I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 11353

The Causal Effect of Education on Chronic Health Conditions

Katharina Janke
David W. Johnston
Carol Propper
Michael A. Shields

FEBRUARY 2018

I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 11353

The Causal Effect of Education on Chronic Health Conditions

Katharina Janke
Lancaster University

David W. Johnston
Monash University

Carol Propper
Imperial College London

Michael A. Shields
Monash University and IZA

FEBRUARY 2018

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

The Causal Effect of Education on Chronic Health Conditions

Studies using education policy reforms to isolate causal effects of education on health produce mixed evidence. We analyse an unusually large sample and study chronic health conditions. For identification, we use two major education reforms, one that raised the minimum school leaving age and one that affected the broader educational attainment distribution. This method generated precise estimates of the impact of education on a comprehensive range of health conditions. Our results indicate that extra education, at the lowest end or higher up the attainment distribution, has little impact on the prevalence of chronic illness. The one interesting exception is diabetes.

JEL Classification: I14, I24, I26

Keywords: education reform, health conditions, causality

Corresponding author:
Carol Propper
Business School
Imperial College London
South Kensington Campus
London SW7 2AZ
United Kingdom
E-mail: c.propper@imperial.ac.uk

I. Introduction

There is a strong positive correlation between education and health at all ages. Many models predict such an association, because the same observed and unobserved factors that determine educational attainment (e.g. parental socioeconomic position, genetic ability, discount rates) also determine health at birth and the subsequent returns to investments in health. For this reason, establishing causality in the education-health relationship has been a major focus of research. One prominent strand in this literature identifies policy reforms that have increased education, and then traces through the effect of these increases on various measures of health and health behaviours.

The most commonly used education policy reform is the increase in the minimum school leaving age (MSLA), which has been progressively introduced in many countries.[1] The results from this literature are very mixed (see Galama et al., 2018). Some studies have found that the additional schooling induced by these reforms drives significant improvements in health or health behaviours (e.g. Lleras-Muney, 2005; Oreopoulos, 2006; Siles, 2009; Powdthavee, 2010; van Kippersluis et al., 2011; Kemptner et al., 2011; Banks and Mazzonna, 2012; Brunello et al., 2013, 2016; Crespo et al., 2014; Fischer et al., 2013; Fletcher, 2015; Li and Powdthavee, 2015). Other studies, sometimes using the same data but different model specifications, have found little or no evidence of a causal effect (Mazumder, 2008; Albouy and Lequien, 2009; Lindeboom et al., 2009; Jürges et al., 2013; Clark and Royer, 2013; Johnston et al., 2015; Siles, 2015; Meghir et al., 2017).[2] The studies use a variety of measures of ill-health, including mortality, biomarkers of chronic health conditions, and self-assessed health measures, but this does not appear to account for the mixed findings. Further, the finding from a number of studies of no effect of education on health is all the more striking because these reforms have been shown to have significantly increased wages (for UK evidence see Harmon and Walker, 1995; Oreopoulos, 2006; Grenet, 2013). Thus, the results from some studies imply that income buys little health (Clark and Royer, 2013).[3] These results stand, therefore, in stark contrast to the conventional wisdom, and the assumptions of many economic models (e.g. Grossman, 1972), that both income and education are significant drivers of health outcomes.

We advance the literature by examining a large range of chronic health conditions, a very large sample, and educational reforms that increase education both at the bottom end and further up the

[1] Evidence from such reforms has now been provided for many developed countries including Australia, Canada, Denmark, France, Germany, Italy, the Netherlands, Norway, Sweden, the UK, and the US.

[2] To add to the uncertainty, many studies also find different results by gender, which is difficult to reconcile with efficiency-type arguments from models of investment in human capital (e.g. Grossman, 1972). For example, see: Kemptner et al., 2011; Banks and Mazzonna, 2012; Brunello et al., 2013; Gathmann et al., 2015).

[3] Similarly, a number of recent studies have found that increased income brings little in the way of adult or child health in developed countries (see, for example, Cesarini et al., 2016; Muris, 2017).

educational distribution, to provide evidence on the causal link between education and health. This unified approach builds on the existing literature in three important ways, and addresses some of the limitations of the literature noted by Galama et al. (2018).

First, most of the existing literature exploits schooling reforms that raise the minimum school leaving age. These reforms target those at the bottom end of the educational distribution. Furthermore, the dates at which the MSLA reforms occurred in high-income countries means that affected individuals are now approaching their 60s and 70s. This often means that we know much less about the effect of education on younger people (Clark and Royer, 2013). We rectify this issue by exploiting two reforms within the same country (the UK) that affected individuals who are now primarily in their 40s and 50s. The first of these is the raising of the MSLA from age 15 to 16. This reform was introduced on the 1st September 1972 and impacted on cohorts born in or after September 1957. It affected around 25 percent of individuals, all at the lower end of the educational attainment distribution. The second reform is the large expansion in British educational attainment that occurred around 1990, driven by an expansion in the provision of higher education and a change in the qualification systems allowing more 18 year olds to attend further and higher education. This affected the opportunities of those born from around 1970 (Devereux and Fan, 2011). Using the two reforms allows us to examine the effect of extra education both at the bottom end and higher up the distribution. The second reform also allows us to look at outcomes at younger ages than are generally studied.

A second contribution is that we explore a much wider definition of health than has generally been undertaken, by examining a range of chronic health conditions. Chronic diseases are conditions for which there is currently no cure, and which are managed with drugs and other treatment. Examples include diabetes, chronic obstructive pulmonary disease, arthritis and hypertension. Treatment of these conditions accounts for a large and growing share of healthcare budgets in all developed countries.[4] The dataset we use contains information on 17 different chronic conditions, including common muscular-skeletal conditions, respiratory conditions, heart conditions, diabetes, and mental health conditions. Examination of specific chronic health conditions also allows us to examine whether education has greater impact for conditions that have been strongly linked to lifestyle choices.

Third, we employ a much larger sample than has previously been used to study morbidity. The full data set includes around 5 million observations from 1.5 million individuals, and therefore provides a large subset of observations proximal to the relevant reform dates. This size allows us to

[4] A recent report for the UK estimated that individuals with long-term conditions now account for about 50 per cent of all family doctor appointments, 64 per cent of all outpatient appointments and over 70 per cent of all inpatient bed days. Treatment and care for people with long-term conditions is estimated to take up around £7 in every £10 of total health and social care expenditure (Department of Health 2012).

provide precise estimates of the causal effect of education for specific chronic health conditions, whilst implementing tight bandwidths, thus increasing the likelihood that we are identifying causal effects. We are also able to test robustness by several means, and to allow the effect of education to change with age.

Using this large sample and exploiting two different reforms, we find robust evidence that increased educational attainment does not reduce the probability of chronic illness, with one interesting exception. The finding of no substantive causal effect of education holds despite significant increases in wages from both reforms. The one exception is diabetes, a disease which is highly linked to life-style and which is increasing rapidly worldwide.

II. UK Evidence

The existing literature from the UK uses a diverse range of data sources, measures of health, sample sizes and methodologies. The apparent consequence is a very mixed set of findings for the effect of education on chronic conditions. Siles (2009) simultaneously exploits the two changes in the MSLA that occurred in the UK in 1947 (age 14 to 15) and 1972 (age 15 to 16), and uses pooled data from the 1980-2004 General Household Surveys (GHS) on around 200,000 individuals aged 25-60 (born 1915 to 1979). She finds that an extra year of education reduces the probability of reporting a long-term illness by 5.5 percentage points. Several other studies use pooled data from the Health Survey for England (HSE), which even if pooled over many years, has a limited sample for the cohorts around the two reform dates. For the earlier of the two MSLA reforms, Powdthavee (2010) finds that an extra year of schooling significantly reduces the probability of adulthood hypertension, but finds no effect using the 1972 reform. Jürges et al. (2013), using HSE data from 1993 to 2006, find no significant effects on biomarkers relevant for heart disease (C-reactive protein and blood fibrinogen). More recently, Davies et al. (2018) look at the effect of the 1972 reform on a wide range of health measures, using 22,000 respondents from the UK Bio-Bank who turned 15 in the year before or after the reform. Notably, individuals volunteer or self-select into the Biobank, and are more likely to be highly educated. The results are mixed across the measures, but consistent evidence is found that staying in school reduces the risk of diabetes and mortality.

The only paper that we are aware of that has used the second reform (the large expansion in British educational attainment in the 1990s) to identify the effect of education on morbidity is James (2015). He uses HSE data for 1991-2012, resulting in small sample sizes and imprecisely estimated effects. He finds that an additional year of education reduces the probability of having an early adulthood limiting illness by 5.7 percentage points (the average age of the sample is 34). Education is also found to significantly reduce the probability of being overweight. In contrast, estimates for self-assessed general health, having a long-standing illness, and having hypertension are statistically insignificant.

The closest paper to our analysis, and the paper with the most robust design and expansive analyses is Clark and Royer (2013). Using two sets of data, they also examine the two MSLA reforms that occurred in the UK. The first data set is from the 2001 Census, which is linked to data on all-cause mortality, and which contains two self-reported measures of health: a self-assessed rating of health as good, fair or bad, and a binary indicator of having a limiting chronic condition. The Census is very large but contains no measure of education, and so from these data they provide reduced-form estimates of the effect of education on mortality and self-assessed health. They complement these estimates with 2SLS estimates using pooled survey data from the General Household Survey (1986–1996) and the HSE (1991–2004), both of which contain education and several measures of health. The pooling of data results in relatively large sample sizes of up to 200,000 individuals; though the cohorts span a long-time period, including individuals born between January 1920 and December 1969. They find no effect of education using the Census data, and some contradictory and imprecisely estimated results using GHS and HSE data. But overall, they conclude that the sharp increases in educational attainment in Britain in 1947 and 1972 led to small and statistically insignificant changes in health.

It appears that there is little consensus for the UK on the effect of education on morbidity even using the same data and the same set of reforms, a finding that is similar to that for other countries and other reforms. This can be attributed to differences in sample size, methodology and possibly also the use of different survey instruments. There is therefore a need to provide evidence using larger samples and for a set of outcomes which are derived from individuals' responses to a single survey instrument. Large samples allow more precision while the use of a single survey eliminates any systematic differences in reporting that arise from different methods of asking questions about health.

III. The Data

To gain precise estimates of the causal effects of increased education on chronic health conditions, we require observations on a very large nationally representative sample of UK-born individuals, together with detailed measures of heath conditions. We use the UK Quarterly Labour Force Survey (QLFS), pooling the data over 15 years (2001q3-2015q4). The QLFS is a nationally representative sample of households in the UK. It has a longitudinal element, re-interviewing respondents for up to 5 quarters.[5]

[5] Our estimates are robust to using only the first observation per individual (see Appendix Table A2).

The data provides a total sample of approximately 5 million observations and 1.5 million individuals.[6] When we limit our samples to narrow age ranges around the reforms, we retain around 300,000 observations per reform. Table 1 provides sample sizes, birth years, survey years, and summary statistics for each reform.

While the QLFS is not a health survey, it collects detailed information from respondents about chronic health conditions. Respondents are asked, "Do you have any health problem or disability that you expect will last for more than a year?" If they answer "yes", they are given a list of 17 types of conditions of which they can select any number. We use the responses to the first question as our measure of having any chronic condition and focus on the 10 most common conditions on the list of 17.[7] We also sum the responses to get a count of conditions. The conditions and sample means are shown in Tables 2 and 3.[8]

Using self-reported conditions as the object of analysis raises the issue of reporting error. To self-report, individuals have to know they have the condition and be willing to report it. However, as the literature suggests that better educated individuals are more likely to recognise they have a condition and report it (for example, Johnston et al., 2011), such reporting error will bias our estimates away from finding a negative relationship of education on chronic illness.

IV. Reform 1: Raising the school leaving age

In September 1972, the minimum school leaving age in Britain was raised from 15 to 16 years. This forced a significant proportion of students who previously would have left school at age 15 to attend for one extra year. Figure 1 demonstrates the effect of the reform by displaying the proportion of students leaving school at age $\geq$ 16 years for each month-year of birth. To estimate the effects of the additional schooling induced by this reform, we use a fuzzy regression discontinuity framework, largely following Clark and Royer (2013). This approach involves the estimation of a two-stage least squares model, using as the instrumental variable an indicator that the individual was born after 1st September 1957 (turned 15 years old after 1st September 1972).

In the first-stage equation we estimate the effects of the reform on educational attainment:

[6] We cluster the standard errors in all models by individual.

[7] The conditions we exclude have low prevalence in our combined sample of 30-60 year olds: vision (1.5%), hearing (2.1%) and speech problems (0.3%), epilepsy (0.8%), learning difficulties (0.6%), and progressive illness, such as M.S. (1.8%). We also exclude a catch-all 'Other condition'. The ranking of the prevalence of conditions in our QLFS sample broadly match the ranking of prevalence estimates for the English population.

[8] Respondents are also asked to identify their 'main' condition from the list of 17 specific conditions. We do not use this as our primary measure as identifying the main condition is difficult when some diseases affect several physical functions.

$$E_{ict} = \alpha_0 + \alpha_1 D_{ic} + f(R_{ic}) + \mathbf{X}'_{ict}\alpha_2 + \varepsilon_{ict} \qquad (1)$$

where E_{ict} is the age at which individual i born in cohort c surveyed at time t completed full-time education. The covariate set includes a binary variable indicating whether the individual was born after 1st September 1957 (D_{ic}), a 'running' variable measuring month-year of birth (R_{ic}), and a vector of exogenous characteristics ($\mathbf{X}_{ict}$). This vector includes a third-order polynomial in age, and dummy variables for gender, year of survey (2002-2015), quarter of survey, month of birth (Jan-Dec), and interactions between month-of-birth dummies and being born after 1st September 1957.

We use a sample of individuals born up to 30 months before and up to 30 months after the 1st September 1957 birth date threshold (March 1955 to February 1960). This narrow window is made feasible by the large sample size of the QLFS. The function $f(R_{ic})$ in equation (1) is a linear function of month-year of birth R_{ic} with different slopes on either side of the birth date threshold (local linear approach). Given the use of a narrow bandwidth, this function should adequately capture any relationship between month-year of birth and educational attainment (and later life health outcomes). Our estimates are comparable if $f(R_{ic})$ is instead a high-order polynomial function (not presented here).

The estimated effect of the change in the minimum school leaving age on educational attainment – α_1 from equation (1) – equals 0.344 (F-statistic = 21.4). This estimate implies that the reform increased average years of education by 0.344 years. As shown in Appendix Table A1 this increase in years of education was generated primarily by a 23 percentage point increase in the proportion of students leaving school at age ≥ 16 years. Appendix Table A1 also shows that the proportion of students obtaining General Certificate of Education Ordinary Levels (O-levels) increased by 4.6 percentage points. In contrast, the reform had near zero effects on leaving school at age ≥ 17 years and obtaining A-level qualifications. Therefore, we can be confident the reforms impacted upon those individuals that were its intended focus.

In the second-stage equation we estimate the effect of educational attainment (defined as years of full-time education) on health:

$$H_{ict} = \beta_0 + \beta_1 E_{ict} + g(R_{ic}) + \mathbf{X}'_{ict}\beta_2 + u_{ict} \qquad (2)$$

where H_{ict} is a binary variable representing a chronic health problem of any type or a particular chronic health problem. As in equation (1), $g(R_{ic})$ is a is a linear function of month-year of birth with different slopes on either side of the threshold.

Table 2 presents our first set of results, the estimated effects of educational attainment on the likelihood of reporting any chronic health problem or disability and the likelihood of reporting particular chronic conditions. We present OLS, reduced form (RF) and IV estimates. The OLS estimates suggest that education has large positive effects on health. All the OLS estimates are statistically significant and some are large in relation to the sample mean. For example, increasing educational attainment by one year is estimated to reduce the likelihood of reporting any chronic health problem or disability by 7.2 percentage points, relative to a sample mean of 40%. For depression, the OLS estimate is 3.5 percentage points compared to a mean of 7%.

The corresponding IV estimates are generally smaller in magnitude and not statistically significant at conventional levels. The IV estimate for reporting any chronic health problem or disability is 1.2 percentage points, 6 times smaller than the OLS estimate. An even more substantive decline is found for the estimates on the number of conditions (0.317 to -0.005). This lack of effect can also be seen in Figure 1, which plots the proportion of respondents with any chronic condition against year of birth for the estimation sample, and shows no clear break at the school leaving age. However, as the IV estimates are less precisely estimated, we cannot rule out large education effects: the 95% confidence interval for any chronic health problem or disability equals [-.067, .091]).[9]

The one statistically significant IV estimate, and the largest effect relative to the sample mean, is for diabetes. This estimate suggests that increasing educational attainment by one year reduces the likelihood of reporting diabetes by 5.8 percentage points (95% CI = [-.099, -.015]). This is 10 times the size of the OLS estimate and is large in relation to the sample mean of 4.6 percent. This effect remains large and statistically significant when only the first observation per respondent is used (see Appendix Table A2). There is also a large and statistically significant education effect on the binary indicator that diabetes is the respondent's 'main' chronic condition (coefficient equals -0.035 with standard error 0.015).

The large size of our sample allows us to examine robustness to different bandwidths and to estimate separate effects for men and women. Our estimates are little changed when the bandwidth is narrowed (see Appendix Table A3). Results for men and women are very similar: small and statistically insignificant effects for the overall measure and for most specific chronic conditions, but a negative effect of education on diabetes, which is large and statistically significant at conventional levels for women (coefficient equals -0.068 with standard error 0.024) though not for men (coefficient equals -0.036 with standard error 0.044).

[9] Our findings for a health problem or disability of any type in Table 2 are similar to those in Clark and Royer (2013) for a limiting long-term illness in their Table 5A, Panel B. Their IV estimate for this outcome, using a bandwidth of 91 months, is 0.000 (se = 0.038).

V. Reform 2: Education Expansion

During the early 1990s a large increase in educational attainment occurred in Britain. Between 1988 and 1994 the proportion of young people entering full-time higher education institutions increased from around 15% to 33% (Devereux and Fan, 2011). There were several simultaneous drivers of this expansion. First, a large increase in the supply of degree-level places occurred when the Further and Higher Education Act 1992 enacted changes in higher education funding and administration, allowing 35 polytechnic institutions to become universities and offer degree-level qualifications. Second, higher education institutions were incentivised to increase enrolment by the Government's decision to reduce the amount of money paid to Universities per student while simultaneously relaxing limits on student recruitment. Third, in 1986 the school leaving (age 16) qualifications, known as Certificates of Secondary Education and O-Levels, were replaced by the single General Certificates of Secondary Education. This change included changes in performance assessment and grading and had the effect of increasing the proportion of students attaining high-grade passes. This, in turn, encouraged students to stay in school beyond the compulsory leaving age of 16 (Blanden and Machin, 2004). Therefore, this set of reforms affected both those who might have left school at the MSLA of 16 and those who progressed to finishing high school but who would not, before the reform, have entered tertiary education.

This increase in participation in higher education can be seen in the increase in the average age at which QLFS survey respondents aged ≥ 30 completed full time education as shown in Figure 1. This Figure, and others like it in the literature[10], demonstrate the large growth in educational attainment that began with cohorts born around 1970 and ended with cohorts born around 1976.

We exploit this growth to estimate the effect of education on health by modelling the relationship between the health of individual *i* born in cohort *c* surveyed at time t (H_{ict}) and the average educational attainment (age left full-time education) of i's cohort $(\overline{EA}_c)$:

$$H_{ict} = \gamma_0 + \gamma_1 \overline{EA}_c + a_{ct} + h(c) + \mathbf{W}'_{ict}\gamma_2 + \nu_{ict} \qquad (3)$$

where a_{ct} are age fixed-effects, $h(c)$ is a linear function of year of birth (cohort), and $\mathbf{W}_{ict}$ includes dummy variables for gender, year of survey, quarter of survey, month of birth, and wave number.[11]

[10] See Blanden and Machin (2004), Devereux and Fan (2011), Machin et al. (2012), and James (2015).

[11] This approach is analogous to the methodology used by Bedard and Deschênes (2006) in their evaluation of the long-term health impacts of military service. In their study, birth-cohort-specific percentage of veterans is the main variable of interest, rather than birth-cohort-specific educational attainment, as in equation (3).

We also include the unemployment rate experienced by different cohorts at age 18, which controls for potential correlation between labour market conditions and the education expansion.[12]

The parameter of primary interest in equation (3), γ_1, is identified from across cohort covariation in educational attainment and health. The validity of our modelling approach therefore depends on whether there exist important cohort health effects that are correlated with the across-cohort growth in educational attainment, after accounting for age and time fixed effects, a cohort-specific unemployment rate, and a linear cohort trend. The cohort trend term will control for observable and unobservable time-invariant cohort-specific predictors of health under the assumption that these factors evolve linearly across cohorts (Bedard and Deschênes, 2006). We also limit the potential for confounding cohort effects by restricting the included birth cohorts to a relatively narrow 10-year window (1966 to 1975). Notably, $\hat{\gamma}_1$ is little changed when we exclude the cohort trend term, although standard errors are smaller.

The estimated education effects for Reform 2 are presented in Table 3. The patterns are very similar to those for Reform 1, although the magnitudes of both sets of estimates are smaller than those in Table 2, driven by the sample's younger age and thus lower frequency of chronic health conditions.[13] The estimated OLS coefficients again clearly suggest that education improves health while the reform-based estimates are generally smaller in magnitude than the OLS estimates and less precisely estimated. The lack of relationship between education and having any chronic condition is also shown in Figure 1, which shows the proportion of respondents with any chronic condition against age for the estimation sample. The Figure shows no clear evidence of a break at any point in the sample window.

However, there is one reform-based estimate that is both larger than the OLS estimate and statistically significant at conventional levels. This is again for diabetes. It suggests that a one-year increase in education reduces diabetes risk by 0.8 percentage points compared to the mean of 1.9%. This estimate is robust to widening the sample window of birth cohorts and to using an indicator that the respondent's main condition is diabetes.[14] Broken down by gender there are no statistically significant associations between education and chronic conditions, though the estimates for diabetes are negative for both genders.

[12] The average educational attainment of i's cohort could be used as an instrumental variable for i's educational attainment in a 2SLS model of health outcomes. The 2SLS estimated education effects from such a model are nearly identical to the estimates of γ_1 in equation (3).

[13] The reduced form (RF) estimates in Table 3, which are generated from equation 3, should be compared to the IV estimates in Table 2.

[14] Estimated effects for 'main condition is diabetes' equals -0.006 with standard error 0.003.

VI. Further examination of diabetes

Diabetes is a disorder characterised by chronic high blood glucose levels (hyperglycemia) due to the body's failure to produce any or enough insulin to regulate high glucose levels. Diabetes and its related complications can have large negative effects on health. For example, it can cause health complications in limbs, eyes, internal organs and premature death (Public Health England, 2014). There are two main types of diabetes. Type 1 is caused by the body's inability to produce insulin and typically occurs in childhood or adolescence. Type 2 is caused by the body's inability to react properly to insulin, and is often diagnosed after age 40. Type 2 diabetes is more prevalent than Type 1 diabetes, and currently accounts for 85-90% of all diabetes cases in the UK and USA.

The world-wide escalation of Type 2 diabetes is linked to the rise in obesity; estimates from the UK indicate that around 85 percent of people with type 2 diabetes are overweight or obese (Public Health England, 2014). Notably, Type 2 diabetes is largely preventable: small amounts of weight loss, in the range of 5-10%, can prevent or delay the development of Type 2 diabetes among high-risk populations, and reductions in weight can also reduce the severity of Type 2 diabetes after its onset (Public Health England, 2014). Sustained weight loss of around 15 kg can even reverse Type 2 diabetes (McCombie et al., 2017).

Our finding that education lowers the probability of having diabetes, derived using different samples of individuals and different identification strategies, may signal a protective effect of education. More educated individuals might be better at processing and acting on information about risk factors and the benefits of lifestyle changes, which in turn might contribute to preventing the onset, lowering the severity or achieving remission of diabetes.

We therefore subjected our findings for diabetes to an additional analysis. We examined whether the effect of education changes with age, given that diabetes onset is strongly age related (see Appendix Figure A1, which shows the prevalence of diabetes across ages for each reform estimation sample). For each reform, we estimate the reduced form model with an additional covariate representing the interaction between the education reform variable and age.[15] To further isolate the age-education interaction effect, in some specifications we additionally include interactions with the business cycle and with a linear time trend.

The results are presented in Table 4. The first set of columns (1-3) presents results for the raising of the MSLA reform, the second (4-6) for the later education expansion. The first row presents

[15] In the raising of the school leaving age sample, the age range is 41 to 60, with the 10th, 50th and 90th percentiles equalling 44, 50 and 56, respectively. In the education expansion sample, the age range is 30 to 49, with the 10th, 50th and 90th percentiles equalling 34, 40 and 45.

the reduced form education effect for individuals with mean age and the second row presents the age-reform interaction effect. The third and fourth add the interactions with the business cycle and time-trend respectively. The first row shows the protective impact of education, which is as shown in the RF columns of Tables 2 and 3. The second row shows that this effect increases with age for the older cohort, and that this increase is robust to added controls in columns 2 and 3. The interaction effect is large: the Column 3 estimate suggests that each extra year of age increases the effect of education by 1/5 of the effect at the mean age. There is no age interaction for the younger individuals who are affected by the second reform. The finding may be explained by the lower prevalence of diabetes and the smaller age gradient among this age group (see Appendix Figure A1).

VII. Discussion

The relationship between health and education is central to models of human capital formation and to policy-making aimed at improving health and reducing health inequalities. Yet the causal evidence on this relationship is very mixed (Galama et al., 2018). This is partly the result of lack of data of sufficient size that can be used when exploiting natural experiments and the wide variety of health measures used across different studies. This has resulted in poorly estimated coefficients and inconclusive results. It has also meant that robust findings are available primarily for individuals affected by the raising of the minimum school leaving age.

In this paper, we contribute to the health-education literature by using a very large data set, which results in greater precision, and which allows us to examine reforms that affected individuals higher up the educational attainment distribution. We focus on common chronic conditions, which are precisely those health measures where education might be expected to play a large role, as they often involve adherence to medication and can be avoided or mitigated by health investments that require knowledge.

Overall, our findings support those of Clark and Royer (2013) and suggest that for most conditions there is no causal effect. Naïve OLS estimates would strongly suggest otherwise. This finding extends to ages when chronic conditions may onset (early 30s onwards) as well as to ages when these conditions are more prevalent (for individuals in their 50s and 60s), and is true for policies which increase the minimum school leaving age as well as ones which increase education more broadly across the educational distribution. Moreover, the results imply that the increased wages associated with both reforms did not lead to better health outcomes.

However, there is one exception to this conclusion. For diabetes, education appears protective, and for age groups in which diabetes is more common, the protective effect increases with age. Our result for diabetes is consistent with that found by Davies et al. (2018) who used a much smaller sample drawn from the UK Biobank.

Diabetes is a condition in which self-management is very important and, if Type 2, can be mitigated and even reversed with input by the individual (for example, by losing weight). Our results therefore suggest that education may reduce the probability of suffering from chronic conditions over which individuals have a high level of control, and consequently, where health-related knowledge may be more important. Exploring this is important to understand how to target policy to reduce the prevalence of such illnesses. This is therefore a valuable avenue for future research: the challenge will be finding causal designs with sufficient power.

References

Albouy, V. and Lequien, L. (2009). Does compulsory education lower mortality? *Journal of Health Economics*, vol. 28, pp. 155-168.

Avendano, M., de Coulon, A. and Nafilyan, V. (2017). Does more education always improve mental health? Evidence from a British compulsory schooling reform. HEGD Working Paper 1710, University of York.

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

Bedard, K. and and Deschênes, O. (2006). The long-term impact of military service on health: Evidence from World War II and Korean War veterans. *American Economic Review*, vol. 96, pp. 176-194.

Blanden, J. and Machin, S. (2004), Educational inequality and the expansion of UK higher education. *Scottish Journal of Political Economy*, vol. 51, pp. 230-249.

Braakmann, N. (2011). The causal relationship between education, health and health related behaviour: Evidence from a natural experiment in England. *Journal of Health Economics*, vol. 30, pp. 753-763.

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

Brunello, G., Foert, M. Schneeweis, N. and Winter-Ebmer, R. (2016). The causal effect of education on health: What is the role of health behaviours? *Health Economics*, vol. 25, pp. 314-336.

Cesarini, D., Lindqvist, E., Ostling, R. and Wallace, B. (2016). Wealth, health, and child development: Evidence from administrative data on Swedish lottery players. *Quarterly Journal of Economics*, vol. 131, pp. 687-738.

Clark, D. and Royer, H. (2013). The effect of education on adult health and mortality: Evidence from Britain. *American Economic Review*, vol. 103, pp. 2087-2120.

Crespo, L., López-Noval, B. and Mira, P. (2014). Compulsory schooling, education, depression and memory: New evidence from SHARELIFE. *Economics of Education Review*, vol. 43, pp. 36-46.

Cutler, D. and Lleras-Muney, A. (2010). Understanding differences in health behaviors by education. *Journal of Health Economics*, vol. 29, pp. 1-18.

Cutler, D. and Lleras-Muney, A. (2012). Education and health: Insights from international comparisons. NBER Working Paper 17738, January.

Davies, N., Dickson, M., Davey Smith, G., van den Berg, G. and Windmeijer, F. (2018). The causal effects of education on health outcomes in the UK Biobank. *Nature Human Behaviour,* vol. 2, pp. 117-125.

Deaton, A. and Paxson, C. (2004). Mortality, income, and income inequality over time in Britain and the United States. In *Perspectives on the Economics of* Aging, by D.A, Wise, Chapter 6. NBER.

Department of Health (2012). Long-term conditions compendium of Information: 3rd edition

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

Devereux, P. and Fan, W. (2011). Earnings returns to the British education expansion. *Economics of Education Review*, vol. 30, pp. 1153-1166.

Fischer, M., Karlsson, M. and Nilsson, T. (2013). Effects of compulsory schooling on mortality: Evidence from Sweden. *International Journal of Environmental Research and Public Health*, vol. 10, pp. 3596-3618.

Fletcher, J. (2015). New evidence on the effects of education on health in the US: Compulsory school laws revisited. *Social Science and Medicine*, vol. 127, pp. 101-107.

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. NBER Working Paper 24225, Cambridge MA.

Gathmann, C., Jürges, H. and Reinhold, S. (2015). Compulsory schooling reforms, education and mortality in twentieth century Europe. *Social Science and Medicine*, vol. 127, pp. 74-82.

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

Grossman, M. (1972). On the concept of health capital and the demand for health. *Journal of Political Economy*, vol. 80, pp. 223-255.

Grossman, M. (2008). The relationship between health and schooling. *Eastern Economic Journal*, vol. 34, pp. 281-292.

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

James, J. (2015). Health and education expansion. *Economics of Education Review*, vol. 49, pp. 193-215.

Ji, J. and Powdthavee, N. (2015). Does more education lead to better health habits? Evidence from the school reforms in Australia. *Social Science and Medicine*, vol. 127, pp. 83-91.

Johnston, D., Lordan, G., Shields, M. and Suziedelyte, A. (2015). Education and health knowledge: Evidence from UK compulsory school reform. *Social Science and Medicine*, vol. 127, pp. 92-100.

Johnston, D., Propper C., Shields, M (2009). Comparing subjective and objective measures of health: evidence from hypertension for the income/health gradient), *Journal of Health Economics* 28:540-522

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

Kemptner, D., Jurges, H. and Reinhold, S. (2011). Changes in compulsory schooling and the causal effect of education on health: Evidence from Germany. *Journal of Health Economics*, vol. 30, pp. 340-354.

Kenkel, D. (1991). Health behaviour, health knowledge, and schooling. *Journal of Political Economy,* vol. 99, pp. 287-305.

Li, J. and Powdthavee, N. (2015). Does more education lead to better health habits? Evidence from the school reforms in Australia. *Social Science & Medicine*, vol. 127, pp. 83-91.

Lindeboom, M., Llena-Nozal, A. and van der Klaauw, B. (2009). Parental education and child health: Evidence from a schooling reform. *Journal of Health Economics*, vol. 28, pp. 109-131.

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

Machin, S., Marie, O. and Vujić, S. (2012). Youth crime and education expansion. *German Economic Review*, vol. 13, pp. 366-384.

Mazumder, B. (2008). Does education improve health? A re-examination of the evidence from compulsory schooling laws. *Economic Perspectives*, vol. 2, pp. 2-16.

McCombie, L., Leslie, W., Taylor, R., Kennon, B., Sattar N. and Leanm M. (2017). Beating type 2 diabetes into remission. *BMJ*, vol. 358, j4030.

Meghir, C., Palme, M. and Simeonova, E. (2017). Education and mortality: Evidence from a social experiment. *American Economics Journal: Applied Economics*, forthcoming.

Muris, C. (2017). Estimation in the fixed-effects ordered logit model. *Review of Economics and Statistics*, vol. 99, pp. 465-477.

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

Powdthavee, N. (2010). Does education reduce the risk of hypertension? Estimating the biomarker effect of compulsory schooling in England. *Journal of Human Capital*, vol. 4, pp. 323-329.

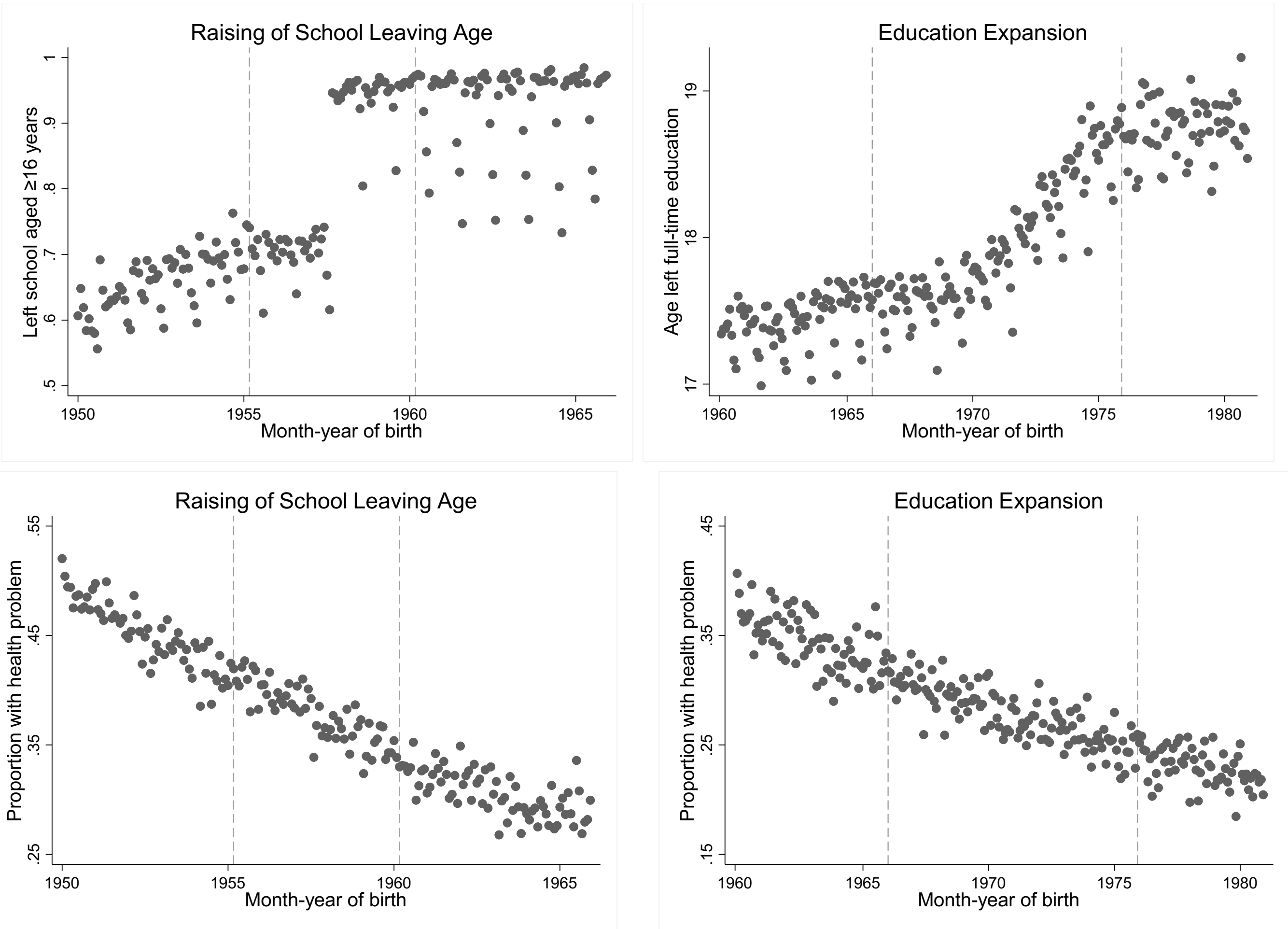
Public Health England. (2014). *Adult Obesity and Type 2 Diabetes* https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/338934/Adult_obesity_and_type_2_diabetes_.pdf

Siles, M. (2009). The causal effect of education on health: Evidence from the United Kingdom. *Economics of Education Review*, vol. 28, pp. 122-128.

Siles, M. (2015). The causal effect of schooling on smoking behaviour. *Economics of Education Review*, vol. 48, pp. 102-116.

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

Figure 1: Impact of Reforms on Educational Attainment and Chronic Health by Month of Birth



Notes: Each point represents the sample mean for a month-year of birth. Dotted vertical lines provide minimum and maximum date-of-births included in the estimation sample.

Table 1: Summary Statistics for each Reform Estimation Sample

	Raising of School Leaving Age Sample	Education Expansion Sample
Male	0.43	0.40
Age	49.90	39.52
Married or cohabitating	0.70	0.68
Non-white	0.02	0.05
Age completed full-time education	17.22	17.94
Employee	0.67	0.70
Self-employed	0.11	0.10
Unemployed	0.03	0.04
Health problem or disability	0.40	0.28
Survey years included	2001 - 2015	2006-2015
Years of birth included	1955-1960	1966-1975
Sample size	261,796	303,450

Note: Figures are sample means.

Table 2: Impact of Raising the Compulsory School Leaving Age on the Likelihood of having a Health Problem and/or Disability

	Mean	OLS	RF	IV
Health problem or disability of any type	0.398	-0.072***	0.004	0.012
		(0.006)	(0.014)	(0.041)
Number of health problems or disabilities	0.746	-0.317***	-0.002	-0.005
		(0.020)	(0.044)	(0.127)
Heart, blood pressure or blood circulation problems	0.130	-0.036***	-0.008	-0.024
		(0.004)	(0.010)	(0.028)
Problems or disabilities with back or neck	0.128	-0.054***	0.007	0.020
		(0.004)	(0.009)	(0.028)
Problems or disabilities with legs or feet	0.126	-0.053***	-0.001	-0.002
		(0.004)	(0.009)	(0.027)
Problems or disabilities with arms or hands	0.096	-0.045***	0.004	0.011
		(0.004)	(0.008)	(0.025)
Chest or breathing problems, asthma, bronchitis	0.083	-0.032***	0.012	0.034
		(0.004)	(0.008)	(0.025)
Depression, bad nerves or anxiety	0.070	-0.035***	0.001	0.003
		(0.003)	(0.007)	(0.021)
Stomach, liver, kidney or digestive problems	0.061	-0.024***	-0.002	-0.006
		(0.003)	(0.007)	(0.020)
Diabetes	0.046	-0.005*	-0.020***	-0.058***
		(0.003)	(0.006)	(0.022)
Severe disfigurements, skin conditions, allergies	0.036	-0.013***	0.008	0.024
		(0.002)	(0.005)	(0.016)
Mental illness, phobias, panics, other nervous disorders	0.033	-0.022***	-0.002	-0.007
		(0.002)	(0.005)	(0.015)

Notes: The OLS figures are coefficient estimates on age completed full-time education using only the cohorts born prior to the reform (1 September 1957) and individuals who completed full-time education ≤ aged 16 years. The RF figures are coefficient estimates on a binary variable indicating whether the individual was born after 1 September 1957 (treatment indicator). The IV figures are coefficient estimates on age completed full-time education. Control variables include a linear function of month-year of birth, a linear function of month-year of birth interacted with the treatment indicator, a third-order polynomial in age, and dummy variables for gender, year of survey (2001-2015), quarter of survey, wave number, month of birth (Jan-Dec), and interactions between month-of-birth dummies and the treatment indicator. Robust standard errors clustered at the individual level are presented in parentheses. *, ** and *** signify p-values less than 0.10, 0.05 and 0.01, respectively. Sample size for OLS models equals 74,624. Sample size for RF and IV models equals 261,796.

Table 3: Impact of Education Expansion on the Likelihood of having a Health Problem and/or Disability

	Mean	OLS	RF
Health problem or disability of any type	0.280	-0.017***	-0.007
		(0.001)	(0.012)
Number of health problems or disabilities	0.471	-0.051***	0.018
		(0.002)	(0.030)
Heart, blood pressure or blood circulation problems	0.045	-0.005***	0.002
		(0.000)	(0.006)
Problems or disabilities with back or neck	0.073	-0.009***	0.001
		(0.000)	(0.007)
Problems or disabilities with legs or feet	0.060	-0.008***	0.007
		(0.000)	(0.006)
Problems or disabilities with arms or hands	0.043	-0.006***	-0.001
		(0.000)	(0.005)
Chest or breathing problems, asthma, bronchitis	0.071	-0.004***	0.003
		(0.000)	(0.007)
Depression, bad nerves or anxiety	0.062	-0.008***	0.008
		(0.000)	(0.006)
Stomach, liver, kidney or digestive problems	0.040	-0.003***	-0.003
		(0.000)	(0.005)
Diabetes	0.019	-0.002***	-0.008**
		(0.000)	(0.004)
Severe disfigurements, skin conditions, allergies	0.029	-0.001***	0.006
		(0.000)	(0.004)
Mental illness, phobias, panics, other nervous disorders	0.030	-0.005***	0.003
		(0.000)	(0.005)

Notes: The OLS figures are coefficient estimates on age completed full-time education using only the cohorts born prior to 1970. The RF figures are coefficient estimates on average age completed full time education for an individual's cohort. Control variables include a linear function of month-year of birth, unemployment rate experienced at age 18, and dummy variables for age surveyed, gender, year of survey, quarter of survey, month of birth, and wave number. Robust standard errors clustered at the individual level are presented in parentheses. *, ** and *** signify p-values less than 0.10, 0.05 and 0.01, respectively. Sample size for OLS models equals 134,035. Sample size for RF models equals 303,450.

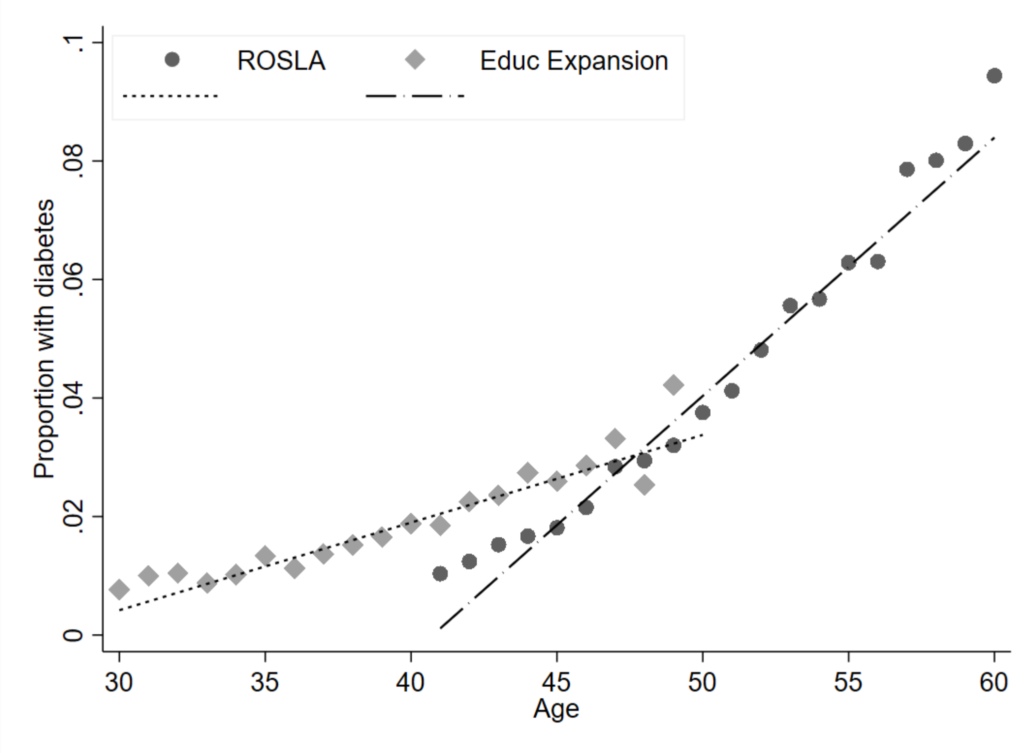
Table 4: Reduced-Form Models of the Impact of Education on the Likelihood of Having Diabetes across Ages

	Raising of School Leaving Age			Education Expansion		
	(1)	(2)	(3)	(4)	(5)	(6)
Mean reform effect	-0.020***	-0.020***	-0.020***	-0.008**	-0.008**	-0.009*
	(0.006)	(0.006)	(0.006)	(0.004)	(0.004)	(0.005)
Age interaction	-0.001**	-0.002***	-0.004**	-0.000	-0.000	-0.000
	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)
UR interaction		0.002	0.002		0.001	0.001
		(0.002)	(0.002)		(0.001)	(0.001)
Trend interaction			0.002			0.000
			(0.002)			(0.001)

Notes: The 'age interaction', 'UR interaction' and 'trend interaction' variables are interaction between age, the national unemployment rate, and a linear time trend, respectively, and the reform variable. The reform variable for columns 1-3 is a binary variable indicating the individual was born after 1st September 1957. The reform variable for columns 4-6 is average age completed full-time education for an individual's cohort. The estimates presented in Columns 1-3 were generated using the same regression specification and estimation sample as the RF column in Table 2; see the Table 2 note for details. Similarly, the estimates presented in Columns 4-6 were generated using the same regression specification and estimation sample as the RF column in Table 3.

<u>**On-line Appendix**</u>

Appendix Figure A1: Relationship between diabetes prevalence and age for each reform estimation sample (ROSLA = raising of minimum school leaving age)



Appendix Table A1: Impact of Raising the Minimum School Leaving Age on Educational Attainment

	Completed Aged ≥ 16	Completed Aged ≥ 17	Highest Qual ≥ GCSE	Highest Qual ≥ A-Level
All Persons				
Estimate	0.228***	0.014	0.046***	-0.000
	(0.004)	(0.014)	(0.009)	(0.022)
F-Statistic	3819.17	1.03	24.25	0.00
Outcome mean	0.70	0.39	0.71	0.52
Men Only				
Estimate	0.224***	-0.012	0.050***	0.012
	(0.018)	(0.026)	(0.008)	(0.017)
F-Statistic	158.56	0.23	43.28	0.46
Outcome mean	0.70	0.39	0.75	0.61
Women Only				
Estimate	0.233***	0.037**	0.042**	-0.012
	(0.017)	(0.018)	(0.020)	(0.028)
F-Statistic	177.01	4.23	4.29	0.18
Outcome mean	0.69	0.39	0.68	0.44

Notes: The figures are OLS coefficient estimates on a binary variable indicating whether the individual was born after 1 September 1957 (treatment indicator). Control variables include a linear function of month-year of birth, a linear function of month-year of birth interacted with the treatment indicator, a third-order polynomial in age, and dummy variables for gender, year of survey (2001-20015), quarter of survey, wave number, month of birth (Jan-Dec), and interactions between month-of-birth dummies and the treatment indicator. Sample only includes the first report from each individual. Sample sizes for all, men and women equal 85083, 38348 and 46735, respectively. Standard errors clustered at the month-year level. *, ** and *** signify p-values less than 0.10, 0.05 and 0.01, respectively.

Appendix Table A2: Estimated Health Impacts of Education Reforms using First Observation per Individual

	Reform 1	Reform 2
Health problem or disability of any type	0.023	0.002
	(0.026)	(0.008)
Number of health problems or disabilities	0.070	0.010
	(0.107)	(0.023)
Heart, blood pressure or blood circulation problems	-0.037	0.001
	(0.024)	(0.004)
Problems or disabilities with back or neck	0.032	0.005
	(0.021)	(0.006)
Problems or disabilities with legs or feet	0.033	0.005
	(0.039)	(0.006)
Problems or disabilities with arms or hands	0.024	-0.007
	(0.016)	(0.005)
Chest or breathing problems, asthma, bronchitis	0.059***	0.003
	(0.012)	(0.005)
Depression, bad nerves or anxiety	0.022	0.006
	(0.020)	(0.005)
Stomach, liver, kidney or digestive problems	0.017	-0.004
	(0.016)	(0.004)
Diabetes	-0.078***	-0.008**
	(0.021)	(0.003)
Severe disfigurements, skin conditions, allergies	-0.008	0.005
	(0.014)	(0.004)
Mental illness, phobias, panics, other nervous disorders	0.005	0.002
	(0.008)	(0.003)
Sample size	85083	110214

Notes: The sample uses wave 1 observations only for each individual. The Reform 1 figures correspond to the estimates contained in the IV Column of Table 2, and the Reform 2 figures correspond to the estimates contained in the RF Column of Table 3. Robust standard errors clustered at the month-year level are presented in parentheses. *, ** and *** signify p-values less than 0.10, 0.05 and 0.01, respectively.

Appendix Table A3: Estimates of the Impact of Raising the Minimum School Leaving Age on Health using Smaller Bandwidths

	30-month window: June 1956 - Dec 1958	40-month window: Jan 1956 - May 1959	50-month window: Aug 1955 - Oct 1959
Health problem or disability of any type	-0.018	-0.014	0.019
	(0.053)	(0.057)	(0.052)
Number of health problems or disabilities	-0.013	0.060	0.054
	(0.164)	(0.182)	(0.165)
Heart, blood pressure or blood circulation problems	-0.035	-0.052	-0.023
	(0.038)	(0.042)	(0.035)
Problems or disabilities with back or neck	-0.011	0.011	0.032
	(0.036)	(0.040)	(0.037)
Problems or disabilities with legs or feet	-0.004	0.018	0.009
	(0.036)	(0.040)	(0.035)
Problems or disabilities with arms or hands	0.008	0.050	0.040
	(0.032)	(0.039)	(0.034)
Chest or breathing problems, asthma, bronchitis	0.048	0.035	0.036
	(0.035)	(0.036)	(0.032)
Depression, bad nerves or anxiety	0.012	0.019	0.004
	(0.028)	(0.031)	(0.027)
Stomach, liver, kidney or digestive problems	0.005	0.008	0.012
	(0.027)	(0.029)	(0.026)
Diabetes	-0.062**	-0.074**	-0.071**
	(0.029)	(0.034)	(0.029)
Severe disfigurements, skin conditions, allergies	0.028	0.042*	0.028
	(0.021)	(0.025)	(0.022)
Mental illness, phobias, panics, other nervous disorders	-0.002	0.003	-0.013
	(0.019)	(0.020)	(0.018)
Sample size	132629	177422	218530