.005]***
Controls?	No	Yes	No	Yes	No	Yes
Mean of Y	0.827		0.113		0.205	
	O-level		*A-level*		*College degree*	
Post	0.038	0.035	0.016	0.015	-0.003	-0.005
	[0.006]***	[0.006]***	[0.006]***	[0.006]**	[0.006]	[0.006]
Controls?	No	Yes	No	Yes	No	Yes
Mean of Y	0.513		0.325		0.368	

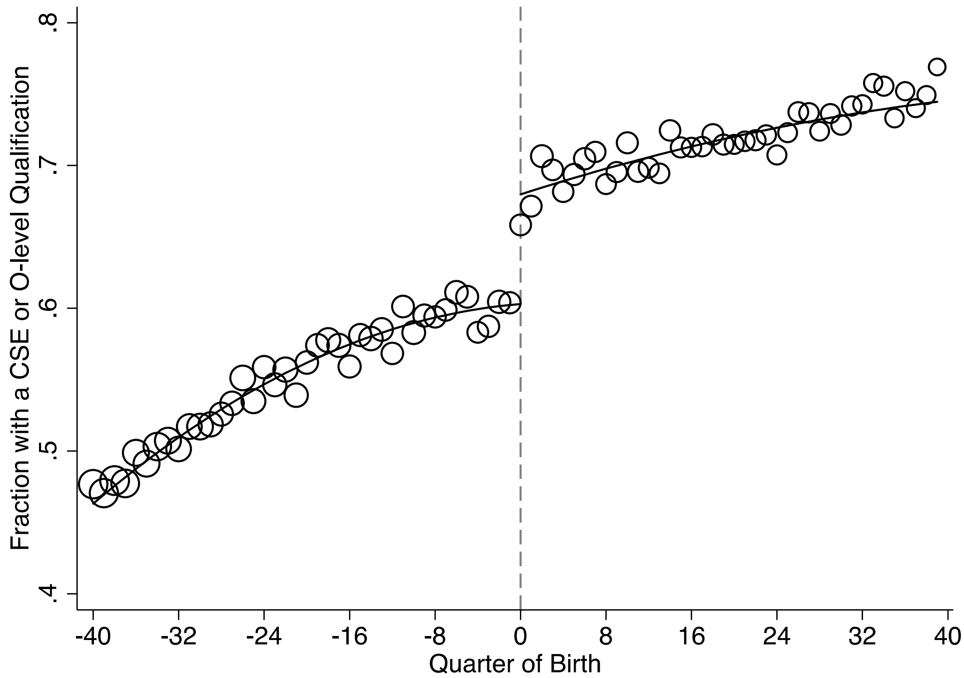
Notes: The table shows the effects of the school reform on education. Each cell corresponds to a separate regression. We report the coefficient on the indicator variable for being born on or after September 1, 1957 (i.e., "Post"). The dependent variable mean in the bottom row is the weighted mean among those born in the 12 months before September 1, 1957. Controls include male, age in days and age squared, dummies for calendar month of birth, dummies for ethnicity, and dummies for country of birth. Robust standard errors. N = 271,082 for "*Stayed in school until 16*" and N = 268,551 for all other outcomes.

One may worry that these students forced to stay in school an extra year did not learn much if they did not put effort into it. The evidence does not support this hypothesis. By the 70's high schools offered a series of two-year courses that ran through grades nine and ten and required students to take exams at the end of grade ten (the grade they are typically in by age 16): Certificate of Secondary Education (CSE) or a General Certificate of Education (GCE)

and are generally not robust to the inclusion of controls – see Appendix Figure D2 and Appendix Table D1.

Ordinary Level (also known as an O-level). By compelling students to stay in school until grade ten, the 1972 ROSLA gave students an incentive to complete these courses and get these qualifications, which are valued in the labor market (Dickson and Smith 2011).

Figure 2: Fraction with a CSE or O-level by Quarter of Birth



Notes: The figure shows the fraction of study participants with a CSE or O-level qualification by quarter of birth. The dashed vertical line marks the introduction of the policy. Cohorts born to the right of the line had to stay in school until age 16 while cohorts born before could leave at age 15. The curves show quadratic polynomials in quarter of birth that capture birth cohort trends. The circumference of each circle reflects the number of participants born in that quarter. *N* = 268,551.

Figure 2 shows that the policy generated a discontinuous increase in the fraction of study participants with these qualifications. In Table 1 we estimate that the policy increased the fraction of study participants with a CSE by 6-7 percentage points and the fraction with an O-level by 3-4 percentage points. Interestingly, the fraction with an A-level, an exam typically taken at age 18 and used for college admissions, increased by 1-2 percentage points. The fraction without any formal qualification dropped by 5 percentage points. All of these reduced-form estimates are statistically significant at 1%. We find no effect of the policy on having a college degree. Consistent with these results, we document that the policy increased household income[18], especially at lower income levels, and enabled workers to get "better jobs", that is, to have occupations with higher socioeconomic status[19]– see Appendix Table E1 and Appendix Table E2.

B. Effects on Average Health

We now turn to the effects of the 1972 ROSLA on average health. We are interested in the relationship between health and education:

$$Health_i = \beta_0 + \beta_1 Educ16_i + g(DoB_i) + \boldsymbol{x}_i'\boldsymbol{\beta}_2 + u_i, \qquad (2)$$

where $Health_i$ is a health measure for individual i. $Educ16_i$, an indicator for whether individual i stayed in school until age 16, is our endogenous measure of education. The function $g(\cdot)$ captures birth cohort trends in health and is allowed to differ on either side of the September 1, 1957 cutoff. We substitute (1) into (2) to get the "reduced-form" effect of the 1972 ROSLA on *average* health:

$$Health_i = \gamma_0 + \gamma_1 Post_i + j(DoB_i) + \boldsymbol{x}_i'\boldsymbol{\gamma}_2 + v_i. \qquad (3)$$

The coefficient γ_1 gives the effect of the school leaving-age reform on average health. The RDD identifying assumption is that, in the absence of the reform, our outcomes of interest would have been smooth across the September 1, 1957 threshold. This assumption is violated if determinants of health are discontinuous at the cutoff (Lee 2008). In Appendix A we partially test for such violations by investigating whether the average (or the cumulative distribution

[18] These results are broadly consistent with Grenet (2013) that finds that the extra year of schooling induced by the 1972 ROSLA increased wages.
[19] Respondents who were employed or self-employed were asked in a verbal interview to describe their jobs. Respondent's answers were coded following the Standard Occupation Classification 2000. We classify the socioeconomic status of the occupations using the 2000 National Statistics Socio-economic Classification (NS-SEC), the primary social classification in the United Kingdom. See Appendix Table D2.

function) of predetermined characteristics, such as gender and place of birth, are discontinuous around September 1, 1957. Since the UK Biobank genotyped the full sample, we also test for the smoothness of a pair of genetic variables, which are determined at conception and are objectively measured.[20] Our analyses indicate that these characteristics are smooth across the September 1 1957 threshold, which strengthens our confidence that the RDD results provide unbiased estimates of the causal effects of education on the health of UK Biobank participants.

We estimate the causal effect of staying in school until age 16 on average health, β_1, through two stages least squares (2SLS), using the indicator for being born on or after September 1, 1957 (i.e., $Post_i$) to instrument for staying in school until age 16 (i.e., $Educ16_i$) in equation (2). We adopt the same specifications used to estimate the effects on education (see section 2.A), namely: 10-year bandwidths, quadratic polynomials to capture birth cohort trends, triangular kernel weights, and the same set of controls. Appendix B shows our results are robust to linear cohort trends and smaller bandwidths.

Figure 3 examines the effects of the 1972 compulsory schooling change on average health. The graphs show average health (y-axis) by quarter of birth (x-axis), where health is measured by the four health indices: the body size index (top left); the lung function index (top right); the blood pressure index (bottom left); and the summary index (bottom right).

Table 2 shows regression estimates of the effects of the 1972 compulsory schooling change on average health.[21] The first rows show the coefficients on the indicator variable for being born on or after September 1, 1957, γ_1 in equation (3), from reduced-form estimates. The third row shows the coefficients on staying in school until age 16 from 2SLS estimates, β_1 in equation (2), where the indicator variable for being born on or after September 1, 1957 is used to

[20] See Online Appendix of Barcellos, Carvalho and Turley (2018) for details about the construction of genetic variables.

[21] Notice that while Figure 3 uses quarter of birth Table 2 uses day of birth (e.g., September 1, 1957). The same distinction applies to Figure 1 and Table 1.

instrument for staying in school until age 16. Again, the health indices were constructed such that higher values correspond to worse health.

Overall Figure 3 suggests education may lead to small average improvements in health, with minor discontinuous decreases in the body size, lung function, and summary indices at the cutoff. One noteworthy exception is blood pressure. There is a discontinuous *increase* in the blood pressure index at the cutoff, suggesting that education may *worsen* this particular dimension of health. Appendix Figures B1-B4 assess the sensitivity of Figure 3 to changes in the bandwidth and to using linear trends.

Table 2 shows that the effects on body size and blood pressure are statistically significant at the 10% significance level. The 2SLS point estimates imply that staying in school until age 16 decreases the body size, the lung function, and the summary indices respectively by 0.15-0.16, 0.17, and 0.12 of a standard deviation. At the same time, staying in school until age 16 *increases* the blood pressure index by 0.15 of a standard deviation.

The p-value of a test of the difference between the effects on the body size and the blood pressure indices is 0.004. The difference between the effects on the lung function and the blood pressure indices has a p-value of 0.069. The difference between the effects on the body size and the lung function indices has a p-value of 0.906. These results, notably the difference between blood pressure and body size, point to the importance of treating health as multi-dimensional and considering the effects of education on different dimensions separately. Focusing on the analysis of summary measures of health can lead to misleading conclusions of no health impact if effects going on opposite directions cancel out, as is the case in Table 2.

Our estimates lie within the 95% confidence intervals of Clark and Royer (2010) – see Appendix Table D3 and Appendix Figure D7.[22] Clark and Royer

[22] In contrast, our results lie outside the confidence intervals of Davies et al. (2017), which study the same reform and data (UK Biobank). We believe this is due to differences in the weighting procedure they use, their specification (bandwidth and polynomial choice) and sample selection (they do not include Wales and Scotland).

(2010) do not present the effects on systolic blood pressure or on lung function. As discussed above, even though we have a larger sample than Clark and Royer (2010), we have a smaller first stage, which explains why our standard errors are not substantially smaller than theirs.

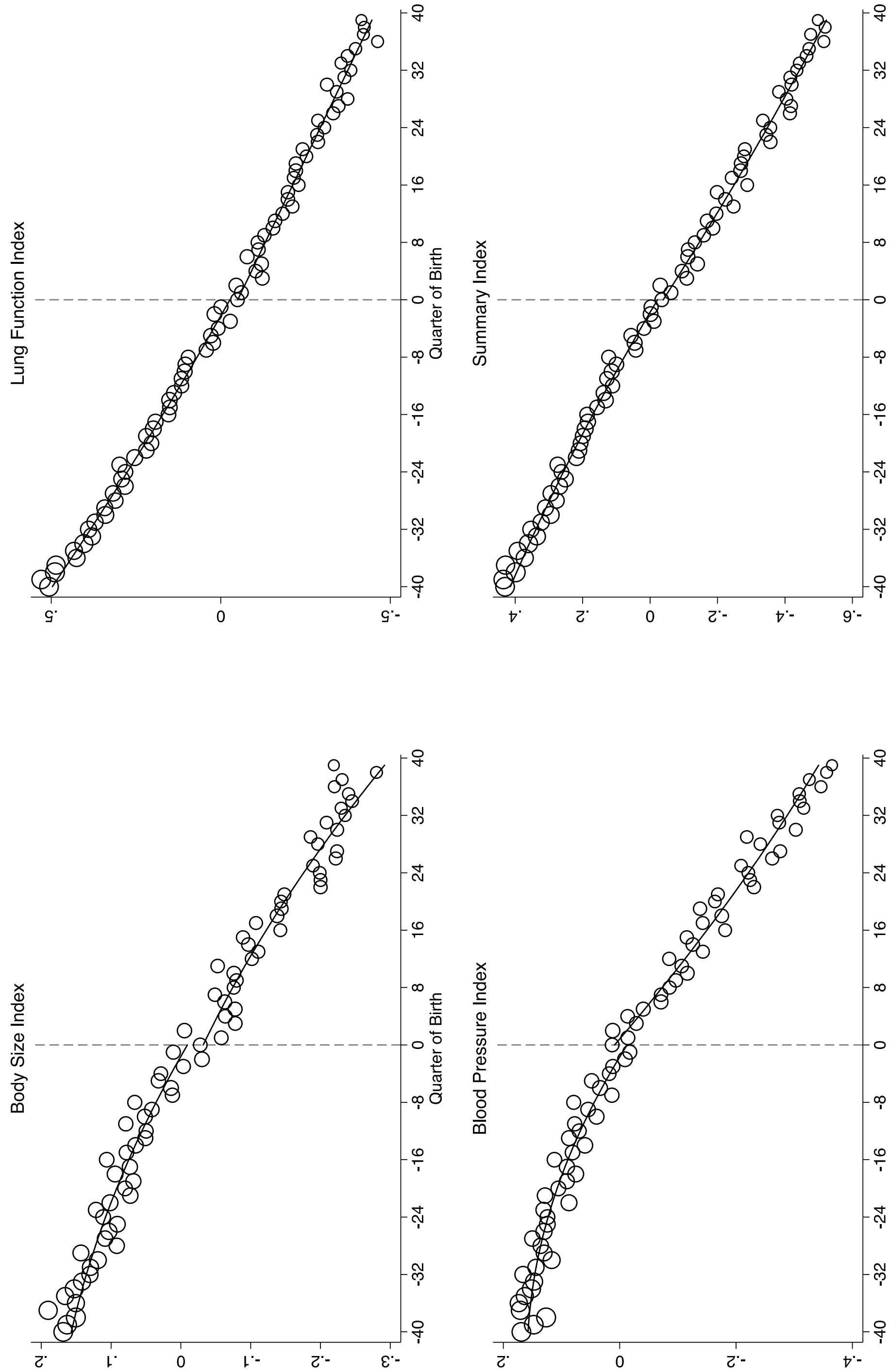
Even though the effects on average health are informative, they may conceal larger effects on particular parts of the health distribution with important policy implications. Moreover, the average effects documented here are suggestive at best, being only significant at the 10% level. As illustrated in Appendix F, if the effects of education on health are concentrated at particular parts of the health distribution, a distributional test may be better powered than a test of difference in means. In the next section, we describe the methods we use to estimate how education affects the distribution of health.

Table 2: Effects on Average Health

	Body Size		*Lung Function*		*Blood Pressure*		*Summary*	
Reduced-form								
Post	-0.023	-0.023	-0.024	-0.022	0.023	0.021	-0.016	-0.016
	[0.013]*	[0.013]*	[0.014]*	[0.014]	[0.013]*	[0.013]*	[0.014]	[0.014]
Two stages least squares								
Stayed in school until 16	-0.154	-0.163	-0.175	-0.174	0.151	0.151	-0.120	-0.125
	[0.083]*	[0.091]*	[0.103]*	[0.112]	[0.084]*	[0.091]*	[0.103]	[0.112]
Controls?	No	Yes	No	Yes	No	Yes	No	Yes
N Observations	266,525	266,525	215,536	215,536	270,647	270,647	212,689	212,689

Notes: The table shows the effects on average health. The first two rows show reduced-form effects of the 1972 Raising of the School Leaving Age. The last two rows show two stages least squares estimates of the effect of staying in school until age 16 obtained by using an indicator for being born on or after September 1, 1957 to instrument for staying in school until age 16. Robust standard errors. Controls include male, age in days and age squared, dummies for calendar month of birth, dummies for ethnicity, and dummies for country of birth.

Figure 3: Effects on Average Health



Notes: These figures show average health by quarter of birth. See Table 2 for number of observations.

3. Methods for Distributional Effects Estimates

In Section 2.B, we estimated the effect of education on *average health* (of compliers) by investigating if there was a discontinuity in the relationship between *average health* and date of birth at the September 1, 1957 cutoff. Here we estimate the effect of education on the *health distribution* (of compliers) by investigating if, at the September 1, 1957 cutoff, there is a discontinuity in the relationship between the *cumulative distribution function (CDF) of health* and date of birth.[23]

We want to estimate the local distributional treatment effect (LDTE) for compliers. Let the *pre-reform CDF* be the CDF for compliers in the limit when date of birth is converging to September 1, 1957 from the left (i.e., $DoB < 0$):

$$F_{pre}(h) = \lim_{DoB \to 0^-} \Pr(Health \leq h | DoB).$$

Similarly, the *post-reform CDF* is defined as the CDF for compliers in the limit when date of birth is converging to September 1, 1957 from the right (i.e., $DoB > 0$):

$$F_{post}(h) = \lim_{DoB \to 0^+} \Pr(Health \leq h | DoB).$$

The LDTE, which is the discontinuity in the CDF at September 1, 1957 (i.e., $DoB = 0$), is estimated as the difference between $F_{post}(h)$ and $F_{pre}(h)$ at a given h:

$$\mu(h) = F_{post}(h) - F_{pre}(h). \tag{4}$$

To estimate $\mu(h)$, in practice we discretize the support of the distribution of health and then estimate (5) for each grid point h:[24]

[23] The RDD identifies differences in the marginal distributions of cohorts affected and unaffected by the reform. Stronger assumptions (such as rank preservation) would be needed to estimate the distribution of treatment effects.

[24] For each one of the three health indexes, we first calculate the 1st and the 99th percentiles among those born before September 1, 1957:

$$\mathrm{I}(Health_i \le h) = \theta_0(h) + \theta_1(h)Educ_i + l(DoB_i; h) + \boldsymbol{x}_i'\boldsymbol{\theta}_2(h) + \varepsilon_i(h), \quad (5)$$

where $\mathrm{I}(Health_i \le h)$ is an indicator variable for whether the health index of individual i is smaller or equal to h. The function $l(\cdot\,; h)$ capture birth cohort trends, which are allowed to differ on either side of the cutoff date *and to vary with* h. It is approximated by a quadratic polynomial of date of birth in days.[25] We estimate (5) through 2SLS using the indicator variable for being born on or after September 1, 1957 ($Post_i$) to instrument for staying in school until age 16 ($Educ_i$). The coefficient on the latter, $\theta_1(h)$, estimates the discontinuity in $\Pr(Health_i \le h)$ at $DoB_i = 0$ and thus provides an estimate of $\mu(h)$.[26]

We find it easier to visualize $\hat{\mu}(\cdot)$ by plotting in the same graph the pre-reform CDF $F_{pre}(\cdot)$ and the post-reform CDF $F_{post}(\cdot)$.[27] For a given grid point h, we proceed in three steps. First, we estimate $\mu(h)$. Second, we estimate $F_{pre}(h)$ – see next paragraph. Third, we estimate $F_{post}(h)$ as the sum of $\hat{F}_{pre}(h)$ and $\hat{\mu}(h)$. We repeat this procedure for each one of the 21 grid points. Finally, we draw the estimated pre-reform and post-reform CDFs by drawing $\hat{F}_{pre}(h)$

$$0.01 = \Pr(Body\ Size_i \le q_1^{bs} | DoB < 0) = \Pr\left(Lung\ Function_i \le q_1^{lf} \middle| DoB < 0\right) = \\ = \Pr\left(Blood\ Pressure_i \le q_1^{bp} \middle| DoB < 0\right)$$
$$0.99 = \Pr(Body\ Size_i \le q_{99}^{bs} | DoB < 0) = \Pr\left(Lung\ Function_i \le q_{99}^{lf} \middle| DoB < 0\right) \\ = \Pr\left(Blood\ Pressure_i \le q_{99}^{bp} \middle| DoB < 0\right)$$

Next, we define the starting point of the grid as the minimum of the 1st percentiles among the health indexes and the endpoint as the maximum of the 99th percentiles among the health indexes:

$$h_1 = min\{q_1^{bs}, q_1^{lf}, q_1^{bp}\}$$
$$h_{21} = max\{q_{99}^{bs}, q_{99}^{lf}, q_{99}^{bp}\}$$

The grid consists of 21 points uniformly distributed between h_1 and h_{21}.

[25] In our main specification with a quadratic polynomial:

$$l(DoB_i; h) = \lambda_1(h)DoB_i + \lambda_2(h)[DoB_i]^2 + \lambda_3(h)I\{DoB_i \ge 0\}DoB_i + \lambda_4(h)I\{DoB_i \ge 0\}[DoB_i]^2.$$

[26] Consider (7), which is the reduced-form version of (5). There is a corresponding RD graph for each grid point h for a total of 21 graphs (per health index). See here for slides with these graphs:
https://www.dropbox.com/s/pcp7kh9h3rw523t/Blood_Pressure.pdf?dl=0 (blood pressure)
https://www.dropbox.com/s/l97my6vkf2e5asw/Body_Size.pdf?dl=0 (body size)
https://www.dropbox.com/s/l9mbdl58wgvkt7x/Lung_Function.pdf?dl=0 (lung function)

[27] In the bottom panel of Appendix Figures B14-B16 we plot $\hat{\mu}(h)$ against h.

and $\hat{F}_{post}(h)$ against h. For any given h, the vertical distance between $\hat{F}_{post}(h)$ and $\hat{F}_{pre}(h)$ is equal to $\hat{\mu}(h)$.

To estimate the pre-reform CDF $F_{pre}(\cdot)$, we restrict the sample to respondents born before September 1, 1957 and who left school at age 15 or younger (i.e., "the compliers") and estimate equation (6) for each grid point h:

$$\mathrm{I}(Health_i \leq h) = \delta_0(h) + k(DoB_i; h) + \xi_i(h), \quad (6)$$

where the function $k(\cdot\,; h)$ captures *pre-reform* birth cohort trends.[28] We can closely represent compliers born before September 1, 1957 by making this sample restriction because there are very few never-takers in our sample (i.e., individuals who would leave school before age 16 whether they were born before or after September 1, 1957). The coefficient on the constant, $\delta_0(h)$, estimates $\Pr(Health_i \leq h)$ as $DoB \to 0^-$ and thus provides an estimate of $F_{pre}(h)$.

Inference based on the standard errors generated by 2SLS estimates of (5) is problematic because it leads to a large number of highly correlated statistical tests, raising concerns about multiple hypothesis testing. We, therefore, use a single distributional test based on Shen and Zhang (2016) to formally investigate whether education changes the distribution of health. Our test compares the pre- and post-reform CDFs *of the whole population*. Under the assumptions of Shen and Zhang (2016), however, any discontinuity in the CDF of the population necessarily implies that there is a discontinuity in the CDF of compliers. This test is therefore based on the reduced-form specification:

$$\mathrm{I}(Health_i \leq h) = \kappa_0(h) + \kappa_1(h)Post_i + l(DoB_i|h) + \boldsymbol{x}_i'\boldsymbol{\kappa}_2(h) + \eta_i(h). \quad (7)$$

The basis of our test is that—under the null hypothesis of no effect on the health distribution—the function of estimates $\hat{\kappa}_1[h(\tau)]$, where $h(\tau)$ is the value

[28] In our main specification with a quadratic polynomial: $k(DoB_i; h) = \phi_1(h)DoB_i + \phi_2(h)[DoB_i]^2$.

corresponding to the τ^{th} quantile of $Health_i$, is a Brownian bridge (Shen and Zhang 2016).[29] In contrast to Shen and Zhang (2016), which implements a Kolmogorov-Smirnov test, we perform an Anderson-Darling test (Anderson and Darling 1952) using the following weighted integral as our test statistic:[30]

$$T = \int_0^1 \frac{\hat{\kappa}_1[h(\tau)]^2}{\tau(1-\tau)} d\tau. \tag{8}$$

Average treatment effects may not be well-powered to detect effects of education on health that are concentrated on the tails of the health distribution (see Appendix F). We chose the Anderson-Darling test because it is uniformly powered for the whole range $\tau \in [0, 1]$ (Stephens 1974). In contrast, the Kolmogorov-Smirnov test is better powered to detect deviations of the distribution near the median. To test for differences in the bottom half of the health distribution, we use a modified version of (8), integrating only from zero to 0.5. Similarly, we test differences in the top half by integrating from 0.5 to 1.[31]

The p-values for the test are calculated by simulation. Specifically, we generate an independent, standard normally distributed outcome for each individual (such that there is no discontinuous change in distribution at the discontinuity), and evaluate T (or the upper and lower distribution analogue) for this simulated variable. By Shen and Zhang (2016), this is equivalent to drawing from the test statistic distribution under the null. This is repeated 5,000

[29] More precisely, the difference in the empirical CDFs estimated in this way is a standard Brownian bridge times a scalar. See Shen and Zhang (2016) for details on calculating the scalar which allows us to transform the difference into a standard Brownian bridge.

[30] Shen and Zhang (2016) use $\max_\tau |\hat{\kappa}_1[h(\tau)]|$ as their test statistic, which corresponds to a Kolmogorov-Smirnov test. The Kolmogorov-Smirnov test has been shown to be well-powered for deviations in the distribution near the median of the distribution, but is poorly powered to detect differences in the distribution in the tails (Stephens 1974).

[31] In practice, we calculate the integral T numerically, using the approximation

$$T \approx \sum_j \frac{1}{100} \frac{\hat{\kappa}_1\left[h(\tau_j)\right]^2}{\tau_j(1-\tau_j)}$$

where $\{\tau_j\}$ is a set of discrete points in 0.01 unit increments. When testing the full distribution we sum from 0.01 to 0.99, inclusive. For the lower or upper portion of the distribution, we sum from 0.01 to 0.50 or 0.50 to 0.99 inclusive, respectively.

times. As the p-value, we report the fraction of times our simulated values of T are greater than our estimated value of T.

The CDF approach described above is closely related to a quantile IV approach. The CDF approach is based on the vertical distance between the pre- and post-reform CDFs whereas a quantile approach is based on the horizontal distance between these two CDFs. Therefore, either approach would lead us to the same substantive conclusions. We opted to present the CDF approach because it is the framework used by Shen and Zhang (2016), whose results we use in our distributional tests. Nevertheless, when we present our results, we also discuss the effects on some particular quantiles of interest.

4. Distributional Effects of Education on Health

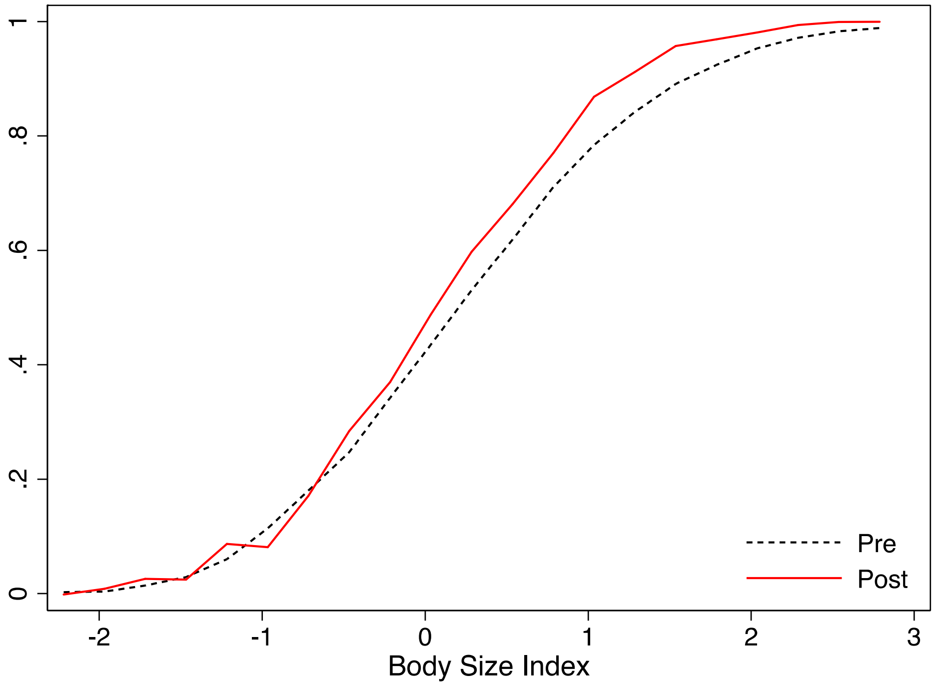
Figure 4 shows the distributional treatment effects of education on body size. It shows the pre- and post-reform CDFs of the body size index for compliers. As explained in Section 3, the pre-reform CDF is obtained by estimating (6) for each grid point h and then plotting $\hat{\delta}_0(h)$ against h.[32] The discontinuity in the CDF, $\mu(h)$, is obtained by estimating (5) through 2SLS for each grid point h. The post-reform CDF at a given grid point h is obtained by adding $\hat{\mu}(h)$ "vertically" to $\hat{\delta}_0(h)$.

Figure 4 shows that education reduces body size: The post-reform CDF is shifted to the left relative to the pre-reform CDF. Importantly, the shift is not parallel; the gains are concentrated at the top of the distribution, among the least healthy. Staying in school until age 16 increases the fraction of study participants with a body size index smaller than 1 standard deviation from 77.5% to 84.4%. Similarly, the 90th percentile of the body size distribution decreases from 1.58 to 1.2 standard deviations. This effect is 2.5 times the

[32] Compliers are less healthy than always takers but not dramatically so. Appendix Figures D4, D5, and D6 compare the pre-reform CDFs for compliers and the whole population (both estimated using equation (5)) for our 3 indices.

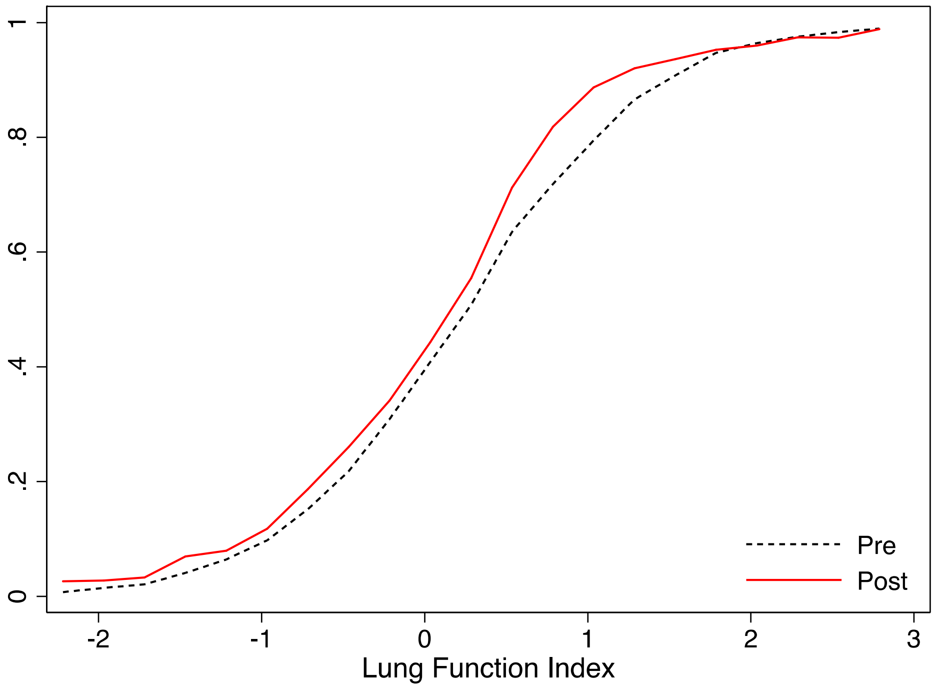
average treatment effect (on the treated) of -0.15 standard deviations estimated in Table 2.

Figure 4: Distributional Effects on Body Size Index



Notes: The figure shows the pre- and post-reform CDFs of the body size index for compliers. $N = 266{,}525$.

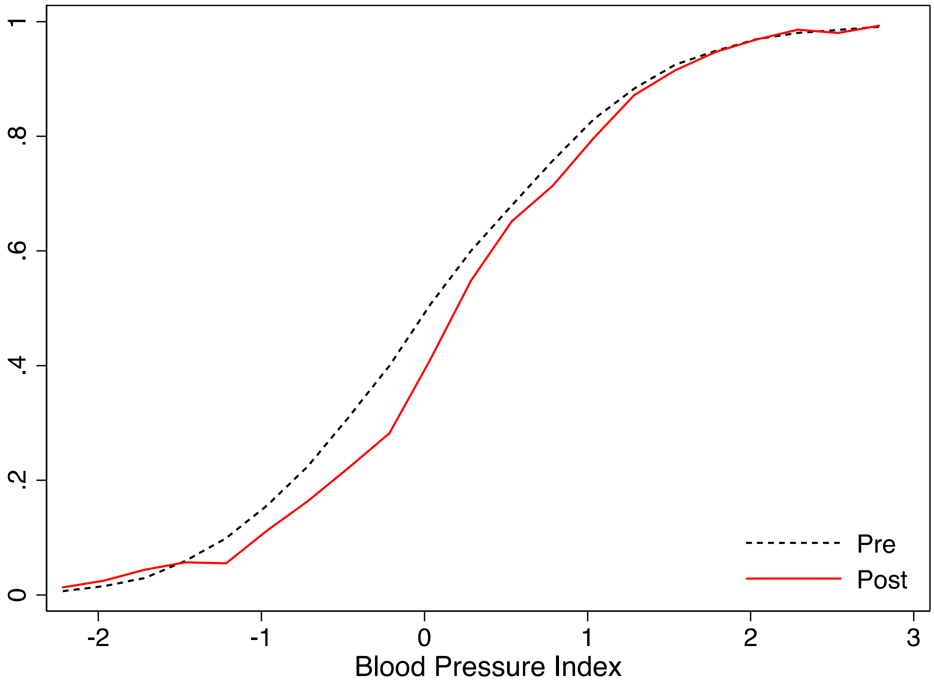
Figure 5: Distributional Effects on Lung Function Index



Notes: The figure shows the pre- and post-reform CDFs of the lung function index for compliers. N = 215,366.

Figure 5 shows that education also improves lung function: The post-reform CDF is shifted to the left relative to the pre-reform CDF. Staying in school until age 16 increases the fraction of study participants with a lung function index smaller than 1 standard deviation from 78.1% to 87.6%. Similarly, the 90th percentile of the lung function distribution decreases from 1.48 to 1.14 standard deviations. This effect is 2 times the average treatment effect (on the treated) of -0.17 standard deviations estimated in Table 2.

Figure 6: Distributional Effects on Blood Pressure Index



Notes: The figure shows the pre- and post-reform CDFs of the blood pressure index for compliers. N = 270,647.

While Figures 4 and 5 show that education improves body size and lung function, Figure 6 shows that education *worsens* one dimension of health: it increases blood pressure.[33] The post-reform CDF lies to the *right* of the pre-reform CDF. Staying in school until age 16 decreases the fraction of study participants with a blood pressure index smaller than 0 from 49.4% to 39.3%. Similarly, the 30th percentile of the blood pressure index distribution increases from -0.49 to -0.16 standard deviations. This effect is 2.2 times the average treatment effect (on the treated) of 0.15 standard deviations estimated in Table 2. This result is particularly striking because blood pressure can be controlled through medication, diet, and exercise (Chobanian et al. 2003), and there is a positive association between education and these healthy behaviors (Park and Kang 2008; Conti, Heckman, and Urzua 2010; Cutler and Lleras-Muney 2010).

[33] The fraction of people with low blood pressure in our sample is negligible; in contrast, 30% are hypertensive (see Figure 8). Therefore we interpret an increase in blood pressure as a worsening of health.

A comparison of Figures 4 and 6 shows that not only the effect on the body size and blood pressure indices have different signs, but the effects also happen in different parts of the respective distributions. While the effects on body size occur in the upper part of the body size distribution, the effects on blood pressure occur in the lower part of the blood pressure distribution among the healthiest. Appendix Figures B5-B10 assess the sensitivity of Figures 5-7 to changing the bandwidth, using linear trends, and the inclusion of controls.

To test whether these shifts in our health indices CDFs are significant and where they are concentrated, we use distributional tests as described in section 3 above. The first row in Table 3 shows p-values of tests of the equality of the pre- and post-reform CDFs. The middle and bottom rows show p-values of tests of the equality of the bottom half (i.e., the healthiest) and the top half (i.e., the least healthy) of pre- and post-reform CDFs.

We can reject the null for the top half of the body size distribution and for the bottom half of the blood pressure distribution (at the 5% significance level). The p-value for the top half of the lung function distribution is 0.0618. Appendix Tables B1-B3 assess the sensitivity of these results to changing the bandwidth, using linear trends, and including controls. While the p-values change across specifications, the main patterns remain: in most cases, we can reject the null of equality for the top half of the body size distribution and for the bottom half of the blood pressure distribution.

Table 3: P-values of Distributional Tests

	Body Size	*Lung Function*	*Blood Pressure*
Full Distribution	0.0896	0.1712	0.0362
Bottom Half	0.9526	0.5962	0.0102
Top Half	0.0126	0.0618	0.1502

Notes: The table shows the p-values of tests of the equality of the full distribution, the bottom and top halves of the pre- and post-reform CDFs.

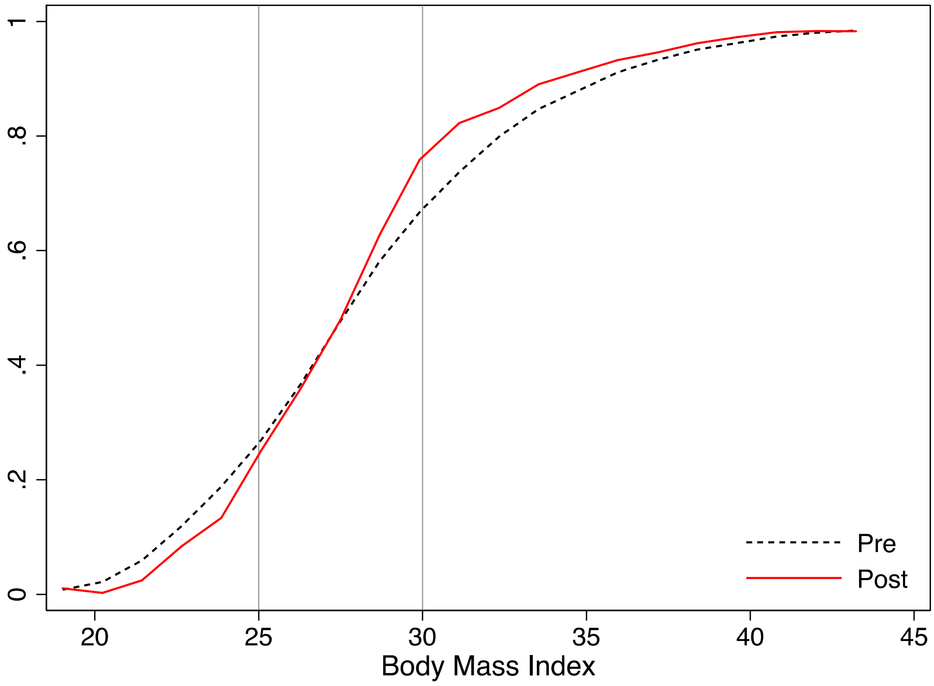
Figures 7 and 8 shed light on these findings by plotting results for measures with clinical thresholds. Figure 7 shows the pre- and post-reform CDFs of body mass index (for compliers). Figure 8 shows the pre- and post-reform CDFs of diastolic blood pressure (for compliers).[34]

Figure 7 shows that the reductions in BMI caused by more education occur where they matter the most: Staying in school until age 16 *reduces* obesity rates (i.e., the fraction of study participants with a BMI below 30) by 7.5 percentage points. In contrast, Figure 8 shows that the increase in blood pressure does not affect the prevalence of stage 2 hypertension (classified as having a diastolic blood pressure above 90 mmHg). Staying in school until age 16 *increases* the probability of stage 1 hypertension (defined as having diastolic blood pressure between 80 mmHg and 90 mmHg; known as "prehypertension" before the 2017 redefinition) by 7.9 percentage points.[35]

[34] Results for systolic blood pressure, omitted due to space constraints, are similar.

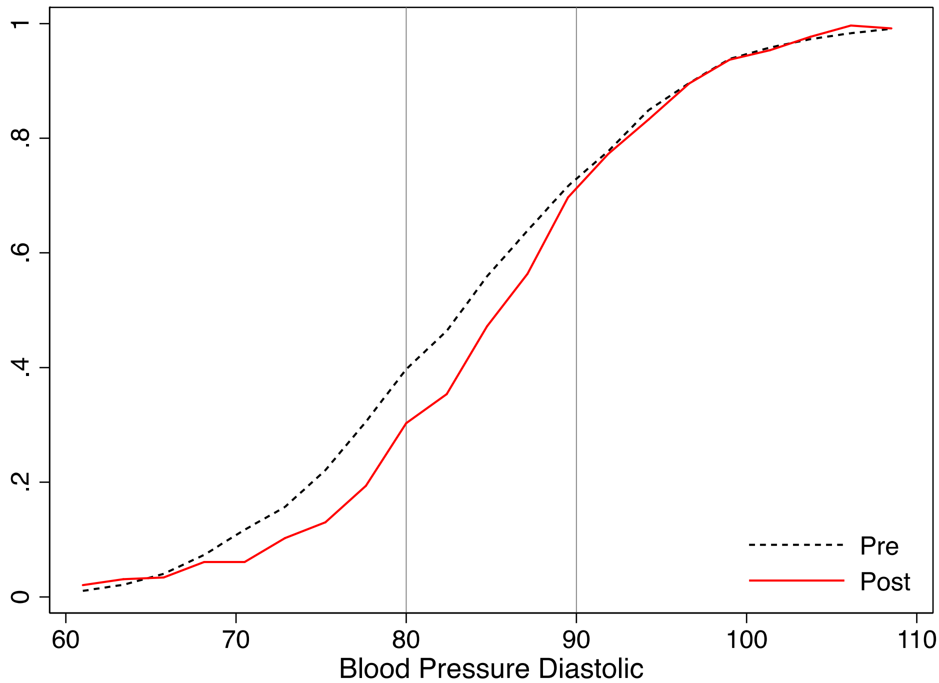
[35] Despite no change in stage 2 hypertension, the increase in the fraction of people in the stage 1 (formerly pre-high) range might still mean a worsening of health, as observational studies indicate that death from both ischemic heart disease (IHD) and stroke increases progressively and linearly from levels as low as 75mmHg DBP (Lewington et al. 2002). In addition, longitudinal data have indicated that DBP between 85 and 89 mmHg are associated with a more than twofold increase in relative risk from cardiovascular disease as compared to those with DBP below 80mmHg (Vasan et al. 2001). However, stage 2 hypertension is more serious than stage 1: while medications are the main treatment for stage 2 hypertension, for stage 1 the focus is on lifestyle changes (Bakris and Sorrentino 2018).

Figure 7: Distributional Effects on Body Mass Index



Notes: The figure shows the pre- and post-reform CDFs of body mass index for compliers. N = 270,019.

Figure 8: Distributional Effects on Diastolic Blood Pressure



Notes: The figure shows the pre- and post-reform CDFs of diastolic blood pressure for compliers. N = 270,647.

These results illustrate the importance of studying distributional effects. While the average treatment effects show improvements in body size and deterioration in blood pressure, the distributional effects reveal in which part of the distributions these changes occur. The deterioration of blood pressure occurs with no observed consequences for the prevalence of stage 2 hypertension. In contrast, the improvements in anthropometrics are concentrated at the right tail, with a large reduction in obesity rates. These effects offer important information that policy-makers might wish to trade-off when considering the health consequences of educational policies.

5. Channels

One of the channels through which education may affect health is health behaviors. Correlational evidence shows that the more educated are more likely to use preventive care, that they manage chronic conditions more effectively, and that they are less likely to smoke and drink heavily (Cutler and Lleras-Muney 2008; Goldman and Smith 2002).

Taking advantage of the richness of the UK Biobank data, we investigate whether education has a causal effect on three types of health behaviors: diet, smoking, and physical activity. Diet was measured using a 24-h dietary assessment tool self-completed through the Internet (Galante et al. 2016).[36] Accelerometers worn for 7 days were used to measure physical activity.[37] Smoking was self-reported.

---

[36] The Oxford WebQ collects information on the quantities of all foods and beverages consumed over the previous day. Respondents are asked whether they consumed any of 21 food groups over the previous day. A positive response results in the screen expanding to reveal a list of commonly consumed foods in the corresponding category. Respondents then select the amount of each food consumed using standard categories to indicate the amount consumed. Energy and nutrient values are generated by multiplying the quantity of each food or drink consumed by its nutrient composition. The Oxford WebQ was included at the assessment visit of the baseline measures for the last 70,724 participants and administered over the Internet to all UK Biobanks participants with a known email address, who were invited to complete the Oxford WebQ on four separate occasions over a 16-month period.

[37] Accelerometer data were collected from May 2013 until December 2015 from 103,720 UK Biobank participants. Our outcome of interest is the average acceleration adjusted for no-wear bias (UKB field 90087): http://biobank.ctsu.ox.ac.uk/crystal/field.cgi?id=90087

Appendix Table E4 shows the effects on diet. Staying in school until age 16 reduces the intake of fat and saturated fat (as a fraction of total energy intake). There are, however, no effects on total caloric intake, sugars, or carbohydrates. Appendix Table E5 shows no effects on the measures of smoking and physical activity we have available.[38]

This analysis suggests that improved diet is an important channel through which education reduces body size. Those who stayed in school until age 16 had better diets in middle age – about 10% lower in fat and 15% lower in saturated fat. Even if the energy content of one's diet is held constant, changes in diet composition can affect body weight (Hall et al. 2012).

The pathways are less clear for the harmful effect of education on blood pressure. We find no effects on a (self-reported) measure of hypertension diagnosis[39] and on (current) blood pressure medication (Appendix Table E5). One alternative hypothesis is that, by changing the types of occupations and careers individuals have, education might have an effect on job responsibilities, expectations, and work-related stress with negative implications for blood pressure.[40] In the U.S. context, for example, academically successful African Americans have higher biomarkers related to cardiometabolic risk (i.e. blood pressure and stress hormones) than other groups (Brody et al. 2013, Miller et al. 2015, Chen et al. 2015). This is potentially driven by stressors related to upward mobility, which could also be playing a role in the U.K context. We have no credible data to test this hypothesis in the UK Biobank so we leave it for future work.

6. Conclusions

[38] These are consistent with Clark and Royer (2013), who also find no effects of the 1972 ROSLA on self-reported smoking or physical activity.

[39] The wording of the question was "Have you been told by a doctor that you have high blood pressure"?

[40] Appendix Table E2 presents evidence that education increases the socioeconomic class of occupations participants hold in middle age. They are less likely to work on semi-routine and lower supervisory occupations, for example.

In this paper, we investigate how education affects the distribution of health along three dimensions: body size, lung function, and blood pressure. We find that an increase in secondary education has important implications for health in middle age. Our results offer three takeaways that add new insights to our understanding about the relationship between education and health.

First, *it is important to consider different health dimensions separately, as education might improve some dimensions while worsening others*. Our findings show that the additional schooling generated by the 1972 ROSLA improved body size while increasing blood pressure. These findings emerge from our mean analysis as well as our distributional analysis. While education seems to have reduced body size through improvements in SES and diet, it is not clear the channels through which education increased blood pressure. Replicating such finding in other contexts and exploring the channels through which education can harm some health dimensions while improving others is an important avenue for future research.

Second, *the focus of the current literature on mean effects can lead to misleading conclusions.* To illustrate, we conducted back-of-envelope calculations of the reduction in mortality implied by either the *average* treatment effect or the *distributional* treatment effects on BMI (see Appendix G for details). Comparing the relative risk of death of an individual with average BMI before and after the reform, we only see a reduction of 0.85%. By contrast, using our estimates of the distributional treatment effects implies a 3.27% reduction in the relative risk of death. This discrepancy is due to the effects being concentrated in parts of the distribution of BMI where the risks of death are higher (Aune et al. 2016).

Even though we cannot estimate treatment effects at the individual level, our results *suggest* that there is substantial variation in the effects across people. While the effects may be large for some, others may not be affected at all. In related work, we present evidence that genetic makeup partly explains why different people are affected differently. Using the same data and natural experiment, we show that the improvements in body size and lung function were larger for individuals with high genetic predisposition to obesity (Barcellos, Carvalho, and Turley 2018).

Third, *the distributional effects of education on health might be different for different health dimensions*. We present evidence that while the improvement in body size is concentrated at the top half of the body size distribution (among the least healthy), the worsening in blood pressure is concentrated at the bottom half of the blood pressure distribution (among the most healthy).

Overall, our results suggest that – *despite the increase in blood pressure* – schooling may be an effective policy tool to improve health. Because the reductions in BMI occur where they matter the most, staying in school until age 16 reduces obesity rates by 7.5 percentage points. In contrast, the increase in blood pressure is concentrated below the clinical threshold for stage 2 hypertension (i.e., diastolic blood pressure above 90 mmHg or systolic blood pressure above 140) with no consequences for prevalence.

REFERENCES

Al Snih, S., Ottenbacher, K.J., Markides, K.S., Kuo, Y.F., Eschbach, K. and Goodwin, J.S., 2007. The effect of obesity on disability vs mortality in older Americans. *Archives of internal medicine*, 167(8), pp.774-780.

Allen et al. 2012: UK Biobank: Current status and what it means for epidemiology

Albouy, Valerie and Laurent Lequien. 2009. "Does compulsory education lower mortality?" *Journal of Health Economics*, 28,155–168.

Anderson, Michael L. 2008. "Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects." Journal of the American statistical Association 103.484, 1481-1495.

Anderson, Theodore W., and Donald A. Darling. 1952. "Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes." The annals of mathematical statistics. 193-212.

Anderson, T.W.; Darling, D.A. 1954. "A Test of Goodness-of-Fit". Journal of the American Statistical Association. 49: 765–769. doi:10.2307/2281537.

Aune, D., Sen, A., Prasad, M., Norat, T., Janszky, I., Tonstad, S., Romundstad, P. and Vatten, L.J., 2016. BMI and all cause mortality: systematic review and non-linear dose-response meta-analysis of 230 cohort studies with 3.74 million deaths among 30.3 million participants. *Bmj*, 353, p.i2156.

Bakris, G., & Sorrentino, M. (2018). Redefining Hypertension—Assessing the New Blood-Pressure Guidelines. *New England Journal of Medicine*.

Barcellos, Silvia Helena, Leandro Carvalho, and Patrick Turley. 2018. "Education can Reduce Health Disparities Related to Genetic Risk of Obesity: Evidence from a British Reform." Unpublished manuscript.

Becker, G.S., Mulligan, C.B. 1997. "The endogenous determination of time preference." *Quarterly Journal of Economics* 112 (August (3)), 729–758.

Bitler, Marianne P., Jonah B. Gelbach, and Hilary W. Hoynes. 2006. "What Mean Impacts Miss: Distributional Effects of Welfare Reform Experiments." *American Economic Review*, 96(4): 988-1012.

Bitler, M. P., Hoynes, H. W., & Domina, T. (2014). Experimental evidence on distributional effects of Head Start (No. w20434). National Bureau of Economic Research.

BLACK, D.A., J.A. SMITH, M.C. BERGER AND B. J. NOEL (2003): "Is the threat of reemployment services more effective than the services themselves? Experimental evidence from the UI system," *American Economic Review* 93(3), 1313–1327.

Brody, G. H., Yu, T., Chen, E., Miller, G. E., Kogan, S. M., & Beach, S. R. (2013). Is resilience only skin deep?: Rural African Americans' socioeconomic status-related risk and competence in preadolescence and psychological adjustment and allostatic load at age 19. *Psychological Science, 24,* 1285–1293. http://dx.doi.org/10.1177/0956797612471954

Chen, E., Miller, G. E., Brody, G. H., & Lei, M. K. (2015). Neighborhood poverty, college attendance, and diverging profiles of substance use and allostatic load in rural African American youth. *Clinical Psychological Science, 3,* 675–685.http://dx.doi.org/10.1177/2167702614546639

Chobanian, A. V., Bakris, G. L., Black, H. R., Cushman, W. C., Green, L. A., Izzo Jr, J. L., ... and Roccella, E. J. 2003. "The seventh report of the joint national committee on prevention, detection, evaluation, and treatment of high blood pressure: the JNC 7 report." **Jama**, 289(19), 2560-2571.

Clark, Damon and Heather Royer. 2010. "The Effect of Education on Adult Health and Mortality: Evidence from Britain." NBER Working Paper No. 16013.

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.

Conti, G., Heckman, J., and Urzua, S. 2010. "The education-health gradient." *The American economic review*, 100(2), 234.

Conti, G., Heckman, J. J., & Urzua, S. (2011). Early endowments, education, and health. Unpublished manuscript, University of Chicago, Department of Economics.

Cutler, David M., Adriana Lleras-Muney, and Tom Vogl. 2008. Socioeconomic Status and Health: Dimensions and Mechanisms. Working Paper 14333. http://www.nber.org/papers/w14333

Cutler, David M., and Adriana Lleras-Muney. 2010. "Understanding differences in health behaviors by education." *J Health Econ*. 29(1):1-28. PMCID: PMC2824018

Cutler, David M., and Adriana Lleras-Muney. 2010. "The education gradient in old age disability." *Research Findings in the Economics of Aging.* Chicago: University of Chicago Press, 101-120. http://www.nber.org/chapters/c8194

Davies, Neil M., Matt Dickson, George Davey Smith, Gerard van den Berg, and Frank Windmeijer. 2016. "The Causal Effects of Education on Health, Mortality, Cognition, Well-being, and Income in the UK Biobank." bioRxiv doi: https://doi.org/10.1101/074815

De Matteis, S., Jarvis, D., Hutchings, S., Darnton, A., Fishwick, D., Sadhra, S., ... & Cullinan, P. (2016). Occupations associated with COPD risk in the large population-based UK Biobank cohort study. *Occupational and environmental medicine*, oemed-2015.

Dickson, M. and Smith, S., 2011. What determines the return to education: An extra year or a hurdle cleared?. *Economics of education review*, *30*(6), pp.1167-1176.

DJEBBARI, H. AND J.A. SMITH (2008): "Heterogeneous impacts in PROGRESA," *Journal of Econometrics* 145(1), 64–80.

FIRPO, S. (2007): "Efficient semiparametric estimation of quantile treatment effects," *Econometrica*, 75(1), 259–276.

Fogel, R.W., 1994. Economic Growth, Population Theory, and Physiology: The Bearing of Long-Term Processes on the Making of Economic Policy. *American Economic Review*, 84(3), pp.369-395.

Galante, J., Adamska, L., Young, A., Young, H., Littlejohns, T. J., Gallacher, J., & Allen, N. (2016). The acceptability of repeat Internet-based hybrid diet assessment of previous 24-h dietary intake: administration of the Oxford WebQ in UK Biobank. *British Journal of Nutrition*, 115(04), 681-686.

Gelman, Andrew and Guido Imbens. 2016. "Why high-order polynomials should not be used in regression discontinuity designs." Unpublished Manuscript.

Grenet, Julien. 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.

Goldman, D.P., Smith, J.P. 2002. "Can patient self-management help explain the SES health gradient?" *Proceedings of the National Academy of Science* 99 (16), 10929–10934.

Grossman, Michael. 2015. "The Relationship between Health and Schooling: What's New?" No. w21609. National Bureau of Economic Research. http://www.nber.org/papers/w21609

Hall, Kevin D., Steven B. Heymsfield, Joseph W. Kemnitz, Samuel Klein, Dale A. Schoeller, and

John R. Speakman. 2012. “Energy Balance and its Components: Implications for Body Weight
Regulation—Consensus Statement from the American Society for Nutrition.” *American Journal
of Clinical Nutrition*, 95, 989–994.

HECKMAN, J.J., J. SMITH AND N. CLEMENTS (1997): “Making the most out of programme evaluations and social exper-iments: Accounting for heterogeneity in programme impacts,” *Review of Economic Studies* 64(4), 487–535.

JACKSON, R., & BEAGLEHOLE, R. (1985). Secular trends in underreporting of cigarette consumption. *American journal of epidemiology*, *122*(2), 341-344.

Janke, K., Johnston, D.W., Propper, C. and Shields, M.A.. 2018. “The Causal Effect of Education on Chronic Health Conditions.” Unpublished manuscript.

Jensen, R., & Lleras-Muney, A. (2012). Does staying in school (and not working) prevent teen smoking and drinking?. *Journal of Health Economics*, 31(4), 644-657.

Johnston, David W., Carol Propper, and Michael A. Shields. 2009. "Comparing subjective and objective measures of health: Evidence from hypertension for the income/health gradient." *Journal of health economics* 28.3, 540-552.

Jürges, Hendrik, Eberhard Kruk and Steffen Reinhold. 2013. “The effect of compulsory schooling on health—evidence from biomarkers.” *J Popul Econ*, 26:645–672.

Kemptner, Daniel, Hendrik Jürgesb and Steffen Reinholdc. 2011. “Changes in compulsory schooling and the causal effect of education on health: Evidence from Germany.” *Journal of Health Economics*, 30, 340–354.

Jensen, R. and Lleras-Muney, A., 2012. “Does staying in school (and not working) prevent teen smoking and drinking?” *Journal of Health Economics*, 31(4): 644-657.

Lee, D. S. (2008). Randomized experiments from non-random selection in u.s. house elections. *Journal of Econometrics*, 142 675 – 697. The regression discontinuity design: Theory and applications

Lee, D.S. and Card, D. 2008. “Regression Discontinuity Inference with Specification Error.” *Journal of Econometrics*, *142*(2): 655-674.

Lee, David S., and Thomas Lemieux. 2010. "Regression Discontinuity Designs in Economics." *Journal of Economic Literature*, 48(2): 281-355.

Lewington S, Clarke R, Qizilbash N, Peto R, Collins R. Age-specific relevance of usual blood pressure to vascular mortality: A meta-analysis of individual data for one million adults in 61 prospective studies. Prospective Studies Collaboration. *Lancet* 2002;360:1903-13.

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

Månsson, N.O., Eriksson, K.F., Israelsson, B., Ranstam, J., Melander, A. and Råstam, L., 1996. Body mass index and disability pension in middle-aged men—non-linear relations. *International journal of epidemiology*, 25(1), pp.80-85.

Meghir, C., M. Palme and E. Simeonova, forthcoming, “Education, Health and Mortality: Evidence from a Social Experiment”, American Economic Journal: Applied Economics.

Miller, G. E., Yu, T., Chen, E., & Brody, G. H. (2015). Self control forecasts better psychosocial outcomes but faster epigenetic aging in low-SES youth. *Proceedings of the National Academy of Sciences of the United States of America, 112,* 10325–10330. http://dx.doi.org/10.1073/ pnas.1505063112

Office of National Statistics. 2005. The National Statistics Socio-Economic Classification User Manual. http://webarchive.nationalarchives.gov.uk/20160105160709/http://www.ons.gov.uk/ons/guide-method/classifications/current-standard-classifications/soc2010/soc2010-volume-3-ns-sec--rebased-on-soc2010--user-manual/index.html

Park, C., and Kang, C. 2008. “Does education induce healthy lifestyle?” *Journal of Health Economics*, 27(6), 1516-1531.

Perez-Arce, Francisco. 2017. “The effect of education on time preferences.” *Economics of Education Review.* Vol. 56 (2), 52-64.

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.
http://www.jstor.org/stable/10.1086/657020

Shen, Shu, and Xiaohan Zhang. 2016. "Distributional tests for regression discontinuity: Theory and empirical examples." Review of Economics and Statistics 98.4: 685-700.

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

Stephens, Michael A. "EDF statistics for goodness of fit and some comparisons." *Journal of the American statistical Association* 69.347 (1974): 730-737.

Sudlow, Cathie, John Gallacher, Naomi Allen, Valerie Beral, Paul Burton, John Danesh, Paul Downey, Paul Elliott, Jane Green, Martin Landray, Bette Liu, Paul Matthews, Giok Ong, Jill Pell, Alan Silman, Alan Young, Tim Sprosen, Tim Peakman, Rory Collins. 2015. UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age. *PLoS Med* 12(3): e1001779. PMCID:PMC4380465

Thomas, P. S., Cowen, E. R., Hulands, G., & Milledge, J. S. (1989). Respiratory function in the morbidly obese before and after weight loss. *Thorax*, *44*(5), 382-386.

Vasan RS, Larson MG, Leip EP, Evans JC, O'Donnell CJ, Kannel WB, et al. Impact of highnormal blood pressure on the risk of cardiovascular disease. *N Engl J Med* 2001;345:1291-7.