ARTICLE IN PRESS

Labour Economics 48 (2017) 87–104

Contents lists available at ScienceDirect

Labour Economics

journal homepage: www.elsevier.com/locate/labeco

# Returning to returns: Revisiting the British education evidence ☆

Peter Dolton [a,c,d], Matteo Sandi [*,a,b]

[a] Department of Economics, University of Sussex, Jubilee Building, Falmer, Brighton BN1 9SL, UK
[b] Centre for Economic Performance, London School of Economics, Houghton Street, London WC2A 2AE, UK
[c] LEMMA, Universitè de Paris II, Pantheon Assas
[d] National Institute of Economics and Social Research, 2 Dean Trench St, London SW1P 3HE, UK

ARTICLE INFO

*Article history:*
Received 25 September 2016
Received in revised form
3 July 2017
Accepted 4 July 2017

*Keywords:*
Returns to education
Schooling
Instrumental variables

*JEL classification:*
H52
I21
I26
I28

ABSTRACT

We revisit the question of what is the rate of return to education in Great Britain. We make two contributions. Firstly, we re-assess the robustness of Harmon and Walker (1995), Oreopoulos (2006) and Devereux and Hart (2010) to equation specification and estimation method. Secondly, we generalize the previous IV approaches by using the month of birth in the calculation of a more accurate IV exploiting the 1947, 1963 and 1972 UK School Leaving Age reforms. Our results highlight the importance of equation specification and they provide a robust case for a 6% Rate of Return to Education for men.



The rate of return to education (RoRtE) is of central importance to education policy. It has variously been estimated that this rate of return could be as low as zero (e.g., Pischke and von Wachter, 2008) or as high as over 15% (e.g., Harmon and Walker, 1995; Buscha and Dickson, 2012). If the RoRtE is zero then a continued policy of high level incentives to acquire more education is difficult to justify on financial grounds. In contrast, if the RoRtE is 15% per annum or greater, then parents and education authorities should provide strong incentives for staying at school for longer. Many papers have estimated this parameter (e.g., Angrist and Krueger, 1991; Card, 1995; Harmon and Walker, 1995; Acemoglu and Angrist, 2001; Kling, 2001; Oreopoulos, 2006; Devereux and Hart, 2010; Carneiro et al., 2011). It is therefore not surprising that there is considerable heterogeneity in the estimates retrieved for this single parameter. While the large number of estimates reflects the attention devoted by economists to this parameter, the heterogeneity of these results makes it difficult to inform education policy.

A number of reasons may lie behind these heterogeneous estimates of the RoRtE. First, returns to education are likely to differ across different types of individuals (see, e.g., Kling 2001, Koop and Tobias 2004; Carneiro et al., 2011). Second, heterogeneity in previous estimates may simply result from the fact that the RoRtE may differ across countries. For example, Pischke and von Wachter (2008) document a zero return to education in Germany; Grenet (2013) finds a return to education close to zero in France; Devereux and Hart (2010) report a 6% return to education in Britain. In part, heterogeneity in these estimates of the RoRtE may also derive from the fact that these were estimated for different labor markets and at different points in time. For example, Devereux and Hart (2010) exploit an education reform that took place in 1947 in Britain while Grenet (2013), instead, examines an education reform that took place in 1967 in France. Third, the RoRtE may also differ at different margins of education: the return to one more year of schooling at, e.g., age 14–15, may not be the same as the return to one more year of university.

What is less well-known is the extent to which estimates may vary due to econometric specification and the estimation techniques employed. This can be investigated using consistent data over one period of time in a single country. In particular, in this paper

☆The authors wish to thank the Editor of Labour Economics, two anonymous referees, Eric Hanushek, Stephen Machin, Paolo Masella, Vikram Pathania, Steve Pischke, Erik Plug, Jeff Smith, John Van Reenen, Richard Tol, Ian Walker, the participants at the CESifo Education conference 2014, and the participants at the RES conference 2015, for helpful comments on our early results. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.
* Corresponding author. Present address: Centre for Economic Performance, London School of Economics, Houghton Street, London WC2A 2AE, UK.
*E-mail addresses:* p.dolton@sussex.ac.uk (P. Dolton), m.sandi@lse.ac.uk (M. Sandi).

http://dx.doi.org/10.1016/j.labeco.2017.07.001
0927-5371/

we examine the sensitivity of some previous 2SLS estimates of the RoRtE to equation specification and to the definition of the instrumental variable (IV) used. Since the estimation of the earnings returns to education involves a well-known problem of endogeneity, much of the RoRtE literature has relied on IV estimation to retrieve unbiased and consistent treatment effects (e.g., Angrist and Krueger, 1991; Card, 1995; Harmon and Walker, 1995; Acemoglu and Angrist, 2001; Kling, 2001; Carneiro et al., 2011). In settings where the treatment is determined partly by whether the assignment variable crosses a cutoff point, i.e., “fuzzy” regression discontinuity (RD) designs, two-stage IV estimation strategies have been commonly employed (e.g., Oreopoulos, 2006).

In order to conduct this analysis we focus on the empirical literature on the RoRtE from Great Britain. In this literature, two Raising of the School-Leaving Age (ROSLA) changes have received particular attention. The first took place in 1947, when the minimum school leaving age was raised from 14 to 15. The second reform took place in 1972, when the age of compulsory schooling increased from 15 to 16. An additional reform, that did not receive as much attention in this literature, also took place in Great Britain in 1963[1]; in that year, the 1962 Education Act came into force, providing, among other things, a modification in the actual school leaving dates for pupils born in certain months of the calendar year. Unlike compulsory schooling law changes in the United States, that affected only 5% of the relevant cohorts (Lleras-Muney, 2005),[2] a high fraction of population was affected by the ROSLA changes in Great Britain (Oreopoulos, 2006). The large fraction of population affected by the ROSLA reforms arguably makes the Local Average Treatment Effect (LATE) of these ROSLA reforms likely to be closer to an Average Treatment Effect (ATE) of the RoRtE (Oreopoulos, 2006); this, in turn, made Great Britain a particularly interesting context for the RoRtE literature.

Key contributions to the RoRtE literature in Great Britain are Harmon and Walker (1995), Oreopoulos (2006), and Devereux and Hart (2010). Harmon and Walker (1995) calculate a 15% RoRtE; Oreopoulos (2006) finds a 10–14% RoRtE; finally, Devereux and Hart (2010) conclude that the RoRtE is 6% or lower. All these studies focus only on men. While, in principle, all these studies analysed the RoRtE in similar labour markets and with a comparable institutional structure, nonetheless they retrieved different RoRtE estimates. Harmon and Walker (1995) was in fact criticized by Card (1999) for not adequately controlling for the underlying trends in schooling achievement and earnings across the different cohorts in their sample; Oreopoulos (2006) and Devereux and Hart (2010) both estimate a fuzzy RD model using a fourth-order polynomial of year of birth to control for the underlying heterogeneity (in both schooling achievement and earnings) of the sample across the years of the data.[3] Following Lee and Lemieux's (2010) and Gelman and Imbens' (2014) recommendations, we test the robustness of RD analysis to the choice of alternative polynomials in the running variable. Failing to use a regression specification which fully nets out for underlying trends and extraneous influences on the schooling decision – and its impact on future earnings – might explain the divergent findings observed in this literature.

This study makes two contributions. Firstly we re-assess the robustness of the key contributions from Great Britain, namely Harmon and Walker (1995), Oreopoulos (2006) and Devereux and Hart (2010). Due to the historically high dropout rates in Great Britain, and the remarkable effect of these policies on overall schooling attainment, this constitutes a particularly suitable context for the investigation of how robust existing RoRtE estimates are to the choice of the polynomial to specify trends in education and earnings over time. Using all the available survey years from the General Household Survey (GHS) and the Family Expenditure Survey (FES) data, which were used in the reviewed studies, we examine the sensitivity of the results to the specification of the relevant earnings function and the sample used. In contrast to previous papers, we devote particular attention to controlling for the underlying trends in schooling and earnings. We seek the optimal specification of the relevant earnings function using Akaike's criterion. We also test the robustness of previous results to the use of nonparametric estimation techniques in the presence of an RD design. Our replication suggests how these heterogeneous results can be reconciled. We find that previous estimates of the returns to compulsory education are sensitive to the functional form chosen. The use of different-order polynomials results in significant differences in the estimated RoRtE.

Having documented the sensitivity of previous estimates to trend specification, in the second part of our analysis we define a more precise 2SLS strategy that takes into account all ROSLA reforms implemented in Great Britain in the post-World War II period. We do this in two ways: first, we calculate a multivalued IV that reflects the exogenously induced extra compulsory schooling that different cohorts of pupils actually faced. Second, we instrument the schooling decision of pupils using a set of mutually exclusive dummies for the 1947, 1963 and 1972 reforms, again reflecting the exogenously induced extra compulsory schooling faced by different pupils over time. Compared to our first approach, use of mutually exclusive dummies for different ROSLA reforms appears complementary, as it does not constrain the effects on schooling and earnings to be the same for the different ROSLA reforms. These reforms affected different complier groups, both in terms of size and therefore, potentially, also in terms of unobserved characteristics and tastes for education. The 1947 ROSLA affected half the relevant cohort, whereas the 1972 ROSLA affected between a quarter and a third of the cohort; this implies that the complier group to this latter reform only includes the third of the cohort who do not have the ability or desire to continue in school beyond 15. Since the relevant compulsory schooling changes depended on month-year-of-birth, in contrast with the previous literature, our approach also incorporates the additional variation within year-of-birth cohorts in the school leaving dates implied by these reforms. Our IVs attempt to add precision to the description of the exogenously induced amount of compulsory schooling implied by all ROSLA reforms in 1947, 1963 and 1972. We use all of these three reforms in order to exploit the entire variation in the data (in order to retrieve more robust and precise estimates), and we control in several different ways for the underlying trends in education and earnings.

The main conclusion of our empirical research is that the RoRtE based on the ROSLA reforms in the UK is 6% for males; the choice of alternative polynomial terms to measure the underlying trends in education and earnings appears important in the determination of the results. This conclusion holds also when we use our new instrumental variable or the set of mutually exclusive dummies in our analysis. However, our new estimates do appear more robust compared to those we replicate from previous studies. Since we use different ROSLA reforms to instrument the schooling decision of pupils, we interpret our calculated RoRtE as a weighted average of the Marginal Policy Relevant Treatment Effect (MPRTE) estimates of these different ROSLA reforms. In the next section we define our terms and review the estimation problem involved in determining the RoRtE. In Section 2 we examine the ROSLA reforms implemented in Britain from the World War II onwards. In Section 3 we review the

[1] We are aware of only three studies that incorporate the 1963 reform in their analyses, namely Del Bono and Galindo-Rueda (2004, 2007), and Dickson and Smith (2011).

[2] The effects in other countries – e.g., Norway, Canada, and France – are similarly small (Albouy and Lequien 2009; Black et al., 2008; Lleras-Muney 2005; Oreopoulos 2006).

[3] Harmon and Walker (1995), in contrast, do not control for any polynomial in the date of birth.

existing literature on the returns to compulsory education in Britain. In Section 4 we describe the strategy employed in our replication analysis and demonstrate the specification sensitivity of some previous estimates of the returns to education. In Section 5 we present our proposed 2SLS alternative, with our redefined instrumental variables. In Section 6 we conclude.

## 1. Analysis of the rate of return to education

Following the notation of Carneiro et al. (2011), for ease of exposition we define the standard simplified Mincer earnings equation as:

$$Y = \alpha + \beta S + \epsilon \tag{1}$$

where $Y$ is the log wage, $S$ indicates schooling years and $\varepsilon$ is a residual. Define potential outcomes $Y_1$ and $Y_0$ as earnings with $s=1$ if an extra year of schooling is undertaken, and $s=0$ otherwise:

$$Y_0 = \mu_0(X) + u_0 \text{ and } Y_1 = \mu_1(X) + u_1 \tag{2}$$

where $X$ is the set of relevant exogenous covariates and

$$\mu_s(X) = E\{Y_s|X = x\} \quad \forall \quad s = 0, 1 \tag{3}$$

The return to schooling is:

$$Y_1 - Y_0 = \beta = \mu_1(X) - \mu_0(X) + u_1 - u_0 \tag{4}$$

and we can define the Average Treatment Effect (ATE) of a year of schooling as:

$$\bar{\beta}(x) = E(\beta|X = x) = \mu_1(X) - \mu_0(X) \tag{5}$$

Assuming that $R=1$ if the individual is subject to a ROSLA reform (of an additional year) and $R=0$ otherwise, we can specify the LATE(ROSLA) as:

$$E\{\beta|X = x, \ R = 1\} = \bar{\beta}\ (x) + E(u_1 - u_0|R = 1, X = x) \tag{6}$$

under the assumption that $u_s \perp R$ and $R=S+1$

$$E\{\beta|X = x, \ R = 1\} = \mu_1(X) - \mu_0(X) \tag{7}$$

Carneiro et al. (2011) also define a marginal treatment effect (MTE) of education as:

$$MTE(X, V_s) = E(\beta|X = x, V_s = v_s), \tag{8}$$

where $V_s$ is the cumulative distribution function of the stochastic errors in the schooling choice equation. In Carneiro et al.'s (2011) study, this schooling choice is endogenous and they examine the use of a set of IVs, Z, and use techniques analogous to propensity score matching to allow them to retrieve this MTE:

$$MTE(x, p) = \frac{\partial E\left(Y|X = x, \ P(Z) = p\right)}{\partial p} \tag{9}$$

where $p$ is the probability of a discrete change in the schooling decision function. In the case of Carneiro et al. (2011), their estimation strategy is predicated on the presence of valid IVs and the conditional independence of $X$ and $Z$ given ($U_0$, $U_1$,$V_s$). This set up allows them to retrieve the MTE as a function of the unobserved heterogeneity type $V_s$ (scaled from 0–1), where $V_s=1$ are the least likely to undertake extra schooling and $V_s=0$ are the most likely to undertake extra schooling.

It is then possible to consider different policy relevant education treatments that act on $Z$ or $p$ and retrieve their marginal policy relevant treatment effects (MRPTE). In our case, our $Z$ is $R$ – the ROSLA reform, which is exogenous and not related to the schooling decision of the individual. Since we can observe

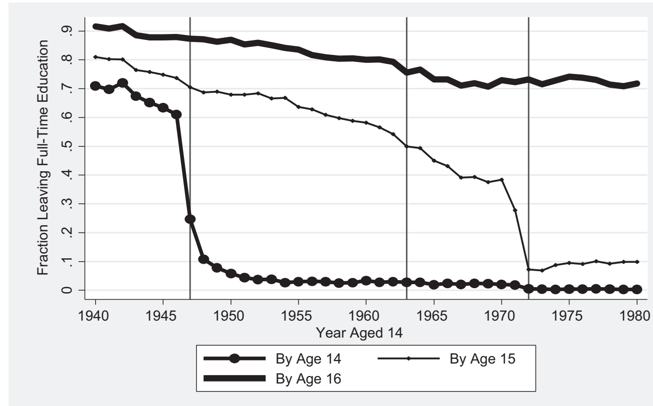


**Fig. 1.** Fraction left full-time education by year aged 14, 15 and 16 *(pooled sample, GHS 1983–98).*

different discrete values of $R$, $r_1$,$r_2$…..,…, $r_\tau$ then we can approximate the MPRTE at the different values of the discrete values of R:

$$MPRTE(x, r_\tau) = \frac{\partial E\left(Y|X = x, \ \ R = r_1, r_2 \ldots . r_\tau\right)}{\partial r_\tau} \tag{10}$$

It is worth noting that, in the presence of an exogenous change in the ROSLA, the estimand of a fuzzy RD design can be interpreted as a weighted version of Eq. (7), i.e., a weighted local average treatment effect (LATE), where the weights reflect the ex-ante probability of the observation's assignment variable being near the threshold (see Lee and Lemieux, 2010). By defining $R$ not just as a binary variable, and by allowing $R$ to take up more than two values, we can potentially retrieve not only the LATE(ROSLA) but also the MPRTE of education – on the assumption that different values of the extra, discrete elements of $R$ are independent of $u_s$.

A second important contribution of Carneiro et al. (2011) was to estimate the MTE function as a function of the unobserved heterogeneity in the schooling equation. We do not do this, as we are not attempting to model the endogenous schooling decision directly. However we can potentially estimate the MPRTE for specific discrete values of $R$. Moreover, the relevance of the discrete policy changes to school leaving ages that we exploit in our analysis implies that we also avoid the problems of weak or invalid instruments often encountered in this literature.[4]

## 2. Contextual background and the compulsory schooling law changes

The first major change to the age of compulsory schooling was enacted in Britain in 1947 as a result of the 1944 Education Act which raised the minimum school leaving age from 14 to 15. The 1944 Education Act also gave power to the Minister of Education to raise the age of compulsory schooling from 15 to 16, at the earliest possible convenience. The Education Minister did so in March 1972, and the age of compulsory schooling was raised to 16 on September 1st 1972. Fig. 1 was calculated on the entire British-born sample available from the 1983–1998 GHS survey years. It illustrates the impact of these policies on the fraction of pupils leaving school by age 14, 15 and 16 by year of birth. The reason for not using later GHS survey years is that no survey was held in 1999, and in that year the survey was redeveloped and relaunched in 2000 with a different design.[5] This is similar to what

[4] See for example Angrist and Krueger (1991), and Clark and Royer (2013).

[5] After 1998, information on the month of birth is only reported in the GHS 2000–01.

Oreopoulos (2006) and Devereux and Hart (2010) have done, and in this, as for most of our analysis, we follow their approach. Finally, information on month of birth is only available starting from the 1983 GHS survey - which explains why we cannot use earlier GHS survey years.[6]

It is apparent from Fig. 1 that both laws had a strong and clear impact on school leaving behavior. In both 1947 and 1972, the fraction of pupils leaving school before the (new) minimum school leaving age dropped sharply; the fraction leaving school before age 15 fell from roughly 60% for the cohort that turned 14 just before April 1947 to roughly 10% for the cohort that turned 14 immediately after the cutoff point. In 1972, for pupils that turned 15 at around the cutoff point, the change in minimum school leaving age decreased the proportion of pupils leaving school by 15 years of age from 35% to less than 10%. Both policies resulted in many British pupils still leaving education at the earliest possible convenience.[7] These large, exogenously induced, discrete changes in the fraction of children staying on at school were first used by Harmon and Walker (1995) in an IV estimate of the RoRtE.

A careful analysis of the evolution of the regulations that determined the period of compulsory schooling in Great Britain over the post-World War II period reveals that the exposure to compulsory full-time education did not only vary as a result of these two reforms. There was a third reform that varied the term a pupil was entitled to leave school. The British education system has three terms that run September–December, January–April, and April–July, with precise dates varying by school and Local Education Authority (LEA). Until 1963, students had to stay in school until the end of the term in which they reached the minimum school leaving age. In 1962 a new Education Act was passed; as a result, starting from 1963 these laws changed, and pupils born September–January had to attend school until Easter, whereas those born February–August had to attend school until June. The first cohort to be directly affected by the 1963 reform were pupils who reached the minimum school leaving age in September 1963 (i.e., when the new law came into force).[8] Fig. 1 shows that the effect of this reform was only about a five percent increase in the overall staying-on rate for 15 and 16 year olds. However, this effect is partly understated as Fig. 1 looks at the whole year effects and not at the within-school-year term effects.

In comparison with the cohort of pupils born before April 1st 1933, the 1947 reform resulted in 12 additional months of compulsory schooling for pupils born from April 1st 1933 up till August 31st 1948. Starting from those born on September 1st 1948, the interaction of the 1947 reform and 1963 reform implied 15 additional months of compulsory schooling for the cohorts born September–December and February–March (i.e., one more year, plus one term). Starting from September 1st 1957, 12 additional months of compulsory schooling should be added to this: starting from those born on September 1st 1957, the 1972 reform implied 24 additional months of compulsory schooling for the cohorts born in January and April–August; for the cohorts born September–December and February–March, the interaction of the 1963 reform and 1972 reform implied 27 additional months of compulsory schooling (i.e., two more years, plus one term).[9] In all these figures

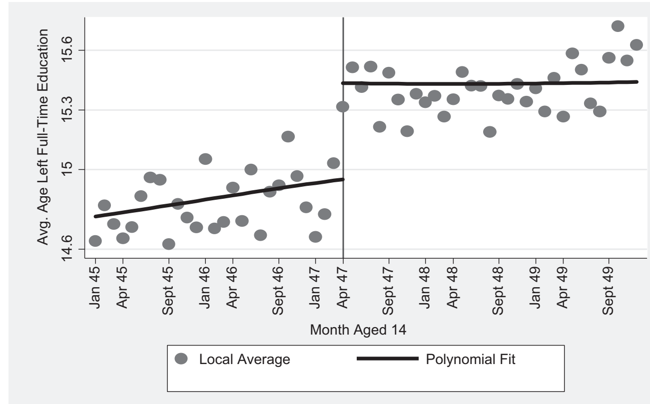


**Fig. 2.** 1947 Reform: average age left full-time education by month aged 14 *(1931–35 birth cohorts – male sample, GHS 1986–98).*

the comparison is with the birth cohorts prior to April 1st 1933.

The 1962 Education Act did not only vary the length of compulsory schooling for pupils born in certain months of the calendar year. Rather, it also required LEAs to provide a mandatory awards system for student maintenance grant for first degree university courses and for other courses of further education; it made parents legally responsible for ensuring that their children received a suitable education at school or otherwise, with failure to comply possibly resulting in prosecution; finally, it placed a legal obligation on the LEAs to ensure that pupils attended school. In part, these provisions may have been motivated by the willingness to induce more students to attend university. In part, these provisions may have also been motivated by the imperfect compliance to the 1947 reform that appears from the visual inspection of Fig. 1. While the 1947 reform raised the compulsory schooling age to 15, 10% of British pupils still left school by age 14 after the reform. As discussed in Harmon and Walker (1995), the 1947 reform imposed important constraints on behavior, which were documented by contemporaneous reports of overcrowding in schools and labour-market shortages at the time. Nickell (1993) and Halsey et al. (1980) also discuss the difficulties related to the implementation of the 1947 reform. Finally, Oreopoulos (2006) also discusses the possible noncompliance, or delayed enforcement of the law, to explain the fact that in Britain a fraction of pupils was still found to leave school by age 14 after 1947.

Figs. 2–4 use the GHS 1986–98 survey years and they show, respectively, the effects of the 1947, 1963 and 1972 reforms on the average age when pupils in our sample left full-time education. Results are shown for the male cohorts that reached the minimum school leaving age close to the ROSLA reform of interest. The effects of the 1947 and 1972 reforms on schooling have been studied by previous authors. In contrast, Fig. 3 newly shows the average school-leaving age of pupils across September 1963, when the 1962 Education Act came into force. A discontinuous increase appears in the school-leaving age starting from the cohort that turned 15 in September 1963. This discontinuity is smaller than the discontinuities associated with the 1947 and 1972 reforms. Given that this reform newly mandated only a few additional months of compulsory schooling, this is perhaps not surprising. Moreover, since the cutoff in Fig. 3 is calculated in September, this discontinuity may derive, in part, from the asymmetry between pupils born in late July and August, who were deemed to have attained the minimum school leaving age at the end of the

[6] In the GHS 1983, 1984 and 1985 surveys, information on the month-year of birth was asked only to women.

[7] This is relevant as it implies that the RoRtE retrieved from using these reforms is unlikely to capture the later effect of enrolment into higher levels of education.

[8] Fig. A.1 in Appendix A offers a stylized description of the implications of these law changes for actual compulsory school attendance.

[9] The 1996 Education Act further modified this, introducing a new unique minimum-school-leaving date for all pupils in the school year in which the pupil turns 16. Compared to the pre-April 1933 cohorts, this reform implied 27 additional months of compulsory school for those born January–March, 24 additional months for those born April–August, and 30 additional months for those born September–December (i.e., two more years, plus two terms). However, since this came into force from 1998 onwards, our sample is not affected by the provision in the 1996 Education Act.

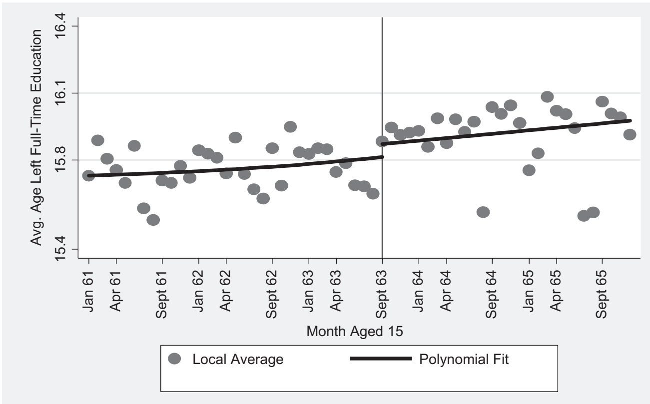


**Fig. 3.** 1963 Reform: average age left full-time education by month aged 15 *(1946–50 birth cohorts – male sample, GHS 1986–98).*

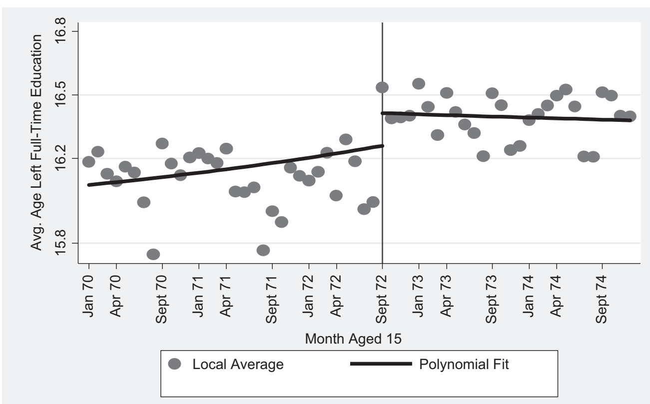


**Fig. 4.** 1972 Reform: average age left full-time education by month aged 15 *(1955–59 birth cohorts – male sample, GHS 1986–98).*

summer term, i.e., before their actual birthday, and their counterpart born in other months.

Nonetheless, it also appears from Fig. 3 that, starting from September 1963, a higher average school leaving age appears to be observed in our data. It is possible that the 1963 reform, by mandating a few months of additional compulsory schooling for pupils born in certain months of the calendar year, by making parents and LEAs legally responsible for school attendance of pupils, and by providing a series of student grants, may have induced pupils to stay in school for longer. Although, in our data, we do not observe the number of terms attended by a pupil, but only the age when pupils left full-time education, the set of provisions in the 1962 Education Act may have resulted in greater compliance to the 1947 ROSLA reform, in longer time spent in school, and in greater desire to attend university. Moreover, it is important to notice that compulsory schooling laws constitute an Intention-To-Treat (ITT); in other words, they only define the minimum age required to leave school, but there may be non-compliance of pupils facing the law, and due consideration needs to be accorded to those who are deciding to stay on in school irrespective of the ROSLA. Therefore, it is possible that this reform may have had a detectable impact on the average school leaving age of British pupils that is observable in our data. This is ultimately an empirical question, which no study has investigated before. We rectify this omission explicitly in our regression analysis in Section 5.

Figs. 5, 6 and 7 also use the GHS 1986–98 survey years, and they display, respectively, the effects of the 1947, 1963 and 1972 reforms on future earnings. The 1947 and 1972 reform effects on earnings have also been discussed by previous authors. Less familiar is Fig. 6 that explores the effect of the 1963 reform on earnings. Although a

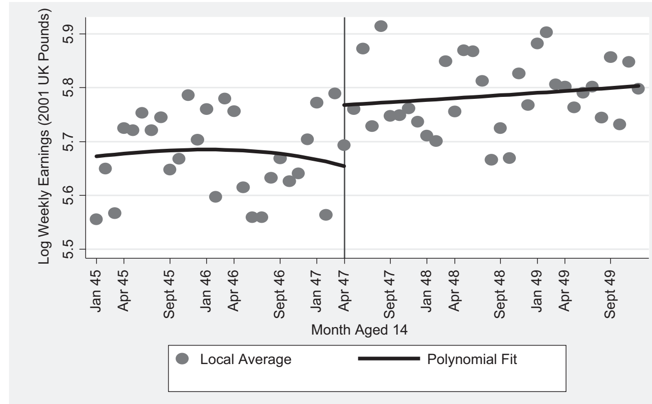


**Fig. 5.** 1947 Reform: average annual log earnings by month aged 14 *(1931–35 birth cohorts – male sample, GHS 1986–98).*

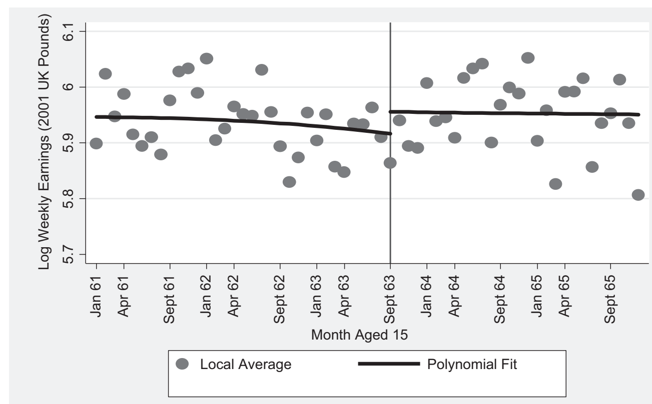


**Fig. 6.** 1963 Reform: average annual log earnings by month aged 15 *(1946–50 birth cohorts – male sample, GHS 1986–98).*

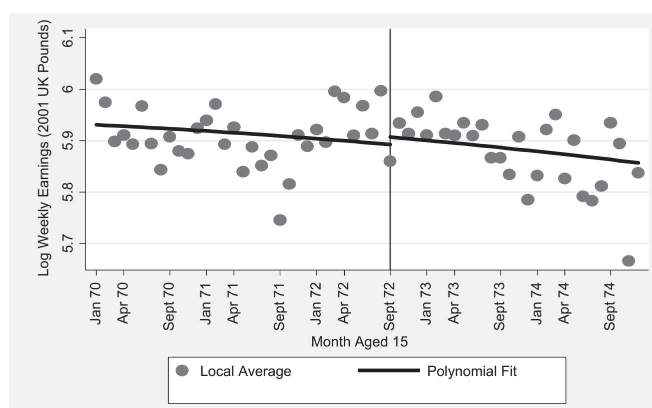


**Fig. 7.** 1972 reform: average annual log earnings by month aged 15 *(1955–59 birth cohorts – male sample, GHS 1986–98).*

discontinuous trend appears across the September 1963 cutoff, the distribution of average earnings does not seem to differ significantly across the two sides of this discontinuity. This would suggest that the 1963 reform may have not had a significant effect on the future earnings of pupils that reached the minimum school leaving age close to the September 1963 cutoff. This is again an empirical question, which we investigate in this study using regression analysis.

In conclusion, variation in the compulsory schooling regime occurred both across year-of-birth cohorts and across month-of-birth cohorts, i.e., both between- and within-year-of-birth cohorts. Although the 1963 reform received less attention than the 1947 or the 1972 reform, it plausibly introduced further exogenous variation in the schooling behavior of British pupils, especially for those born in specific months of the calendar year. This implies that a comprehensive analysis of the returns to compulsory education in Britain exploiting the observed policy changes ought to take into account all the changes in compulsory schooling that future generations faced, in comparison with the cohorts born before April 1933. Further, this also implies that the returns to schooling due to these reforms ought to be analyzed and estimated with month-of-birth comparisons. We exploit these legislative changes in the definition of our new instrumental variables in Section 5.

## 3. Previous estimates of the compulsory schooling law changes

A number of studies have analyzed the effects of the 1947 and 1972 UK ROSLA reforms on the earnings returns to education: namely, Harmon and Walker (1995), Oreopoulos (2006), Devereux and Hart (2010), Dickson and Smith (2011), Buscha and Dickson (2012), Grenet (2013), Clark and Royer (2013) and Dickson (2013). All these studies have adopted an IV approach exploiting UK government's ROSLA reforms.

Harmon and Walker (1995) attempt to capture the effect of both the 1947 and 1972 reforms on future earnings. They use data from the Family Expenditure Survey (FES) for the survey years 1978–86, focusing on males aged 18–64. They adopt an IV methodology; using the cohorts of males born before 1933 as omitted category, they define one dummy variable for pupils who entered their 14th year between 1947 and 1971 (therefore facing a compulsory school age of 15), and one for pupils entering their 14th year after 1971, who therefore faced a compulsory school age of 16. They control for a quadratic of age, for the administrative region and for survey year, without imposing any further restrictions to the sample used in the estimation. Their IV estimates suggest that the effect of an additional year of compulsory education on log earnings is 0.15, much larger than the OLS estimate of 0.06.[10]

Oreopoulos (2006) uses the 1983–98 survey years from the General Household Survey (GHS) and includes in the analysis all British-born individuals from 1921–51. He focuses on the 1947 reform using a regression discontinuity (RD) approach and includes in the specification a fourth-order polynomial in year of birth. By including the quartic of year of birth, he attempts to control for cohort trends. His results are consistent with Harmon and Walker (1995), since he finds IV estimates of the annual gain in earnings ranging from 10% to 14%, irrespective of the proportion of population affected by the compulsory school policies and whether the result is calculated for men or for women.

Devereux and Hart (2010) also focus on the 1947 reform; they follow Oreopoulos (2006) by estimating monetary returns to compulsory schooling including a quartic of year of birth and using similar estimation samples. Like Oreopoulos (2006), they include individuals who were born between 1921 and 1951 and are aged between 28 and 64. They complement the GHS with the New Earnings Survey Panel Dataset (NESPD), which offers a large sample of high-quality administrative earnings data. Devereux and Hart (2010) do not find any significant returns to compulsory schooling for the pooled sample, whereas they find IV estimates for men that are much closer to the conventional OLS estimates.

Dickson and Smith (2011) use the Quarterly Labour Force Survey (LFS) data to exploit separately the 1963 reform and the 1972 reform to investigate whether an additional year of schooling is beneficial *per se* or whether qualifications are key. They focus on men, and they restrict their analysis to the cohorts born between September 1947 and August 1967, i.e., ten years before and after the 1972 reform. They argue that the 1963 reform increased the likelihood of achieving academic qualifications, while it only affected marginally the amount of schooling received by British pupils. In contrast, the 1972 reform resulted both in one additional year of schooling and in an increased likelihood of achieving academic qualifications. Since the first cohort affected by the 1963 reform were pupils born in September 1948, in fact their sample restriction does not allow them to test whether the introduction of the 1963 reform had any effects on the average length of schooling of pupils. Using the different reforms, they retrieve IV estimates that are not statistically different, suggesting that returns to academic qualifications, rather than an additional year of schooling, determine most of the estimated return to additional years of compulsory schooling.

Del Bono and Galindo-Rueda (2004, 2007) also use the 1963 reform to examine the impact of qualifications on future labour market outcomes. They use data from the 1993–2003 LFS, from the Youth Cohort Study and a dataset that combines information from the New Earnings Study and the Joint Unemployment and Vacancies Operating System Cohort. They focus on pupils born after the introduction of the 1963 reform, and they document the importance of academic qualifications on future labour market outcomes, especially for women. Buscha and Dickson (2012) use the UK Household Longitudinal Study (UKHLS) and they focus on the effects of the 1972 reform on the earnings for individuals in their early 50s. They estimate the returns to secondary education on (log) hourly pay to be at 18.4% for the pooled sample, 22.6% for men, and 16.9% for women. Grenet (2013) focuses on the 1972 reform and compares the effects for future earnings of the 1972 reform in England and Wales with the 1967 Berthoin reform in France. He uses large samples from the UK Labour Force Survey (LFS) and, like previous RD analyses, includes in the specification a fourth-order polynomial of year of birth. He concludes that, unlike the 1967 Berthoin reform in France, the ROSLA intervention in England and Wales resulted in significant increases in future earnings for pupils forced to stay in school. He attributes this discrepancy mostly to the fact that the new school-leaving age implied the obtainment of a certificate in Britain, whereas the same was not true in France.

Clark and Royer (2013) investigate the health returns to compulsory schooling. Focusing on both the 1947 and 1972 law changes, they fail to find any significant impact of the compulsory schooling reforms on health outcomes; however, they also conclude that the 1947 reform increased earnings by roughly 8 log points. Unlike previous studies, they use month-of-birth comparisons and include a quartic of month-year of birth. However, in the calculation of their IV, they do not take into account the within-year-of-birth variation in the exposure to the 1963 reform.[11] Finally, using data from the British Household Panel Survey (BHPS), Dickson (2013) compares RoRtE results retrieved by exploiting the 1972 ROSLA reform with results retrieved using variations in schooling associated with early smoking behavior. He finds IV estimates of 10.2% when he uses the 1972 ROSLA reform, 12.9% when using early smoking and 12.5% when both instruments are used simultaneously.

## 4. Replication analysis

Among the reviewed studies, all those that estimate an RD design specify a quartic polynomial of the assignment variable, but only Grenet (2013) and Clark and Royer (2013) test it against

[10] Harmon and Walker (1995) were criticized by Card (1999) for not including cohort controls in their specification. However, since they control for survey year, the linear age variable is in fact a linear cohort variable (Devereux and Hart, 2010).

[11] In other words, these authors still define a dichotomous IV that only takes the value 1 or 0.

alternative specifications, such as a local linear regression. Based on the suggestion in Lee and Lemieux (2010), in this section we test the sensitivity of the results reported in Harmon and Walker (1995), Oreopoulos (2006) and Devereux and Hart (2010) to two important factors: namely, their sensitivity across alternative specifications and across comparable datasets.

For this replication analysis we use both the GHS and the FES data. From the GHS, we use four separate subsets of data: namely, the GHS 1979–1998 survey years as in Devereux and Hart (2010), the GHS 1983–1998 survey years as in Oreopoulos (2006), the GHS 1979–1986, similarly to Harmon and Walker (1995)[12] and, finally, the 1979–2006 GHS survey years (i.e., all the GHS survey years publicly available at the time of writing) .[13] In our replication analysis we analyze the effect of the 1947 reform. For this analysis, we follow Oreopoulos (2006) and Devereux and Hart (2010) by including British-born individuals who were born between 1921 and 1951 and are aged between 28 and 64. When we introduce our new 2SLS strategy in the next section, since this exploits the exogenous variation in schooling induced by the 1947, 1963 and 1972 reforms, we use British-born individuals who were born between 1921 and 1965 and are aged between 28 and 64. In contrast, when we do the analysis on the GHS 1979–1986, in order to produce estimates comparable to Harmon and Walker (1995) we follow these authors, i.e., no cohort restrictions are imposed in this case, and we include all available individuals aged 18–64.[14] Finally, results are presented and discussed for the male subsample.[15] In our main specification, the first stage equation can be written as follows:

$$EdAge_i = g_0 + g_1 Law_i + g_2(YoB_i)^n + \varepsilon_i, \tag{11}$$

where $i$ indexes individuals, $EdAge_i$ represents age left school, $Law$ is a dummy variable indicating if the ROSLA has changed, and $g_2(YoB_i)$ is a polynomial function of order $n$ of year of birth. In the reduced form specification we model log weekly earnings on the $Law$ variable and a polynomial function of order $n$ of year of birth. Formally,

$$\ln Y_i = \nu_0 + \nu_1 Law_i + \nu_2(YoB_i)^n + \theta_i, \tag{12}$$

where $i$ indexes individuals, $Y_i$ represents weekly earnings and $\nu_2(YoB_i)$ is a polynomial function of order $n$ of year of birth. This polynomial is necessary to control for the underlying heterogeneity (in both schooling achievement and earnings) of the sample across the years of the data. Finally, 2SLS estimates are derived in order to retrieve the impact of compulsory schooling on earnings. The only exception to our main specification applies to the analysis of the GHS 1979–86 survey years; since here we present estimates comparable to Harmon and Walker (1995), we accordingly add regional dummies and survey year dummies to the specifications above.

In our specifications we replicate Harmon and Walker (i.e., using the GHS 1979–86), we set to missing cases for which hourly wage observations are less than £1 or more than £150 (in December 2001 pounds), we exclude cases where weekly hours are greater than 84, less than 1, or missing, and we estimate robust standard errors, allowing for clustering by birth cohort. All specifications also include controls for age; similarly to Oreopoulos (2006) and Devereux and Hart (2010), some specifications include a quartic function of age, while some include age dummies. Following Harmon and Walker (1995) and Heckman and Polachek (1974), we also estimate the analysis including a quadratic of age. Unlike previous parametric RD estimates of the returns to compulsory schooling, we do not restrict the polynomial function of year of birth to be a quartic, i.e., a fourth-order polynomial. Rather, following the recommendation of Lee and Lemieux (2010), we perform the IV analysis for polynomials of up to order six. Moreover, unlike the previous parametric RD estimates that we replicate here, we also interact the $Law$ dummy with the $(YoB_i)$ polynomial for each order $n$ of year of birth to allow for heterogeneous functions either sides of the discontinuity. This is in fact one further test of the sensitivity of previous RoRtE estimates to equation specification.[16] The estimates are very similar to those shown here and are available from the authors on request.

Tables 1 and 2 report the IV estimates from our replication analysis using the GHS data for the male sample.[17] Table 1 presents the IV estimates from our parametric analysis; Table 2 presents the estimates from our local regression approach, where we use only observations close to the threshold. In our parametric analysis, using the GHS 1979–1998 survey years and including a quartic function of year-of-birth in the analysis, we are able to replicate the analysis of Devereux and Hart (2010). Using the 1983–98 GHS survey years as in Oreopoulos (2006) and controlling for a quartic function of year-of-birth, we can replicate the results reported in Devereux and Hart (2010) when using Oreopoulos' (2006) sample data.[18] Using the 1979–86 GHS survey years, we implement the analysis using the same econometric strategy as in Harmon and Walker (1995), in order to test the robustness of their findings to the use of different sample data. For reasons of space, we show the results of this last exercise, as well as the results when using our new larger GHS dataset in Table C.1 in Appendix C.

The 2SLS estimates in the upper panel of Table 1 appear sensitive to the trend specification. Using higher-order polynomial controls results in both higher point estimates and greater statistical significance in the estimated effects. One plausible interpretation of this might be that, by estimating higher-order polynomials, greater weight is being placed on later cohorts of individuals. This would be in line with Gelman and Imbens' (2014) suggestion that use of higher-order polynomials in RD analysis places greater weight on observations away from the discontinuity, and it may also result in a greater likelihood of finding statistical significance in the estimated effects.

Nonetheless, an analysis of our first stage estimates suggests that both low and high order polynomials allow us to robustly isolate the effect of this reform. These are reported in the lower panel of Table 1. In all cases, our first stage estimates suggest that using low- or high-order polynomials would yield reliable and statistically-significant estimates. This reflects the relevance of the 1947 reform, but it also shows that the sensitivity in the 2SLS estimates that we find cannot be explained by problems of weak or invalid instruments. Regardless of whether higher- or lower-order polynomials are chosen to describe the underlying heterogeneity in schooling and earnings of the sample across the years of the data, starting from the quartic function chosen by the previous literature, relatively modest deviations from the chosen order of the polynomial seem to affect the calculation of the RoRtE. This applies to all datasets considered – including our own.[19]

In light of these results, and in an attempt to understand which model is preferable, we retrieved the Akaike Information Criterion

[12] We also use the FES 1978-86 survey years in order to replicate exactly Harmon and Walker's (1995) results. The results of this exercise are discussed later in this section.

[13] However, as Devereux and Hart (2010) notice, no survey was held in 1999. In that year, the survey was redeveloped and relaunched in 2000 with a different design. For this reason, when we define our new 2SLS strategy and compare it to previous estimates in Section 5, we only focus on GHS survey years until 1998.

[14] Table A.1 in Appendix A reports the descriptive statistics from the GHS on the key variables used in the econometric analysis. Table A.2 in Appendix A reports descriptive statistics from the FES 1978–86 survey years, similarly to those reported in Harmon and Walker (1995).

[15] Results for the pooled sample are available from the authors upon request.

[16] We thank the editor of Labour Economics for suggesting this.

[17] Table C.1 in Appendix C reports additional estimates for the male sample.

[18] We do not report directly the replication of the results in Oreopoulos (2006) because this author later wrote a corrigendum for Oreopoulos (2006).

[19] SeeTable C.1 in Appendix C.

**Table 1**
1st stage and 2SLS effects of ROSLA laws on schooling and log weekly earnings – male sample.

| 2SLS estimates | Order YoB polynomial | GHS 1979–1998 (D&H) | | | GHS 1983–1998 (O) | | |
|---|---|---|---|---|---|---|---|
| | | (1) | (2) | (3) | (4) | (5) | (6) |
| | One | −0.020 | −0.017 | −0.017 | −0.042 | −0.041 | −0.054 |
| | | (0.026) | (0.027) | (0.027) | (0.034) | (0.037) | (0.040) |
| | Two | 0.031* | 0.034* | 0.039* | 0.005 | 0.012 | 0.006 |
| | | (0.017) | (0.017) | (0.022) | (0.033) | (0.032) | (0.035) |
| | Three | 0.019 | 0.020 | 0.023 | 0.006 | 0.007 | 0.003 |
| | | (0.021) | (0.021) | (0.026) | (0.033) | (0.033) | (0.035) |
| | Four | 0.063** | 0.061** | 0.067** | 0.062 | 0.060 | 0.064 |
| | | (0.026) | (0.026) | (0.031) | (0.051) | (0.052) | (0.054) |
| | Five | 0.061** | 0.060** | 0.067** | 0.064 | 0.062 | 0.069 |
| | | (0.028) | (0.029) | (0.033) | (0.048) | (0.049) | (0.049) |
| | Six | 0.096*** | 0.097*** | 0.102*** | 0.114* | 0.115* | 0.125** |
| | | (0.034) | (0.035) | (0.036) | (0.056) | (0.058) | (0.058) |
| First-stage estimates | | | | | | | |
| | One | 0.546*** | 0.545*** | 0.546*** | 0.496*** | 0.485*** | 0.484*** |
| | | (0.028) | (0.028) | (0.028) | (0.042) | (0.041) | (0.042) |
| | Two | 0.553*** | 0.551*** | 0.553*** | 0.500*** | 0.493*** | 0.492*** |
| | | (0.037) | (0.037) | (0.038) | (0.058) | (0.057) | (0.057) |
| | Three | 0.572*** | 0.575*** | 0.577*** | 0.535*** | 0.526*** | 0.525*** |
| | | (0.039) | (0.040) | (0.040) | (0.051) | (0.051) | (0.050) |
| | Four | 0.470*** | 0.471*** | 0.469*** | 0.427*** | 0.429*** | 0.433*** |
| | | (0.025) | (0.024) | (0.024) | (0.043) | (0.043) | (0.045) |
| | Five | 0.471*** | 0.471*** | 0.469*** | 0.425*** | 0.427*** | 0.432*** |
| | | (0.024) | (0.024) | (0.023) | (0.041) | (0.040) | (0.043) |
| | Six | 0.459*** | 0.458*** | 0.462*** | 0.423*** | 0.424*** | 0.434*** |
| | | (0.032) | (0.031) | (0.031) | (0.050) | (0.050) | (0.053) |
| Age controls | | Quadratic | Quartic | Dummies | Quadratic | Quartic | Dummies |

NOTE – D&H is Devereux and Hart (2010) and O is Oreopoulos (2006). Estimates in columns (2) and (3) when using a 4th order YoB polynomial are those replicated from previous papers.
*** Significant at the 1% level.
** Significant at the 5% level.
* Significant at the 10% level.

**Table 2**
AIC calculated for IV models – male sample.

| Order YoB Polynomial | | Devereux and Hart (2010) AIC | | | Oreopoulos (2006) AIC | | |
|---|---|---|---|---|---|---|---|
| **1947** reform | 1 | 74,844 | 74,648 | 74,611 | 52,277 | 52,204 | 52,841 |
| | 2 | 71,475 | 71,288 | 71,077 | 50,322 | 50,192 | 50,310 |
| | 3 | 72,183 | 72,140 | 71,946 | 50,298 | 50,267 | 50,430 |
| | 4 | 69,893 | **69,965** | **69,716** | 48,540 | 48,650 | 48,510 |
| | 5 | 69,973 | 70,045 | 69,712 | 48,485 | 48,590 | 48,378 |
| | 6 | 68,863 | 68,824 | 68,289 | 47,701 | 47,269 | 47,387 |
| | Age controls | Quadratic | Quartic | Dummies | Quadratic | Quartic | Dummies |

Note: Estimates in bold are AIC estimates that apply to those replicated from Devereux and Hart (2010).

(AIC) for each of these models, and the result of this exercise is shown in Table 2.[20] If one model looked clearly superior to others, then the choice of one specification (and in turn, one estimated RoRtE) over the others would appear justified. However, from the results in Table 2 considerable uncertainty still remains, as no model appears clearly superior to others. Differences in the AIC do emerge across models, but not in an order of magnitude necessary to confidently choose one model and rule out alternative options. Starting again from the quartic function chosen by the previous literature, relatively modest deviations from the chosen order of the polynomial seemingly result in different answers in the IV estimates, but not in the AIC values. This makes it hard to determine what RoRtE estimate can be believed.

Table 3 shows our Local Average RD Wald estimates. Heteroskedasticity-robust standard errors were clustered by year-of-birth, and 1000 bootstrap replications were conducted for inference. Starting from an optimal bandwidth (in the sense of Imbens and Kalyanaraman, 2012), results are also shown for observations within half the optimal bandwidth and observations within twice the optimal bandwidth. The running variable is still the year of birth, and the optimal bandwidth was at least greater than two year-of-birth cohorts in almost all cases.[21] Local regression estimates appear partially consistent with the results in Table 1. In fact, we retrieve a 6% RoRtE estimate, both when we use our larger dataset and when we use Devereux and Hart's (2010) data. Use of the Oreopoulos' (2006) dataset results in a 20% RoRtE estimate, while use of GHS survey years similar to Harmon and

[20] Table C.2 in Appendix C shows the AIC results for the 2SLS estimates in Table C.1 in Appendix C.

[21] The only case in which the optimal bandwidth turned out to be marginally smaller than two year-of-birth cohorts is the Local RD analysis of the 1947 ROSLA using the GHS sample similar to Harmon and Walker's (1995) sample.

**Table 3**
Local average RD effects of ROSLA laws on schooling and log weekly earnings – male sample.

| | | (Our own) | (D&H) | (O) | (H&W) |
|---|---|---|---|---|---|
| 1947 reform | Wald – optimal bandw. | 0.066** | 0.066** | 0.202** | −0.033 |
| | | (0.031) | (0.031) | (0.102) | (0.034) |
| | Wald – 0.5 * optimal bandw. | 0.050 | 0.050 | 0.000 | 0.000 |
| | | (0.035) | (0.035) | (0.000) | (0.000) |
| | Wald – 2 * optimal bandw. | 0.068* | 0.066* | 0.128 | −0.058 |
| | | (0.040) | (0.037) | (0.080) | (0.049) |

NOTE – D&H is Devereux and Hart (2010), O is Oreopoulos (2006) and H&W is Harmon and Walker (1995).
*** Significant at the 1% level.
** Significant at the 5% level.
* Significant at the 10% level.

**Table 4**
1st stage and 2SLS effects of ROSLA laws on schooling and log hourly earnings – male sample.

| | | 1st stage: schooling | | | 2SLS: hourly earnings | | |
|---|---|---|---|---|---|---|---|
| | ROSLA reform | (1) | (2) | (3) | (4) | (5) | (6) |
| Row 1 | 1947 | 0.539*** | 0.673*** | 0.675*** | 0.154*** | 0.091*** | 0.087*** |
| H&W | | (0.055) | (0.064) | (0.073) | (0.016) | (0.018) | (0.020) |
| | 1972 | 0.109 | 0.649*** | 0.687*** | | | |
| | | (0.077) | (0.089) | (0.099) | | | |
| Row 2 | 1947 | 0.539*** | 0.673*** | 0.675*** | 0.154*** | 0.091*** | 0.087*** |
| Clustered standard errors | | (0.105) | (0.132) | (0.142) | (0.030) | (0.018) | (0.016) |
| | 1972 | 0.109 | 0.649*** | 0.687*** | | | |
| | | (0.180) | (0.152) | (0.164) | | | |
| Row 3 | 1947 | 0.607*** | 0.665*** | 0.655*** | 0.019 | 0.044* | 0.031 |
| Quartic of year of birth | | (0.170) | (0.169) | (0.169) | (0.032) | (0.022) | (0.027) |
| | 1972 | 0.504*** | 0.654*** | 0.693*** | | | |
| | | (0.163) | (0.163) | (0.173) | | | |
| Row 4 | 1947 | 0.609*** | 0.668*** | 0.641*** | 0.024 | 0.052** | 0.033 |
| Regional and year dummies | | (0.173) | (0.175) | (0.173) | (0.031) | (0.021) | (0.026) |
| Omitted | 1972 | 0.531*** | 0.687*** | 0.714*** | | | |
| | | (0.168) | (0.167) | (0.174) | | | |
| Age controls | | Quadratic | Quartic | Dummies | Quadratic | Quartic | Dummies |

NOTE – Estimates in row 1 and columns (1) and (4) are those replicated from their paper.
*** Significant at the 1% level.
** Significant at the 5% level.
* Significant at the 10% level.

Walker's (1995) dataset results in a different RoRtE estimate for the compliers to the 1947 ROSLA (although in this case the small size of the optimal bandwidth may explain the lack of significance in the result). Local Average RD Wald estimates do not appear robust to modifications of the chosen bandwidth; however, caution is needed in the interpretation of the estimates derived using observations within half the optimal bandwidth in Table 3. This is because the small number of observations may result in lack of power, or even introduce bias due to the presence of seasonality by month of birth. It is also important to notice that we have only a few clusters here, therefore our cluster-robust standard errors may be underestimated and this may potentially explain some of the significance in the estimates in Table 3.

Finally, using the FES 1978–86 survey years, we replicate exactly Harmon and Walker's (1995) results. Since these authors use both the 1947 and 1972 ROSLA reforms to instrument the school-leaving age of pupils, Table 4 displays the first stage estimates for both reforms and the resulting 2SLS estimates of education on earnings. In addition to replicating Harmon and Walker (1995), we also move progressively towards the specification in Devereux and Hart (2010), to test the stability of their results not only across sampling datasets, but also across different specifications. In the analysis of the FES data, we follow Harmon and Walker (1995) and we use hourly earnings (rather than weekly earnings) to measure the returns to schooling. Using a quadratic of age and focusing exclusively on the male sample, an accurate replication of their estimate is reported in Table 4, row 1, columns (1) and (4). However, this result critically depends on the age controls used, and on the inclusion of controls for the birth cohort. In fact, when the specification of Devereux and Hart (2010) is applied to the FES 1978–86 survey data, the results change significantly; once again, the 2SLS estimates appear sensitive to the chosen functional form. This provides support to Card's (1999) criticism of Harmon and Walker (1995) regarding adequate controls for systematic inter-cohort changes in educational attainment and earnings.

Reflecting on these results, we suggest that trend specification can play an important role in the estimation of the RoRtE. Our analysis suggests that previous LATE estimates are especially sensitive which indicates that, even in contexts where clear and persuasive sources of identification are available, trend specification can play an important role in the determination of the IV results. Since we do not have a satisfactory way of choosing the optimal order of the estimated polynomial (Gelman and Imbens, 2014), our analysis demonstrates the importance of reporting goodness-of-fit tests and checking the robustness of RD analysis to alternative polynomials.

## 5. Alternative 2SLS strategies

In this section we attempt to retrieve more robust estimates of the RoRtE by exploiting the 1947 and 1972 reforms in conjunction with one additional reform, namely the 1963 reform. As discussed in

Section 2, the 1963 reform introduced a number of relevant provisions that may have exogenously influenced the schooling decisions of British pupils. Namely, it required LEAs to provide a set of student maintenance grants for first degree university courses and for other courses of further education, it made parents legally accountable for school attendance of their children, and it made the LEAs legally responsible for school attendance of resident pupils. Moreover, by varying the school-leaving dates for pupils born in certain months of the calendar year, *de facto* the 1963 reform exposed certain pupils to greater length of compulsory schooling than others. The 1963 reform introduced further exogenous variation in the *de facto* exposure to compulsory schooling for British pupils – both between- and within-year-of-birth cohorts. By making month-year-of-birth comparisons, rather than year-of-birth comparisons,[22] we attempt to define a more accurate IV which incorporates the variation both between- and within-year-of-birth implied by the school-leaving dates discussed above. Compared to the instruments in the existing literature, our IV reflects more precisely the amount of extra-compulsory schooling that different cohorts of pupils faced in Britain in our sample period.

In our analysis we adopt two 2SLS approaches based on a cardinal measure of how much 'extra' compulsory education later cohorts of pupils were mandated relative to the pre-April 1933 cohorts. Firstly, we define a new multivalued IV that takes the value 0 for pupils born before April 1st 1933 (i.e., our control group); it takes the value 1 for pupils born from April 1st 1933 until August 31st 1948, since these pupils were forced to attend one more year of school. For pupils born from September 1st 1948 up until August 31st 1957, the instrument takes the value 1.33 for those born September–December, as well as for those born February–March; instead, it still takes up value 1 for those born in the remainder of the calendar year. This is because, as a result of the 1947 reform and the 1963 reform, pupils born September–December and pupils born February–March were forced to attend one more year of school plus one term, in comparison with the control group. On the contrary, the length of compulsory schooling faced by pupils born in the remainder of the calendar year was not directly affected by the 1963 reform. Starting from September 1st 1957, our instrument takes the value 2; however, for those born September–December, as well as for those born February–March, the instrument takes the value 2.33; this is because, as a result of the 1972 reform and the 1963 reform, these pupils were forced to attend two more years of school plus one term, in comparison with the control group.

Secondly, we define a set of mutually exclusive dummies for each value of our new multivalued IV. This results in the use of four separate dichotomous IVs to instrument the endogenous schooling decision. As discussed in our introduction, definition of one unique IV assumes an homogeneous effect of our ROSLA reforms on the schooling decisions and future earnings of British pupils. Yet the 1947 reform affected approximately half the relevant cohort; the 1963 reform affected a smaller fraction of population, especially those born in specific months of the calendar year; the 1972 reform took place 25 years after the 1947 reform, and it affected approximately a third of the relevant cohort. Given that these reforms took place at different points in time, and in all likelihood they affected different fractions of population, the assumption of homogeneity in their first-stage and reduced-form effects may be questionable. By defining a set of mutually exclusive dummies for each value of our multivalued IV, we relax this assumption and we allow for heterogeneous effects of these reforms on education and earnings. This, in turn, also allows us to newly test explicitly if the 1963 reform had any detectable impact

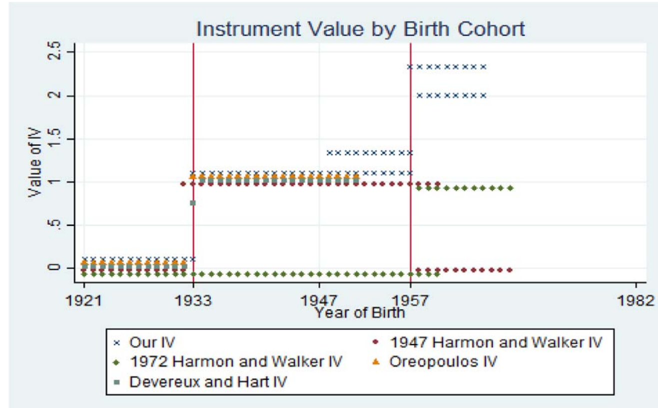


**Fig. 8.** Stylized description of our instrument vis-à-vis instruments used in reviewed literature.

on the school leaving behavior of British pupils. The reason for showing results using both approaches is twofold; first, both approaches seem sensible *a priori*, while the heterogeneity in the effects of the ROSLA reforms on schooling and earnings is an empirical question. Secondly, a comparison of our 2SLS estimates using both approaches allows us to assess the relevance of this potential heterogeneity for the retrieved RoRtE estimates.

At this juncture, it should be noticed that our 2SLS approaches aim to describe more accurately than previous studies the length of compulsory schooling faced by different cohorts of pupils. For this reason, we only use the 1963 reform to describe the further between- and within-year of birth variation in the length of compulsory schooling introduced by this reform. An alternative strategy would have been to include in the same group all pupils born between September 1948 and August 1957, i.e., all those exposed to the 1947 and 1963 reform, but not to the 1972 reform. We did not do this because our interest in this study lies in the implications for the RoRtE estimates of a more precise description of the length of compulsory schooling as an instrument for the educational choice – not in the effect of the 1963 reform *per se*. Therefore, we thought it preferable to describe with our IV the effect that this reform had on the length of compulsory schooling. Conceptually, this also provides more direct comparability of our new results in this section with our replication analysis in Section 4.[23]

Fig. 8 graphs our instrument against those used in Harmon and Walker (1995), Oreopoulos (2006) and Devereux and Hart (2010). There is no major difference between our instrument and those of Oreopoulos (2006) and Devereux and Hart (2010) for the 1921–1947 birth cohorts; however, for the 1948–1951 cohorts, neither Oreopoulos (2006) nor Devereux and Hart (2010) take into account the provisions of the 1963 reform – since we do, for these years our instrument diverges from theirs.[24] Harmon and Walker (1995) capture the changes in compulsory schooling with two dummies[25]; by doing so, they also omit to account for the 1963 reform. This may have introduced measurement error in the calculation of the RoRtE,

[22] Clark and Royer (2013) is the first study to do so, although their focus is on health outcomes.

[23] When we define a dummy variable for all cohorts that reached the minimum school leaving age after the introduction of the 1963 reform, the first-stage effect of the 1963 reform on schooling appears even stronger than the first-stage estimates of the 1963 reform presented in this study. The first-stage estimates of the 1963 reform presented in this study only apply to pupils born in September–December and February–March of the year-of-birth cohorts affected by the 1963 reform.

[24] After the 1951 cohort, a comparison between our IV and their IVs is not possible since they do not include later cohorts in the analysis.

[25] In Harmon and Walker's (1995) IVs some overlapping is observed close to the threshold years because of the inclusion of the Scottish sample in their analysis; since Scotland came under a different regulation, and the 1947 and 1972 policies were implemented, respectively, in 1946 and 1976, Harmon and Walker (1995) attribute the policy changes to 1946 and 1976 for the Scottish sample.

**Table 5**
2SLS effects of ROSLA laws on log weekly earnings – male sample, 4th order yob/myob polynomials.

| Controls | *IV* Used *N Obs* | RoRtE estimates | | | |
|---|---|---|---|---|---|
| Y-o-B polyn., years of age, as per D&H | Old IV | 0.033 | 0.032 | 0.032 | 0.033 |
| | $N$ =63,746 | (0.022) | (0.022) | (0.022) | (0.025) |
| | New IV | 0.055*** | 0.057*** | 0.056*** | 0.057*** |
| | $N$ = 63,084 | (0.017) | (0.017) | (0.017) | (0.018) |
| | Set of IVs | 0.050*** | 0.051*** | 0.050*** | 0.052*** |
| | $N$ = 63,084 | (0.015) | (0.015) | (0.015) | (0.016) |
| Y-o-B polyn., years of age, as per D&H | Old IV | 0.080* | 0.075 | 0.074 | 0.055 |
| | $N$ = 37,963 | (0.047) | (0.049) | (0.049) | (0.055) |
| | New IV | 0.053* | 0.049 | 0.049 | 0.037 |
| | $N$ = 37,963 | (0.030) | (0.031) | (0.031) | (0.033) |
| | Set of IVs | 0.050* | 0.046 | 0.046 | 0.031 |
| | $N$ = 37,963 | (0.028) | (0.028) | (0.028) | (0.031) |
| Y-o-B polyn., years of age, survey year FE, survey month FE, month of birth FE | Old IV | 0.082* | 0.076 | 0.074 | 0.056 |
| | $N$ = 37,963 | (0.046) | (0.047) | (0.048) | (0.053) |
| | New IV | 0.068** | 0.062* | 0.062* | 0.049 |
| | $N$ = 37,963 | (0.034) | (0.035) | (0.036) | (0.039) |
| | Set of IVs | 0.066** | 0.061* | 0.060* | 0.044 |
| | $N$ = 37,963 | (0.030) | (0.031) | (0.031) | (0.035) |
| M-Y-o-B polyn., years of age, survey year FE, survey month FE, month of birth FE | Old IV | 0.086** | 0.080* | 0.078* | 0.060 |
| | $N$ = 37,963 | (0.040) | (0.041) | (0.041) | (0.044) |
| | New IV | 0.069* | 0.062* | 0.063* | 0.050 |
| | $N$ = 37,963 | (0.036) | (0.037) | (0.037) | (0.039) |
| | Set of IVs | 0.068** | 0.063* | 0.062* | 0.047 |
| | $N$ = 37,963 | (0.032) | (0.033) | (0.033) | (0.035) |
| M-Y-o-B polyn., months of age, survey year FE, survey month FE, month of birth FE, as per C&R | Old IV | 0.086** | 0.079* | 0.077* | |
| | $N$ = 37,963 | (0.040) | (0.041) | (0.041) | |
| | New IV | 0.068* | 0.061* | 0.062* | |
| | $N$ = 37,963 | (0.036) | (0.037) | (0.037) | |
| | Set of IVs | 0.068** | 0.062* | 0.061* | |
| | $N$ = 37,963 | (0.032) | (0.033) | (0.033) | |
| | Age controls | Quadratic | Cubic | Quartic | Dummies |

NOTE – D&H is Devereux and Hart (2010) and C&R is Clark and Royer (2013).
*** Significant at the 1% level.
** Significant at the 5% level.
* Significant at the 10% level.

and our instrument attempts to explore the relevance of this. Finally, it is also important to emphasize that we are not, *per se*, using the month of birth as our IV – rather our IV is derived by the conjunction of a month and year of birth and exactly what ROSLA reform may have affected each individual.[26] The strength of this approach is to emphasize the continuous nature of years of education and its separate terms as partly completed exogenous years of education. We are aware of the limitations of adding extra years or terms of education together from different eras, as they may be very heterogeneous across time – i.e., a term of extra education in 1963 may not be the same as one third of a year in 1947 or 1972. However, the inclusion of a variety of polynomials at the year-of-birth or at the month-year-of-birth level aims to control precisely for these heterogeneous effects. In our econometric analysis, the first stage equation can be written as follows:

$$EdAge_i = \gamma_0 + \gamma_1 X_i + \gamma_2 Laws_i + \gamma_3 (DoB_i)^n + u_i, \tag{13}$$

where $i$ indexes individuals, $EdAge_i$ represents age left school, $Laws_i$ is our new multivalued IV (set of IVs) that captures all ROSLA reforms and $\gamma_3(DoB_i)$ is a polynomial function of order $n$ of date of birth; we estimate polynomials at the year-of-birth level, as well as at the month-year-of-birth level. Age controls are also included among the covariates. Finally, $X_i$ represents varying sets of additional controls at the individual level that are included to test for the sensitivity of our estimates to equation specification. In the 2SLS specification we model log weekly earnings on the instrumented age left school, a polynomial function of order $n$ of date of birth, controls for age, and a set of additional controls at the individual level. Formally:

$$ln\ Y_i = \delta_0 + \delta_1 X_i + \delta_2 \left( EdAge_i = Laws_i \right) + \delta_3 (DoB_i)^n + w_i, \tag{14}$$

where $i$ indexes individuals, $Y_i$ represents weekly earnings, $EdAge_i = Laws_i$ represents the instrumented level of education and $\delta_3(DoB_i)$ is a polynomial function of order $n$ of date of birth; as we do in the first stage, we estimate polynomials at the year-of-birth level, as well as at the month-year-of-birth level. In our specifications, we also include varying sets of additional controls $X_i$ at the individual level to test for the sensitivity of our estimates to equation specification.

Since information on month of birth was not available from the FES, here we only use data from the 1983–1998 GHS surveys. Our sample comprises all individuals born in Britain from 1921–1965, and aged 28–64 at the time of the survey. Also, we still set to missing cases for which hourly wage observations are less than £1 or more than £150 (in December 2001 pounds), we exclude cases where weekly hours are greater than 84, less than 1, or missing, and we estimate robust standard errors, allowing for clustering by year- or by month-year of birth (according to which polynomial is estimated). This results in a sample of 117,144 individuals; 76,935 of these reported complete information on the month-year of birth, as well as on the month-year of the interview.

Table 5 reports our IV estimates for the male sample. All results in Table 5 are calculated using (year-of-birth or month-year-of-birth) polynomials of order four. We respectively report the estimations

[26] Following Angrist and Krueger (1991) and Acemoglu and Angrist (2001) there is a large literature now which uses month of birth directly as an IV in education studies as well as other fields. The logical potential endogeneity problem with this literature is that there may be non-random seasonality in this month of birth which induces effects on education outcomes or even on earnings directly. See Buckles and Hungerman (2013). Despite the fact that we do not use the month of birth directly as an IV, we show in Appendix B that our data does not exhibit this seasonality (despite the significance of the single month of August).

using the old dichotomous IV from Harmon and Walker (1995), that omits to incorporate the 1963 reform, our new multivalued IV, which takes account of the number of each term of extra schooling a pupil was mandated, and finally our set of IVs, that incorporates the effect of the 1963 reform on compulsory schooling and allows for heterogeneous effects of the ROSLA reforms.

Similarly to Table 1, in Table 5 different columns display RoRtE estimates using different controls for age. In the first three rows we specify the same functional form as Devereux and Hart (2010) and Oreopoulos (2006). Therefore, we control for a polynomial of year of birth and for age. In rows four, five and six we apply the same functional form only to the subsample of observations that reported complete information on the month-year of birth, as well as on the month-year of the interview. Starting from row seven, we move progressively towards the specification in Clark and Royer (2013) (i.e., the latest contribution to this literature), including additional covariates to the 2SLS equation and restricting the analysis to those individuals that reported complete information on all the covariates included in the analysis. In an attempt to better control for the presence of seasonality, in rows seven, eight and nine, we also include fixed effects, namely, at the survey-year level, at the survey-month level, and at the month-of-birth level. Inclusion of month-of-birth fixed effects should mitigate concerns regarding the presence of seasonality by month of birth in schooling achievement and future earnings. In rows ten, eleven and twelve, we also control for polynomials at the month-year of birth level, still including fixed effects at the survey-year level, at the survey-month level, and at the month-of-birth level. Finally, in rows thirteen, fourteen and fifteen, we re-define age in months (i.e., not in years), therefore controlling for within-year-of-birth heterogeneity in age. We still include fixed effects at the survey-year level, at the survey-month level, and at the month-of-birth level.[27]

The 2SLS estimates in Table 5 appear robust across different specifications and they suggest that the average MPRTE of the ROSLA reforms in Great Britain is 6%. Moreover, despite the wide array of controls included in the analysis, no major differences emerge in the estimates across alternative sets of controls. Our IV estimates of the returns to education no longer seem to be as sensitive as previous estimates to the order of the polynomial of month-year of birth chosen.

In Table 6 we vary the order of the estimated polynomials at the year-of-birth and month-year-of-birth level respectively. By doing this, we subject our estimates to a comparable degree of scrutiny as we did in Table 1 for those in previous studies.[28] To be precise, we apply a variety of polynomials to the estimates in row fifteen in Table 5, where we modeled age in months, and we included fixed effects at the survey-year level, at the survey-month level, and at the month-of-birth level. For each order $n$ of our month-year-of-birth polynomials, Table 6 also reports our first-stage estimates of each ROSLA reform used in our analysis. This is important as it shows that, regardless of the order of the polynomial chosen in our analysis, the 1963 reform had a positive and significant impact on the schooling decisions of British pupils. For all the reported first-stage coefficients, we also conducted a series of Wald tests to check whether the coefficient associated with the first-stage effect of the 1947 reform and the coefficient associated with the additional first-stage effect of the 1963 reform are statistically different. In all cases we could reject the null hypothesis of equality of these first-stage coefficients before the introduction of the 1972 reform.[29] These results provide evidence that the 1963 reform had a detectable effect on the schooling trajectories of British pupils – over and above the effect of the 1947 reform. After the introduction of the 1972 reform, we could reject in almost all our regressions the null that the coefficient associated with the 1963 and 1972 reform is equal to the coefficient associated with the 1972 reform only. This is again suggestive of a first-stage effect of the 1963 reform over and above the effect of the 1972 reform for the pupils born in the months that the 1963 reform applied to.

The point estimates in Table 6 show substantial consistency across different orders of the month-year of birth polynomials used. In most cases, our results suggest that the RoRtE for men lies between 6% and 7%. The only exceptions are the 2SLS estimates retrieved using month-year of birth polynomials of order one and two. One plausible explanation for this may be that, given the large number of cohorts included in our regressions (i.e., 1921–1965), first- and second-order polynomials in month-year-of-birth may not suffice to adequately control for the underlying trends in schooling and earnings over time. Therefore, we again sought the optimal specification using the Akaike's criterion. In this case, the AIC appeared more informative, suggesting that the models using polynomials of order three and four in Table 6 are preferable to those using lower order polynomials. Nonetheless, results from polynomials of order one and two in Table 6 further corroborate the conclusion in our previous section on the importance of the trend specification when conducting IV analysis and when important underlying heterogeneity is present in the sample.

Interestingly, in a comparison of the IV estimates in Table 5, the introduction of our new IV, that also incorporates the effect of the 1963 reform, appears to result in smaller point estimates of the RoRtE compared to the "Old IV" approach. Not controlling for the 1963 reform results in higher and slightly more unstable estimates – in line with the suggestion from our replication analysis. The 2SLS estimates that utilize our IV instead appear more precise, and more robust to equation specification. Use of our set of IVs seems to bring further benefit to our estimates; in particular, point estimates appear always smaller and standard errors appear slightly smaller as well. This suggests that allowing for more flexible first-stage and reduced-form estimates of the effects of our ROSLA reforms results in a higher degree of precision in our estimates. As shown in Table 6, the definition of a more precise IV also resulted in greater robustness in the estimated parameter across different orders of the estimated polynomial.

Given the importance of the age of measurement of earnings in the determination of the RoRtE estimates, we also re-estimated Table 5 focusing only on individuals aged 28–40, as the bias in the RoRtE estimates should be minimized around this age range (see Bhuller et al., 2017). While results using 4th order polynomials of year-of-birth and month-year of birth appear smaller than those in Table 5, when using lower order polynomials, i.e., of order three or two, the estimated RoRtE still clusters around 7% for men. Given that fewer age groups are included in this analysis, we regard these results as consistent with our main conclusion of a 6% RoRtE for men.

Finally, although in our narrative we focused entirely on men, we also estimated our results for the pooled sample.[30] Regardless of the underlying functional form chosen, the estimated coefficients on returns to education for the pooled sample generally go to zero. It is only when second order polynomials are used that estimates appear significant. However, they also appear unstable across alternative sets of controls. We do not regard these pooled coefficients as meaningful since the inclusion of women in the pooled sample introduces all the other problems of women selecting themselves into work or making dynamic life-cycle participation and family decisions. The inherent difficulty of including women into the analysis without modeling the selectivity of

[27] When we re-define age in months, we do not include dummies for months of age, since this results into too many controls and overfitting of the model.

[28] We thank an anonymous referee for suggesting this.

[29] Results for these Wald tests are available on request from the authors.

[30] Our results are available on request to the authors.

**Table 6**
1st stage and 2SLS effects of ROSLA laws on schooling and log weekly earnings – male sample.

| Order MYoB Polynomial | | ROSLA Reform | First stage estimates (1) | (2) | (3) | 2SLS estimates (4) | (5) | (6) |
|---|---|---|---|---|---|---|---|---|
| M-Y-o-B polyn., months of age, survey year FE, survey month FE, month of birth FE, as per C&R. $N = 37{,}963$ | One | 47 Reform | 0.525*** | 0.528*** | 0.528*** | −0.000 | −0.003 | −0.004 |
| | | | (0.034) | (0.038) | (0.037) | (0.030) | (0.032) | (0.032) |
| | | 47/63 Reform | 0.661*** | 0.663*** | 0.661*** | | | |
| | | | (0.044) | (0.046) | (0.045) | | | |
| | | 72 Reform | 0.635*** | 0.639*** | 0.638*** | | | |
| | | | (0.053) | (0.056) | (0.056) | | | |
| | | 72/63 Reform | 0.712*** | 0.716*** | 0.714*** | | | |
| | | | (0.052) | (0.053) | (0.053) | | | |
| | Two | 47 Reform | 0.439*** | 0.446*** | 0.446*** | 0.032 | 0.028 | 0.027 |
| | | | (0.040) | (0.042) | (0.042) | (0.034) | (0.036) | (0.036) |
| | | 47/63 Reform | 0.582*** | 0.587*** | 0.584*** | | | |
| | | | (0.048) | (0.049) | (0.049) | | | |
| | | 72 Reform | 0.619*** | 0.631*** | 0.630*** | | | |
| | | | (0.052) | (0.055) | (0.054) | | | |
| | | 72/63 Reform | 0.698*** | 0.708*** | 0.706*** | | | |
| | | | (0.051) | (0.053) | (0.052) | | | |
| | Three | 47 Reform | 0.502*** | 0.497*** | 0.500*** | 0.064* | 0.058* | 0.059* |
| | | | (0.046) | (0.046) | (0.046) | (0.033) | (0.034) | (0.034) |
| | | 47/63 Reform | 0.625*** | 0.621*** | 0.620*** | | | |
| | | | (0.051) | (0.051) | (0.051) | | | |
| | | 72 Reform | 0.722*** | 0.712*** | 0.715*** | | | |
| | | | (0.058) | (0.058) | (0.057) | | | |
| | | 72/63 Reform | 0.791*** | 0.781*** | 0.782*** | | | |
| | | | (0.056) | (0.055) | (0.055) | | | |
| | Four | 47 Reform | 0.537*** | 0.532*** | 0.532*** | 0.068** | 0.062* | 0.061* |
| | | | (0.046) | (0.046) | (0.046) | (0.032) | (0.033) | (0.033) |
| | | 47/63 Reform | 0.643*** | 0.639*** | 0.640*** | | | |
| | | | (0.050) | (0.050) | (0.050) | | | |
| | | 72 Reform | 0.735*** | 0.725*** | 0.725*** | | | |
| | | | (0.056) | (0.056) | (0.056) | | | |
| | | 72/63 Reform | 0.796*** | 0.787*** | 0.787*** | | | |
| | | | (0.054) | (0.054) | (0.054) | | | |
| | Five | 47 Reform | 0.501*** | 0.502*** | 0.502*** | 0.068 | 0.068 | 0.067 |
| | | | (0.053) | (0.053) | (0.053) | (0.043) | (0.043) | (0.043) |
| | | 47/63 Reform | 0.612*** | 0.613*** | 0.614*** | | | |
| | | | (0.053) | (0.053) | (0.053) | | | |
| | | 72 Reform | 0.678*** | 0.678*** | 0.678*** | | | |
| | | | (0.070) | (0.070) | (0.070) | | | |
| | | 72/63 Reform | 0.742*** | 0.742*** | 0.742*** | | | |
| | | | (0.068) | (0.068) | (0.068) | | | |
| | Six | 47 Reform | 0.433*** | 0.433*** | 0.434*** | 0.070 | 0.087* | 0.072* |
| | | | (0.053) | (0.053) | (0.053) | (0.045) | (0.050) | (0.043) |
| | | 47/63 Reform | 0.524*** | 0.525*** | 0.526*** | | | |
| | | | (0.055) | (0.055) | (0.055) | | | |
| | | 72 Reform | 0.694*** | 0.694*** | 0.693*** | | | |
| | | | (0.068) | (0.068) | (0.068) | | | |
| | | 72/63 Reform | 0.745*** | 0.745*** | 0.745*** | | | |
| | | | (0.066) | (0.066) | (0.066) | | | |
| | Age controls | | Quadratic | Quartic | Dummies | Quadraticatic | Quartic | Dummies |

NOTE – C&R is Clark and Royer (2013).
*** Significant at the 1% level.
** Significant at the 5% level.
* Significant at the 10% level.

women into work, as well as the problem of fertility and the dynamics of female labour supply, all conspire to make this problem intractable for a pooled sample. How much of our estimates for the pooled sample is due to education *per se*, rather than selection, is a moot point. Confining our interest to men only, our results suggest that introducing this methodology provides us with more stable coefficients. However, even when we use our multivalued IV, as well as our new set of IVs, we still find the specification of the underlying trends in education and earnings to play an important role in the determination of the RoRtE estimates.

## 6. Discussion and conclusion

This paper has revisited the empirical literature on the returns to education in the UK. Courtesy of exogenous UK government school leaving age reforms, this is a particularly interesting setting to investigate this important parameter. Not surprisingly, a rich literature has flourished in this context. Our contribution to the debate relating to the rate of return to education is twofold. First, we examine the robustness of the papers by Harmon and Walker (1995), Oreopoulos (2006) and Devereux and Hart (2010). We do this by using all the available data and examining the sensitivity of the results to the specification of the polynomial used to describe the assignment variable (i.e., date of birth) away from the threshold. Our replication analysis attests to the sensitivity of previous RoRtE estimates to the specific functional form chosen. Since in reality the 'true function' is not known, our analysis highlights the importance of reporting goodness-of-fit tests and checking the robustness of RD analysis to alternative polynomials of the controls. Since we conduct our analysis in a context where multiple informative instrumental variables were available to retrieve the RoRtE, this conclusion is particularly relevant.

Secondly, we generalize the IV approach of the previous papers by using the month of birth in conjunction with the ROSLA in the calculation of a more accurate IV. We do this in two ways: firstly, we define a new multivalued IV that reflects the extra fractions-of-a-year of schooling that later cohorts of pupils were mandated in comparison with the pre-April 1933 cohorts. Secondly, we also instrument the school leaving age of our sample observations with a set of mutually exclusive dummies for each value of our new multivalued IV. With the latter approach we allow the first-stage and reduced form effects of the ROSLA reforms to differ. Nonetheless, both approaches retrieve very similar results. Our analysis provides more consistency in the results with RoRtE estimates generally found at 6%. By redefining the instrument to directly reflect the extra exogenous education administered to the treated population, we find estimates of the RoRtE that are close to the estimates in Devereux and Hart (2010).

The main conclusion of our empirical research is that the RoRtE based on the ROSLA policies in the UK is 6% for males. In contrast, we also find that the RoRtE for the pooled sample is generally zero. Compared to previous estimates, our results appear more robust to the inclusion of alternative controls; in comparison with estimates using the binary IV in Harmon and Walker (1995), use of our IVs shows the estimates of the RoRtE to be smaller and more precise. Nonetheless, also our RoRtE estimates appear sensitive to the choice of the polynomial used to describe the underlying unobservable trends in education and earnings in the sample: our estimates range from 5–6% and statistically significant when using polynomials of order three or four to 0–3% and non-statistically significant when using polynomials of order one and two, further reflecting the importance of reporting goodness-of-fit tests and checking the robustness of RD analysis to alternative polynomials.

## Appendix A. Descriptive statistics

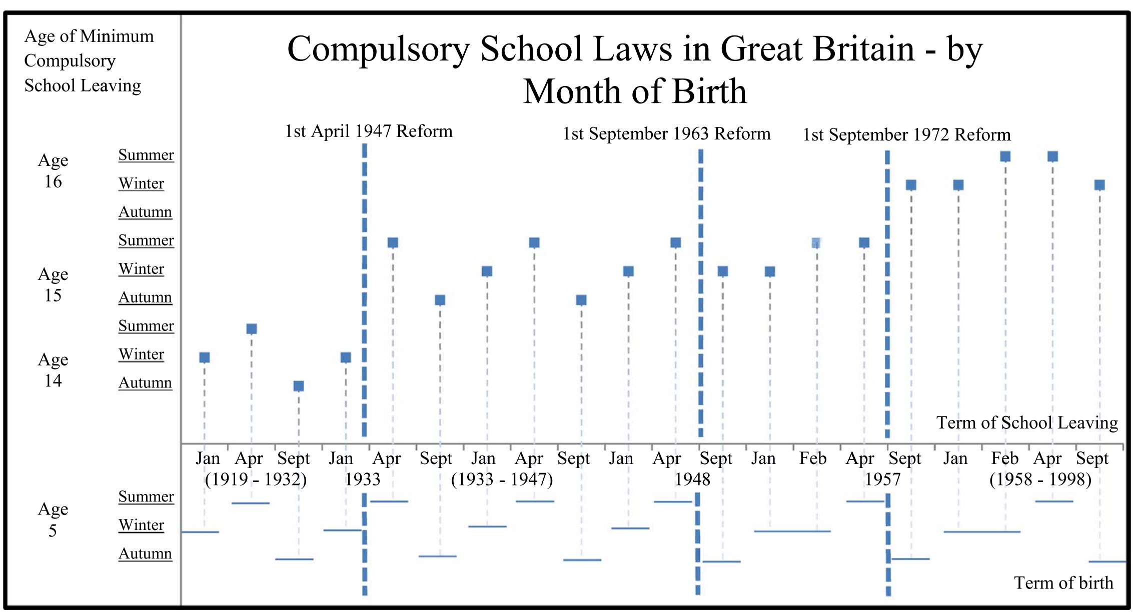


**Fig. A.1.** Stylized description of the effects of compulsory school laws. Clark and Royer (2013) note that not all local schools authorities adhered to the same school entry rules: some admitted all students at the beginning of the academic year in which they reached the age of five (i.e., in September); others had two rather than three entry points.

**Table A.1**
GHS descriptive statistics for the study of 1947 SLA reform.

| 1947 SLA reform | GHS 1979–2006 | | | GHS 1979–1998 | | | GHS 1983–1998 | | | GHS 1979–1986 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Variable | Obs. | Mean | Std. dev. | Obs. | Mean | Std. dev. | Obs. | Mean | Std dev | Obs. | Mean | Std. dev. |
| Survey year | 96,549 | 88.13 | 7.24 | 85,766 | 86.29 | 5.30 | 53,502 | 89.57 | 3.90 | 71,388 | 82.26 | 2.32 |
| Cohort | 96,549 | 20.20 | 8.06 | 85,766 | 19.42 | 8.13 | 53,502 | 20.90 | 7.30 | 71,388 | 23.23 | 12.89 |
| Female | 96,549 | 0.45 | 0.50 | 85,766 | 0.45 | 0.50 | 53,502 | 0.46 | 0.50 | 71,388 | 0.44 | 0.50 |
| Age | 96,549 | 47.59 | 8.57 | 85,766 | 46.47 | 8.35 | 53,502 | 48.34 | 7.42 | 71,388 | 38.53 | 12.60 |
| Employee | 96,549 | 0.90 | 0.30 | 85,766 | 0.91 | 0.29 | 53,502 | 0.89 | 0.31 | 71,388 | 0.94 | 0.24 |
| Hours worked | 96,549 | 34.85 | 12.97 | 85,766 | 34.82 | 12.88 | 53,502 | 35.03 | 13.33 | 71,388 | 35.18 | 11.53 |

**Table A.1** (*continued*)

| 1947 SLA reform | GHS 1979–2006 | | | GHS 1979–1998 | | | GHS 1983–1998 | | | GHS 1979–1986 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Variable | Obs. | Mean | Std. dev. | Obs. | Mean | Std. dev. | Obs. | Mean | Std dev | Obs. | Mean | Std. dev. |
| Log (hourly wage) | 96,549 | 1.91 | 0.59 | 85,766 | 1.88 | 0.58 | 53,502 | 1.94 | 0.60 | 71,388 | 1.77 | 0.52 |
| Log (weekly earnings) | 96,549 | 5.35 | 0.87 | 85,766 | 5.33 | 0.87 | 53,502 | 5.39 | 0.88 | 71,388 | 5.24 | 0.79 |
| Age left school | 96,549 | 15.58 | 1.43 | 85,766 | 15.46 | 1.22 | 53,502 | 15.58 | 1.21 | 71,388 | 15.56 | 1.24 |
| Left school by age 14 | 96,549 | 0.17 | 0.37 | 85,766 | 0.19 | 0.39 | 53,502 | 0.13 | 0.34 | 71,388 | 0.19 | 0.40 |
| Left school by age 15 | 96,549 | 0.62 | 0.49 | 85,766 | 0.63 | 0.48 | 53,502 | 0.60 | 0.49 | 71,388 | 0.54 | 0.50 |
| Left school by age 16 | 96,549 | 0.81 | 0.39 | 85,766 | 0.82 | 0.38 | 53,502 | 0.81 | 0.40 | 71,388 | 0.81 | 0.39 |
| Law mandates school until 15 | 96,549 | 0.79 | 0.40 | 85,766 | 0.77 | 0.42 | 53,502 | 0.84 | 0.36 | 71,388 | 0.75 | 0.43 |
| Law mandates school until 16 | 96,549 | 0 | 0 | 85,766 | 0 | 0 | 53,502 | 0 | 0 | 71,388 | 0.17 | 0.37 |
| Law mandates school until 15 – different Calc for Scotland | 96,549 | 0.79 | 0.40 | 85,766 | 0.77 | 0.42 | 53,502 | 0.84 | 0.36 | 71,388 | 0.75 | 0.43 |
| Law mandates school until 16 – different Calc for Scotland | 96,549 | 0 | 0 | 85,766 | 0 | 0 | 53,502 | 0 | 0 | 71,388 | 0.16 | 0.36 |

**Table A.2**
FES descriptive statistics for the joint study of 1947 and 1972 SLA reforms.

| Variable | Pooled sample | | | Male sample | | |
|---|---|---|---|---|---|---|
| | Obs. | Mean | Std. dev. | Obs. | Mean | Std. dev. |
| Ln(wage) | 61,019 | 1.751 | 0.345 | 34,335 | 1.931 | 0.198 |
| Years of schooling | 61,019 | 16.159 | 2.136 | 34,335 | 16.159 | 2.203 |
| Age | 61,019 | 38.447 | 12.562 | 34,335 | 38.739 | 12.667 |
| Yorkshire | 61,019 | 0.088 | 0.283 | 34,335 | 0.088 | 0.283 |
| Northwest | 61,019 | 0.112 | 0.315 | 34,335 | 0.110 | 0.312 |
| East Midlands | 61,019 | 0.073 | 0.261 | 34,335 | 0.075 | 0.264 |
| West Midlands | 61,019 | 0.098 | 0.297 | 34,335 | 0.099 | 0.298 |
| East Anglia | 61,019 | 0.036 | 0.186 | 34,335 | 0.037 | 0.188 |
| Southeast | 61,019 | 0.310 | 0.463 | 34,335 | 0.306 | 0.461 |
| Southwest | 61,019 | 0.072 | 0.259 | 34,335 | 0.074 | 0.262 |
| Scotland | 61,019 | 0.090 | 0.286 | 34,335 | 0.089 | 0.285 |
| Northern Ireland | 61,019 | 0.014 | 0.118 | 34,335 | 0.014 | 0.115 |
| Wales | 61,019 | 0.049 | 0.217 | 34,335 | 0.051 | 0.219 |
| Year = 1979 | 61,019 | 0.114 | 0.318 | 34,335 | 0.116 | 0.321 |
| Year = 1980 | 61,019 | 0.116 | 0.320 | 34,335 | 0.116 | 0.321 |
| Year = 1981 | 61,019 | 0.120 | 0.325 | 34,335 | 0.121 | 0.326 |
| Year = 1982 | 61,019 | 0.116 | 0.320 | 34,335 | 0.117 | 0.322 |
| Year = 1983 | 61,019 | 0.102 | 0.302 | 34,335 | 0.101 | 0.301 |
| Year = 1984 | 61,019 | 0.107 | 0.310 | 34,335 | 0.104 | 0.306 |
| Year = 1985 | 61,019 | 0.103 | 0.303 | 34,335 | 0.101 | 0.302 |
| Year = 1986 | 61,019 | 0.105 | 0.307 | 34,335 | 0.102 | 0.303 |

## Appendix B. Analysis of presence of seasonality – by month of birth

**Table B.1**
*T*-tests for significance of differences in age left full-time education for cohorts born in 1921–41. Male sample.

| Month of birth (vs all other months) | $t$ | $\Pr(\lvert T\rvert > \lvert t\rvert)$ |
|---|---|---|
| January | −0.0749 | 0.9409 |
| February | −1.2105 | 0.2382 |
| March | −0.8256 | 0.4172 |
| April | −0.1797 | 0.8590 |
| May | 0.3722 | 0.7131 |
| June | 0.4859 | 0.6315 |
| July | −0.0867 | 0.9316 |
| August | −0.4755 | 0.6386 |
| September | 1.5459 | 0.1346 |
| October | 0.4021 | 0.6913 |
| November | 0.1951 | 0.8470 |
| December | 0.0086 | 0.9932 |

**Table B.2**
*T*-tests for significance of differences in age left full-time education for cohorts born in 1941–61. Male sample.

| Month of birth (vs all other months) | $t$ | $\Pr(\lvert T\rvert > \lvert t\rvert)$ |
|---|---|---|
| January | 0.4499 | 0.6569 |
| February | 0.1484 | 0.8832 |
| March | 0.7110 | 0.4840 |
| April | 0.3068 | 0.7617 |
| May | 0.1555 | 0.8777 |
| June | 0.1049 | 0.9173 |
| July | −2.2182 | 0.0360** |
| August | −5.0371 | 0.0000*** |
| September | 1.3791 | 0.1812 |
| October | 1.4906 | 0.1486 |
| November | 1.2834 | 0.2111 |
| December | 0.7511 | 0.4596 |

**Table B.3**
*T*-tests for significance of differences in earnings for cohorts born in 1921–41. Male sample.

| Month of birth (vs all other months) | *t* | Pr(\|*T*\| > \|*t*\|) |
|---|---|---|
| January | 0.0645 | 0.9491 |
| February | −0.1956 | 0.8465 |
| March | −0.5846 | 0.5645 |
| April | 1.3882 | 0.1751 |
| May | −0.4907 | 0.6284 |
| June | 1.4979 | 0.1422 |
| July | −0.5500 | 0.5882 |
| August | 0.4496 | 0.6567 |
| September | 0.0190 | 0.9850 |
| October | 1.5436 | 0.1319 |
| November | −0.0481 | 0.9620 |
| December | −0.5449 | 0.5904 |

**Table B.4**
*T*-tests for significance of differences in earnings for cohorts born in 1941–61. Male sample.

| Month of Birth (vs all other months) | *t* | Pr(\|*T*\| > \|*t*\|) |
|---|---|---|
| January | 1.5823 | 0.1262 |
| February | 2.1104 | 0.0437** |
| March | −0.2324 | 0.8177 |
| April | 0.1572 | 0.8763 |
| May | 0.1411 | 0.8888 |
| June | 1.1857 | 0.2475 |
| July | −0.9570 | 0.3490 |
| August | −2.6389 | 0.0143** |
| September | −0.7718 | 0.4476 |
| October | 0.3591 | 0.7229 |
| November | −0.2111 | 0.8347 |
| December | −0.0117 | 0.9908 |

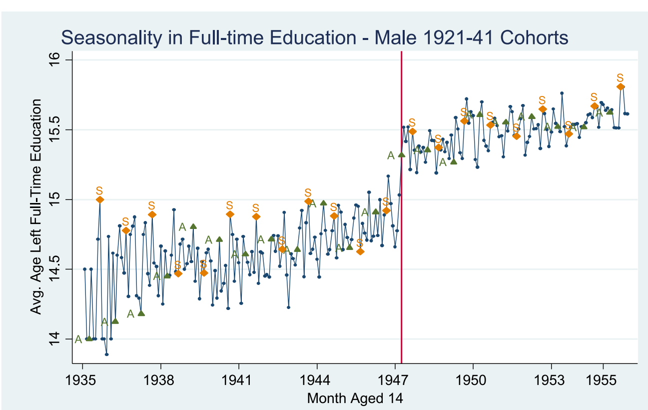


**Fig. B.1.** Seasonality in age left full-time education for cohorts born in 1921–41. Male sample *(A is April and S is September).*

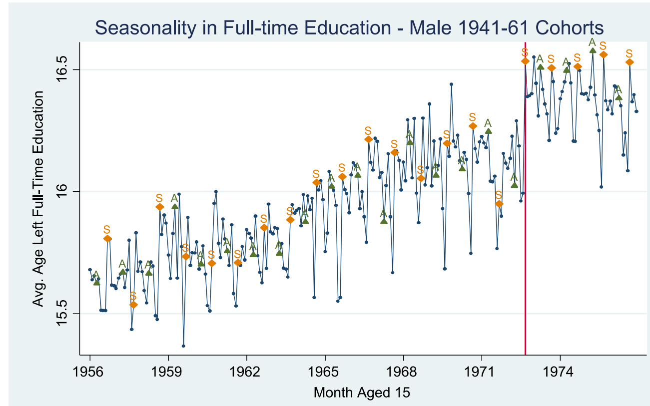


**Fig. B.2.** Seasonality in age left full-time education for cohorts born in 1941–61. Male sample *(A is April and S is September).*

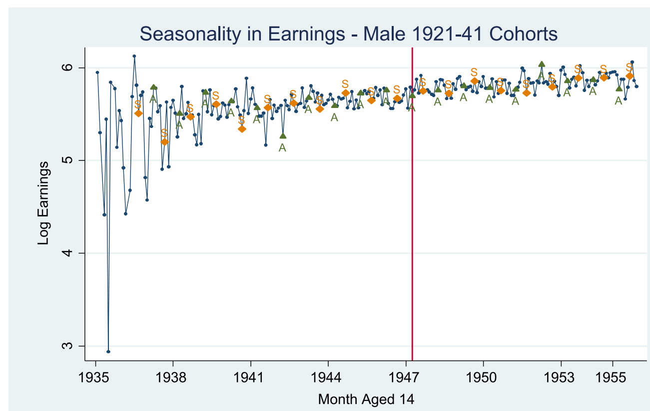


**Fig. B.3.** Seasonality in earnings for cohorts born in 1921–41. Male sample *(A is April and S is September).*

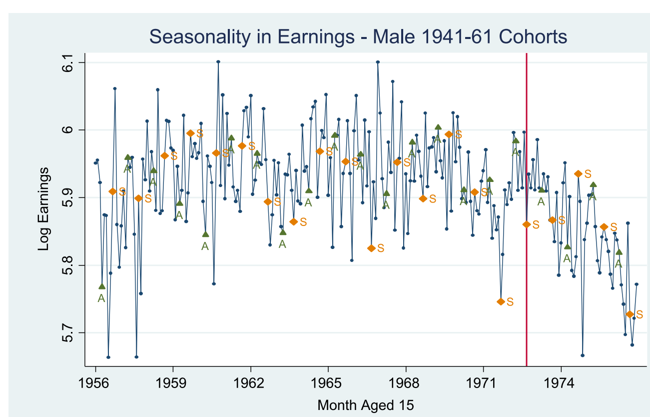


**Fig. B.4.** Seasonality in earnings for cohorts born in 1941–61. Male sample *(A is April and S is September).*

## Appendix C. Replication analysis – additional 2SLS estimates and AIC

**Table C.1**
1st stage and 2SLS effects of ROSLA laws on log weekly earnings – male sample.

| Order YoB Polynomial | | GHS 1979–2006 (1) | (2) | (3) | GHS 1979–1986 (H&W) (4) | (5) | (6) |
|---|---|---|---|---|---|---|---|
| 2SLS estimates | One | −0.075* | −0.057 | −0.054 | 0.032* | 0.107*** | 0.106*** |
| | | (0.043) | (0.039) | (0.039) | (0.018) | (0.021) | (0.026) |
| | Two | 0.029* | 0.035** | 0.040* | 0.002 | 0.066*** | 0.046* |
| | | (0.017) | (0.017) | (0.022) | (0.019) | (0.021) | (0.027) |
| | Three | 0.039** | 0.040** | 0.043* | −0.120*** | 0.041* | 0.007 |
| | | (0.018) | (0.018) | (0.022) | (0.023) | (0.022) | (0.029) |
| | Four | 0.069*** | 0.067** | 0.072** | 0.032 | 0.065*** | 0.040 |
| | | (0.024) | (0.024) | (0.029) | (0.023) | (0.022) | (0.029) |
| | Five | 0.068** | 0.066** | 0.072** | 0.018 | 0.042* | 0.035 |
| | | (0.025) | (0.026) | (0.030) | (0.025) | (0.025) | (0.030) |
| | Six | 0.095*** | 0.095** | 0.104** | 0.021 | 0.010 | 0.027 |
| | | (0.034) | (0.036) | (0.039) | (0.029) | (0.029) | (0.033) |
| First-stage estimates | | | | | | | |
| | One | 0.433*** | 0.461*** | 0.466*** | 0.598*** | 0.606*** | 0.563*** |
| | | (0.036) | (0.033) | (0.033) | (0.027) | (0.033) | (0.038) |
| | Two | 0.535*** | 0.553*** | 0.563*** | 0.607*** | 0.617*** | 0.576*** |
| | | (0.040) | (0.042) | (0.047) | (0.028) | (0.034) | (0.039) |
| | Three | 0.584*** | 0.597*** | 0.601*** | 0.550*** | 0.601*** | 0.552*** |
| | | (0.040) | (0.044) | (0.046) | (0.029) | (0.034) | (0.040) |
| | Four | 0.490*** | 0.482*** | 0.477*** | 0.620*** | 0.626*** | 0.580*** |
| | | (0.028) | (0.025) | (0.027) | (0.036) | (0.036) | (0.043) |
| | Five | 0.494*** | 0.485*** | 0.482*** | 0.577*** | 0.579*** | 0.563*** |
| | | (0.024) | (0.023) | (0.023) | (0.036) | (0.037) | (0.043) |
| | Six | 0.454*** | 0.444*** | 0.451*** | 0.562*** | 0.556*** | 0.550*** |
| | | (0.023) | (0.023) | (0.025) | (0.040) | (0.040) | (0.046) |
| | Age controls | Quadratic | Quartic | Dummies | Quadratic | Quartic | Dummies |

NOTE – H&W is Harmon and Walker (1995).
*** Significant at the 1% level.
** Significant at the 5% level.
* Significant at the 10% level.

**Table C.2**
AIC calculated for IV models – male sample.

| Order of polynomial | | Our own dataset AIC | | | Harmon and Walker (1995) AIC | | |
|---|---|---|---|---|---|---|---|
| **1947** reform | 1 | 93,491 | 91,446 | 91,158 | 47,282 | 45,105 | 45,078 |
| | 2 | 84,048 | 83,651 | 83,419 | 48,550 | 45,467 | 45,974 |
| | 3 | 83,470 | 83,372 | 83,212 | 56,935 | 46,084 | 47,466 |
| | 4 | 82,010 | 82,066 | 81,893 | 46,476 | 45,423 | 46,068 |
| | 5 | 82,059 | 82,170 | 81,923 | 47,058 | 46,028 | 46,256 |
| | 6 | 81,054 | 81,015 | 81,005 | 46,980 | 47,380 | 46,594 |
| | Age controls | Quadratic | Quartic | Dummies | Quadratic | Quartic | Dummies |

*Note* – Estimates in bold are AIC estimates that apply to those replicated from Devereux and Hart (2010).

## Appendix D. AIC values for our new 2SLS results

**Table D.1**
2SLS effects of ROSLA laws on log weekly earnings – male sample, 4th order yob/myob polynomials.

| Controls | *IV* used *N obs* | AIC values | | | |
|---|---|---|---|---|---|
| Y-o-B polyn., years of age | Old IV $N = 63{,}746$ | 97478.47 | 97576.42 | 97557.35 | 97524.02 |
| | New IV $N = 63{,}084$ | 95125.89 | 95,010 | 95045.78 | 95014.58 |
| | Set of IVs $N = 63{,}084$ | 95425.22 | 95365.39 | 95384.26 | 95311.52 |
| Y-o-B polyn., years of age | Old IV $N = 37{,}963$ | 63246.28 | 63370.55 | 63391.39 | 63990.74 |
| | New IV $N = 37{,}963$ | 63999.03 | 64155.46 | 64148.94 | 64608.14 |
| | Set of IVs $N = 37{,}963$ | 64105.66 | 64250.68 | 64262.73 | 64835.88 |
| Y-o-B polyn., years of age, survey year FE, survey month FE, month of birth FE | Old IV $N = 37{,}963$ | 63110.61 | 63256.61 | 63300.94 | 63801.48 |
| | New IV $N = 37{,}963$ | 63476.6 | 63652.97 | 63631.93 | 64030.26 |
| | Set of IVs $N = 37{,}963$ | 63520.7 | 63671.62 | 63697.17 | 64188.42 |
| M-Y-o-B polyn., years of age, survey year FE, survey month FE, month of birth FE | Old IV $N = 37{,}963$ | 63007.45 | 63157.48 | 63202.82 | 63729.19 |
| | New IV $N = 37{,}963$ | 63442.39 | 63627.35 | 63601.18 | 64,050 |
| | Set of IVs $N = 37{,}963$ | 63451.23 | 63608.6 | 63629.93 | 64165.85 |
| M-Y-o-B polyn., months of age, survey year FE, survey month FE, month of birth FE | Old IV $N = 37{,}963$ | 63002.48 | 63172.21 | 63217.66 | |
| | New IV $N = 37{,}963$ | 63446.75 | 63658.26 | 63626.14 | |
| | Set of IVs $N = 37{,}963$ | 63447.11 | 63622.1 | 63647.42 | |
| | Age controls | Quadratic | Cubic | Quartic | Dummies |

## References


Acemoglu, D., Angrist, J.D., 2001. How large are human-capital externalities? Evidence from compulsory schooling laws eds.. Bernanke, Ben S., Rogoff, K. (Eds.), N*BER Macroeconomics Annual* 2000. MIT Press, Cambridge, MA, pp. 9–59.

Albouy, V., Lequien, L., 2009. Does compulsory education lower mortality? J. Health Econ. 28 (1), 155–168.

Angrist, J.D., Krueger, A.B., 1991. Does compulsory school attendance affect schooling and earnings? Q. J. Econ. 106 (4), 979–1014.

Bhuller, M., Mogstad, M., Salvanes, K., 2017. Life-cycle earnings, education premiums, and internal rates of return. J. Labor Econ. 35 (4), October 2017. http://www.journals.uchicago.edu/journals/jole/forthcoming.

Black, S. E., P. J. Devereux, and K. G. Salvanes. (2008). "Staying in the classroom and out of the maternity ward? The effect of compulsory schooling laws on teenage births." *Econ. J.* 118 (530): 1025–54.

Buckles, K., Hungerman, D., 2013. Season of birth and later outcomes: old questions, new answers. The Review of Economics and Statistics 95. MIT Press, pp. 711–724, July.

Buscha, F., Dickson, M., 2012. The raising of the school leaving age: returns in later life. Econ. Lett. 117 (2), 389–393.

Card, D., 1995. Using geographic variation in college proximity to estimate the return to schooling. In: Christofides, Louis N., Grant, E.Kenneth, Swidinsky, Robert (Eds.), Aspects of Labour Market behaviour: Essays in Honour of John Vanderkamp. University of Toronto Press, Toronto, pp. 201–222.

Card, D., 1999. The causal effect of education on earnings. In: Orley, Ashenfelter, David, Card (Eds.), Handbook of Labor Economics 3A. Elsevier, Amsterdam, pp. 1801–1863, edited by.

Carneiro, P., Heckman, J., Vytlacil, E.J., 2011. Estimating marginal returns to education. Am. Econ. Rev. 101, 2754–2781, October 2011.

Clark, D., Royer, H., 2013. The Effect of education on adult mortality and health: evidence from Britain. Am. Econ. Rev. 103 (6), 2087–2120.

Del Bono, E., Galindo-Rueda, F., 2004. Do a Few Months of Compulsory Schooling Matter? The Education and Labour Market Impact of School Leaving Rules. Institute for the Study of Labor (IZA), p. 1233, Discussion Paper.

Del Bono, E., Galindo-Rueda, F., 2007. The Long Term Impacts of Compulsory Schooling: Evidence from a Natural Experiment in School Leaving Dates. Centre for the Economics of Education, LSE, CEE Discussion Papers 0074.

Devereux, P.J., Hart, R.A., 2010. Forced to be rich? Returns to compulsory schooling in Britain. Econ. J. 120 (549), 1345–1364.

Dickson, M., 2013. The causal effect of education on wages revisited. Oxf. Bull. Econ. Stat. 75 (4), 477–498.

Dickson, M., Smith, S., 2011. What determines the return to education: an extra year or a hurdle cleared? Econ. Edu. Rev. 30 (6), 1167–1176.

Gelman, A., Imbens, G., 2014. Why High-Order Polynomials Should not be Used in Regression Discontinuity Designs.. National Bureau of Economic Research, Working Paper 20405.

Grenet, J., 2013. Is extending compulsory schooling alone enough to raise earnings? Evidence from French and British compulsory schooling laws. Scand. J. Econ. 115 (1), 176–210.

Halsey, A.H., Heath, A., Ridge, J., 1980. Origins and Destinations-Family, Class and Education in Modem Britain. Oxford University Press, Oxford.

Harmon, C., Walker, I., 1995. Estimates of the economic return to schooling for the United Kingdom. Am. Econ. Rev. 85 (5), 1278–1286.

Heckman, J., Polachek, S., 1974. Empirical evidence on the functional form of the earnings-schooling relationship. J. Am. Stat. Assoc. LXIX, 350–354.

Imbens, G., Kalyanaraman, K., 2012. Optimal bandwidth choice for the regression discontinuity estimator. Review of Economic Studies 79. Oxford University Press, pp. 933–959.

Kling, J.R., 2001. Interpreting instrumental variables estimates of the returns to schooling. J. Bus. Econ. Stat. 19 (3), 358–364.

Koop, G., Tobias, J.L., 2004. Learning about heterogeneity in returns to schooling. J. Appl. Econom. 19, 827–849.

Lee, Lemieux, T., 2010. Regression discontinuity designs in economics. J. Econ. Lit. 48 (2), 281–355.

Lleras-Muney, A., 2005. The relationships between education and adult mortality in the United States. Rev. Econ. Stud. 72 (250), 189–221.

Nickell, S., 1993. Cohort Size and Earnings. Mimeo, Oxford University.

Oreopoulos, P., 2006. Estimating average and local average treatment effects of education when compulsory schooling laws really matter. Am. Econ. Rev. 96 (1), 152–175.

Pischke, J.-S., von Wachter, T., 2008. Zero returns to compulsory schooling in Germany: evidence and interpretation. Rev. Econ. Stat. 90, 592–598, 2008.