Online Appendix

The Effect of Education on Adult Mortality and Health: Evidence from Britain

Clark and Royer

Appendix A: Additional Robustness Checks

In this appendix, we present a number of additional tables and figures as robustness checks. We discuss these tables and figures in the main text.

Appendix B: Measurement Issues

A Measuring Years of Education

Throughout the paper we refer to the health impacts of "years of education." This is defined as the age at which someone left full-time education minus five. In this Appendix we describe why years of education is measured in this way.

The starting point is the two survey datasets used in this paper: the Health Survey of England (HSE) and the General Household Survey (GHS). These surveys ask respondents the following questions:

HSE "At what age did you finish your continuous full-time education at school or college?"

GHS "How old were you when you left there, or when you finished or stopped your course?" [*In response to "Now thinking just of your full-time education, what type of school or college did you last attend full-time?"*]

The relationship between "age left full-time education" and "years of education" is governed by English laws that determine the period of compulsory schooling (i.e., compulsory full-time education). These laws state that students must have started school by the term after they turn five, although students have, traditionally, started school in the term in which they turned five (Sharp, 1997).[1] Before 1962, these laws stated that students could not leave until the end of the term in which they reached the minimum school leaving age. The UK education system has three terms that run September-December, January-April, and April-July; exact dates vary by school district. Hence, students born in October could not leave until Christmas, students born in February could not leave until Easter and students born in May could not leave until July. After 1962, these laws stated that students born September-January could not leave until April, and students born February-August could not leave until June. We illustrate the structure of these laws in Appendix Figure B1a.

Based on these laws, we can calculate the *minimum number of years and school terms of full-time education* that must have been completed by someone reporting leaving at the minimum leaving age and the *maximum number of years and school terms of full-time education* that must have been completed by someone reporting leaving at the minimum leaving age. This is illustrated in Appendix Figure B1b. The minimum is always nine years and one term for those able to leave before the 1947

[1] Sharp (1997) claims that these laws changed to accommodate earlier starting as pupil numbers fell in the 1970s. This is confirmed by the survey evidence in Crawford, Dearden and Meghir (2007).

change, ten years and one or two terms for those able to leave before the 1972 change and eleven years and one or two terms for the rest. Because respondents that report leaving at the minimum leaving age may not have left at the first available opportunity, the maximum is greater than this. In particular, as seen in Appendix Figure B1b, respondents that report leaving at the minimum leaving age could have completed one or two terms more than the minimum years and school terms of full-time education that is completed by those that leave at the first available opportunity.

With this discussion in mind, we calculate "years of education" as the age left full-time education minus five. That X years of education then refers to X years and one term, X years and two terms or X+1 years is not a problem: standard measures of completed years or grades of education share the same property (i.e., we do not observe if fractions of grades are completed). In addition, any measurement error in years of education will bias downwards the estimated impact of the 1947 and 1972 law changes on educational attainment. This will bias upwards our two-stage least squares estimates of the health effects of education. Since these estimates are consistently small, this phenomenon is likely of little importance here.

Compliance post-1972

As seen in Figure 1, non-compliance among summer-borns increases after 1972. That is, after 1972, a large fraction of students born in June, July and August appear to leave full-time education at age fifteen or younger. In fact, while there may be a small amount of non-compliance throughout the entire period studied, this post-1972 phenomenon is likely another consequence of the structure of the laws.

Before 1972, when the minimum leaving age was fifteen, students were required to stay in full-time education until part way through grade nine. Like all grades except grade ten, grade nine finishes in mid-July. Hence only students born in late July and August could have left at fourteen (recall that students can leave at the end of the term in which they turn fifteen, where the summer term ends in July but for the purposes of the minimum leaving age is defined to run until September). After 1972, when the minimum leaving age became age sixteen, students had to stay until part way through grade ten. Grade ten 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 fifteen, technically younger than the minimum leaving age (sixteen). This can account for the apparent increase in post-1972 non-compliance. Since the exam period is an unusual period of schooling, even respondents that were in school throughout this period may report leaving in April. This would further contribute to the apparent increase in non-compliance.

While this phenomenon could introduce some error into our measurement of years of education, since all of our models control for month-of-birth dummy variables, we think it will have little impact on our analysis. Provided these month-of-birth patterns are consistent over time in the post-reform period, they will be controlled for in our regressions. For the 1972 reform, we interact the month-of-birth dummies with the reform dummy, allowing the seasonal patterns to be different pre- and post-reform since the seasonal patterns appear to change with the reform. As expected, the non-compliance phenomenon is much less striking once we regression adjust for month of birth.

B Measurement Error in Mortality Rates

i Sources of Measurement Errors

For each birth cohort, our analysis requires a count of the population at risk of dying in each month. As noted in the main text, for months before April 1991, we infer the population size by taking the population of those born in England and Wales and resident in England and Wales at the time of 1991 Census (the earliest available, enumerated April 1991) and adding deaths occurring between the month of interest and April 1991. For months after April 1991, we subtract deaths occurring between April 1991 and the month of interest. In combination with the population at risk, we need a measure of the number of deaths. Our measure of deaths is the death count of individuals dying in England and Wales by month-year of death and month-year of birth.

There are various sources of measurement error that will drive differences between measured mortality and true mortality:

Issues with Estimating Population Alive

1. **Emigration from England and Wales:** The 1991 Census will provide an inaccurate count of the number alive in 1991 because some people will have emigrated out of England and Wales. As we show below, this is not a first-order issue.
2. **Census undercounting:** The 1991 Census will provide an inaccurate count of the number alive in 1991 because of Census under-counting. According to Dale and Marsh (1993), evidence from a post-census validation study suggests that among cohorts between ages 50 and 65 in 1991 (i.e., the cohorts in our 1947 analysis), more than 99 percent were enumerated in the 1991 Census (Table 6.11).
3. **Misreporting of age in 1991 Census:** The 1991 Census will provide an inaccurate count of the number alive in 1991 because of misreporting of age information. In part, this is because the head of household may provide age information for other household members, and is less likely to provide accurate information. According to Heady, Smith and Avery (1996), evidence from a post-census validation study suggests that the error rate on the age variable is on the order of two percent for the cohorts that we study (Table 8).

Issues with Estimating Deaths

1. **Exclusion of deaths among emigrants:** The death counts will provide an inaccurate count of the number of deaths in a particular period because they will exclude deaths among emigrants.
2. **Misreporting of date of birth on death records:** The death counts will provide an inaccurate count of the number of deaths in a particular period because the death certificate may contain inaccurate date of birth information. To assess the magnitude of this problem, we used data from the Longitudinal Study, a 1 percent sample of the 1971 Census linked to the 1981, 1991, 2001 Censuses and other vital statistics data including death records. We considered individuals who died after 2001, and for whom the date of birth on all four Censuses agree. For less than 1.5% of these individuals did the date of birth on the death record disagree with the date of birth on the Censuses, and in most cases these dates were off by one day. Note that while there may be some inaccuracies in the date of birth information on death records, there is virtually no undercounting of deaths in England and Wales (Charlton and Murphy, 1997).
3. **Inclusion of deaths among immigrants:** The death counts include immigrants who died in England and Wales and as such, the death counts do not measure deaths among the population of interest (i.e., those born in England and Wales).

ii Simulated Impacts of Measurement Error

In this section, we use simulations to assess the consequences of these various measurement errors. We only consider the 1947 change although our conclusions will also apply to the 1972 change. The basic strategy is as follows. First, we assume reduced-form mortality effects of the ROSLA of a particular size and generate data for UK- and non-UK born populations that are consistent with these ROSLA effects (ROSLA effects for non-UK born populations are assumed to be zero). Second, we add measurement error to these data (e.g., by deleting a fraction of the observations equal to the assumed rate of Census under-counting). We do not factor emigration into these simulations since emigration impacts both the population and death counts. Hence, provided the ROSLAs do not affect the composition of emigrants, emigration should not make our estimates of the mortality effects of the ROSLAs inconsistent. As argued in the main text, however, we suspect the ROSLAs had little impact on the composition of emigrants. Third, we compare the assumed true effects and the effects estimated with the error-ridden data and infer the likely size of any measurement error biases.

We now describe the steps of the simulation in more detail:

1. **Create hazard rates of death for cohorts born between 1925 and March 1933 (pre-reform cohorts).** The Human Mortality Database provides hazard rates of death by age and year. For our purposes, we need these expressed by month-year of birth and month-year of death (the level of aggregation of our data). To do this, we convert the hazard rates to survival rates for integer age a and month $m \in \{1, 12\}$ that aggregates to the age-in-years survival rate and evolves according to $ln(\frac{s_{a,m+1}}{s_{a,m}}) = \frac{1}{12}\frac{s_{a+1}-s_a}{s_a}$ where s_a is the survival rate for integer age a and $s_{a,m}$ is the survival rate at age a and month m. The survival rate is equal to (1-hazard rate). Since we do not observe monthly survival rates in these data, we could have made any number of assumptions about the evolution of the monthly survival rates (subject to them aggregating to the correct age-in-years survival rates). This assumption ensures there is no discontinuity in the monthly hazard from month 12 at age a to month 1 at age $a+1$. To get the hazards for cohort month-year of birth, we assume the year-of-birth hazards refer to cohorts born in June of each year and we linearly interpolate (estimated separately at every age level) to generate hazards for other months.
2. **Create hazard rates of death for cohorts born between April 1933 and 1940**. We use the trends in age-specific mortality rates across these cohorts to simulate the hazards for every month-year of birth cohort. Specifically, we use these estimated trends in age-specific mortality and, for ages 16 and older, we assume a shift in the hazard rate driven by exposure to the ROSLA. The assumed hazard effects are derived from the assumed log-odds effects (i.e., the assumed effect of the ROSLA).
3. **Use the hazard rates from steps 1 and 2 to generate the fraction dying at each point in time.** The hazard rates are used as parameters in the binomial distribution. We assume each cohort starts with 40,000 members, roughly the average cohort size across these cohorts.
4. **Factor in measurement error from (a) undercounting in the Census, (b) misreporting of age information in the Census, and (c) misreporting of age information on the death certificate.** For (a), we assume that some fraction of individuals are missing from the Census count. For (b) and (c), we model reported age at death as follows: the reported age is the true age with probability equal to 1-measurement error rate, and errors are then spread uniformly over the 6 months before and after the true age. We assume that the errors in the Census and death certificates are independent.
5. **Follow steps 1.-3. to simulate the size of the immigrant population.** We use a similar procedure to simulate cohorts of immigrants, the only difference being that these do not experience any ROSLA effect. We make an assumption about the immigration rate in each cohort (see below).
6. **Calculate the size of the population of natives in the 1991 Census using results from steps 3 and 4.**
7. **Calculate the number of deaths in each month (natives and immigrants).**
8. **Combine data from steps 6 and 7 to calculate Census and death counts to estimate mortality effects from the two-step procedure described in the text.**

There are five parameters to be chosen here: (1) the ROSLA effect on mortality (2) the fraction of immigrants in the population (3) the Census undercount rate (4) reported age errors on the Census and (5) reported age errors on the Death Certificates. We constrain (3)-(5) to be the same as we expect them to be of roughly equal magnitudes. We consider ROSLA effects of 0 (baseline), -0.05, -0.1, -0.15 and -0.2, where these numbers are effects on the log-odds ratio. The last of these is comparable to the smallest effect implied by the results of Lleras-Muney (2005). We consider immigration rates of 0 (baseline), 0.1 (most likely) and 0.2 (worst-case). Data from the 1991 Census suggests that for the cohorts relevant to the 1947 reform, the immigration rate (i.e., fraction of the UK population that is an immigrant) is around 0.1. We consider undercount and age error rates of 0% (baseline) and 2% (worst-case).

For each combination of assumptions, the following Table provides mean errors (i.e., estimated effect less true effect) based on 100 replications (sampling variation is very small). Thus, because we consider all true effects to be negative, negative error rates imply that the measured effect is further away from zero than the true effect.

	Immigration=0%		Immigration=10%		Immigration=20%	
	Error=0%	Error=2%	Error=0%	Error=2%	Error=0%	Error=2%
Effect=0	0.0016	0.0027	0.0004	0.0004	0.0007	0.0008
Effect=-0.05	0.0014	0.0017	-0.0039	-0.0039	-0.0074	-0.0074
Effect=-0.10	-0.0004	0.0003	-0.0082	-0.0083	-0.0153	-0.0155
Effect=-0.15	-0.0063	-0.0066	-0.0127	-0.0130	-0.0241	-0.0244
Effect=-0.20	0.0112	0.0082	-0.0181	-0.0186	-0.0325	-0.0332

Note: the error percentages (e.g., Error=0%) refer to the error rate of the parameters (3)-(5) discussed above.

The numbers in this Table point to three conclusions.

1. **Error rates in date of birth and Census undercounting do not appear to generate significant bias.** This can be seen in every pair of cells that vary the error rate holding the effect size and the immigration rate constant.
2. **The percent bias is increasing negatively in the immigration rate.** For example, looking at the last row under a zero error rate assumption, the bias is 5.6% (0.0112) for an immigration rate of 0%, 9.1% (-0.0181) for an immigration rate of 10% and 16.25% (-0.0325) for an immigration rate of 20%. Thus, the larger the immigration rate the larger the bias towards finding a protective effect of education.
3. **The percent bias is less than the immigration rate.** For example, looking at the last column, the bias for effect sizes of -0.05, -0.1, -0.15 and -0.2 are -0.00738 (14.8%), -0.0155 (15.5%), -0.0244 (16.3%) and -0.0332 (16.6%), all less than the assumed immigration rate of 20%.

iii Intuition

The previous simulation results suggest that measurement issues are unlikely to bias our estimates. We now complement these simulation exercises with a mathematical analysis of the biases driven by measurement error. In particular, under some simplifying assumptions, we derive bounds on the size of these biases.

The key to this analysis is the observation that the measured log-odds ratio for cohort c in period t is approximately equal the true log-odds ratio for cohort c in period t plus six error terms:

$$ln(\widehat{O}_{ct}) \sim ln(O^*_{ct}) + e_{1ct} + e_{2ct} + e_{3ct} + e_{4ct} + e_{5ct} + e_{6ct}$$

where the measured log-odds ratio is $ln(\widehat{O}_{ct})$ and the true log-odds ratio is $ln(O^*_{ct})$.

We prove this fact in the "Derivations" section below. This observation is useful for three reasons. First, the coefficients estimated in the first step of our hazard analysis (i.e., the cohort fixed effects) are closely related to the <u>measured</u> period-t log-odds ratio (i.e., $ln(\widehat{O}_{ct})$). Specifically, although the first step of estimating the mortality effects involves the use of data pooled across several time periods and a model that includes age fixed effects, the dependent variable above (i.e., $ln(\widehat{O}_{ct})$) is equivalent to the set of birth cohort fixed effects that would be generated by a first step logit model estimated on data for a single period t without age fixed effects. Second, the true mortality effects of the ROSLA should correspond to the discontinuity in the <u>true</u> period-t log-odds ratio (i.e., the discontinuity in $ln(O^*_{ct})$). That is because other forces impacting the true period-t log-odds ratios (i.e., cohort trends and age trends) should be smooth through the ROSLA threshold. Third, it follows that if we can bound the discontinuities in the error terms (i.e., difference in the average error term for those born immediately after the 1933 April threshold and the average error term for those born immediately before the 1933 April threshold) in the equation above, then we can bound the extent to which our estimates of the ROSLA effect on mortality (i.e., the discontinuity in $ln(\widehat{O}_{ct})$) differ from the true ROSLA effect on mortality (i.e., the discontinuity in $ln(O^*_{ct})$).

As shown in the "Derivations" section below, under some mild assumptions additional to those made in the simulations section, we can derive the bounds for each of the discontinuities in the e_{jct} terms where $j = 1, 2, ..., 6$ from the equation above. We denote a discontinuity in e_{jct} as e^{RD}_{jt}. We divide each of these discontinuities by $|\alpha^{RD}_t|$, the absolute value of the discontinuity in the true log-odds mortality ratio, to express each of the discontinuities in the error terms as a percentage of the true ROSLA mortality effect. In particular, we show that:

$$\begin{aligned}
\frac{|e^{RD}_{1t}|}{|\alpha^{RD}_t|} &< e_d(1-m)|(1+|\alpha^{RD}_t|)(1+\frac{|\alpha^{RD}_t|}{2}) \\
\frac{|e^{RD}_{2t}|}{|\alpha^{RD}_t|} &< i(1-m)[1+|\alpha^{RD}_t|] \\
\frac{|e^{RD}_{3t}|}{|\alpha^{RD}_t|} &= 0 \\
\frac{|e^{RD}_{4t}|}{|\alpha^{RD}_t|} &< e_d[1+\frac{|\alpha^{RD}_t|}{2}] \\
\frac{|e^{RD}_{5t}|}{|\alpha^{RD}_t|} &< i(1-m) \\
\frac{|e^{RD}_{6t}|}{|\alpha^{RD}_t|} &< 2u_s + e_s(1-u_s)(1+\frac{|\alpha^{RD}_t|}{2})
\end{aligned}$$

where e_d and e_s are the rates of age misclassification on the death certificates and in the Census, respectively, and u_s is the rate of undercounting in the Census. The parameters i and m are the immigration rate (ratio of immigrants to total population in England and Wales) and emigration rate (the fraction of native-born outside of England and Wales), respectively.[2]

Before deriving these bounds, we use them to provide some intuition for the three conclusions that we drew from the simulation exercise:

1. **Error rates in date of birth and Census undercounting do not appear to generate significant bias.** If the error rates (i.e., e_d, e_s, and u_s are all zero, the biases associated with the first, fourth and sixth terms disappear. But if these error rates are 2%, the biases associated with these error terms are still tiny.
2. **The bias is increasing negatively in the immigration rate.** The only non-trivial biases here are those associated with the second and fifth error terms. Since these are proportional to the immigration rate, this conclusion is not surprising. Note that while we only bound the absolute value of the bias here, the simulations suggested that the bias will be towards finding a more protective effect of education.
3. **The bias is less than the immigration rate.** Since the second and fifth terms are both roughly bounded by the immigration rate itself, and since the other terms are small, an upper

[2] One might wonder why the emigration rate appears here, given that we have argued that emigration will have equal impacts on population and death counts and hence no effect on our estimates. The reason is that while this is true in the absence of any other measurement errors, there are interactions between emigration and these errors. Still, we expect these effects to be small so we did not model them in our simulation analyses. We assume that the emigration rate is the same among pre- and post-ROSLA cohorts.

bound on the overall percent bias is roughly twice the immigration rate. Although this is greater than the immigration rate, it is still not surprising that the simulations suggested that the percent bias is less than the immigration rate. First, these are bounds, hence the actual biases may be much smaller. Second, we are considering the sum of the absolute value of each bias. Since they need not take the same sign, the absolute value of the sum could be much smaller.

iv Derivations

We now derive this measured log-odds equation and these six error bounds.

Measured log-odds equation

Our measured log-odds ratio is defined as follows:

$$ln(\widehat{O}_t) = ln(\frac{\widehat{D}_t}{\widehat{P}_t - \widehat{D}_t})$$

where $\widehat{O}_t$ is the measured log-odds ratio in period t, $\widehat{D}_t$ is the measured number of deaths, and $\widehat{P}_t$ is the measured population size. We can express the difference between the measured log-odds of death and the true log-odds of deaths as

$$ln(\widehat{O}_t) - ln(O^*_t) = ln(\frac{\widehat{D}_t}{D^*_t}) - ln(\frac{\widehat{P}_t - \widehat{D}_t}{P^*_t - D^*_t})$$

where the $*$ are used to denote the true values. As discussed in the main text, our estimate of deaths at time t is

$$\widehat{D}_t = D^*_t - D^{E*}_t + D^I_t + D^{EW\#}_t$$

where D^*_t represents all deaths to individuals who were born in England and Wales, D^E_t represents deaths to emigrants who were born in England and Wales but died elsewhere, D^I_t represents deaths to immigrants who died in England and Wales, and $D^{EW\#}_t$ is the measurement error in the deaths measured in England and Wales (i.e., measurement error related to date of birth). For any point in time t, let s be the relevant Census year used in the calculation of the population size at time t (see text for more details). Then, analogously, our estimate of the population at time t is

$$\widehat{P}_t = P^*_t - P^{E*}_t + P^{EW\#}_s + \mathbf{1}_{t<s}\sum_{i=t}^{s-1} D^{EW\#}_i + \mathbf{1}_{t<s}\sum_{i=t}^{s-1} D^I_i - \mathbf{1}_{t>s}\sum_{i=s}^{t-1} D^{EW\#}_i - \mathbf{1}_{t>s}\sum_{i=s}^{t-1} D^I_i$$

where P^*_t represents the population count of individuals alive at time t who were born in England and Wales, P^{E*}_t represents population of emigrants who were born in England and Wales but live elsewhere, $P^{EW\#}_s$ is the measurement error in the population measured in England and Wales (i.e., measurement error related to date of birth and Census undercounting), and $\mathbf{1}$ represents the indicator function. Then, we can substitute these expressions for $\widehat{P}_t$ and $\widehat{D}_t$ into the equation for the difference between the measured log-odds of death and the true log-odds of death above. The resulting equation is:

$$\begin{aligned} ln(\widehat{O}_t) - ln(O^*_t) \quad &= \quad ln(1 + \frac{-D^{E*}_t + D^I_t + D^{EW\#}_t}{D^*_t}) \\ &\quad - ln(1 + \frac{-(P^{E*}_t - D^{E*}_t) + P^{EW\#}_s + f(D^{EW\#}) + f(D^I)}{P^*_t - D^*_t}) \end{aligned}$$

where

$$f(D^{EW\#}) = \mathbf{1}_{t<s}\sum_{i=t}^{s-1} D^{EW\#}_i - \mathbf{1}_{t>s}\sum_{i=s}^{t-1} D^{EW\#}_i - D^{EW\#}_t$$

and

$$f(D^I) = \mathbf{1}_{t<s}\sum_{i=t}^{s-1} D_i^I - \mathbf{1}_{t>s}\sum_{i=s}^{t-1} D_i^I - D_t^I$$

Using the facts that $ln(1+y) \sim y$ for small y, $P_t^* - D_t^* = P_{t+1}^*$, and $P_t^{E*} - D_t^{E*} = P_{t+1}^{E*}$,

$$ln(\widehat{O}_t) - ln(O_t^*) \sim \frac{D_t^{EW\#}}{D_t^*} + \frac{D_t^I}{D_t^*} + \frac{P_{t+1}^{E*}}{P_{t+1}^*} - \frac{D_t^{E*}}{D_t^*} - \frac{f(D^{EW\#})}{P_{t+1}^*} - \frac{f(D^I)}{P_{t+1}^*} - \frac{P_s^{EW\#}}{P_{t+1}^*}$$

Define each part of the right-hand side of the equation above as follows:

$$\begin{aligned}
e_{1t} &= \frac{D_t^{EW\#}}{D_t^*} \\
e_{2t} &= \frac{D_t^I}{D_t^*} \\
e_{3t} &= \frac{P_{t+1}^{E*}}{P_{t+1}^*} - \frac{D_t^{E*}}{D_t^*} \\
e_{4t} &= \frac{f(D^{EW\#})}{P_{t+1}^*} \\
e_{5t} &= \frac{f(D^I)}{P_{t+1}^*} \\
e_{6t} &= \frac{P_s^{EW\#}}{P_{t+1}^*}
\end{aligned}$$

Biases

Discontinuities in these six error terms will bias our estimates of the ROSLA mortality effects (i.e., drive a wedge between the regression discontinuity in the true log-odds mortality ratio and the regression discontinuity in the measured log-odds mortality ratio). To simplify, we define ROSLA discontinuities in each of the error terms as the difference between the value of the error term for the first post-ROSLA cohort (i.e., April 1933) and the value for the first pre-ROSLA cohort (i.e., March 1933).

Definitions

Before deriving expressions for the terms, it is useful to define the following expressions:

$$\begin{aligned}
x_{Lt} &\equiv (x_{ct}|c = March\,1933) \\
x_{Rt} &\equiv (x_{ct}|c = April\,1933) \\
x_t^{RD} &\equiv x_{Rt} - x_{Lt} \\
v(x_t) &\equiv \frac{x_{Rt} - x_{Lt}}{x_{Lt}}
\end{aligned}$$

where the first two equivalence statements are conditional statements. We also define:

$$\alpha^{RD} \equiv ln(O)^*_{Rt} - ln(O)^*_{Lt}$$

where $ln(O)^*$ is the true log odds ratio.[3]

We make several assumptions to derive bounds for each of the terms of the bias expression. First, as in the simulation, we assume that a fraction u_c of residents are not counted at the Census, and that this fraction is constant across cohorts. Second, we assume that because of age misreporting, the Census assigns only a fraction $1-e_s$ of those that belong to cohort c to cohort c, and assigns a fraction $\frac{e_s}{2}$ of those in cohorts $c-1$ and $c+1$ to cohort c. We make an equivalent assumption for death counts; the analogous error rate for death counts is e_d. Third, we assume that true death and population counts among the last two pre-ROSLA are equal and that the true death and population counts among the first two post-ROSLA cohorts are equal. While one can imagine more complicated structures for the measurement error (e.g., that used in the simulations), it seems reasonable to assume that misclassification is symmetric (i.e., misclassified deaths are as likely to be assigned to older as to younger cohorts) and "local" (i.e., misclassified deaths to cohort c are more likely to be assigned to cohort c+1 than to cohort c+10). Fourth, we assume that the initial population sizes of the pre- and post-ROSLA cohorts near the threshold are the same; this assumption is consistent with our simulation assumptions. Fifth, we assume that for any time t, the ratio of true deaths in England and Wales to the true population size in t+1 is less than $\frac{1}{256}$.[4,5] Sixth, we assume that the sum of annual hazards is less than 1.[6]

These assumptions give rise to the following expressions for measured variables:

$$\begin{aligned}
\widehat{D}^{EW}_{Lt} &= [(1-\frac{e_d}{2})D^{EW*}_{Lt} + \frac{e_d}{2}D^{EW*}_{Rt}] \\
\widehat{D}^{EW}_{Rt} &= [(1-\frac{e_d}{2})D^{EW*}_{Rt} + \frac{e_d}{2}D^{EW*}_{Lt}] \\
\widehat{P}^{EW}_{Ls} &= (1-u_s)[(1-\frac{e_s}{2})P^{EW*}_{Ls} + \frac{e_s}{2}P^{EW*}_{Rs}] \\
\widehat{P}^{EW}_{Rs} &= (1-u_s)[(1-\frac{e_s}{2})P^{EW*}_{Rs} + \frac{e_s}{2}P^{EW*}_{Ls}]
\end{aligned}$$

where P^{EW*} is the true population of native-borns living in England and Wales and a similar definition apples to D^{EW*} In deriving bounds on the biases, we make repeated use of the following inequalities:

Inequality 1:

$$\boxed{0 < v(P^*_t) < |\alpha^{RD}|}$$

Inequality 2:

$$\boxed{\frac{D^*_{Lt}}{D^*_{Rt}} < \frac{H^*_{Lt}}{H^*_{Rt}} < 1 + |\alpha^{RD}|}$$

[3]Here we exclude the t subscript as we implicitly assume that the treatment has the same effect on the log-odds ratios at all points in time.

[4]The earliest data point in our sample is 1970 and we have population counts from the April 1991 Census. Thus, the gap in months between the 1991 Census and January 1970 is 256 months.

[5]In the Human Mortality Database, the maximum ratio of annual deaths to annual population is 0.019. Dividing this ratio by 12 to derive a monthly death rate gives 0.001583, considerably less than $\frac{1}{256}$.

[6]The sum of the annual hazards to age 72 is 0.38 according to Human Mortality Database. We would expect that the sum of monthly hazards (not available from the Human Mortality Database) to be similar.

Inequality 3:

$$\boxed{0 \geq |v(D_t^*)| < |\alpha^{RD}|}$$

where H_{Rt}^* is the true hazard rate (i.e., ratio of deaths to population) among the post-reform cohort. We prove each of these in turn:

Proof of Inequality 1
To begin, note $v(P_t^*) = \frac{P_{Rt}^* - P_{Lt}^*}{P_{Lt}^*}$. Thus, since $ln(1+y) \sim y$ for small y,

$$v(P_t^*) \sim ln(1 + \frac{P_{Rt}^*}{P_{Lt}^*} - 1) = ln(P_{Rt}^*) - ln(P_{Lt}^*)$$

Also, by definition,

$$\alpha^{RD} = ln(\frac{H_{Rt}^*}{1 - H_{Rt}^*}) - ln(\frac{H_{Lt}^*}{1 - H_{Lt}^*})$$

After some algebra along with invoking the ln approximation, it can be shown that

$$\alpha^{RD} \sim \frac{H_{Rt}^* - H_{Lt}^*}{(1 - H_{Rt}^*)H_{Lt}^*}$$

Now we derive expressions for P_{Lt}^* and P_{Rt}^*

$$P_t^* = P_1^* \prod_{i=1}^{t-1}(1 - H_i^*)$$

where H_t^* is the true hazard of death at time t and P_1^* is the initial population size. Then, it follows that

$$ln(P_{Lt}^*) = ln(P_{L1}^* \prod_{i=1}^{t-1}(1 - H_{Li}^*))$$

A similar expression follows for $ln(P_{Rt}^*)$. Invoking the fact that $ln(1+y) \sim y$ for small y, it is true that

$$ln(P_{Lt}^*) \sim ln(P_{L1}^*) - \sum_{i=1}^{t-1} H_{Li}^*$$

Using these expressions for P_{Lt}^* and P_{Rt}^*, $ln(1+y) \sim y$ for small y, and the assumption that $P_{R1}^* = P_{L1}^*$, we have

$$v(P_t^*) \sim \sum_{i=1}^{t-1}(H_{Li}^* - H_{Ri}^*)$$

Thus, combining the expressions for α^{RD} and $v(P_t^*)$, we have

$$v(P_t^*) \sim \sum_{i=1}^{t-1} -\alpha^{RD}(1 - H_{Ri}^*)H_{Li}^*$$

Since the hazard rates are bounded by 0 and 1 and assuming $\alpha^{RD} < 0$,

$$v(P_t^*) < -\alpha^{RD} \sum_{i=1}^{t-1} H_{Li}^*$$

Since in our data $\sum_{i=1}^{t-1} H_{Li}^* < 1$[7] thus $v(P_t^*) > 0+$,

$$v(P_t^*) < -\alpha^{RD} \leq |\alpha^{RD}|$$

Proof of Inequality 2
To show $\frac{H_{Lt}^*}{H_{Rt}^*} < 1 + |\alpha^{RD}|$, assuming that education reduces the hazard of death and using the ln approximation, it follows that

$$|\alpha^{RD}| \quad = \quad ln(\frac{H_{Lt}^*}{H_{Rt}^*}\frac{1 - H_{Rt}^*}{1 - H_{Lt}^*}) > ln(\frac{H_{Lt}^*}{H_{Rt}^*}) \sim \frac{H_{Lt}^*}{H_{Rt}^*} - 1$$

To show $\frac{D_{Lt}^*}{D_{Rt}^*} < \frac{H_{Lt}^*}{H_{Rt}^*}$, after using algebra invoking the definitions of $v(D_t^*)$, $v(H_t^*)$, and $v(P_t^*)$, it is true that

$$v(D_t^*) = v(H_t^*) + v(P_t^*)\frac{H_{Rt}^*}{H_{Lt}^*}$$

Since $v(P_t^*) > 0$,

$$v(D_t^*) > v(H_t^*)$$

which implies that

$$\frac{D_{Rt}^*}{D_{Lt}^*} > \frac{H_{Rt}^*}{H_{Lt}^*}$$

Proof of Inequality 3
To show $|v(D_t^*)| < |\alpha^{RD}|$, we consider the cases of $v(D_t^*) < 0$ and $v(D_t^*) > 0$.
$\boxed{\textit{Case I: } v(D_t^*) < 0}$
If $v(D_t^*) < 0$, then

$$|v(D_t^*)| = -v(D_t^*) = \frac{D_{Lt}^* - D_{Rt}^*}{D_{Rt}^*}\frac{D_{Rt}^*}{D_{Lt}^*}$$

[7]From the Human Mortality Database, the sum of the annual hazards to age 72 is 0.38. The sum of the monthly hazards (not available from the HMD) should be similar.

Then because $v(D_t^*) < 0$, it is true that $D_{Lt}^* > D_{Rt}^*$, so

$$|v(D_t^*)| < \frac{D_{Lt}^* - D_{Rt}^*}{D_{Rt}^*}$$

Invoking Inequality 2, $\frac{D_{Lt}^*}{D_{Rt}^*} < 1 + |\alpha^{RD}|$,

$$|v(D_t^*)| < |\alpha^{RD}|$$

$\boxed{\textit{Case II: } v(D_t^*) > 0}$

If $v(D_t^*) > 0$, note that as shown earlier, $v(D_t^*) = v(H_t^*) + v(P_t^*)\frac{H_{Rt}^*}{H_{Lt}^*}$. $v(H_t^*) < 0$ and subsequently $\frac{H_{Rt}^*}{H_{Lt}^*} < 1$ because the assumption that education reduces mortality. From inequality 1, it is true that $v(P_t^*) < |\alpha^{RD}|$. Hence

$$|v(D_t^*)| < |\alpha^{RD}|$$

Biases Associated with the Error Terms
We can now derive the biases associated with the error terms.

Discontinuity in Error Term 1: $e_{1t} = \frac{D_t^{EW\#}}{D_t^*}$
Following the definition of this error term, the discontinuity in the error term can be expressed as:

$$e_{1t}^{RD} = \frac{D_{Rt}^{EW\#}}{D_{Rt}^*} - \frac{D_{Lt}^{EW\#}}{D_{Lt}^*}$$

It follows that

$$e_{1t}^{RD} = \frac{D_{Lt}^{EW\#}}{D_{Rt}^*}[v(D_t^{EW\#}) - v(D_t^*)]$$

Given our measurement error assumptions, it is true that $v(D_t^{EW\#}) = -2$. Using this fact along with the definition of $D_{Lt}^{EW\#}$, which simplifies to $\frac{e_d}{2}(D_{Rt}^{EW*} - D_{Lt}^{EW*})$,

$$\begin{aligned} e_{1t}^{RD} &= \frac{e_d}{2}(\frac{D_{Rt}^{EW*} - D_{Lt}^{EW*}}{D_{Lt}^{EW*}})(\frac{D_{Lt}^{EW*}}{D_{Rt}^*})(-2 - v(D_t^*)) \\ &= e_d[-v(D_t^{EW*})](\frac{D_{Lt}^{EW*}}{D_{Lt}^*})(\frac{D_{Lt}^*}{D_{Rt}^*})[1 + \frac{v(D_t^*)}{2}] \end{aligned}$$

Since we only assume that immigrants of post-reform cohorts have different mortality rates than natives, $\frac{D_{Lt}^{EW*}}{D_{Lt}^*} = (1-m)$, where m is the fraction of immigrants in the England/Wales population. Thus,

$$e_{1t}^{RD} = e_d[-v(D_t^*)](1-m)(\frac{D_{Lt}^*}{D_{Rt}^*})[1 + \frac{v(D_t^*)}{2}]$$

We cannot sign this expression because we cannot sign $v(D_t^{EW*})$: the hazard of death decreases with the reform but the at-risk population increases through the ROSLA threshold. However, taking the absolute value of the expression above,

$$|e_{1t}^{RD}| \quad \leq \quad e_d|v(D_t^*)|(1-m)(\frac{D_{Lt}^*}{D_{Rt}^*})[1+\frac{|v(D_t^*)|}{2}]$$

Invoking Inequality 2 and Inequality 3,

$$\frac{|e_{1t}^{RD}|}{|\alpha^{RD}|} \quad < \quad e_d(1-m)(1+|\alpha^{RD}|)(1+\frac{|\alpha^{RD}|}{2})$$

Discontinuity in Error Term 2: $e_{2t} = \frac{D_t^I}{D_t^*}$

Following the definition of this second error and recognizing the $D_{Rt}^I = D_{Lt}^I$ since we assume the population of immigrants is equal on both sides of the threshold:

$$e_{2t}^{RD} \quad = - \quad \frac{D_{Lt}^I}{D_{Rt}^*} v(D^*)$$

Taking the absolute value of this expression and using Inequality 3,

$$\frac{|e_{2t}^{RD}|}{|\alpha^{RD}|} \quad < \quad \frac{D_{Lt}^I}{D_{Rt}^*}$$

Using Inequality 2 and the fact that $D_{Lt}^I < (1-m)iD_{Lt}^*$ where i is the fraction of the immigrants in the England/Wales population,

$$\frac{|e_{2t}^{RD}|}{|\alpha^{RD}|} \quad < \quad i(1-m)(1+|\alpha^{RD}|)$$

Discontinuity in Error Term 3: $e_{3t} = \frac{P_{t+1}^{E*}}{P_{t+1}^*} - \frac{D_t^{E*}}{D_t^*}$

Provided the ROSLA does not change the composition of emigrants, e_{3t} will be smooth even if the ROSLA changes the probability of emigration. In contrast, if the ROSLA "improved" the composition of emigrants, then e_{3t}^{RD} would be positive. The intuition follows: a change in emigration decreases measured deaths but also the measured population. The impacts cancel out provided the composition of emigrants is unchanged.

Algebraically, given that $P_{t+1}^{E*} = mP_{t+1}^*$, $D_t^* = P_{t-1}^* - P_t^* = P_t^*(1-H_t^*)$, and $D_t^{E*} = mD_t^*$, then

$$\begin{aligned} e_{3t} &= (1-m)(1-i) - (1-m)(1-i) \\ &= 0 \end{aligned}$$

Thus,

$$\frac{|e_{3t}^{RD}|}{|\alpha^{RD}|} = 0$$

Discontinuity in Error Term 4: $\frac{f(D^{EW\#})}{P_{t+1}^*}$

This error term is the sum of a number of terms of the form:

$$\frac{D_i^{EW\#}}{P_{t+1}^*}$$

The discontinuity in this term can be expressed, after some algebra, as:

$$e_{4t}^{RD} = \frac{D_{Li}^{EW\#}}{P_{Rt+1}^*}[v(D_i^{EW\#}) - v(P_{t+1}^*)]$$

Since $v(D_i^{EW\#}) = -2$, we have:

$$e_{4t}^{RD} = [-D_{Li}^{EW\#}][\frac{2 + v(P_{t+1}^*)}{P_{Rt+1}^*}]$$

Taking the absolute value and recognizing that $P_{Rt+1}^* > P_{Lt+1}^*$ because we assume that the reform reduces mortality:

$$|e_{4t}^{RD}| \quad < \quad |D_{Li}^{EW\#}|[\frac{2 + v(P_{t+1}^*)}{P_{Lt+1}^*}]$$

Since $D_{Li}^{EW\#} = \frac{e_d}{2} D_{Li}^{EW*} v(D_i^{EW*})$,

$$|e_{4t}^{RD}| \quad < \quad e_d|v(D_i^{EW*})|\frac{D_{Li}^{EW*}}{P_{Lt+1}^*}[1 + \frac{v(P_{t+1}^*)}{2}]$$

Since $v(D_i^{EW*}) = v(D_i^*)$,

$$|e_{4t}^{RD}| < e_d|v(D_i^*)|\frac{D_{Li}^{EW*}}{P_{Lt+1}^*}[1 + \frac{v(P_{t+1}^*)}{2}]$$

Invoking Inequality 3,

$$|e_{4t}^{RD}| < e_d|\alpha^{RD}|\frac{D_{Li}^{EW*}}{P_{Lt+1}^{EW*}}(1-m)|[1 + \frac{|\alpha^{RD}|}{2}]$$

The discontinuity in e_4 is the sum of many such discontinuities, depending how far period t is from the Census period. Although this sum could be as large as 256 terms (21 years (1970-1991) x 12 months + 4 months),[8] our data suggest that $\frac{D_{Li}^{EW*}}{P_{Lt+1}^{EW*}} < \frac{1}{256}$ for any periods t and i.[9] This implies that:

$$\frac{|e_{4t}^{RD}|}{|\alpha^{RD}|} < e_d(1-m)[1 + \frac{|\alpha^{RD}|}{2}]$$

[8]The Census is conducted in April.

[9]Using the Human Mortality Database, the maximum ratio of annual deaths to population is 0.019. Dividing by 12 for a monthly death rate gives 0.00158, well below $\frac{1}{256}$. Therefore, the sum of these terms will be less than 1.

Discontinuity in Error Term 5: $\frac{f(D^I)}{P^*_{t+1}}$

The fifth error term e_{5t} is the sum of a number of terms of the form:

$$\frac{D^I_i}{P^*_{t+1}}$$

Like the fourth term, we can write the discontinuity of the error term, after some algebra, as:

$$e^{RD}_{5t} = \frac{D^I_{Li}}{P^*_{Rt+1}}[v(D^I_i) - v(P^*_{t+1})]$$

Since $v(D^I_i) = 0$ as we assume deaths to immigrants are smooth through the discontinuity,

$$e^{RD}_{5t} = \frac{D^I_{Li}}{P^*_{Rt+1}}[-v(P^*_{t+1})]$$

Thus, since $P^*_{Rt+1} < P^*_{Lt+1}$ as we assume that hazard rates decline with education:

$$|e^{RD}_{5t}| < \frac{D^I_{Li}}{P^*_{Lt+1}}v(P^*_{t+1})$$

Invoking Inequality 1 followed by the fact that $\frac{D^{EW*}_{Li}}{D^*_{Li}} = \frac{P^{EW*}_{Lt+1}}{P^*_{Lt+1}}$,

$$\begin{aligned}
|e^{RD}_{5t}| &< \frac{D^I_{Li}}{P^*_{Lt+1}}|\alpha^{RD}| \\
&< D^I_{Li}|\alpha^{RD}|\frac{D^{EW*}_{Li}}{D^*_{Li}}\frac{1}{P^{EW*}_{Lt+1}} \\
&= \frac{D^I_{Li}}{D^{EW*}_{Li}}\frac{D^{EW*}_{Li}}{D^*_{Li}}\frac{D^{EW*}_{Li}}{P^{EW*}_{Lt+1}}|\alpha^{RD}|
\end{aligned}$$

Since $\frac{D^{EW*}_{Li}}{D^*_{Li}} = (1-m)$ and $\frac{D^I_{Li}}{D^{EW*}_{Li}} = i$,

$$|e^{RD}_{5t}| < i(1-m)\frac{D^{EW*}_{Li}}{P^{EW*}_{Lt+1}}|\alpha^{RD}|$$

Although the discontinuity in e_5 could consist of a sum of up to 256 such terms, $\frac{D^{EW*}_{Li}}{P^{EW*}_{Lt+1}} < \frac{1}{256}$ for any t and i (using the logic invoked earlier for the 4th term) hence:

$$\frac{|e^{RD}_{5t}|}{|\alpha^{RD}|} < i(1-m)$$

Discontinuity in Error Term 6: $\frac{P^{EW\#}_s}{P^*_{t+1}}$

Following the algebra like we did for the 4th and 5th error terms, the discontinuity in the 6th error term can be written as:

$$e_6^{RD} \;=\; -\frac{P_{Ls}^{EW\#}}{P_{Rt+1}^{*}}[v(P_s^{EW\#}) - v(P_{t+1}^{*})]$$

Using the definition of $v(P_s^{EW\#})$ along with the definitions of $P_{Rs}^{EW\#}$ and $P_{Ls}^{EW\#}$, we have

$$e_6^{RD} \;=\; \frac{P_{Ls}^{EW\#}}{P_{Rt+1}^{*}}[\frac{[(1-u_s)e_s + u_s](P_{RS}^{EW*} - P_{LS}^{EW*})}{P_{Ls}^{EW\#}} + v(P_{t+1}^{*})]$$

Multiplying by $\frac{P_{Ls}^{EW*}}{P_{Ls}^{EW*}}$ and then later using the definition of $v(P_s^{*})$,

$$\begin{aligned} e_6^{RD} &= \frac{P_{Ls}^{EW*}}{P_{Rt+1}^{*}}\frac{[(1-u_s)e_s + u_s](P_{RS}^{EW*} - P_{LS}^{EW*})}{P_{Ls}^{EW*}} + \frac{P_{Ls}^{EW\#}}{P_{Rt+1}^{*}}v(P_{t+1}^{*}) \\ &= \frac{P_{Ls}^{EW*}}{P_{Rt+1}^{*}}[(1-u_s)e_s + u_s]v(P_s^{*}) + \frac{P_{Ls}^{EW\#}}{P_{Rt+1}^{*}}v(P_{t+1}^{*}) \end{aligned}$$

Now consider two cases for $P_{Ls}^{EW\#}$: One where $P_{Ls}^{EW\#} > 0$ and the other where $P_{Ls}^{EW\#} < 0$.

Case I: $P_{Ls}^{EW\#} > 0$

We know from previous calculations that $P_{Ls}^{EW\#} = (1-u_s)[(1-\frac{e_s}{2} - \frac{1}{1-u_s})P_{Ls}^{EW*} + \frac{e_s}{2}P_{Rs}^{EW*}]$. Thus, it follows that then $P_{Ls}^{EW\#} < (1-u_s)\frac{e_s}{2}(P_{Rs}^{EW*} - P_{Ls}^{EW*})$. Hence, after some simplification and using the definition of $v(P_s^{*})$, it is true that

$$e_6^{RD} < \frac{P_{Ls}^{EW*}}{P_{Rt+1}^{*}}[(1-u_s)e_s + u_s]v(P_s^{*}) + \frac{P_{Ls}^{EW*}}{P_{Rt+1}^{*}}(1-u_s)\frac{e_s}{2}v(P_s^{*})v(P_{t+1}^{*})$$

Case II: $P_{Ls}^{EW\#} < 0$

In this case, then the second term in the expression for e_6^{RD} will be negative. Thus,

$$e_6^{RD} \;<\; \frac{P_{Ls}^{EW*}}{P_{Rt+1}^{*}}[(1-u_s)e_s + u_s]v(P_s^{*})$$

After simplifying the formula for $P_{Ls}^{EW\#}$, it is true that $P_{Ls}^{EW\#} = (1-u_s)[\frac{e_s}{2}(P_{Rs}^{EW*} - P_{Ls}^{EW*})] - u_s P_{Ls}^{EW*}$. Since mortality rates are a declining function of education, it is true that $-u_s P_{Ls}^{EW*} < P_{Ls}^{EW\#} < 0$. Thus,

$$|e_6^{RD}| \;<\; u_s\frac{P_{Ls}^{EW*}}{P_{Rt+1}^{*}}v(P_{t+1}^{*})$$

Combining Case I and Case II together (i.e., if $|e_6^{RD}| < a$ or $|e_6^{RD}| < b$, then $|e_6^{RD}| < a + b$), it follows that:

$$|e_6^{RD}| \;<\; \frac{P_{Ls}^{EW*}}{P_{Rt+1}^{*}}[(1-u_s)e_s + u_s]v(P_s^{*}) + \frac{P_{Ls}^{EW*}}{P_{Rt+1}^{*}}(1-u_s)\frac{e_s}{2}v(P_s^{*})v(P_{t+1}^{*}) + u_s\frac{P_{Ls}^{EW*}}{P_{Rt+1}^{*}}v(P_{t+1}^{*})$$

Dividing both sides by $|\alpha^{RD}|$, invoking Inequality 1, and using the assumption that hazard rates decline with education such that $P^*_{Rt+1} > P^*_{Lt+1}$, we have

$$\frac{|e_6^{RD}|}{|\alpha^{RD}|} < \frac{P_{Ls}^{EW*}}{P^*_{Lt+1}}\{[(1-u_s)e_s + u_s] + (1-u_s)\frac{e_s}{2}|\alpha^{RD}| + u_s\}$$

Since $P_{Lt+1}^{EW*} = (1-m)P^*_{Lt+1}$,

$$\frac{|e_6^{RD}|}{|\alpha^{RD}|} < \frac{P_{Ls}^{EW*}}{P_{Lt+1}^{EW*}}(1-m)\{[(1-u_s)e_s + u_s] + (1-u_s)\frac{e_s}{2}|\alpha^{RD}| + u_s\}$$

If $\frac{P_{Ls}^{EW*}}{P_{Lt+1}^{EW*}}(1-m) < 1$,[10] then

$$\frac{|e_6^{RD}|}{|\alpha^{RD}|} < 2u_s + e_s(1-u_s)(1 + \frac{|\alpha^{RD}|}{2})$$

Appendix C: The Impacts of the Compulsory Schooling Changes on Earnings

In this Appendix, we report estimates of the earnings impacts of the two law changes. Although other studies have estimated these impacts, our empirical specification differs from that used in those studies, hence it is worth using it to re-analyze these impacts. The main difference between our specification and those used previously is the level at which across-cohort comparisons are made. While the previous literature made these comparisons at the year-of-birth level, we make these comparisons mainly at the month-year of birth level.

Previous estimates

Four studies have used these compulsory schooling law changes to estimate the earnings returns to education: Harmon and Walker (1995), Oreopoulos (2006), Devereux and Hart (2010) and Grenet (2010). All use year-of-birth comparisons. Harmon and Walker (1995) was the first study to use the law changes to generate instrumental variables estimates of the earnings returns to education. They use Family Expenditure Survey (FES) data for the years 1978-1986 and focus on males aged 18-64. The instruments in their model are dummy variables indicating the leaving age facing respondents (fifteen or sixteen) and they also control for survey year, age and age squared. They obtain instrumental variables estimates of the effect of an additional year of education on log earnings of 0.15, much larger than the least squares estimate they calculate, 0.06. They obtain similar estimates when they used an ordered probit to model years of education (to account for the categorical nature of the data).

Oreopoulos (2006) notes that while Harmon and Walker (1995) control for age and survey year, they may not take full account of cohort trends. Oreopoulos (2006) instead focuses on the 1947 change and accounts for cohort trends using a regression discontinuity approach that controls for a fourth-order polynomial in year of birth. In addition to exploiting variation induced by the 1947 change, he also uses a difference-in-difference approach to exploit variation in the timing of this British change relative to a comparable change that took place in Northern Ireland in the 1950s. Using both approaches, he obtains large instrumental variables estimates for men and women, comparable to those found by Harmon and Walker (1995). For example, using only British data, his

[10] For the last period in our dataset, the Human Mortality Database suggests that $\frac{P_{Ls}^{EW*}}{P_{Lt+1}^{EW*}} \sim 1.22$, hence this expression will be true when $m \sim 0.19$. Although this is probably too high, multiplying the final expression here by a number slightly greater than 1 will not change the fact that biases driven by this sixth error term are small.

instrumental variables estimate of the effect of an additional year of education on log earnings is 0.15 with a standard error of 0.06 for men and 0.15 with a standard error of 0.13 for women. These estimates are based on General Household Survey (GHS) data for the years 1983-1998; the sample includes those born in Britain between 1921 and 1951 (aged 32-64 at the time of the surveys).

Devereux and Hart (2010) estimate earnings returns using regression discontinuity models based on similar specifications and estimation samples. Again, they only have access to data at the year-of-birth level. The innovation in their study is the use of large samples of high-quality administrative earnings data. For men, their estimates are much closer to conventional least squares estimates (around 0.05-0.06); for women, their estimates are not statistically different from zero. Devereux and Hart (2010) also report estimates using GHS data and GHS samples that correspond closely to the sample based on the administrative data (excluding the self-employed, including foreign-born). For men, they obtain instrumental variables estimates of the effect of an additional year of education on weekly earnings of nearly 6 percent (standard error of 2 percent). For women the corresponding estimates are 1.5 percent (standard error of 3 percent). They conclude that the earnings returns to additional years of education may be closer to least squares estimates than was previously thought.

Grenet (2010) uses large samples of Labor Force Study (LFS) data to estimate the earnings impacts of the 1972 compulsory schooling change. He uses data at the year-of-birth level, after first standardizing year of birth to be relative to the relevant birth threshold. For both men and women, he obtains estimates close to although slightly lower than least squares estimates (7 percent versus 10 percent). These are, however, reasonably precise (standard errors around 0.03) and appear to be robust to age controls and alternative polynomial specifications.

New estimates based on month-year of birth

Both the 1947 and 1972 law changes were introduced part way through the year. As such, we estimate their earnings impacts using models at the month-year of birth level. Our estimation sample is similar to that used by Devereux and Hart (2010), who also focus on British-only data. In particular, we choose the sample of employees that were born in Britain between 1921 and 1945 and that report working between 1 and 84 hours in the previous week. We also discard observations for which real hourly wages (in 2001) are less than one pound or greater than 150 pounds. Unfortunately, month-year of birth is not available in the GHS until 1985. This means that our samples are smaller than those used by Devereux and Hart (2010). It also means that if we used the same age restrictions (between ages 25 and 60), we would not observe outcomes for those born before 1925. We therefore expand the age range to include those up to age 64; Harmon and Walker (1995) and Oreopoulos (2006) focus on similar age ranges.

Appendix Table C1 reports estimates of the reduced-form effects of the 1947 change on both earnings and on the probability of being observed working a certain number of hours (i.e., selecting into the sample). We report separate estimates for men and women and for each outcome we report estimates from two specifications. In column (1) we report the estimate from a specification similar to that used by Devereux and Hart (2010). That is, we use a global polynomial (over the same data window) with a fourth-order polynomial in the running variable (in our case month-year of birth). In column (2) we report estimates based on a local linear regression approach. The log weekly earnings and log hourly earnings are either condition on 1-84 hours of work per week or normal hours (defined as 35-60 hours of work per week).

If the treatment is defined as "years of primary and secondary schooling," as in Devereux and Hart (2010), the instrumental variable estimates will be roughly twice as large as the reported reduced-form estimates; if the treatment is defined as "years of full-time education" (i.e., primary and secondary plus post-secondary education), the instrumental variables estimates will be closer to three times as large. That is because, empirically, the law changes have a larger impact on total years of primary and secondary schooling than they do on total years of full-time education.

Presumably that is because the extra year of secondary schooling is, for some, a substitute for additional post-secondary education that would otherwise have been obtained. In the extreme case, if all students that leave at the compulsory age pursue at least one year of post-secondary education, and if students compelled to stay for an extra year in secondary school reduce by one year the amount of post-secondary education pursued, the compulsory schooling change will have large impacts on total years of primary and secondary schooling and no impact on total years of full-time education.

The estimates presented in this Table suggest that for men, the 1947 change increased earnings by nearly 8 log points. The global polynomial estimates (column (1)) are very close to the preferred local linear estimates (column (2)). The local linear estimates are also fairly robust to bandwidth choice. As Devereux and Hart (2010) also found, estimates for hourly earnings are slightly smaller than those for weekly earnings. Appendix Figure C1 shows the graph corresponding to the weekly earnings estimates. This figure superimposes fitted lines from the local linear approach on the scatterplot of raw data by month-year of birth. The local linear fit covers a bandwidth of 40 months.

It is tempting to use the point estimates in Appendix Table C1 to claim that the 1947 change had earnings effects on the order of those of Harmon and Walker (1995) and Oreopoulos (2006). We think that our samples are too small to pinpoint the precise magnitude of these effects. For example, when we focus on the 90 per cent of this sample that works "normal" hours, the estimates (second set of rows in Appendix Table C1) are much closer to those reported by Devereux and Hart (2010). Instead, we view our results as consistent with all of the previous literature in that we find that this reform had statistically significant effects on male earnings.

Turning to women, we find negative but not statistically significant effects of the 1947 change on earnings. Again, these effects are smaller when we focus on the 90 per cent of this sample with regular hours (10-50, again defined using the data). These findings are broadly in line with those of Devereux and Hart (2010), who also find negative but statistically insignificant estimates. To the extent that we worry about these negative point estimates, it is important to note that the point estimates for the "observed working" outcomes suggest that the 1947 change may have induced some women to enter the labor force. But they are not statistically significant.

In Appendix Table C2, we present estimates of the effects of the 1972 change, not investigated by Oreopoulos (2006) or Devereux and Hart (2010). Because this reform had smaller impacts on years of education, around 0.2-0.3, it is harder to detect earnings effects in samples of this size. For example, a reduced-form earnings effect of 0.015 would scale up to an instrumental variables estimate of 0.06, roughly the size of the estimate found for the 1947 change by Devereux and Hart (2010). For men, the point estimates exceed 0.015 using small bandwidths but not otherwise. For women with regular hours, point estimates of the earnings effects are also larger, although again imprecise. Both are consistent with the estimates of Grenet (2010), but his estimates are much more precise because he uses labor force survey data rather than health data. The obvious conclusion is that the GHS samples are too small to shed light on the earnings effects of the 1972 change. Our data are more powerful for determining the health and mortality effects of the change given the measurement error in earnings in these survey data.

Appendix D: Migration Issues

As emphasized in the data section, our mortality counts will exclude individuals who left England. As there is some evidence from Malamud and Wozniak (2010) that higher-educated individuals have higher propensities to migrate, migration could cause us to undercount deaths among the post-reform cohorts relative to the pre-reform cohorts. But because the population counts are also affected by emigration, differential emigration rates by cohort will not impact our estimates except to the extent that the ROSLAs affect the composition of emigrants.

Nevertheless, we estimate the effect of the 1947 compulsory school reform on migration to the

United States.[11] Of course, this analysis excludes other large receiving countries. In 2007, Australia (59 thousand emigrants), Spain (21 thousand emigrants), France (20 thousand emigrants), the United States (19 thousand emigrants), and Poland (18 thousand emigrants) were the top five intended countries for emigrants from the United Kingdom (ONS, 2008). Birth cohort details needed for the research design are not available in the public-use versions of the Canadian and Australian Censuses. Emigrants to the United States from the United Kingdom accounted for roughly 6 percent of all emigrant flows in 2007 (ONS, 2008).

To test for migration differences by cohort, we use the 1960, 1970, and 1980 US Census data and the 2005-2007 American Community Surveys (the only datasets with information on quarter of birth). For each dataset, we present separate regression discontinuity estimates of the effect of the reform on the fraction of individuals of each birth cohort born in England or Wales appearing in the United States (Appendix Table D1). Consistent with the results of Malamud and Wozniak (2010), nearly all of these estimates are positive, although only 3 out of 15 are statistically significant at the 5 percent level. For the male/female pooled regressions, the magnitudes vary from a 1-percent negative effect (1980 Census) to a 32-percent positive effect (2006 ACS). Since our mortality estimates are small and our migration estimates are, on balance, positive, we do not think that differential migration patterns can explain our mortality results.

References

Charlton, John, and Mike Murphy. 1997. *The Health of Adult Britain, 1841-1994.* London:Stationery Office.

Crawford, Claire, Lorraine Dearden, and Costas Meghir. 2007. "When you are born matters: the impact of date of birth on child cognitive outcomes in England."

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

Grenet, Julien. 2010. "Is it Enough to Increase Compulsory Education to Raise Earnings? Evidence from French and British Compulsory Schooling Laws."

Harmon, Colm, and Ian Walker. 1995. "Estimates of the Economic Return to Schooling for the UK." *American Economic Review*, 85(5): 1278–1296.

Heady, Patrick, Stephen Smith, and Vivienne Avery. 1996. "1991 Census Validation Survey: Quality Report: The Second Report of a Survey Carried out by Social Survey Division of ONS on Behalf of the Census Offices of Great Britain."

Lleras-Muney, Adriana. 2005. "The Relationship Between Education and Adult Mortality in the United States." *Review of Economic Studies*, 72: 189–221.

Malamud, Ofer, and Abigail Wozniak. 2010. "The Impact of College Education on Geographic Mobility: Identifying Education Using Multiple Components of Vietnam Draft Risk." *NBER Working Paper No. 16463.*

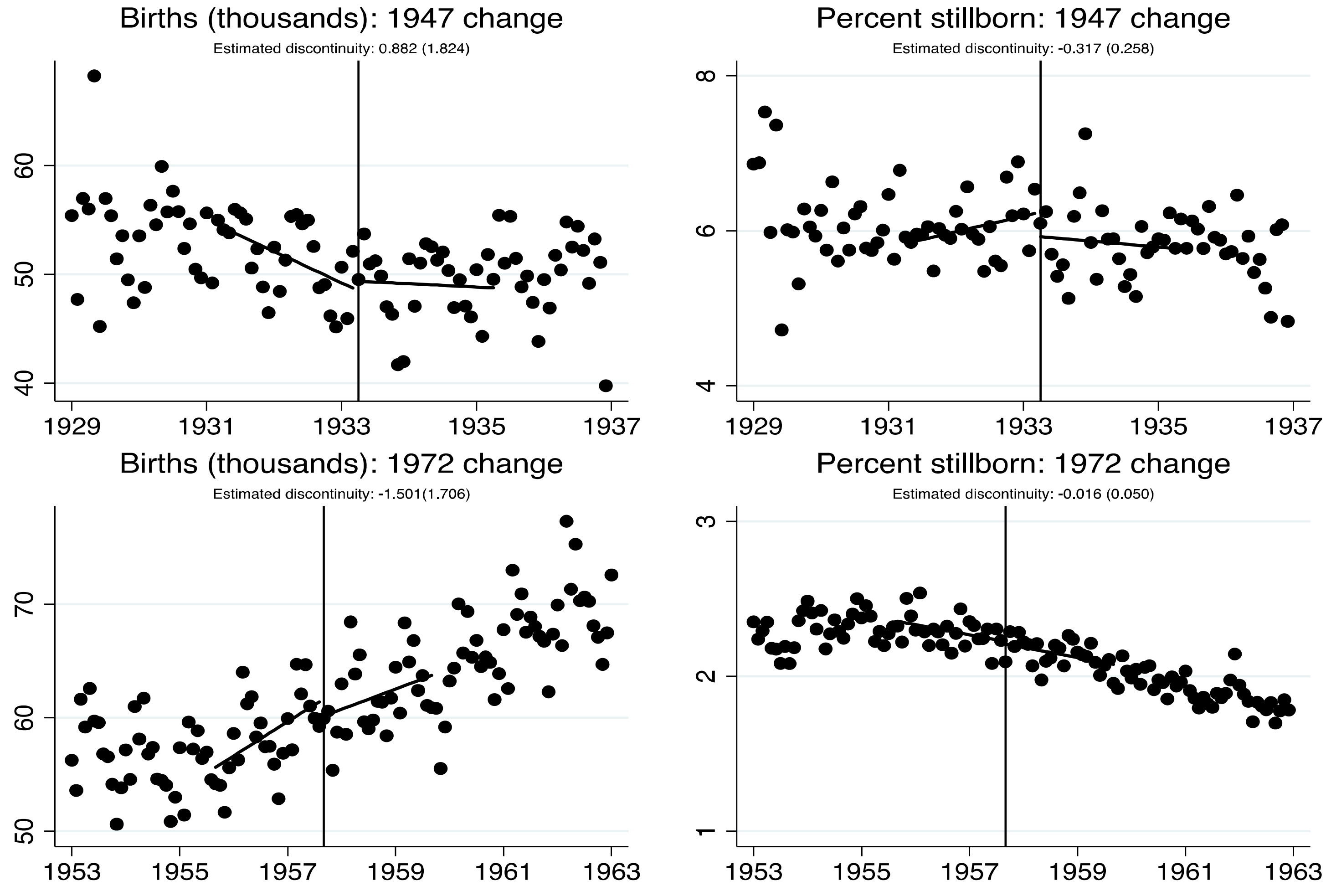
ONS. 2008. "Migration Statistics 2008 Annual Report."

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

Sharp, C. 1997. "Local Education Authority Admission Policies and Practices." In *Four-Year-Olds in School: Policy and Practice: An NFER/SCDC Seminar Report.* , ed. C Sharp and G Turner. London:Newcombe House.

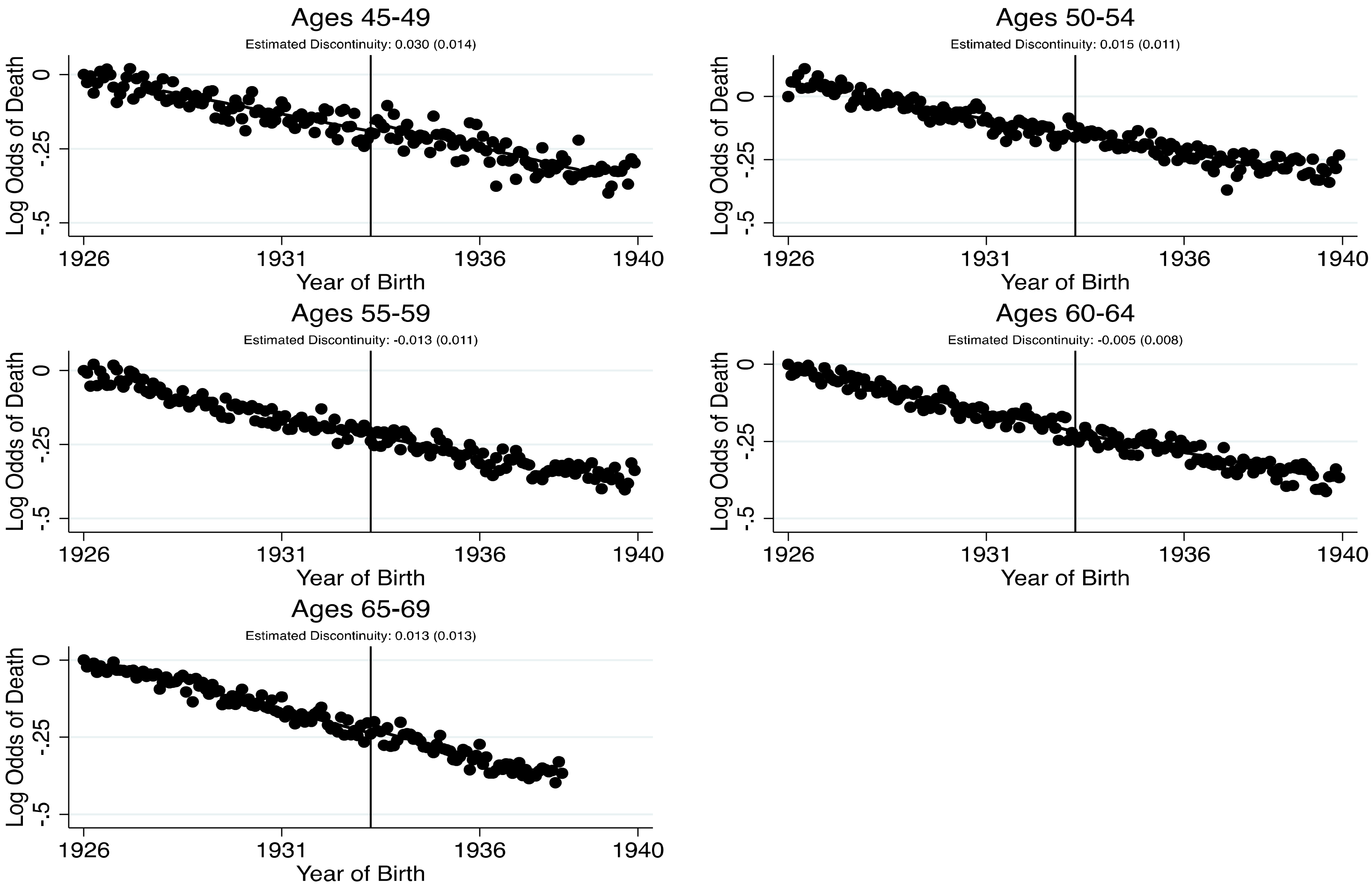
[11] We cannot cleanly identify the effect of the second reform on immigration to the US because of a lack of data by month-year of birth (the 1972 reform affected cohorts born part-way through the third quarter of 1957 - (i.e., September 1957 and later)).

Appendix Figure A1: Cohort size and percent stillborn



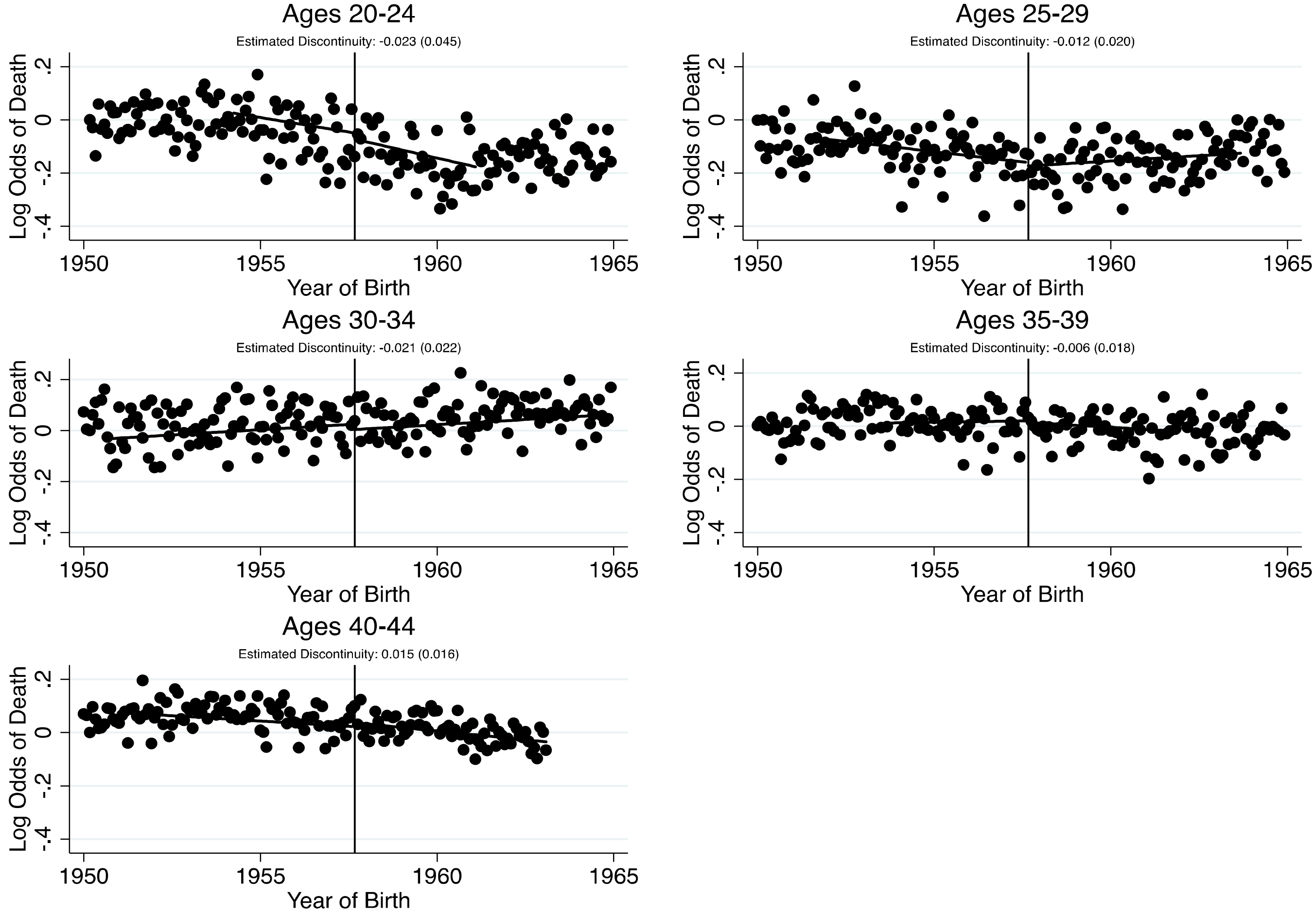
Notes: Estimates reported above panels are based on a regression of the outcome on the relevant ROSLA dummy and a linear cohort trend interacted with this dummy (using 24 observations on either side of the threshold). The estimated discontinuities are based on local linear regressions; the standard errors of the estimates are presented in parentheses. The fitted values of these local linear regressions are also plotted.

Appendix Figure A2: The impact of the 1947 change on mortality by 5-year age group



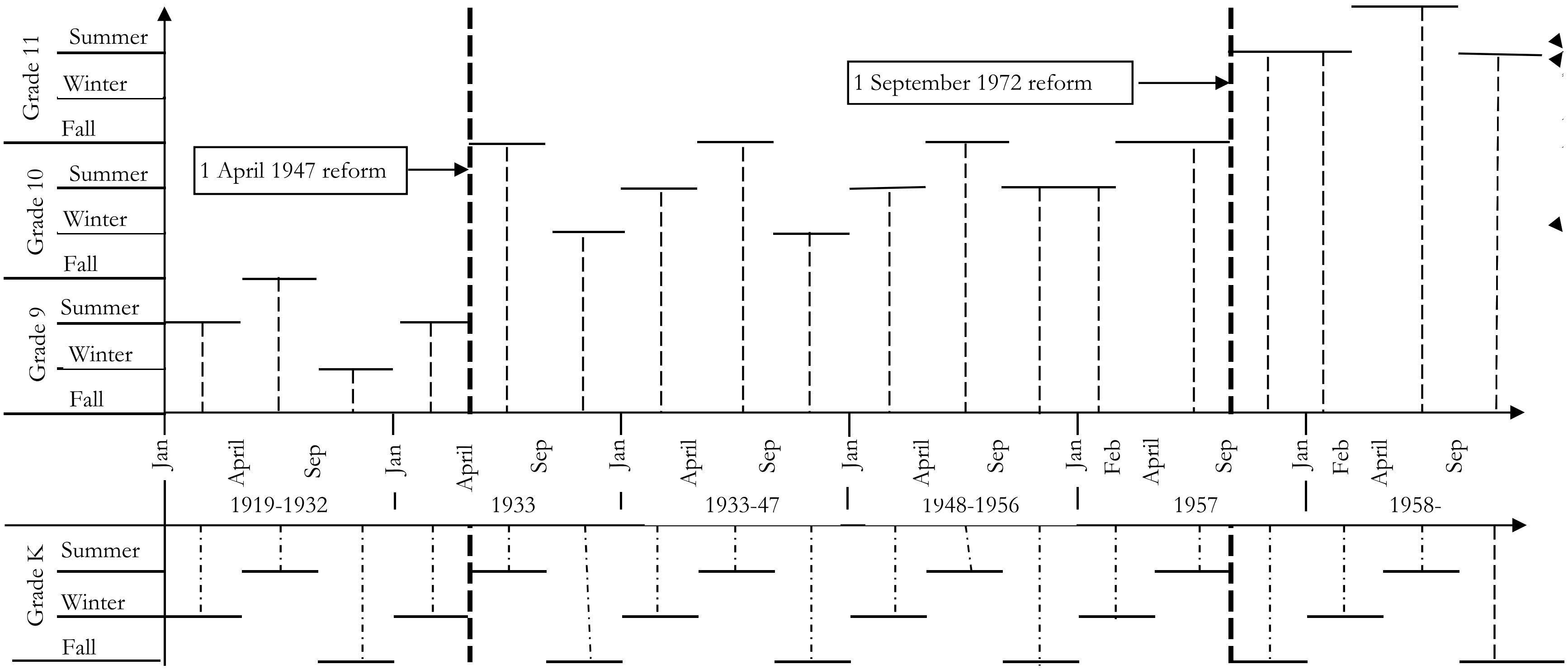
Notes: The log odds ratio is defined as the logarithm of the odds of dying for the relevant cohort relative to the January 1926 cohort. Estimates are based on a balanced cohorts. Points represent the log odds death ratio for each month-year of birth cell. The estimated discontinuities are based on local linear regressions; the standard errors of the estimates are presented in parentheses. The fitted values of these local linear regressions are also plotted.

Appendix Figure A3: The impact of the 1972 change on mortality by 5-year age group



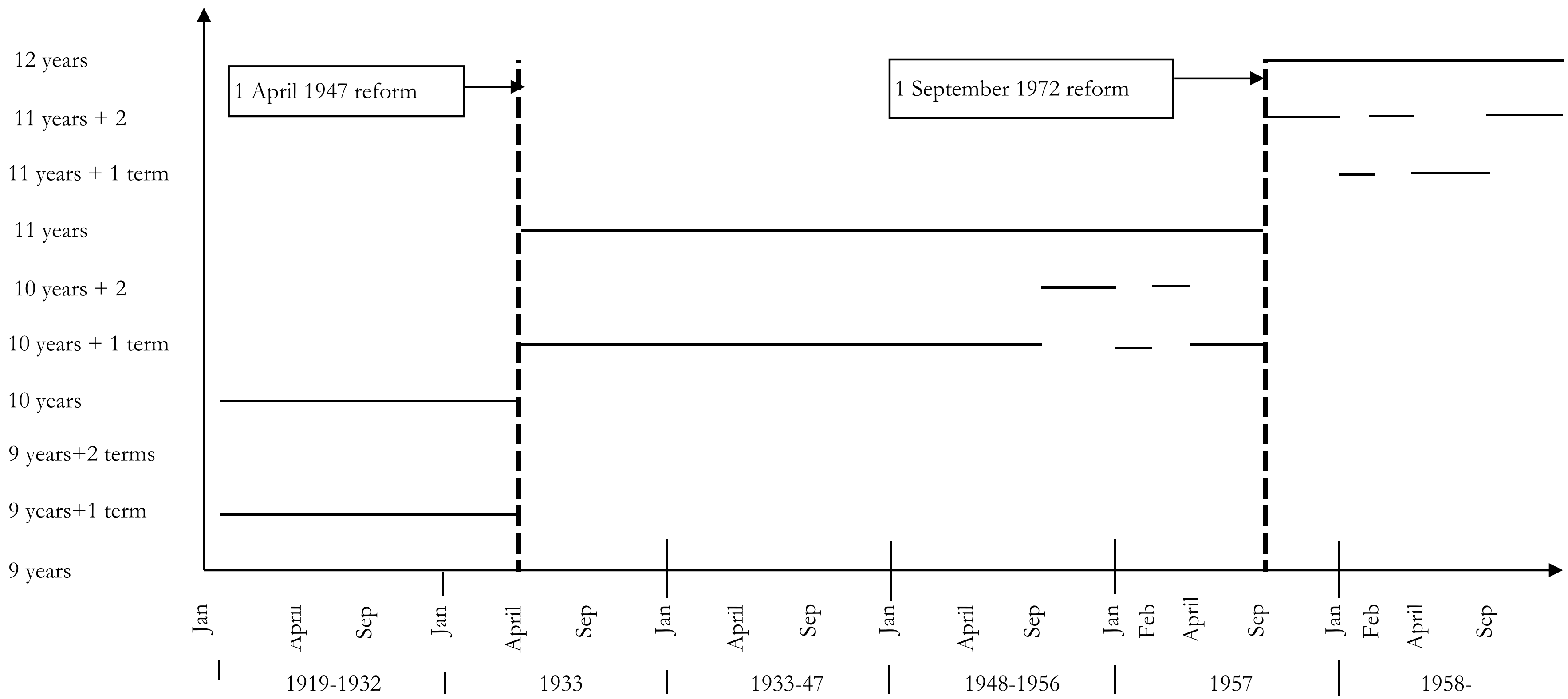
Notes: The log odds ratio is defined as the logarithm of the odds of dying for the relevant cohort relative to the March 1950 cohort. Estimates are based on a balanced cohorts. Points represent the log odds death ratio for each month-year of birth cell. The estimated discontinuities are based on local linear regressions; the standard errors of the estimates are presented in parentheses. The fitted values of these local linear regressions are also plotted.

Appendix Figure B1a: Stylized description of compulsory school laws by month of birth



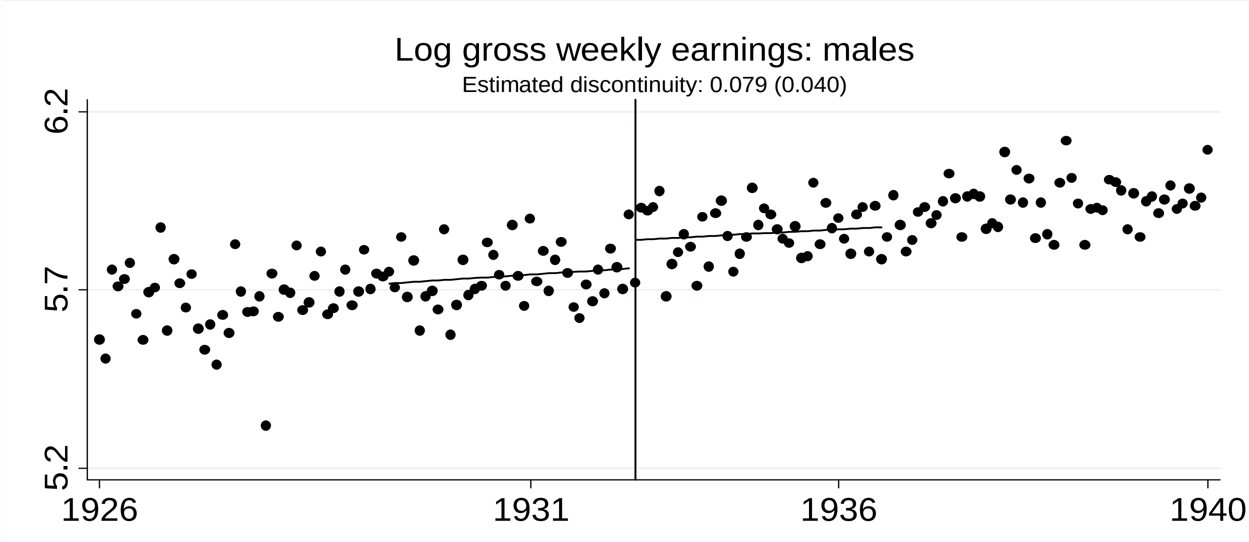
Notes: Not all local schools authorities followed the school entry policy depicted above: some admitted all students at the start of the academic year in which they turned five (i.e., in September); others had two rather than three points of entry.

Appendix Figure B1b: Maximum/Minimum years and terms of full-time education for those leaving at minimum leaving age by birthdate



Notes: The x-axis denotes different birth cohorts. For example, the first part of the x-axis refers to those between January and April over the years 1919-1932. The y-axis displays the minimum and maximum years and terms of full-time education for those who report leaving at the minimum school leaving age. There exists a minimum and a maximum because students that report leaving at the minimum school leaving age (e.g., 15) could have left at the first available opportunity or could have continued for one or two more terms before leaving. For example, people born between January and April in the years 1919-1932 who report leaving at the compulsory school leaving age (14) would have spent 9 years and one term in full-time education had they left at the first opportunity. If they did not leave at the first opportunity, they could have received 9 years and two terms or 10 years of education. Since they report leaving at 14, they could not have received ten years and one term of education (or more) . See Appendix B for more details.

Appendix Figure C1: The impact of the 1947 change on male earnings



Notes: Scatter plot based on sample as described in notes to Appendix Table C1. Fitted solid line based on birth cohort (defined by month-year of birth) interacted with being born after April 1933. Estimated discontintuity refers to coefficient on this dummy variable; its standard error is presented in parentheses.

Appendix Table A2: The impact of the 1947 change on mortality rates by sex

	Overall	Ages 45-49	Ages 50-54	Ages 55-59	Ages 60-64	Ages 65-69
	Panel A: Men					
Reduced-Form Estimate	0.009	0.026	0.031	-0.001	-0.0238	0.021
	(0.006)	(0.015)	(0.016)	(0.011)	(0.0122)	(0.014)
Bandwidth in Months	30	81	51	49	30	29

	Overall	Ages 45-49	Ages 50-54	Ages 55-59	Ages 60-64	Ages 65-69
	Panel B: Women					
Reduced-Form Estimate	0.0116	0.0421	-0.003	-0.006	-0.011	0.022
	(0.0061)	(0.0211)	(0.016)	(0.020)	(0.009)	(0.016)
Bandwidth in Months	41	55	58	32	82	35

Notes: The estimates report the log-odds ratio for the probability of dying for those just to the right of the birth cohort threshold for the 1947 change versus those just to the left of the threshold. The estimates are derived from a two-step procedure described in the text. All regressions use data by month-year of birth cohort from the Office of National Statistics along with Census population counts. Regressions include calendar month-of-birth fixed effects Chosen bandwidths are based on a cross-validation procedure described in the text. Robust standard errors are presented in parentheses.

Appendix Table A3: The impact of the 1972 change on mortality rates by sex

	Panel A: Men					
	Overall	Ages 20-24	Ages 25-29	Ages 30-34	Ages 35-39	Ages 40-44
Reduced-Form Estimate	-0.023	-0.065	0.038	-0.003	0.001	-0.016
	(0.033)	(0.111)	(0.046)	(0.062)	(0.041)	(0.057)
Bandwidth in Months	49	36	69	76	53	42

	Panel B: Women					
	Overall	Ages 20-24	Ages 25-29	Ages 30-34	Ages 35-39	Ages 40-44
Reduced-Form Estimate	0.025	-0.111	-0.122	0.0003	0.097	0.033
	(0.027)	(0.079)	(0.081)	(0.0766)	(0.035)	(0.062)
Bandwidth in Months	43	82	70	76	73	45

Notes: The estimates report the log-odds ratio for the probability of dying for those just to the right of the birth cohort threshold for the 1972 change versus those just to the left of the threshold. The estimates are derived from a two-step procedure described in the text. All regressions use data by month-year of birth cohort from the Office of National Statistics along with Census population counts. Regressions include calendar month-of-birth fixed effects; in the case of the 1972 change, these are allowed to vary on either side of the threshold. Chosen bandwidths are based on a cross-validation procedure described in the text. Robust standard errors are presented in parentheses.

Appendix Table A4: Education effects on summary health measures

Outcomes	OLS	RF	IV	Bandwidth	N
Panel A: 1947 ROSLA					
Health bad, longstanding illness, reduced activity	-0.092	0.025	0.056	60	122,723
	(0.010)	(0.014)	(0.032)		
Currently smoke, ever smoke	-0.030	0.014	0.029	79	77,443
	(0.010)	(0.017)	(0.034)		
Obese, overweight, hypertension	-0.055	0.002	0.005	40	31,177
	(0.018)	(0.025)	(0.050)		
Panel B: 1972 ROSLA					
Health bad, longstanding illness, reduced activity	-0.084	0.001	0.003	43	117,462
	(0.013)	(0.029)	(0.076)		
Currently smoke, ever smoke	-0.187	0.023	0.061	72	94,450
	(0.030)	(0.026)	(0.070)		
Obese, overweight, hypertension	-0.080	0.030	0.089	57	61,896
	(0.015)	(0.048)	(0.151)		

Notes: Each panel presents summary estimates for the group of variables listed in the first column. The estimates are based on standardized versions of these variables hence effect sizes are in standard deviation units (see text for details).

Appendix Table A5: Education effects on clinical health measures

Outcome	OLS	IV-2	IV-3	IV-4	OLS (LM)	N
Obese (BMI>30)	-0.021	0.018	0.016	0.028	-0.016	97,115
(Depvar mean 0.21)	(0.001)	(0.016)	(0.015)	(0.016)	(0.005)	
p-value (overid/LM)		0.27	0.31	0.39	0.00	
Overweight (BMI>25)	-0.024	0.010	0.003	0.000	-0.001	97,115
(Depvar mean = 0.622)	(0.001)	(0.017)	(0.016)	(0.017)	(0.005)	
p-value (overid/LM)		0.22	0.29	0.12	0.95	
BMI	-0.279	0.270	0.205	0.312	-0.081	97,115
(Depvar mean = 26.83)	(0.010)	(0.173)	(0.164)	(0.162)	(0.050)	
p-value (overid/LM)		0.02	0.04	0.43	0.01	
Hypertension	-0.010	-0.029	-0.029	-0.025	-0.018	79,298
(Depvar mean = 0.365)	(0.001)	(0.018)	(0.018)	(0.018)	(0.006)	
p-value (overid/LM)		0.54	0.55	0.97	0.71	
Diastolic blood pressure	-0.058	-0.094	0.038	-0.389	-0.296	79,298
(Depvar mean = 76.01)	(0.027)	(0.445)	(0.423)	(0.423)	(0.146)	
p-value (overid/LM)		0.003	0.01	0.36	0.83	

Notes: This table presents estimates of the effects of years of education on various outcomes using data from the Health Survey of England (pooled waves 1991-2004). See notes to Table 5b.

Appendix Table C1: Earnings impact of 1947 change to compulsory schooling

	Panel A: Men						Panel B: Women					
	Log Weekly Earnings		Log Hourly Earnings		Observed Working		Log Weekly Earnings		Log Hourly Earnings		Observed Working	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Hours: 1-84	0.070	0.081	0.038	0.049	-0.001	-0.006	-0.072	-0.079	-0.044	-0.041	-0.005	0.011
	(0.028)	(0.037)	(0.022)	(0.031)	(0.017)	(0.023)	(0.045)	(0.053)	(0.025)	(0.031)	(0.015)	(0.020)
N	11964	3462	11964	3462	22553	7181	11133	3075	11133	3075	24364	7792
Normal hours	0.047	0.044	0.035	0.028	0.010	0.005	-0.069	-0.049	-0.053	-0.026	-0.016	0.000
	(0.022)	(0.031)	(0.022)	(0.031)	(0.015)	(0.019)	(0.043)	(0.050)	(0.025)	(0.029)	(0.015)	(0.020)
N	11022	3156	11022	3156	22553	7181	10152	2787	10152	2787	24364	7792

Notes: In column (1) we report reduced-form estimates from a global polynomial regression specification similar to that used by Devereux and Hart (2008), the main difference being that the column (1) estimates are generated using a month-year of birth discontinuity design, as opposed to the year-of-birth discontinuity design of Devereux and Hart (2008). In column (2) we report reduced-form estimates from local linear regression models of outcomes on month of birth, a dummy for birth month after April 1933, and an interaction of birth month and this dummy. The bandwidth is 40 months on either side of April 1933. The observed working outcomes refer to the probability of being observed to work 1-84 hours (in the top panel) and working normal hours in the bottom panel. These are not conditional on working at all, hence can be interpreted as the probability of being in the samples used in the earnings analysis. Normal hours, determined from the data, are defined as 35-60 for males and 10-50 for females.

Appendix Table C2: Earnings impact of 1972 change to compulsory schooling laws

	Panel A: Men			Panel B: Women		
	Log Weekly Earnings	Log Hourly Earnings	Observed Working	Log Weekly Earnings	Log Hourly Earnings	Observed Working
Hours: 1-84	0.067	0.097	0.041	0.073	0.044	-0.086
	(0.056)	(0.052)	(0.026)	(0.073)	(0.032)	(0.028)
N	6545	6545	9327	5964	5964	10462
Normal hours	0.104	0.111	0.056	0.113	0.035	-0.105
	(0.057)	(0.057)	(0.020)	(0.064)	(0.038)	(0.038)
N	6168	6168	9327	5359	5359	10462

Notes: We report reduced-form estimates from local linear regression models using the month-year of birth discontinuity design of outcomes on birth month, a dummy for birth month after September 1957, and an interaction of birth month and this dummy. The bandwidth is 40 months on either side of September 1957. Normal hours, determined from the data, are defined as 35-60 for males and 10-50 for females.

Appendix Table D1: Effect of 1947 change to compulsory schooling laws on migration to US

	Panel A: Men					
Reduced-Form Estimate	0.145	-0.039	0.059	0.099	-0.098	0.040
	(0.079)	(0.137)	(0.079)	(0.105)	(0.127)	(0.149)
Percent in US for 1933Q1	0.13	0.79	0.55	0.19	0.46	1.08
Bandwidth in Quarters	20	34	14	27	28	28
Dataset	1960 Census	1970 Census	1980 Census	2005 ACS	2006 ACS	2007 ACS
	Panel B: Women					
Reduced-Form Estimate	0.686	0.649	-0.084	0.0004	0.225	0.357
	(0.523)	(0.216)	(0.127)	(0.2202)	(0.200)	(0.171)
Percent in US for 1933Q1	1.23	1.38	1.49	1.06	0.76	1.23
Bandwidth in Quarters	10	35	14	30	31	30
Dataset	1960 Census	1970 Census	1980 Census	2005 ACS	2006 ACS	2007 ACS
	Panel C: Overall					
Reduced-Form Estimate	0.216	0.332	-0.010	0.041	0.089	0.014
	(0.209)	(0.156)	(0.067)	(0.131)	(0.127)	(0.132)
Percent in US for 1933Q1	0.67	1.08	1.01	0.62	0.60	1.15
Bandwidth in Quarters	12	27	14	28	30	23
Dataset	1960 Census	1970 Census	1980 Census	2005 ACS	2006 ACS	2007 ACS

Notes: This table presents regression discontinuity estimates of the effect of the 1947 reform on immigration to the United States from persons born in England or Wales. The bandwidth is chosen using the cross-validation procedure discussed in the text. These regressions include a fully-interacted linear polynomial in quarter-year of birth cohort relative to April 1947. The migration rate equals the number of persons appearing in the US dataset divided by the size of the quarter-year of birth cohort multiplied by 100. All of these regressions use data by quarter-year of birth cohort data from the Office of National Statistics. Calendar quarter-of-birth fixed effects are included in the regressions. Regressions are weighted by the number of births. Reported bandwidths are in quarters. Robust standard errors are presented in parentheses.