I Z A

DISCUSSION PAPER SERIES

The Wage Returns to Education over the Life-Cycle: Heterogeneity and the Role of Experience

Franz Buscha
Matt Dickson

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

The Wage Returns to Education over the Life-Cycle: Heterogeneity and the Role of Experience

Franz Buscha
University of Westminster

Matt Dickson
University of Bath,
CMPO, University of Bristol and IZA

ABSTRACT

The Wage Returns to Education over the Life-Cycle: Heterogeneity and the Role of Experience[*]

This paper re-examines the wage returns to the 1972 Raising of the School Leaving Age (RoSLA) in England and Wales using a high-quality administrative panel dataset covering the relevant cohorts for almost 40 years of their labour market careers. With best practice regression discontinuity methods we find at best a zero return to the additional education for men. However, we contend that regression discontinuity methods in this context will give unreliable estimates of the return. Using the panel data to correct for this we find a local average treatment effect of 7% over the lifetime for this additional year of education.

JEL Classification: I26, I28, J24, J31

Keywords: returns to education, life-cycle, earnings, experience

Corresponding author:

Matt Dickson
Department of Social and Policy Sciences
University of Bath
Bath BA2 7AY
United Kingdom
E-mail: m.dickson@bath.ac.uk

1. Introduction

The 1990s saw a new wave of UK literature on the returns to education, with studies for the first time exploiting minimum school leaving age reforms to derive causal estimates of the impact of additional schooling on earnings.[1] The consensus across these studies was that the return to education for UK men lay in the range of 15-20%. Whilst this is good news for policy makers and believers in human capital theory – especially in light of the raising of the participation age in England and Wales from 16 to 17 in 2013 and up to 18 in 2015 – these early results have since been called into question. In fact, recent re-appraisals of the causal return to education (Devereux and Hart, 2010; Grenet, 2013) using more sophisticated identification techniques and larger datasets, suggest that the effect of an additional year of schooling is less than half of what was previously estimated. The UK returns literature is therefore in a state where 'causal' returns to an extra year of schooling apparently vary between 3% and 20%.

One possible explanation for this emerging disparity in the returns to education is the role of life-cycle effects. That is, the returns to education could be non-constant over the life-cycle and therefore the varying returns estimated may be explained by the fact that many of these studies use data from different time periods. Recent developments in the literature (see Haider and Solon, 2006; Bhuller, *et al.*, 2011) suggest that life-cycle effects are an important component in the assessment of the return to education and that constant returns should not be taken for granted. Bhuller *et al.* specifically warn of the danger of life-cycle bias and caution regarding the necessity to pay close attention to differences in age composition when comparing estimates of the returns to schooling. However, although life-cycle bias is one possible contributory factor in explaining the varying estimates of the return to education in the UK literature, there are also other aspects to consider: different studies exploit different schooling reforms that affect the behaviour of complier groups differing in size and characteristics; these reforms also work on different educational margins, some of which coincide with major exam taking school years; methodological differences, for example, different functional forms estimated or different econometric techniques applied, and different data sources (some of which are weaker than others).

In this paper we highlight the importance of these factors by examining the wage effect of the 1972 RoSLA in England and Wales over the life-cycle of those affected by it. Following Devereux and Hart (2010), we use high quality administrative data from the New Earnings Survey Panel Dataset, providing a large sample of individuals affected by the 1972 reform. The data covers 38 years, spanning from 1975 when the first affected cohorts were 17/18 years old, through to 2012, when these cohorts were 54/55. The wide range of the data available allows us to examine the impact of the reform pooled over almost the entire career, and importantly, the

---

[1] See Harmon and Walker, 1995; Harmon and Walker, 1999; Chevalier and Walker, 2002.

size of the sample allows us to examine the impact in each *individual year* during the lifecycle. In each case, we make use of regression discontinuity design (RDD), which has become the standard empirical strategy to achieve more robust causal inference. The size of the sample also allows us to focus very narrowly on the cohorts in the immediate vicinity of the policy change. Following Grenet (2013), we also use data from the Labour Force Survey (LFS) 1993 to 2011[2], which allows us to highlight the way in which the picture changes when only the later part of the lifecycle is examined using smaller and less reliable data sources.

Our results suggest that there is significant lifecycle heterogeneity in the return to education. Regression discontinuity analysis of the NESPD data shows that the impact of the 1972 RoSLA was initially *negative*: individuals affected earned lower hourly wages in their late teenage years and into their 20s. This is the case irrespective of the bandwidth chosen for the RDD estimation. From the age of 30 onwards there are very few significant differences in earnings as a result of RoSLA, with approximately equal numbers of negative and positive point estimates, depending on choice of bandwidth.

More generally we make two contributions. Firstly, we show that the now standard approach to assessing the impact of compulsory schooling reforms by pooling data and implementing a RDD will mask lifetime heterogeneity and potentially lead to inaccurate policy inference depending on the age distribution of the estimating data. Secondly, and perhaps more importantly, we show that while the RDD methodology attempts to show the impact of one additional year of schooling and one less year of potential labour market experience (i.e. the 'net' return to a year of education), this is in practice difficult to achieve when identification comes from cohort comparisons, and the institutional set-up dictates that potential experience (unlike age) does not accrue in units of under one year. As such, local linear estimators must attempt to extrapolate longer-run *between*-cohort trends to *within*-cohort differences – something that is difficult to do accurately given the discrete nature of (potential) labour market experience and the reality that relative age-within-school-cohort affects educational and labour market outcomes, meaning that bandwidths of under one year are not advised.[3] The alternative approach to local linear RDD estimation in this case is to compare only those born in the cohorts immediately before and after the reform, the approach we favour here. This will average the age-within-cohort effects that complicate local linear regressions and given that the two cohorts being compared are as close as possible in terms of time of birth, the comparison is not undermined by longer-term inter-cohort trends in educational attainment and wages. However, this approach is not without

---

[2] Terminal education age information is available from 1986 onwards and we also make use of this. Grenet's data was limited to the period 1993 to 2006.

[3] See Gelman and Imbens (2014) for discussion of why local linear regression as opposed to the global polynomial approach is preferred in RD estimation. Lee and Lemieux (2010) and Imbens and Lemieux (2008) discuss the inclusion of covariates in RD designs stating that they can potentially improve precision on the assumption that for each covariate Z the conditional distribution of Z given X is continuous at the discontinuity point. This is not the case when Z is potential experience (X is school cohort).

cost. The comparison now becomes between individuals with one additional year of schooling but *two* fewer years of experience. This is because, as well as the discontinuity in education at the policy reform point, there is also a discontinuity in potential experience. The negative effect of lower experience as a result of additional education (added to the effect of the reform cohort being a year younger) is potentially substantial amongst the complier group for minimum school leaving age reforms, thus the mean differences RDD comparison results in a downwardly biased estimate of the 'net' return to education. Thus best practice RDD estimators are not able to adequately control for the experience differences because of the discrete nature of potential experience and the discontinuity in potential experience at the RoSLA point. As such, they are not able to compare 'like with like' in the way that more traditional parametric approaches attempt to do. We propose a correction to the RDD estimates and demonstrate that taking into account the experience difference between the treated and comparison groups can have a large effect – in this case the zero or negative return to education becomes significant and positive. These findings explain some of the diverging results in the UK literature. We argue then, firstly that lifecycle effects in the returns to education are an important component in the general analysis of returns to education and, secondly that without adjustment RDD estimates exploiting compulsory schooling reforms and relying on cohort comparisons do not give an accurate estimate of the return to education for the complier group.

The remainder of this paper is structured as follows. Section 2 reviews the literature and describes the raising of the school leaving age reform in 1972. Section 3 outlines our data and the approach that we take to overcoming the issues we identify with the current literature. Section 4 describes the methodology and our proposed correction to standard RDD estimates before Section 5 presents the results. Finally section 6 offers some discussion and conclusions.

2.   Literature Review

2.1 The returns to education in the UK

The literature on estimating the returns to education has made great strides forwards since the seminal articles from the 1950s, 1960s and 1970s and this is evidenced by numerous high profile survey and review articles that summarize collective research findings to date (such as Card, 1999 and 2001; Harmon *et al*, 2003; Heckman *et al*. 2003; Lemieux, 2006 and Polacheck, 2007, Oreopoulos and Petronijevic, 2013). Although many issues still remain within this literature, there have been a number of significant advances that have taken place in the last few decades.

At the forefront of such innovation is the empirical implementation of instrumental variable techniques that attempt to causally estimate the effect of additional schooling on earnings. Specifically, papers using the exogenous variation in the school leaving age laws (Angrist and Krueger, 1991 and Harmon and Walker, 1995) began to make convincing

arguments that it is possible to estimate robust causal returns to education. Although various other identification approaches such as twin studies (Ashenfelter and Krueger, 1994), alternative instruments (Card, 1995) or estimators exploiting differences in 2nd moments (Klein and Vella, 2009) have also been proposed, to date instrumental variable regressions using school leaving age laws remain the most widely accepted technique for estimating causal returns to education.[4] As such, there has been a wide range of studies making use of this method to analyse not only wage returns to education but also the causal effect of education on a number of other outcomes *inter alia* health, employment and crime.

Within this instrumental variable literature in the UK, a revisionist view has recently emerged. Early UK evidence on the returns to education by Harmon and Walker (1995), Harmon and Walker (1999) and Chevalier and Walker (2002) suggested that one additional year of schooling is associated with relatively large returns to earnings, in the order of 15-20%. However, more recent studies by Devereux and Hart (2010) and Grenet (2013) appear to contradict such findings and suggests smaller causal estimates, in the range of 3-8% per additional year of schooling. One argument for reconciling these different estimates is that larger and better data sources in addition to newer and more precise estimation techniques (RDD) have allowed for a more accurate picture to emerge – one that significantly reduces the causal return to education.

An additional explanation (see Buscha and Dickson, 2012), is that due to data constraints and cross-sectional pooling, various papers have examined returns to education at different points in the lifecycle. Considering all of these estimates, a potential explanation to reconcile the two strands of the UK returns literature is that returns may vary over the lifecycle and this translates into different causal estimates depending on the age distribution of the sample. The suggestion that lifecycle bias may play an important role when estimating the returns to education is echoed by Haider and Solon (2006) and Bhuller, *et al.*, (2011) who support the notion that constant returns should not be so readily assumed and that it is vital to incorporate potential lifecycle effects into any assessment of the returns to education. Bhuller *et al.* (2011), for example, show clearly that the return to schooling for Norwegian men varies across the lifecycle. Using administrative data and exploiting school leaving age reforms for identification, they show that the return to one additional year of schooling is non-existent at age 28, 10% at age 33, peaks at 15% at age 43 and then diminishes to 13% at age 58. Furthermore, after reviewing the recent literature, Heckman, Lochner and Todd (2006) conclude that wage patterns have changed substantially over time across cohorts and as a result cross-sections no longer approximate the life-cycle earnings or schooling returns of any particular individual. Consequently, they propose that the use of data that follows actual cohorts over the life-cycle is essential to accurately

---

[4] It should be noted that the interpretation of IV estimates must be seen within the context of a Local Average Treatment Effect (LATE) which limits the causal effects to only a subsample of compliers.

measure their true earnings pattern and estimate the education premiums experienced by individuals. This paper is the first to our knowledge to do exactly this for the UK.

2.2 The 1972 Raising of the School Leaving Age

The raising of the school leaving age (RoSLA) in the UK has occurred twice[5] since the end of the Second World War: to some controversy in the immediate post-war period in 1947 and most recently in 1972. The original foundation for both pieces of legislation can be found in the Education Act of 1944[6] which increased the compulsory schooling leaving age from 14 to 15 on the 1st April 1947 whilst at the same time providing the President of the Board of Education with the powers to raise the school leaving age to 16 as soon as it was deemed practicable (Education Act, 1944, section 35). This finally occurred in 1972 and from 1st September that year all children attending schools in England and Wales were required to stay on until the age of 16.[7]

The 1972 RoSLA event therefore affected all individuals who were born on or after the 1st September 1957. Anybody born after this point was subject to a minimum of 11 years of schooling whilst those born before could have received a minimum of 10 years of schooling. In addition, because the nature of the English and Welsh schooling system implies that the minimum entry and exit ages are regulated at a national level there was little scope for regional variations to exist (in contrast the United States education system).

The RoSLA events that took place in 1947 and 1972 provide quasi-experimental variation in the average number of years of education for cohorts born around the discontinuities induced by the policy changes – i.e. the cohorts born either side of 1st April 1933 and 1st September 1957. This variation has been exploited in a number of studies examining not only the impact on hourly wages but also the effect on alternative outcomes such as health, crime and voting behaviour. Table 1 summarises the various findings, the upper panel focusing on earnings effects while the lower panel documents impacts on the additional outcomes examined in the literature. As alluded to above, the estimates of the wage returns to education derived using RoSLA as an instrument range from the very low or zero (Dickson and Smith, 2011), through estimates of 6-7% (Grenet, 2013) to the much higher estimates of 15% to 20% (Harmon and Walker, 1995; Chevalier and Walker, 2002).

---

[5] We do include the most recent raising of the *participation* age in 2013 and 2015 or the change induced in 1962 when the 'Easter Leaving Rule' was introduced.

[6] Also known as the "Butler Act" after the President of the Board of Education, Rab Butler, it was passed on 3rd August 1944.

[7] More of the historical context can be found in McCulloch, *et al.* (2012).

Table 1: A Review of the Impact of the School Leaving Age Reforms of 1947 and 1972 and its effects

Study	Data	RoSLA	Effect of RoSLA on education	Effect of RoSLA on **outcome**
Earnings				
Harmon and Walker (1995)	Pooled Family Expenditure Survey 1978 to 1986	1947 and 1972	0.541 additional years of schooling for the 1947 change and 0.110 additional years for the 1972 change	IV estimate of 15% **on hourly wages** for men
Chevalier and Walker (2002)	British Household Panel Survey 1991 to 1996	1972	Not reported	IV estimate of 17-20% **in hourly wages** for men
Oreopoulos (2006)	Pooled General Household Survey's 1983 to 1998	1947 and 1972	Combined effect of 0.453 additional years of schooling	Combined reduced form effect on **hourly wages** of 6% and an IV effect of 15% for men.
Devereux and Hart (2010)	Pooled General Household Survey's (GHS) 1979 to 1998 and the Pooled New Earnings Survey Panel (NESPD) 1975 to 2001	1947	0.469/0.397 (GHS/NESPD) additional years of schooling for men 0.550/0.511 (GHS/NESPD) additional years of schooling for women	The 1947 reform had a reduced form effect of 2% for **hourly wages** for men. No effect for women.
Dickson and Smith (2011)	Pooled Quarterly Labour Force Survey 1993 to 2010	1972	Proportion with no academic qualifications fell by 0.071; those with level 1 qualifications increased by 0.047; those with level 2 qualifications increased by 0.041. No effect on level 3+ qualifications	No effect on log **hourly wages** A positive (reduced form) **employment** effect of 9%
Buscha and Dickson (2012)	UK Household Longitudinal Study 2011	1972	0.225 additional years of schooling for men 0.356 additional years of schooling for women	**Hourly wages** increased by 5% for men and 6% for women.
Grenet (2013)	Pooled Quarterly Labour Force Survey 1993 to 2004	1972	0.274 additional years of schooling for men 0.317 additional years of schooling for women	The 1972 reform has a reduced from effect of 2-3% on **hourly wages** for men.
Dickson (2013)	British Household Panel Survey	1972	0.564 additional years of schooling for men	An additional year of schooling

Study	Data	RoSLA	Effect of RoSLA on education	Effect of RoSLA on **outcome**
	1991 to 2006		(men only examined)	(IV estimate) increases **hourly wages** by 10% for men.

Other outcomes

Study	Data	RoSLA	Effect of RoSLA on education	Effect of RoSLA on **outcome**
Milligan *et al.* (2004)	Pooled British Election Studies 1964, 74, 79, 83, 87, 92 and 97	1947 and 1972	Relative to those aged 14 the 1947 change increased the average age before drop out by 0.512 and the 1972 change by 0.953	There is no effect on the probability of **voting**
Siles (2009)	Pooled General Household Survey's 1980 to 2004	1947 and 1972	0.593 additional years of schooling for the 1947 change and 0.186 additional years for the 1972 change	One more year of education (IV estimate) increase probability of self-reported **good health** by approximately 7%
Machin *et al.* (2011)	Pooled General Household Survey's 1972 to 1996	1972	5.7% drop in the proportion of individuals with no qualification and an increase of 0.221 years of additional schooling	RoSLA significantly impacted **crime** rates. The reduced from effect on the conviction rate is -5%.
Siles (2011)	Pooled General Household Survey's 1978 to 2004	1947 and 1972	0.466 additional years of schooling for the 1947 change and 0.233 additional years for the 1972 change	No effect on **early childbirth** from 1947 reform. 1972 reform reduced the probability of a teen birth by 15%
Wilson (2012)	Pooled Labour Force Survey's 1975 to 2006	1972	0.506 additional years of schooling for women (women only examined)	Reduction in probability of **a teen birth** by 7%.
Clark and Royer (2013)	Pooled Health Survey of England 1991 to 2004	1947 and 1972	1947 = +0.420yrs of schooling for men 1947 = +0.527yrs of schooling for women 1972 = +0.318yrs of schooling for men 1972 = +0.252yrs of schooling for women	No significant effect on **health** indicators (objective) or self-reported health, mortality, or health-related behaviours

3. Data

3.1 The LFS and NESPD

Our data come from two sources. Firstly, to replicate and examine recent findings from Grenet (2013), we use the British Labour Force Survey. We use Quarterly LFS data, pooled from 1993 quarter one to 2011 quarter four inclusive. The LFS is the largest regular household survey in the UK and is designed to be representative of the population living in private households, with approximately 60,000 households responding each quarter. Following Grenet (2013) we restrict the earnings estimations to include only information from an individual's last interview. We select individuals resident in England and Wales and exclude those living in Scotland or Northern Ireland and those who were born outside of England and Wales unless they moved to Britain prior to commencing their secondary education. We also exclude those who are self-employed and those who first left full-time education after the age of 25. We calculate gross hourly pay excluding overtime and deflate earnings into 2013 £s. Finally, we trim the earnings distribution to remove the top and bottom 1% of the distribution.

Secondly, following Devereux and Hart (2010), we make use of the New Earnings Survey Panel Dataset, which is a 1% sample of employees in Great Britain, based on the last two digits of an individual's National Insurance number. The NESPD follows the earnings of approximately 170,000 individuals over the period 1975 to 2011. The administrative nature of the NESPD provides a high quality companion to LFS data as earnings information is obtained via employer payroll records. Employers are legally obliged to provide this information, which results in a very high response rate. Moreover, because individuals are tracked via their NI number, attrition from the survey due to unemployment, withdrawal from labour force or failure of sample location does not result in permanent attrition from the data since such individuals will likely be captured again in later years.

However, not everyone is included in the sample frame of the NESPD. The self-employed, those who switch jobs around the time of the questionnaire date and individuals who earn less than the minimum Pay As You Earn (PAYE) tax threshold are not included in this survey. Self-employed individuals are excluded from our analysis in any case and those switching jobs at the time of survey are unlikely to significantly impact our estimates. The failure to include individuals who earn below the PAYE threshold is potentially a more serious issue and one which was investigated in some detail by Devereux and Hart (2010) and Dickens (1999). They both concur that this is not a substantial issue and that controlling for this under-sampling has a negligible effect on estimates. In operationalizing this data we use the same dependent variable as Devereux and Hart (2010): gross hourly pay excluding overtime deflated to 2013 £s. We cut the top 1% of the earnings distributions but leave the bottom 1% due to the aforementioned PAYE selection.

Four important differences exist between our study and that of Devereux and Hart (2010). Firstly, Devereux and Hart focus on the 1947 RoSLA which raised the school leaving age from 14 to 15 whilst we analyse the 1972 RoSLA which raised the school leaving age from 15 to 16. Ours is the only study to date to examine the impact of the 1972 RoSLA using data of this sample size and administrative quality. Secondly, Devereux and Hart pool their data together to obtain one combined estimate of the effect of schooling on earnings.[8] In contrast, we focus on the lifecycle heterogeneity in the impact of RoSLA and as such present results year-by-year in addition to providing pooled estimates. Thirdly, whether looking at year-by-year estimates or pooled estimates we use both mean differences and local linear regression to implement the RDD rather than the global polynomial approach used by Devereux and Hart and subsequently critiqued by Gelman and Imbens (2014). Fourthly, we have precise information on the individual's year of birth and the school cohort to which they belonged. In the English schooling system, a 'school year' cohort includes all individuals born between 1st September in year *t* and the 31st August in year *t+1* inclusive. Year of birth can be calculated using year of survey and age at survey, and in most cases this will be able to correctly assign people to pre- or post-reform school years. However for the calendar year that includes parts of both the pre-reform and post-reform cohorts, year of birth alone cannot assign people into the correct cohort. This could introduce a degree of measurement error into estimates (an issue we outline in Appendix C).[9] We side-step this issue by exploiting precise information on the school year that an individual belonged to.

3.2 Descriptive Statistics

This section presents descriptive statistics on our two data sources. Because we are principally interested in the last cohort before the schooling leaving age was raised and the first cohort after the school leaving age was raised we limit our descriptive overview to this range of the data for both the LFS and NESPD.[10] In addition, we forego presenting detailed wage descriptives for the LFS as our preferred data source, the NESPD, has a larger sample size, spans

---

[8] To overcome the fact that some individuals have more responses than other individuals they weight each observation by the inverse of the number of times the person appears in the data. Because the focus of our analysis is specifically on life-cycle effects where we estimate year-by-year effects, we do not generally require this type of re-weighting but where we estimate using multiple observations per person we apply the same IPW procedure.

[9] Devereux and Hart calculated year of birth as (survey year – age – 1). For the years pre-2004, the age variable in the NESPD referred to age on 1st January and as such (survey year – age – 1) will correctly assign year of birth. From 2004 onwards the NESPD age variable refers to age at time of survey, which is in April. In this case, (survey year – age – 1) will assign the correct year of birth for people born in 8 or 9 months of the year (depending on how April births are treated). As the vast majority of individuals in Devereux and Hart's data appear pre-2004 this issue is very minor for their estimations, moreover they show their estimates to be robust to different assumptions regarding those born in 1933 (the year containing individuals from both the pre- and post-reform cohorts).

[10] We thank the Office for National Statistics for providing information allowing us to identify the school cohort that each individual is a member of in the New Earnings Survey Panel Dataset.

more survey years, and provides higher quality wage data.[11] Ultimately, we only use the LFS to create a reference point by replicating Grenet (2013) and to highlight the education discontinuity of the 1972 RoSLA.

In terms of educational differences caused by the 1972 RoSLA event we are only able to use the LFS since only this data contains information on education. However, as we can see from Table 1, it is by now well-established that the 1972 school leaving age reform significantly increased the amount of schooling received for men in the treated cohorts: on average by between one quarter and one third of a year of extra schooling. Unsurprisingly we find a similar result in our LFS data, estimating an average increase in schooling of approximately one third of a year as a result of the 1972 RoSLA (60 month bandwidth estimate: 0.321 [s.e. 0.044]). However, within the context of our study it is worth exploring whether the additional schooling of the treated cohort remains stable over successive LFS waves. To the best of our knowledge no study has yet examined this issue and one would expect, *a priori*, the educational difference to remain stable over time. Figure 1 presents these results.

Figure 1: Average age left school by survey year in the LFS
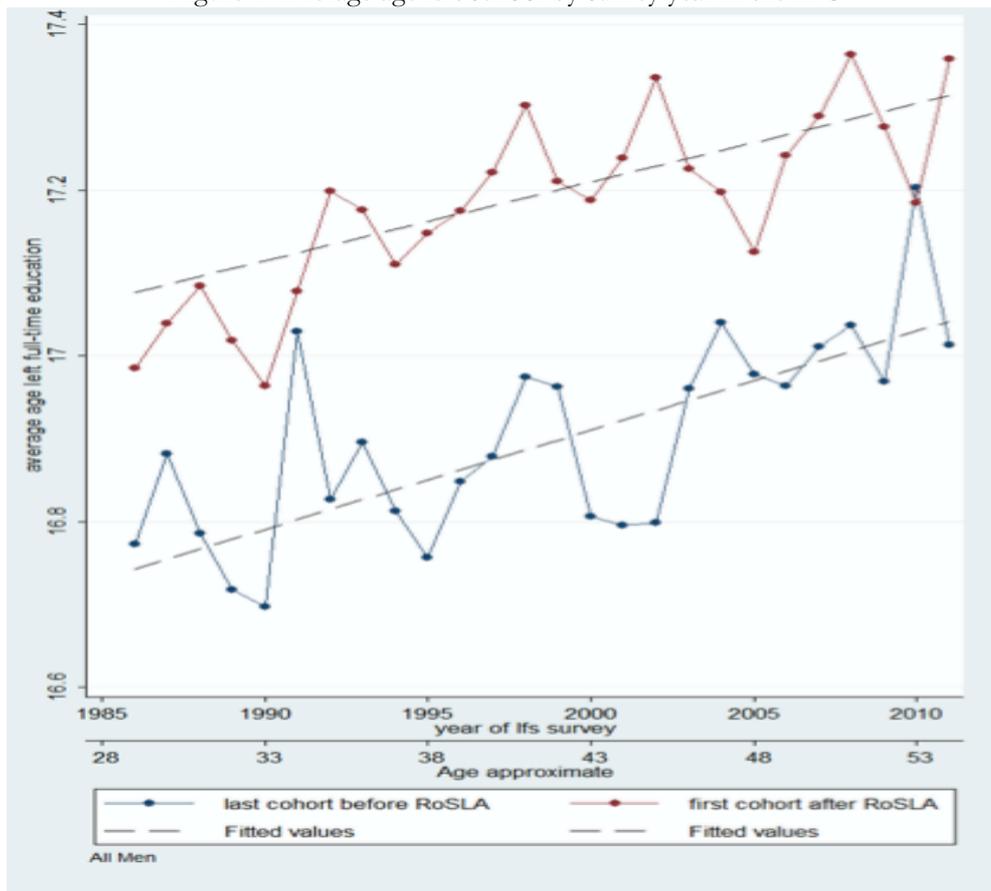
Source: LFS 1986-2011

---

[11] For example, we find that LFS wage results are very sensitive to the inclusion of proxy responses, which comprise almost one third of the data. Results are contained in appendix Table B1.

Figure 1 offers a surprising finding. Even though, on average, the educational distance between the two cohorts remains stable over time, there appears to be substantial year-by-year variation in average reported schooling levels. This is concerning because if estimation is carried out on non-pooled (or relatively small pooled) data then the estimated education effect is likely to vary greatly depending on what portion of the data one is analysing. Moreover, it appears that substantial recall bias is creeping into the LFS over time. Although the LFS is not a true long-run panel data source, the relevant cohorts across the successive waves ought to be identical due to the large sampling framework of the LFS. Moreover, because we are specifically using information on the age at which individual's left full-time education, there is no scope for lifelong learning practices to push up the average age left education over time. This, however, is not the case in the observed data: average levels of age-left full time education are rising over time for both cohorts. Fortunately, the suggestion of recall bias does not invalidate our – or the literature's – results since these are driven by the difference in education which appears to stay stable over time. Nevertheless, the existence of strong fluctuations combined with evidence of recall bias makes the usage of schooling in first-stage estimates a potentially hazardous endeavour and this is one of the reasons why we primarily prefer to focus on reduced form wage effects.

We now turn to the NESPD which provides higher quality wage data. Moreover, unlike the LFS, the NESPD is not limited to earnings information post-1992, which means that we can obtain information about a greater part of the lifecycle and how earnings vary across it. Specifically, we can look at the early part of the age earnings profile that has to date not been examined in any study of the 1972 RoSLA. Figure 2 presents the average log gross hourly pay by cohort and by year of survey:

Figure 2: Average log gross hourly pay by survey year in the NESPD

Men only, RPI adjusted, base = 2013
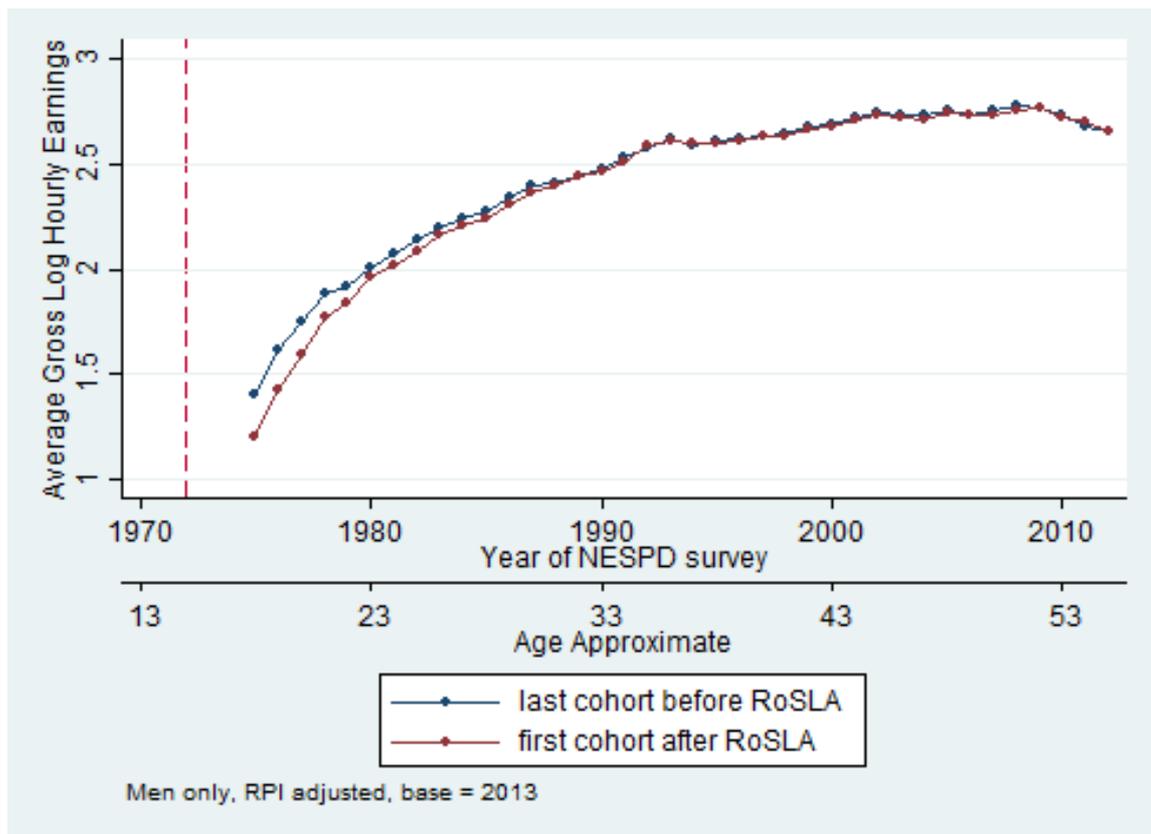
Several things are apparent in examining Figure 2.[12] First, the cohort with more education does not consistently report higher earnings across the lifecycle. Second, the pay gap between cohorts is high during the early parts of the lifecycle (and in favour of the *pre*-reform cohort) and narrows during the middle and later parts. Third, the lifecycle profile of both cohorts, although roughly inverted u-shaped, contain dips in earnings during the 1990 recession and the 2008 recession.

Although this graph is descriptive in nature it foreshadows some of our results. Importantly, analysis of the NESPD suggests that there is a penalty incurred by an additional year of schooling. The first cohort after the 1972 RoSLA reports consistently lower gross hourly pay until they reach their 30s and we conjecture that the most obvious explanation for this phenomenon is the relative lack of labour market experience of this later cohort. Indeed, because the later cohort is on average one year younger and had to stay in school for one additional year they lack on average *two* years of potential labour market experience compared with their peers from the school year above. This is not without consequence; as previously shown, for example, by Angrist (1990), the loss of potential labour market experience for U.S. Vietnam lottery draftees resulted in permanently lower earnings throughout their lives. In other words, the loss of

---

[12] The corresponding age-earnings profile for the LFS is shown in appendix Figure B1.

an additional year's potential labour market experience is likely to have a long-run impact on lifetime earnings.

If this is indeed the case then this suggests that much of the recent literature (Devereux and Hart, 2010 and Grenet, 2013) has not compared 'like-with-like' since at the point of the discontinuity both cohorts differ not only in terms of schooling but also in terms of potential labour market experience. When data is pooled and a regression discontinuity estimate is calculated, the difference in experience between the cohorts at each time point in the data is not accounted for. While a loss of experience – an opportunity cost of the additional education – is part of the 'net' return, the experience difference between the 'compliers' and their equivalents in the pre-reform cohort is not one year but *two*. Regression discontinuity methods attempt to correct for the inevitable one-year difference in experience between cohorts, however this correction is in practice very difficult when the level of measurement in the data is one-year cohort groups, as it involves extrapolating the between cohort differences to a within school-cohort gradient. Due to the nature of the English education system, potential experience in the labour market comes in discrete one-year units, and thus – unlike age – the relationship between potential experience and wages is not a smooth function. Therefore it is necessarily extrapolation based on the longer-run cohort-to-cohort trend that brings the linear trends either side of the discontinuity to the discontinuity point.

Moreover, this extrapolation will also suffer from the fact that comparing those just either side of the discontinuity would mean the youngest within the school cohort just before the reform and the oldest within the school year just after. Given the well-established age-within-cohort effects on education and labour market outcomes (see Crawford et al., 2010), successfully getting arbitrarily close to the cut off would bias the estimate upwards on this account. The upward trend in education as we approach from the left of the discontinuity (i.e. as we move from older to younger cohorts) is actually comprised of a series of downward splines in attainment from the oldest in the year (September born) down to the youngest (August born), with splines shifting upwards for the younger cohorts (see Figure 3, panel A for a stylized picture of this). These within cohort effects are however driven by age (maturity) rather than differences in potential labour market experience, which are constant within a school cohort (see Figure 3, panel B). Thus while local linear methods could successfully account for the age trend in outcomes, the running variable is actually school cohort and *not age* and the non-smooth nature of potential labour market experience which is deterministically related to school cohort makes the local extrapolation problematic in trying to provide an accurate discontinuity estimate. Information on month of birth would to some extent allow age-within-cohort effects to be accounted for, however the only effect of this is to flatten the downward splines of Figure 3 Panel A, to make it more like Panel B. The discrete nature of experience around the cut-off means that for each cohort the potential labour market experience level is flat within the cohort,

hence the relationship with wages is a series of steps – a single (average) point for each cohort. Joining these using local linear methods (or global polynomials) will allow an extrapolation to the discontinuity but this within-cohort extrapolation from a series of between cohort step changes is somewhat arbitrary. It is potential labour market experience (*rather than age*) that should matter for wages and (*unlike age*) this is not equalized as we approach the cut-off from each side. Even had we information only on those individuals born just one day either side of the cut-off, their potential experience difference would still be two years (for 'compliers') even though their age is all but equal and does smoothly converge from either side of this discontinuity.

Consequently, we argue that comparison of cohort averages for the before and after cohort is econometrically the most transparent approach – providing a clear quasi-experimental framework that does not involve any within cohort extrapolation. The proximity of the two cohorts means that the comparison between them is not undermined by longer-term inter-cohort trends in educational attainment and wages. However, when identification comes from cohort comparison with cohorts differing by one year in age, with one cohort (the younger) affected by a compulsory schooling reform, the result is that rather than comparing individuals with one additional year of schooling and consequently one year less of experience, the experience difference is two years at each point in time.[13] This is more obviously highlighted when looking at each survey year in isolation, yet clearly is also true when many years of data are pooled (Devereux and Hart, 2010 and Grenet, 2013).

---

[13] In contrast, if identification derived from a geographic discontinuity affecting educational attainment, the 'treated' and 'untreated' individuals would be within the same school cohort and the RDD estimate *would* provide an estimate of the effect of one additional year of schooling and one year less potential experience, the 'net' return to a year of schooling.

Figure 3: Schematic representation of the School cohort-Log earnings relationship
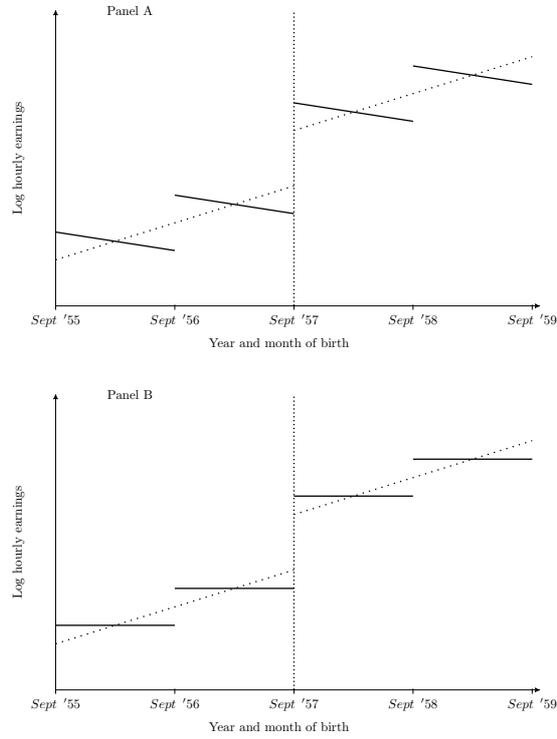
4. Methodology

4.1 Causal Identification

To identify the causal effect of education on wages we follow the standard approach of using instrumental variable (IV) methods (see for example, Machin*, et al.* 2011). In this approach a causal effect of education is estimated via the inclusion of a 1972 RoSLA dummy variable in the first stage education regression in a two-stage least squares framework. By arguing that the 1972 RoSLA event was an exogenous occurrence that increased levels of education randomly we can obtain a causal IV estimate as follows:

$$E_i = \alpha_1 + \beta_1 RoSLA72_i + x_i'\gamma_1 + \varepsilon_{1i} \tag{1}$$

$$Y_i = \alpha_2 + \beta_2 RoSLA72_i + x_i'\gamma_2 + \varepsilon_{2i} \tag{2}$$

15

where (1) and (2) are the reduced form equations for education E_i and log hourly wages Y_i [14] respectively, β_1 is the estimate of the 1972 RoSLA dummy on education (measured in years of schooling) whilst β_2 is the estimated effect of the RoSLA event on log hourly wages. If chosen to be included, x_i can be vector of additional control variables (such as age, gender and regional dummies) with parameter estimates γ. Finally, ε_{1i} and ε_{2i} are two normally distributed error terms with mean zero. The structural form for earnings Y_i is then given by:

$$Y_i = \alpha_3 + \beta_3 E_i + x_i' \gamma_3 + \varepsilon_{3i} \qquad (3)$$

where the IV estimate of β_3 in (3) is given by the ratio of the reduced form coefficients in (1) and (2), $\beta_3 = \beta_2 / \beta_1$.

Although the NESPD does not contain any measures of education, it is possible to obtain an estimated measure of $\hat{\beta}_3$ by estimating the reduced form equation (2) – in order to obtain $\hat{\beta}_2$ – and then using an out-of-sample estimation for β_1 (see Devereux and Hart, 2010). However, because so many studies have consistently estimated $\hat{\beta}_1$ we argue that not even out-of-sample estimates for β_1 are necessary. As highlighted in Table 1, the UK literature has generally agreed that the estimated effect of 1972 RoSLA on education was to raise average years of schooling by approximately one third of a year. Using LFS data, we estimate a similar value: $\hat{\beta}_1 = 0.321$. It is for this reason that we focus much of our attention on estimating a precise value of β_2 in the reduced form equation (2).

Assignment to the 'treatment' of the 1972 RoSLA is based on an individual's date of birth and although the date of birth is exogenous for individual i, it would make little sense to compare individuals who were born too many years apart due to cohort, generational and/or lifecycle effects which contaminate exogenous differences between such individuals. To some extent it is possible to control for such differences by including a series of age and cohort controls in the vector x_i and traditionally, β_1 and β_2 are estimated like this using a 2SLS framework (see *inter alia* Harmon and Walker, 1999). However, a disadvantage of 2SLS is in its linear and parametric restrictions. Given the large sample qualities of our data (particularly the NESPD), we are able to implement a regression discontinuity design in order to estimate β_1 and β_2. Following the approach outlined in Imbens and Lemieux (2008) we use non-parametric techniques where the effects of the 1972 RoSLA on education and earnings are estimated by local linear regression in a local region near the discontinuity.[15] It is worth noting at

---

[14] We first detrend log wages to remove any idiosyncratic time effects and perform all of our analysis on the detrended wage residuals.

[15] Grenet (2013) and Devereux and Hart (2010) implement a parametric approach to RDD in the LFS and NESPD respectively.

this stage that even if we were to try to implement an RDD controlling for age differences between cohorts and therefore try to deal this way with the 'experience penalty' issue outlined above, this can only be done in the context of a global polynomial approach implementation of the RDD.[16] Gelman and Imbens (2014) argue powerfully against such an approach to estimation of the RDD and recent work by Dolton and Sandi (2015) confirms the fragility of RDD estimates of the return to education in the UK when using the global polynomial approach. Moreover, for the reasons outlined above, even controlling for age is inadequate given that age is *no*t a good proxy for experience around the discontinuity.

Rather, in addition to mean difference comparisons, we use the local linear regression approach advocated by Imbens and Lemieux (2008) and Gelman and Imbens (2014). To estimate the values of β_2 we fit a linear regression function to observations within distance h on either side of the discontinuity point (1st September 1957)

$$\min\nolimits_{\alpha_{pre\,1957}:\beta_{pre\,1957}} \sum\nolimits_{i:c-h<X_i<c}\left(Y_i - \alpha_{pre\,1957} - \beta_{pre\,1957}\cdot(X_i - c)\right)^2 \quad (4)$$

and

$$\min\nolimits_{\alpha_{post\,1957}:\beta_{post\,1957}} \sum\nolimits_{i:c-h<X_i<c}\left(Y_i - \alpha_{post\,1957} - \beta_{post\,1957}\cdot(X_i - c)\right)^2 \qquad (5)$$

Where Y_i is log hourly wages, X_i is the number of years (or months) an individual is born before/after the 1st September 1957 discontinuity, the αs and βs are the regression intercepts and slope values computed for data in the region surrounding the discontinuity, *c,* with bandwidth, *h.* These values are calculated twice, once for the left hand side of *c* (born pre-September 1957) and once for the right hand side (born post-September 1957) of *c.* The intercept values at the discontinuity, $\mu_{pre\,1957}(c)$ and $\mu_{post\,1957}(c)$, can then be computed by:

$$\hat{\mu}_{pre\,1957}(c) = \hat{\alpha}_{pre\,1957} + \hat{\beta}_{pre\,1957}\cdot(c - c) = \hat{\alpha}_{pre\,1957} \qquad (6)$$

and

$$\hat{\mu}_{post\,1957}(c) = \hat{\alpha}_{post\,1957} + \hat{\beta}_{post\,1957}\cdot(c - c) = \hat{\alpha}_{post\,1957} \qquad (7)$$

Given these estimates we can then compute the treatment effect of 1972 RoSLA log hourly wages Y_i as follows[17]:

$$\hat{\beta}_2 = \hat{\alpha}_{post\,1957} - \hat{\alpha}_{pre\,1957} \qquad (8)$$

---

[16] Including covariates in the local linear regression setting is also specifically warned against in the Stata rd module used to implement the regression discontinuity estimator (see Nichols, 2011).
[17] The estimate of the impact of the 1972 RoSLA on education uses the analogous procedure.

Imbens and Lemieux (2008) suggest that from a practical point of view a simple rectangular kernel should be used to estimate β_2 and that robustness be verified by using different choice of bandwidths. We therefore report results for varying bandwidths but limit these to multiples of 12 months for LFS data $h = [12, 24, 36, 48, 60, 72]$ to avoid possible contamination by within-cohort month-of-birth effects. For NESPD the absence of precise month-of-birth information implies using yearly bandwidths so that $h = [1, 2, 3, 4, 5, 6]$. In addition, because of the above outlined extrapolation problems for local linear regression techniques in this context, and to be as transparent as possible, we complement our estimation of β_2 with mean-differences for a one year bandwidth around the discontinuity. Finally, we produce the parameters estimates of β_2 for each year of data.[18] This allows us assess to what extent the 1972 RoSLA education reform influenced labour market outcomes across the life-cycle.

4.2 Correcting for the experience loss

However, as mentioned in section 3, there is a strong suggestion that the 1972 RoSLA resulted in negative outcomes, potentially suggesting that additional schooling is associated with *negative* returns to education in the early part of the lifecycle. Clearly, if we assume that additional schooling does not lead to net negative human capital accumulation, this cannot be true and other confounding factors must play a role since the singular impact of additional schooling on wages should at worst be zero.

Although the empirical literature generally argues that at the point of discontinuity the treated and non-treated cohorts are alike in all characteristics except for the additional schooling, as outlined above this is not the case. At best RDD and other discontinuity estimates must rely on cohort averages around the discontinuity, since all RoSLA events are contaminated by within-school year effects if one moves 'too close' to the discontinuity. On top of such intra-year education differences, there are two further possible differences between the pre- and post-RoSLA cohorts: macro-economic circumstances that affected cohorts differentially and the loss of potential experience for the post-RoSLA cohort.

The relevant macro-economic circumstances relate to the 1973-1975 recession caused by the oil crisis and it conceivable that youths entering the labour market post 1972 RoSLA were faced with a tougher labour market when compared to the previous cohort. This may have caused lower average earnings in the post-RoSLA cohort as youth unemployment rose and wages fell. However, this argument is predicated on the assumption that wages for the pre-RoSLA

---

[18] To smooth our results we use rolling 3-year samples from the data, so rather than looking at each year from 1975 to 2012 singly, the 1975 sample contains 1975, 1976 and 1977, the 1976 sample contains 1976, 1977 and 1978 and so forth. We use inverse probability weighting as per Devereux and Hart (2010) within these 3-year bands to ensure that each individual in each 'year' carries the same weight. The NESPD sample size allows single year analysis which confirms the rolling 3-year results, see appendix A.

cohort would have been unaffected, possibly due to their already entrenched position in the labour market. Whilst we are unable to test this hypothesis directly, claimant count statistics for the period suggest that the proportion of individuals seeking work dropped substantially in 1973 and 1974 when compared to 1971 or 1972 (Denman and McDonald, 1995). The economic recession did not 'bite' into unemployment statistics until 1975 onwards and it thus appears that school leavers who were part of the first post 1972 RoSLA cohort (and left in 1974) would have faced a period of relatively good labour market conditions.[19] A pictorial representation of this is shown below in Figure 4.

Figure 4: UK Claimant Count in the 1970s (seasonally adjusted)
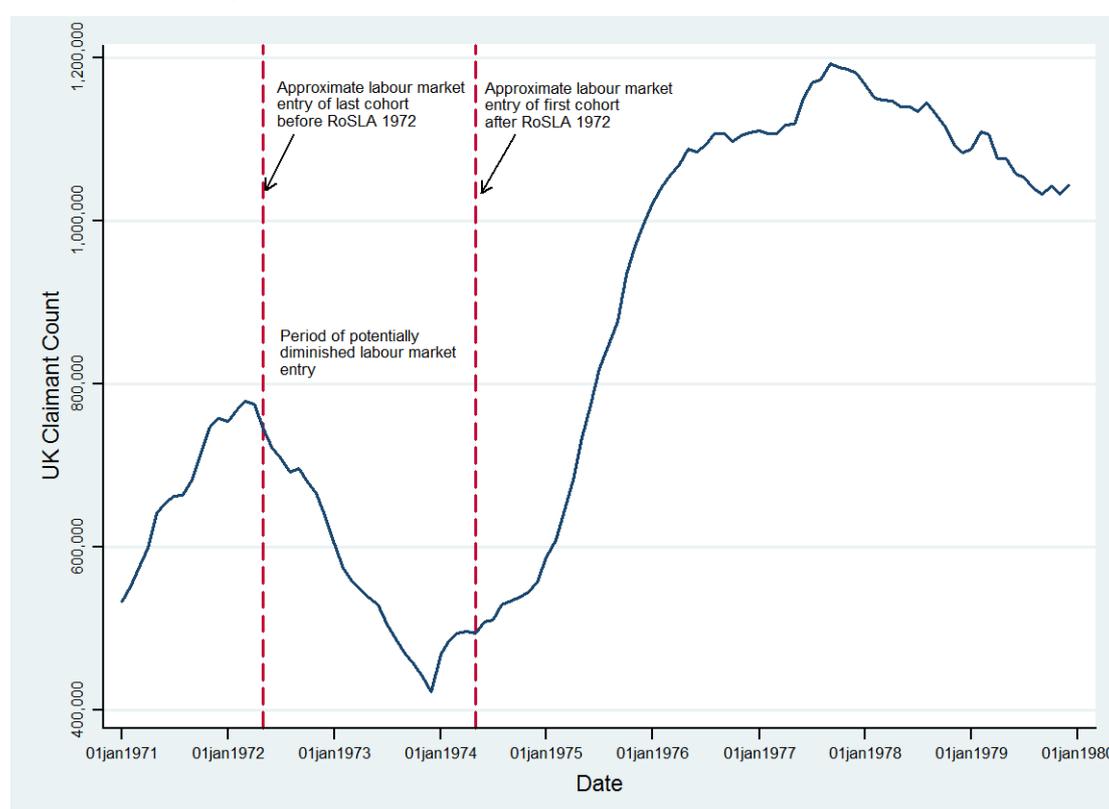
Source: Office of National Statistics

The loss of potential labour market experience is, therefore, likely to be a more significant factor in explaining our negative (descriptive) results. As previously noted by Rosenzweig and Wolpin (2000: 841) "...the additional year of schooling comes at the expense of an additional year of experience,..., what is identified is the difference between returns to schooling and experience". In other words, the experience loss incurred by an extra year of schooling, in addition to the experience difference caused by the raw age difference between the older pre-1972 RoSLA cohort and the younger post-1972 RoSLA cohort, are likely to be major

---

[19] In fact, the lack of minimum age school leavers during the summer of 1973 may have led to increased demand in the subsequent year, thereby artificially increasing wages for post-RoSLA cohorts.

factors in explaining any negative findings. This suggests that sophisticated regression discontinuity designs do not, and cannot, compare 'like with like' when the identification comes from a cohort based approach within a returns to schooling context.

It is notable that the early literature on returns to education was explicitly implementing Mincer's (1974) human capital earnings function, as derived from Becker's (1964) model of investment in human capital. One of the explicit assumptions of the original model is that time in school is independent of time in the work-place – that is, that schooling does not have a deleterious impact on experience or T′(s) =1, where T is the retirement age. Therefore the only opportunity costs of additional schooling are the foregone wage for that year plus any tuition costs. The 'experience opportunity cost' is not allowed for – in fact it is explicitly assumed not to exist. As such, the model aims to control for any experience differences by including a polynomial (traditionally quadratic) in experience in the estimated earnings function and estimate the growth rate of earnings with additional schooling.[20] Thus the estimated return in the model is intended to be the 'gross' return to education, having controlled out differences in experience. This type of argumentation is carefully highlighted in Rosenzweig and Wolpin (2000:p832-853) who demonstrate the importance of omitted experience in identifying Mincer functions when using natural experiments, including school leaving age reforms.

However, and crucially, the best practice RDD methodology does not allow differences in experience to be controlled and thus estimates the 'net' return to education, including the (likely negative) impact of the potential experience lost whilst in school for an additional year. The RDD methodology is therefore estimating a different parameter to the earlier returns to education literature: the estimand is different. This may explain some of the divergence in recent findings where RDD estimates (Devereux and Hart, 2010 and Grenet, 2013) return low values whilst more traditional 2SLS estimates, which attempt to completely control for the experience differences via age polynomials, return high values. Moreover, it is not clear that even local linear methods will be successful in attempting to estimate the 'net' effect of education because of the fact that potential experience is not a smooth within school-cohort function of age as detailed in section 3.

Using the alternative of one-year bandwidth mean differences, the estimates compare otherwise identical individuals with one more year of schooling but *two* fewer years of experience. In a cross-sectional setting there is relatively little that can be done to correct for any experience difference at the discontinuity since identification of the educational discontinuity comes from school cohort, which is perfectly correlated with potential experience. However, with panel data it is possible to attempt a correction. We do this by replacing current values of earnings with

---

[20] With additional assumptions (no tuition costs or psychic costs of education, and no taxes) we can recover the internal rate of return to schooling from the schooling parameter, though these necessary assumptions have been shown not to hold empirically (see Heckman *et al.,* 2006).

future values of earnings for the post-1972 RoSLA cohorts only. Assuming that earnings growth after schooling is only a function of experience and not schooling levels (i.e. there is additive separability in the Mincer framework) then replacing current earnings with future earnings $\left(y_{it_{post1957}} = y_{it+n_{post1957}} \right)$ *should* correct any experience loss present at the identification discontinuity.

In any comparison of successive cohorts, there will always be a 'potential experience' difference of one year given that cohort *t+1* will at each point in time be one year younger than cohort *t*. Therefore, if we wish to compare individuals with exactly the same levels of 'potential experience' then we would need to set $n = 1$ i.e. backshift all post-reform cohorts' wages by one year. With respect to RoSLA, 'always-takers' are not impacted by the reform and therefore this backshift of just one year would be sufficient to equalise their potential experience with the cohort before. However, for the complier group for the RoSLA reform, this would not be sufficient since they are two years less experienced than the minimum school leaving age leavers from the cohorts before – thus $n = 2$ is the correct backshift for them. Given that only one third of the cohort were bound by the reform the correct backshift is two years for one-third of the cohort but only one year for two-thirds of the cohort. Though we have the panel of earnings observations for each individual, we are not able to identify compliers and so we proceed by presenting the results of backshifting all post-reform cohort members' wages by one year and two years. Given the size of the complier group, the implied correct aggregate backshift to equalise potential experience would be $n = 1.33$. Therefore the correct estimate will be between the values we find for the $n = 1$ and $n = 2$ backshifts – this will be the correct 'gross' estimate akin to what the earlier 2SLS literature attempted to estimate.

If we follow the argument that we wish – in the spirit of the more recent RDD approach – to compare individuals with one additional year of schooling but one less year of experience, i.e. the 'net' impact of the RoSLA, then we would still wish to backshift the compliers, though in this case only by one year. At present, a one-year mean differences estimate of the RoSLA effect will compare the earnings of the pre-reform cohort with compliers who have one additional year of schooling but *two* fewer years of experience. Hence to get back to the 'net' impact of the reform we need to backshift one year. This equalizes age for all members of the post-reform cohort with the pre-reform cohort and also equalizes experience for the always-takers – thus assuming no reform effect for the always-takers, the bias implicit from the age/experience difference between these cohorts will be eliminated. For the compliers it equalizes age but now compares compliers with individuals in the previous cohort who have one less year of schooling but one more year of experience – which is the right 'net' comparison.

5. Results

5.1 Pooled LFS and Pooled NESPD

Firstly we present estimates of the impact of the 1972 RoSLA on education and wages by pooling the LFS survey years and NESPD survey years. In part this is to ensure that we can reconcile our estimates with those of Grenet (2013) but also to set a reference point for when we explore lifecycle effects. Table 2 presents our local linear RDD estimates of the reduced form impact of the 1972 RoSLA on years of education and log hourly wages for various relevant bandwidths where we limit the data span such that it begins in 1993 (to mimic Grenet, 2013).[21] Table 3 presents our RDD estimates of the reduced form impact of the 1972 RoSLA when the data span is increased to 1975-2012 (only available for NESPD) and also include a placebo test where the discontinuity year is set to 1964/65.[22]

As can be seen, we find an inconsistent picture of results from the RDD estimates on pooled data, with estimates clearly different depending on the data used. For the LFS, we find a positive, and generally significant, impact of approximately 2% on wages from the education reform.[23] The impact on years of education in the LFS is also statistically significant at around 0.33 years – suggesting that approximately one-third of the men in the immediate post-RoSLA cohorts were bound by the reform. These estimates, using a slightly longer dataset than that used by Grenet (2013), replicate the more recent findings in the UK literature: that the reduced form impact of 1972 RoSLA – averaged across almost two decades of the lifecycle – is approximately a significant 2% increase in hourly wages. As it is estimated that just under one third of the men in these cohorts were in the complier group for this reform, the implied local average treatment effect estimate of the return to an additional year of education is approximately 6%, which is a far cry from earlier estimates of 15-20%.

However, turning to the corresponding NESPD estimates, whose data range and variable setup was designed as to mimic the LFS estimates as closely as possible, we see that the estimated return to the 1972 RoSLA is 0% across the entire bandwidth range.[24] Thanks to the large sample size of the NESPD these estimates are precisely estimated. In seeking to explain

---

[21] To match the majority of the literature, we concentrate on impacts for men, however results for women are available from the authors on request.

[22] We move the policy forward 7 years for this placebo test such that none of the bandwidth estimates overlap with the actual RoSLA policy cohorts. Moving the policy 7 years backwards gives similar results.

[23] When choosing a bandwidth it should be noted that the narrower bandwidths for the LFS data could be in danger of distorting the estimate by including age within school-year effects and are therefore likely to be somewhat upward biased. This is because there is a danger that lower bandwidths give too much weight to a comparison of August born and September born children who are not directly comparable because of within-year birth effects. We therefore favour the higher bandwidths 48, 60 and 72 for the LFS.

[24] In contrast to the LFS, for the NESPD we favour the lowest bandwidth estimate. This is because the large sample size of the NESPD allows us to get 'closer' to the discontinuity without losing precision through higher standard errors. In addition, the lack of month of birth information within years does not force the estimator to accommodate complex within-year variations. As such, our preferred estimate is the simple means-comparison of bandwidth 1.

such contrasting results robustness tests suggested that our LFS estimates are relatively unstable, especially within the context of proxy respondents (who make up approximately one-third of the LFS respondents). Table B1 in the appendix highlights the impact of the removal of proxy respondents in the LFS; as a result, estimates drop down to approximately 1% and become statistically insignificantly different from zero, and are thus congruent with the NESPD's 0% estimates. Our initial results thus suggest that compared to the estimates in the literature that use this approach, a further downward revision is needed.

Table 2: RDD Estimates of the effect of the 1972 RoSLA on log hourly wages and years of schooling, various bandwidths: pooled LFS data and pooled NESPD data; limited data range

Bandwidth (months)		LFS 1993-2011 Log hourly wage	Years of schooling	N	Bandwidth (years)		NESPD 1993-2012 Log hourly wage	N
12	*Coeff.*	0.024	0.392∗∗∗	10,380	1	*Coeff.*	-0.004	64,550
	Std. Err.	(0.020)	(0.102)			*Std. Err.*	(0.006)	
24	*Coeff.*	0.027§	0.427∗∗∗	20,715	2	*Coeff.*	0.005	129,945
	Std. Err.	(0.014)	(0.070)			*Std. Err.*	(0.011)	
36	*Coeff.*	0.020§	0.379∗∗∗	30,875	3	*Coeff.*	-0.007	194,567
	Std. Err.	(0.012)	(0.057)			*Std. Err.*	(0.008)	
48	*Coeff.*	0.023*	0.349∗∗∗	41,324	4	*Coeff.*	-0.007	261,915
	Std. Err.	(0.010)	(0.049)			*Std. Err.*	(0.007)	
60	*Coeff.*	0.018*	0.321∗∗∗	52,158	5	*Coeff.*	0.000	330,058
	Std. Err.	(0.009)	(0.044)			*Std. Err.*	(0.006)	
72	*Coeff.*	0.014§	0.286∗∗∗	62,716	6	*Coeff.*	0.001	395,875
	Std. Err.	(0.008)	(0.040)			*Std. Err.*	(0.005)	
Average		*0.019*	*0.359*				*-0.002*	

§ $p<0.10$ * $p<0.05$, ** $p<0.01$, ***$p<0.001$. Estimates based on local linear regression using rectangular kernel. Bandwidth 1 for NESPD data cannot be estimated using local linear regression since only one data point exists. We complemented this by using mean differences for bandwidth 1. LFS: 1993q1-2011q4, NESPD: 1993-2012. Men only.

Table 3: RDD Estimates of the effect of the 1972 RoSLA on log hourly wages and years of schooling, various bandwidths: pooled NESPD data; whole data range, including placebo result

Bandwidth (years)		NESPD 1975-2012 Log hourly wage	N	Bandwidth (years)		NESPD 1975-2012 (placebo) Log hourly wage	N
1	*Coeff.*	-0.020∗∗∗	113,457	1	*Coeff.*	-0.004	109,323
	Std. Err.	(0.005)			*Std. Err.*	(0.003)	
2	*Coeff.*	-0.003	227,855	2	*Coeff.*	-0.001	216,758
	Std. Err.	(0.009)			*Std. Err.*	(0.008)	
3	*Coeff.*	-0.012	339,227	3	*Coeff.*	0.001	324,311
	Std. Err.	(0.006)			*Std. Err.*	(0.006)	
4	*Coeff.*	-0.012*	453,323	4	*Coeff.*	0.005	429,773
	Std. Err.	(0.005)			*Std. Err.*	(0.005)	
5	*Coeff.*	-0.005	567,115	5	*Coeff.*	0.004	533,047
	Std. Err.	(0.005)			*Std. Err.*	(0.004)	
6	*Coeff.*	-0.004	675,370	6	*Coeff.*	0.006	631,920
	Std. Err.	(0.004)			*Std. Err.*	(0.004)	
Average		*-0.010*				*0.001*	

* $p<0.05$, ** $p<0.01$, ***$p<0.001$. Estimates based on local linear regression using rectangular kernel. Bandwidth 1 for NESPD data cannot be estimated using local linear regression since only one data point exists. We complemented this by using mean differences for bandwidth 1, NESPD: 1975-2012. Men only.

Extending the available data range beyond what is measurable in the LFS and including the years 1975 to 1992 in our analysis of the NESPD (see Table 3), we observe an additional downward change in our estimated coefficients. The estimated reduced form impact on wages from the 1972 RoSLA is now approximately -2% for our preferred one-year bandwidth mean difference. Extrapolating such results would suggest an implied local average treatment effect of approximately -6% for an additional year of education, an estimate the kind of which has never been seen before in any related literature.

Such results clearly raise serious issues in relation to the estimation procedure and the quality of the data. However, several important findings can be taken from Tables 2 and 3. The first is that data quality matters and in our case, we favour the inherent administrative quality of the NESPD compared to the survey responses from the LFS. The second is that any difference in the returns to education between the two data ranges suggests the presence of lifecycle effects that has previously been masked in prior empirical studies (*inter alia* Grenet, 2013; Devereux and Hart, 2010 and Oreopolous, 2006). The change from a reduced form estimate of 0% for 1993-2012 to between -1% and -2% for 1975-2012 in the NESPD suggests that the return to education is not constant over time. Finally, the negative estimate is driven by the aforementioned experience factor as we do not believe that the gross return to education can be less than zero. The higher bandwidth local linear estimates which attempt to take account of experience differences and provide a net estimate of the return suggest between -0.5% and -1.2% return, though generally not significant. Thus this empirical evidence suggest that, ultimately, the net impact of increasing the school leaving from 15 to 16 in 1972 was at best zero and possibly to the *detriment* of those affected by it. The right hand panel of Table 3 contains the results from a falsification exercise in which the impacts of a non-existent RoSLA in 1964/65 were estimated. As we would expect, there are no significant effects of this placebo RoSLA – all estimates are close to zero but with standard errors comparable to the left-side of the table.

5.2 Lifecycle Results
We now turn to exploring the divergence in our pooled results by estimating the reduced form effect of 1972 RoSLA at each point of the lifecycle – using the NESPD from 1975 to 2012 corresponding to the approximate ages of 18 to 54 for the relevant cohorts born just before and just after September 1957.

Table 4 presents estimates of the impact of RoSLA using mean differences based on one-year bandwidths. The NESPD is large enough to support analysis based on one-year bandwidths, which means that these results are not contaminated by other factors that potentially differ between cohorts as we move away from the discontinuity. Moreover, using mean differences within these one-year bandwidths allows us to get as close to the discontinuity as we

can without the danger that age-within-cohort effects will distort the estimated difference between the cohorts.[25]

Table 4: Estimates of the Reduced Form impact of 1972 RoSLA on log hourly wages at different points in the life-cycle, NESPD, one year mean differences

Year	NESPD			NESPD experience adjusted (n = 1)			NESPD experience adjusted (n = 2)		
	Coeff	S.E.	N	Coeff	S.E.	N	Coeff	S.E.	N
1975	-0.162***	(0.009)	6208	0.015	(0.009)	5445	0.160***	(0.009)	5326
1976	-0.141***	(0.008)	6774	0.021*	(0.008)	5927	0.121***	(0.008)	5806
1977	-0.113***	(0.008)	7077	0.007	(0.008)	6198	0.073***	(0.008)	6126
1978	-0.069***	(0.007)	7340	0.010	(0.008)	6521	0.040***	(0.008)	6517
1979	-0.046***	(0.007)	7738	0.001	(0.007)	7026	0.033***	(0.008)	7006
1980	-0.042***	(0.007)	8187	0.000	(0.008)	7519	0.034***	(0.008)	7433
1981	-0.036***	(0.008)	8432	0.011	(0.008)	7772	0.048***	(0.008)	7599
1982	-0.037***	(0.008)	8552	0.016*	(0.008)	7810	0.046***	(0.009)	7706
1983	-0.034***	(0.008)	8443	0.020*	(0.009)	7757	0.041***	(0.009)	7675
1984	-0.032***	(0.009)	8550	0.016	(0.009)	7907	0.042***	(0.010)	7862
1985	-0.029***	(0.009)	8648	0.017	(0.009)	8092	0.041***	(0.010)	7977
1986	-0.025**	(0.010)	8944	0.018	(0.010)	9359	0.047***	(0.010)	8261
1987	-0.021*	(0.010)	9061	0.023*	(0.010)	8495	0.042***	(0.011)	8347
1988	-0.013	(0.010)	9268	0.026*	(0.011)	8667	0.041***	(0.011)	8518
1989	-0.015	(0.011)	9283	0.022*	(0.011)	8695	0.031**	(0.011)	8486
1990	-0.009	(0.011)	9316	0.020	(0.011)	8692	0.039***	(0.012)	8531
1991	-0.009	(0.011)	9173	0.028*	(0.011)	8599	0.046***	(0.012)	8541
1992	0.006	(0.011)	9139	0.035**	(0.012)	8687	0.050***	(0.012)	8677
1993	-0.006	(0.011)	9330	0.021	(0.012)	8919	0.033**	(0.012)	8769
1994	-0.003	(0.012)	9643	0.027*	(0.012)	9089	0.032**	(0.012)	8890
1995	-0.001	(0.012)	9704	0.028*	(0.012)	9028	0.028*	(0.012)	8821
1996	0.005	(0.012)	9760	0.032**	(0.012)	9029	0.028*	(0.012)	8912
1997	-0.003	(0.011)	9764	0.017	(0.012)	9086	0.020	(0.012)	8865
1998	-0.014	(0.011)	9953	0.008	(0.012)	9160	0.007	(0.012)	8978
1999	-0.014	(0.011)	9878	0.013	(0.012)	9086	0.009	(0.012)	8932
2000	-0.007	(0.011)	9866	0.010	(0.012)	9105	0.007	(0.012)	8996
2001	-0.006	(0.011)	9979	0.014	(0.012)	9272	0.023	(0.012)	9140
2002	-0.017	(0.011)	10214	0.013	(0.012)	9456	0.014	(0.012)	9292
2003	-0.018	(0.011)	10472	0.010	(0.012)	9622	0.018	(0.012)	9145
2004	-0.009	(0.011)	10559	0.029*	(0.012)	9377	0.024*	(0.012)	8864
2005	-0.006	(0.011)	9863	0.014	(0.013)	8689	0.024	(0.013)	8203
2006	-0.007	(0.012)	8995	0.016	(0.013)	7886	0.026*	(0.013)	7668
2007	-0.006	(0.012)	8755	0.010	(0.013)	7932	0.024	(0.013)	7736
2008	-0.005	(0.012)	9129	0.019	(0.012)	9305	0.020	(0.013)	8032
2009	0.008	(0.012)	9748	0.028*	(0.012)	8830	0.042**	(0.014)	7274
2010	0.011	(0.012)	9543	0.052***	(0.013)	7308	0.055***	(0.016)	5814
2011	0.010	(0.014)	6376	0.036*	(0.017)	4425	-		
2012	-0.007	(0.018)	3071	-	-		-		
Average (IPW)	-0.020***	(0.005)	113,457	0.022***	(0.005)	102,741	0.038***	(0.005)	99,368

Standard errors in parenthesis. * p<0.05, ** p<0.010, *** p<0.001 Results based on rolling three year data bands except for last two years. Estimates from 1 year mean differences

Results in the left-hand panel of Table 4 confirm our previous descriptive findings; specifically that the NESPD suggests that the 1972 RoSLA was associated with substantially

---

[25] To reduce the year-by-year noise levels we decide to smooth our estimates by estimating the effect of 1972 RoSLA on rolling 3-year data bands (see footnote 18), though estimates are robust to using single-year data bands (see appendix A for a pictorial representation of various specification checks).

lower average earnings for those who were affected by it. The negative effect of additional schooling lasted until 1986 when these men were just under the age of 30. Thereafter there are no significant differences in earnings between the two cohorts, though the point estimates are negative for the majority of years. It is interesting to observe that the pooled estimated return for men (in Table 3) apparently masks significant heterogeneity in the return to education – suggesting that our initial assumption of heterogeneity over the lifecycle is correct. Appendix Table A1 replicates Table 4 but for the local linear regressions with a 5-year bandwidth.[26] As discussed in section 3, the local linear RDD approach attempts to control (albeit imperfectly) for the cohort-to-cohort differences in potential experience that naturally affect successive cohorts, thus we expect the negative impacts in the initial years to be smaller – and this is what we find. However, the pattern of lifecycle heterogeneity is similar to what we observe in Table 4 – in particular the large negative impact in the early years of the career. We test the robustness of our findings by estimating the year-by-year results using a number of different specifications. Firstly we estimate using local linear regressions either side of the discontinuity for bandwidths of 2, 3, 4 and 6 years in addition to the 5 year bandwidth results. Appendix Table A2 contains a pictorial representation of these results highlighting the consistency of the pattern of point estimates attained. Secondly, we replicate this robustness table for all estimates using single year data bands, see Table A3. Again a clear, consistent pattern is found for the point estimates, confirming the estimates from the one-year mean differences approach.

Adjusting for the experience penalty

Results from the right hand columns of Tables 4 show the impact of backshifting individual wages to adjust for the 'experience penalty'. In stark contrast to results from first column, these adjustments see the reduced form impact of 1972 RoSLA become strongly positive at all points over the lifecycle. The highly negative effects in the early part of the lifecycle are replaced by positive coefficients which are in some cases significant. Moreover, there appears much less evidence of the strong lifecycle effects exhibited in the left panel of the table. Men receive an hourly earnings premium of between 1% and 2% in many parts of the lifecycle, with slightly higher estimates where significant in the latter parts. There is a suggestion then from Table 4 that the true returns to education are more homogenous than originally suggested by a naïve analysis.

In light of the discussion in sections 3 and 4, we believe that the naïve mean difference estimates present a downwardly biased estimate of the 'net' return to an additional year of schooling. To truly compare 'like with like' i.e. those with the same labour market experience, we need to backshift the post-reform cohort's wages by two years for the compliers and one year for

---

[26] The point estimates using bandwidth 4 or 6 are very similar and the consistent pattern of results across bandwidths is reflected in Appendix Tables A2 and A3.

the always takers. This would give an estimate of the 'gross' return, akin to early IV literature. As we cannot identify the compliers, we have presented the results when backshifting one year and two years in Tables 4 and 5, the latter table shows the pooled estimates for the backshifting exercise. Backshifting one year equalises the age of the pre- and post-reform cohorts at each point in time, and equalises experience for the always takers, with compliers still one year less experienced than the minimum age leavers in the pre-reform cohort. Backshifting two years equalises experience for the compliers but also 'over equalises' age for all members of the cohort and experience for the always takers. As such, the n = 2 backshift estimates are clearly over-estimates of the true 'gross' return. The unbiased estimate of the 'gross' return lies somewhere between the n = 1 and the n = 2 estimate, closer to the n = 1 figure (n = 1.33 would be approximately correct for the mean difference estimates). Looking at Table 5 and using the one-year bandwidth (mean difference) estimate we could conjecture that this would be approximately 2.75% as the estimate of the 'gross' reduced form effect of RoSLA. This would imply a 'gross' LATE effect of a year of schooling to be 8.25%. In contrast, since the local linear regressions already attempt to account for the inevitable one year cohort-to-cohort difference in potential experience (however inadequately) this means that conceptually the backshift of n = 1 in this case should correspond to n = 2 in the mean difference estimates. The correct 'gross' estimate from the local linear estimates would be found at the n = 0.33 point between the average n = 0 reduced form effect (for bandwidth 5 this is -0.5%) and the n = 1 reduced form effect (3.80%). This would be approximately 0.93%, implying a 'gross' LATE of 2.79%.

Alternatively, from the mean differences, just using the n = 1 backshift gives us the correct 'net' estimate for the compliers as we are now comparing those with one additional year of schooling and one less year of experience – with the experience of always-takers equalized. The one-year bandwidth (mean difference) estimate of 2.20% implies a 'net' LATE effect of a year of schooling to be 7.33%. The bandwidth 5 local linear estimate of the 'net' LATE would be -1.5% though not significantly different from zero.

Table 5: RDD Estimates of the effect of the 1972 RoSLA on log hourly wages and years of schooling, various bandwidths: pooled NESPD data; whole data range, back shifted

Bandwidth (years)		NESPD 1975-2012 (back shift $n = 0$)		NESPD 1975-2012 (back shift $n = 1$)		NESPD 1975-2012 (back shift $n = 2$)	
		Log hourly wage	N	Log hourly wage	N	Log hourly wage	N
1	*Coeff.*	-0.020***	113,457	0.022***	102,741	0.038***	99,368
	Std. Err.	(0.005)		(0.005)		(0.005)	
2	*Coeff.*	-0.003	227,855	0.039***	206,457	0.055***	199,615
	Std. Err.	(0.009)		(0.009)		(0.009)	
3	*Coeff.*	-0.012	339,227	0.031***	306,562	0.048***	296,004
	Std. Err.	(0.006)		(0.006)		(0.007)	
4	*Coeff.*	-0.012*	453,323	0.032***	409,382	0.049***	394,958
	Std. Err.	(0.005)		(0.005)		(0.005)	
5	*Coeff.*	-0.005	567,115	0.038***	511,707	0.056***	493,416
	Std. Err.	(0.005)		(0.005)		(0.005)	
6	*Coeff.*	-0.004	675,370	0.039***	608,723	0.057***	586,783
	Std. Err.	(0.004)		(0.004)		(0.004)	
Average		*-0.010*		*0.033*		*0.050*	

* $p<0.05$, ** $p<0.01$, *** $p<0.001$. Estimates based on local linear regression using rectangular kernel. Bandwidth 1 for NESPD data cannot be estimated using local linear regression since only one data point exists. We complemented this by using mean differences for bandwidth 1, NESPD: 1975-2012. Men only.

6. Discussion and conclusions

In this paper we have re-estimated the impact of the 1972 RoSLA in England and Wales to derive an estimate of the causal effect of education on wages for the compliers affected by the reform. This is the first study to use high quality administrative data to study the impact of this reform. Of particular interest was whether (causal) returns vary over the lifecycle and preliminary results suggested that this *is* the case. Using 38 years of NESPD data we exploit the exogenous variation induced by the 1972 school leaving age reform and using our preferred mean differences specification found an average reduced form effect of -1 to -2 percentage points over the lifecycle (which implies a -3 to -6% LATE effect). The local linear approach finds a reduced form impact of 0% to -1.2%, implying a LATE of between 0% and -3%. These results are even lower than recent findings of Grenet (2013) and Devereux and Hart (2010) who have both suggested that the returns to education are considerably lower than previously thought.

Further analysis of year-by-year returns suggests that these negative returns are not homogenous over the lifecycle: we found significant negative effects in the early part of work-life and non-significant effects in the later part. In rationalising such results we argued that even the most sophisticated RDD designs, where the evaluation of 1972 RoSLA takes places right at the boundary of the discontinuity, cannot overcome the problem of the experience discontinuity induced by a cohort identification approach and the fact that unlike age, potential experience is not a smooth function within cohort, it is flat within cohort. Thus we contend that the local linear methodology is not adequate and in addition that the mean difference negative estimates should be considered downward biased estimates of the 'net' effect of the additional education.

We attempted to correct for the experience bias induced by RDD methods and conclude that a) the corrected returns to education are indeed positive, and b) the variation in the returns to education over the lifecycle is less extreme than initially estimated. We estimated that the net return to an extra year of schooling is approximately 7% and that returns are actually relatively homogenous over time.

In order to better contextualise our results we computed the estimated effect of the 1972 RoSLA on annual wages in 2013 prices. These data are taken from gross weekly wages which are multiplied by 52 and then adjusted by the estimated coefficients in Table 4. Although these figures represent 'back-of-the-envelope' calculations they nonetheless provide a useful contextual narrative in which to view the effect of an additional year of schooling under the various RDD identification regimes.

Table 6: Estimated effect of 1972 RoSLA on total earnings 1975-2010 (2013 prices)

Year	Average pay 1956/1957 cohort	Estimated average pay 1957/58 cohort	Estimated average pay 1957/58 cohort, experience adjusted (n = 1)	Estimated average pay 1957/58 cohort, experience adjusted (n = 2)
Total earnings 1975-2010	£860,000	£840,000	£880,000	£890,000
Effect of RoSLA 1972		-£15,000	£19,000	£29,000
Effect of 1 year of extra schooling		-£45,000	£57,000	£87,000

Estimates are based on hourly pay from NESPD data assuming a 37 hour working week and 52 weeks pay per year. Rounded to nearest thousand £

Table 6 shows that the total lifetime earnings for an individual in the pre-1972 RoSLA cohort, up to 2010, was approximately £860,000. According to our estimates in Table 4, the reduced form effect of the 1972 RoSLA is thus estimated to be approximately £-15,000. Dividing this value by 0.33 years of schooling suggests a LATE effect of £-45,000 over a 35-year work span. Adjusting these estimate for the incurred experience loss gives reduced form values of £19,000 (n=1) and £29,000 (n=2), with the 'true' gross return likely to be somewhere near £22,000. Taking the n=1 backshift value as a corrected estimate of the 'net' effect of the additional schooling, this would then suggest a LATE estimate of approximately £57,000 to one year of additional schooling over a period of 35 years.

Our results are important for a variety of reason. First, we argue that recent estimates of the return to education in the UK underestimate the 'true' return to an additional year of schooling due to their failure to adequately compensate for the discontinuity in experience between the cohorts. This 'bias' will similarly affect any estimates of returns to education that exploit compulsory school leaving age reforms and estimate returns using RDD to compare successive school cohorts.

Arguably the single year of inevitable experience loss due to an additional year of schooling is part of the opportunity cost and should be netted into the estimates of the return to education

i.e. the parameter that RDD attempts to estimate is the correct policy parameter. However, it may also be important to decompose the return into the 'pure' (positive) education effect and the necessary (negative) experience loss effect, recovering a parameter closer to the growth rate of earnings with education that the 1990s literature sought to estimate. Both of these components may change independently over time and it is important for wider policy to know the return to the skills associated with education, abstracted from the experience costs of gaining more education. Moreover, should a reform be considered that increases education without reducing experience – for example by lowering the school start age – it is important to be able to estimate the likely return to this sort of pure education change. Therefore, it is important in any study to be clear about whether it is the 'gross' or 'net' return to education parameter that is being estimated and in the case of the latter, whether the estimation technique will indeed actually provide an unbiased estimate of this.

Second, our results shed light on the variability of returns to education over the lifecycle. Significant differences in returns can be estimated when examining different portions of the lifecycle, especially when using unadjusted RDD methods. Returns to education studies are thus advised to consider the possible impact of their data range on their final estimates; and if not empirically at least contextually.

Third, the evidence presented here provides an important contribution to the recent 'revisionist' returns literature and the older returns literature in the UK. When using unadjusted RDD estimation methods to mimic the recent literature, our results suggest that the wage effect of the 1972 RoSLA were much worse than previously estimated. For the first time in the related literature we would conclude that an entire cohort of people would arguably have been better off (in terms of wages) had they not received additional education. We have however also shown that such estimates are likely to be downward biased and proposed reasonable adjustments that recover a LATE of the additional year of schooling to be approximately 7% estimated over the duration of the working life.

References

Angrist, Joshua D., and Alan B. Krueger. 1991. Does compulsory school attendance affect schooling and earnings? *Quarterly Journal of Economics* 106, no. 4:979-1014.

Ashenfelter, O., and Alan B. Krueger. 1994. Estimates of the economic return to schooling from a new sample of twins. *American Economic Review* 84, no. 5:1157–1174.

Becker, Gary S. 1964. *Human Capital: A Theoretical & Empirical Analysis with Special Reference to Education*. New York: Columbia University Press.

Bhuller, Manudeep, Magne Mogstad, and Kjell G. Salvanes. 2011. Life-cycle bias and the returns to schooling in current and lifetime earnings. Discussion Paper no. 5788, IZA, Bonn.

Buscha, Franz, and Matt Dickson. 2012. The raising of the school leaving age: returns in later life. *Economic Letters* 117:389-393.

Card, David. 1995. Using geographic variation in college proximity to estimate the return to schooling. In *Aspects of labour market behaviour: Essays in honour of John Vanderkamp*, ed. Louis N. Christofides, E. Kenneth Grant, and Robert Swidinsky. Toronto: University of Toronto Press.

Card, David. 1999. The causal effect of education on earnings. *Handbook of Labor Economics* 3, part 1: 1801–1863

Card, David. 2001. Estimating the return to schooling: progress on some persistent econometric problems. *Econometrica* 69, no. 5:1127–1160.

Chevalier, Arnaud, and Ian Walker. 2002. Further estimates of the returns to education in the UK. In *The Returns to Education Across Europe,* ed. Harmon, Colm, Ian Walker, and Niels Westergard-Nielsen. Cheltenham, England: Edward Elgar.

Clark, Damon, and Heather Royer. 2013. The effect of education on adult health and mortality: evidence from Britain. *American Economic Review* 103, no. 6:2087–2120.

Crawford, Claire, Lorraine Dearden, and Costas Meghir. 2010. When you are born matters: the impact of date of birth on educational outcomes in England. Discussion Paper no. W10/06, IFS, London.

Denman, James, and Paul McDonald. 1996. Unemployment statistics from 1881 to the present day. *Labour Market Trends* 104, no. 1: 5-18.

Devereux, Paul J., and Robert A. Hart. 2010. Forced to be rich? Returns to compulsory schooling in Britain. *The Economic Journal* 120:1345-1364.

Dickens, Richard. 2000. The evolution of individual male earnings in Great Britain: 1975-1995. *The Economic Journal* 110:27-49.

Dickson, Matt. 2013. The causal effect of education on wages revisited. *Oxford Bulletin of Economics and Statistics* 75, no. 4:477-498.

Dickson, Matt, and Sarah Smith. 2011. What determines the return to education? An extra year or a hurdle cleared? *Economics of Education Review* 30:1167-1176

Dolton, Peter, and Matteo Sandi. 2015. Returning to returns: revisiting the British education evidence. Unpublished manuscript, Department of Economics, University of Sussex.

Gelman, Andrew, and Guido W. Imbens. 2014. Why high-order polynomials should not be used in regression discontinuity designs. Working Paper no. 20405, National Bureau of Economic Research, Cambridge, Massachsetts.

Grenet, Julien. 2013. Is extending compulsory schooling alone enough to raise earnings? Evidence from French and British compulsory schooling laws. *The Scandinavian Journal of Economics* 115, no. 1:176-210.

Haider, Steven, and Gary Solon. 2006. Life-cycle variation in the association between current and lifetime earnings. *American Economic Review* 96, no. 4:1308–1320.

Harmon, Colm, Hessel Oosterbeek, and Ian Walker. 2003. The returns to education: microeconomics. *Journal of Economic Surveys* 17, no. 2:115–155.

Harmon, Colm, and Ian Walker. 1995. Estimates of the economic return to schooling for the United Kingdom. *American Economic Review* 85, no. 5:1278-1286.

Harmon, Colm, and Ian Walker. 1999. The marginal and average returns to schooling in the UK. *European Economic Review* 43, no. 4-6:879-887.

Heckman, James J., Lance J. Lochner, and Petra E. Todd. 2003. Fifty Years of Mincer Earnings Regressions. Working Paper no. 9732, National Bureau of Economic Research, Cambridge, Massachsetts.

Imbens, Guido W., and Thomas Lemieux. 2008. Regression discontinuity designs: a guide to practice. *Journal of Econometrics* 142:615-635.

Klein, Roger, and Francis Vella. 2009. Estimating the return to endogenous schooling decisions via conditional second moments. *Journal of Human Resources* 44, no. 4:1047-1065.

Lee, David S., and Thomas Lemieux. 2010. Regression discontinuity designs in economics. *Journal of Economic Literature* 48, no. 2:281–355.

Lemieux, Thomas. 2006. Increasing residual wage inequality: composition effects, noisy data, or rising demand for skill? *American Economic Review* 96, no. 3:461-498.

Machin, Stephen, Olivier Marie, and Suncica Vujic 2011. The crime reducing effect of education. *The Economics Journal* 121:463-484.

Milligan, Kevin, Enrico Moretti, and Philip Oreopoulos. 2004. Does education improve citizenship? Evidence from the United States and the United Kingdom. *Journal of Public Economics* 88, no. 9-10:1667-1695.

Mincer, Jacob. 1974. *Schooling, Experience and Earnings*. New York: Columbia University Press.

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

Nichols, Austin. 2011. rd 2.0: Revised Stata module for regression discontinuity estimation. http://ideas.repec.org/c/boc/bocode/s456888.html

Oreopoulos, Philip. 2006. Estimating average and local average treatment effects of education when compulsory schooling laws really matter. *American Economic Review* 96, no. 1:152–75.

Oreopoulos, Philip, and Uros Petronijevic. 2013. Making college worth it: a review of research on the returns to higher education. Working Paper no. 19053, National Bureau of Economic Research, Cambridge, Massachusetts.

Polachek, Soloman. 2007. Earnings over the lifecycle: the Mincer earnings function and its applications. Discussion Paper no. 3181, IZA, Bonn.

Rosenzwieg, Mark R. and Kenneth I. Wolpin. 2000. Natural "natural experiments" in economics. *Journal of Economic Literature* 38:827-874.

Siles, Mary. 2009. The causal effect of education on health: evidence from the United Kingdom. *Economics of Education Review* 28:122-128.

Siles, Mary. 2011. The effect of schooling on teenage childbearing: evidence using changes in compulsory education laws. *Journal of Population Economics* 24, no. 2:761-777.

Wilson, Tanya. 2012. Compulsory education and teenage motherhood. Unpublished Manuscript, Royal Holloway University of London.

Appendix A – NESPD robustness checks

Table A1: Estimates of the Reduced Form impact of 1972 RoSLA on log hourly wages at different points in the life-cycle, NESPD, RDD using local linear regression, bandwidth 5 years

Year	NESPD			NESPD experience adjusted (n=1)		
	Coeff	S.E.	N	Coeff	S.E.	N
1975	-0.113***	(0.009)	63314	0.080***	(0.009)	61667
1976	-0.101***	(0.008)	69193	0.065***	(0.008)	66706
1977	-0.068***	(0.007)	74060	0.056***	(0.007)	70688
1978	-0.013*	(0.007)	78691	0.065***	(0.007)	74522
1979	0.026***	(0.007)	83794	0.068***	(0.007)	79310
1980	0.018**	(0.007)	88588	0.051***	(0.007)	83676
1981	0.012	(0.007)	92677	0.055***	(0.007)	87347
1982	0.005	(0.007)	96271	0.053***	(0.008)	89695
1983	0.002	(0.008)	98922	0.056***	(0.008)	91893
1984	0.002	(0.008)	103608	0.048***	(0.008)	95621
1985	-0.001	(0.008)	108715	0.041***	(0.009)	100536
1986	0.001	(0.009)	116728	0.042***	(0.009)	107354
1987	0.003	(0.009)	123496	0.045***	(0.010)	113359
1988	0.008	(0.009)	130911	0.047***	(0.010)	119431
1989	0.013	(0.010)	135735	0.047***	(0.010)	123589
1990	0.016	(0.010)	138240	0.046***	(0.010)	126416
1991	0.018	(0.010)	139240	0.048***	(0.010)	128097
1992	0.024*	(0.010)	141547	0.052***	(0.011)	131988
1993	0.008	(0.011)	149219	0.036***	(0.011)	138209
1994	0.006	(0.011)	157796	0.034**	(0.011)	144455
1995	0.003	(0.011)	163381	0.033**	(0.011)	147090
1996	0.008	(0.011)	168696	0.030**	(0.011)	150803
1997	-0.001	(0.011)	173674	0.019	(0.011)	153976
1998	-0.011	(0.011)	179858	0.010	(0.011)	157717
1999	-0.012	(0.011)	183164	0.016	(0.011)	159610
2000	-0.003	(0.011)	187436	0.013	(0.011)	164158
2001	-0.006	(0.011)	195210	0.012	(0.011)	171779
2002	-0.020	(0.010)	206271	0.006	(0.011)	180379
2003	-0.018	(0.010)	217161	0.007	(0.011)	186905
2004	-0.005	(0.010)	224643	0.028**	(0.011)	183429
2005	-0.002	(0.011)	214725	0.022	(0.011)	172098
2006	-0.001	(0.011)	202389	0.033**	(0.012)	160795
2007	-0.001	(0.011)	203083	0.018	(0.012)	167779
2008	-0.005	(0.011)	217219	0.018	(0.011)	180676
2009	-0.002	(0.011)	235246	0.018	(0.011)	193617
2010	-0.004	(0.011)	236729	0.032**	(0.012)	143528
2011	-0.008	(0.013)	159237	0.024	(0.015)	79076
2012	-0.019	(0.017)	78084	-		
Average (IPW)	-0.005	-0.005	567,115	0.038***	0.005	511,707

Standard errors in parenthesis. * p<0.05, ** p<0.01, *** p<0.001. Results based on rolling three-year data bands except for the last two years. Estimates from local linear regressions with bandwidth 5 years.

Table A2: Regression Discontinuity Design Estimates of the Reduced Form impact of the 1972 RoSLA on log hourly wages at different points in the life-cycle; NESPD, three year data-bands. Pictorial representation of estimated coefficients.

Year	Mean differences	3-year rolling data bands with IPW				
		RD BW2	RD BW3	RD BW4	RD BW5	RD BW6
1975						
1976						
1977						
1978						
1979						
1980						
1981						
1982						
1983						
1984						
1985						
1986						
1987						
1988						
1989						
1990						
1991						
1992						
1993						
1994						
1995						
1996						
1997						
1998						
1999						
2000						
2001						
2002						
2003						
2004						
2005						
2006						
2007						
2008						
2009						
2010						
2011						
2012						

Legend:
- ■ negative significant (p<0.05) [red]
- ■ negative non-significant [pink]
- ■ positive significant (p<0.05) [black]
- ■ positive non-significant [grey]

Notes: Regression discontinuity estimates from local linear regression with rectangular kernel, various bandwidths.

Table A3: Regression Discontinuity Design Estimates of the Reduced Form impact of the 1972 RoSLA on log hourly wages at different points in the life-cycle; NESPD, one year data-bands. Pictorial representation of estimated coefficients.

Year	Mean differences	1-year data bands				
		RD BW2	RD BW3	RD BW4	RD BW5	RD BW6
1975						
1976						
1977						
1978						
1979						
1980						
1981						
1982						
1983						
1984						
1985						
1986						
1987						
1988						
1989						
1990						
1991						
1992						
1993						
1994						
1995						
1996						
1997						
1998						
1999						
2000						
2001						
2002						
2003						
2004						
2005						
2006						
2007						
2008						
2009						
2010						
2011						
2012						

- negative significant (p<0.05)
- negative non-significant
- positive significant (p<0.05)
- positive non-significant

Notes: Regression discontinuity estimates from local linear regression with rectangular kernel, various bandwidths.

Appendix B – LFS Robustness checks

Table B1: Regression Discontinuity Design Estimates of the Reduced Form impact of the 1972 RoSLA on log hourly wages and years of schooling, various bandwidths; pooled LFS data, 1993q1-2011q4. Men only, excluding proxy respondents

Bandwidth (mths)		LFS Log hourly wage	Years of schooling	N
12	*Coeff.*	-0.004	0.341***	7,017
	Std. Err.	(0.024)	(0.128)	
24	*Coeff.*	0.011	0.419***	14,015
	Std. Err.	(0.017)	(0.087)	
36	*Coeff.*	0.012	0. 396***	20,896
	Std. Err.	(0.014)	(0.071)	
48	*Coeff.*	0.014	0. 360***	27,907
	Std. Err.	(0.012)	(0.061)	
60	*Coeff.*	0.009	0. 320***	35,250
	Std. Err.	(0.011)	(0.054)	
72	*Coeff.*	0.001	0.258***	42,312
	Std. Err.	(0.010)	(0.049)	

* p<0.05, ** p<0.01, *** p<0.001. Estimates based on local linear regression using rectangular kernel.

Figure B1: Average log gross hourly earnings by survey year in the LFS

Men only, RPI adjusted, base = 2013
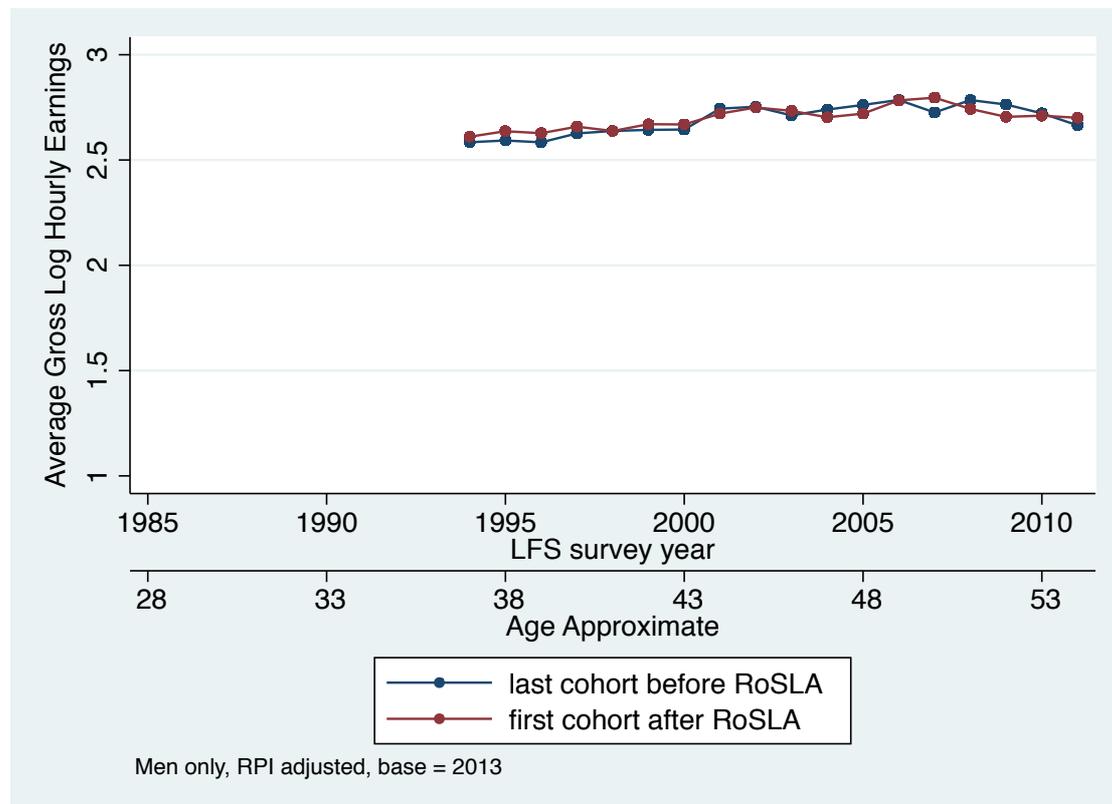
Appendix C – Identifying accurate school cohorts in the NESPD

School cohorts in the English school system do not align with year of birth but rather include all individuals in school year *t* who were born between 1st September in year *t-1* and 31st August in year *t*, thus both year of birth and month of birth information is required to assign individuals correctly to their school cohort. Year of birth is not usually available in the NESPD nor is month of birth ever provided to researchers. However, we were able to make use of information especially provided to us by the UK's Office for National Statistics (ONS) that allows us to assign people to the correct school cohort given their year and month of birth. Without the additional information provided to us by the ONS, we would have had to assign school cohort only on the basis of year of birth, which would immediately introduce measurement error. Previously in the literature using this data (see Devereux and Hart, 2010) constructed year of birth by using the information on the survey year and the age of the individual. For the years prior to 2004, the age variable in the NESPD referred to age on 1st January of the survey year. As such, (survey year – age – 1) will assign individuals to the correct year of birth. This is the situation for the vast majority of individuals in the data used by Devereux and Hart who examine the earlier 1947 RoSLA.

From 2004 onwards the age variable in NESPD refers to age at the time of the survey reference date in April. Constructing year of birth as (survey year – age – 1) for observations from 2004 onwards will inevitably introduce further measurement error depending on whether or not the individual's birthday is before or after the date that the survey is carried out. For those who have a birthday after the date of the survey the correct calculation would be (survey year – age at time of survey – 1). However, for those who have their birthday before the survey date, (survey year – age at time of survey – 1) will give the year of birth as one too early: if someone is born in March and we take their age in at the time of the survey in April and calculate (survey year – age at time of survey – 1), we will give them a year of birth as X-1 when their actual year of birth is X.

The NESPD survey refers to a week in April each year. Therefore, for the majority of a calendar year, individuals will not have had their birthday that year by the time of the survey in April (even if we assume April born individuals have had their birthday at the time of the survey, those born in the eight months from May to December inclusive will not have). Thus the calculation of year of birth as (survey year – age at time of survey – 1) will be correct for two-thirds of individuals (assuming an approximately uniform distribution of births across the year which is a reasonable assumption in the UK).

However, as month of birth is required to assign individuals to their correct school cohort – which is key here as the RoSLA affected the September 1957-August 1958 born cohort – and only (calculated) year of birth is available, estimation would need to proceed by excluding anyone with year of birth 1957 as it would not be known which school they should be assigned

to i.e. whether born January to August and so in the last cohort before the RoSLA school year (1956-57) or born September to December and so in the first cohort after RoSLA (1957-58). This would mean that the last 'cohort' before the reform would be those whose calculated year of birth is 1956 and the first 'cohort' after the reform would be those whose calculated year of birth is 1958.

Even if individual's year of birth can be accurately calculated from pre-2004 observations, calculating year of birth and then excluding those born in 1957 introduces substantial measurement error. Individuals born in the first four months of the calendar year 1958 (i.e. in the first cohort post-reform) will erroneously be assigned 1957 as their year of birth and will be excluded. Individuals born September to December 1958 i.e. in the *second* cohort post reform, will be included as part of the *first* cohort post reform and so would erroneously be included in any calculations comparing the first cohort after reform and the last cohort before. Similarly, those born from May to August 1956 i.e. in the *second* last cohort pre reform, will be included as part of the *last* cohort pre-reform and so would erroneously be included in any calculations comparing the first cohort after reform and the last cohort before.

Note that the alternative, calculating year of birth as (survey year – age) which will be correct for individuals born in the first four months of the year, creates an even more serious measurement error problem. Anyone born in May, June, July or August 1957, will be incorrectly assigned to birth year 1958 and on this basis would be considered as part of the first cohort *post* reform whereas they were actually part of the last cohort *pre* reform. To see this consider: an individual born in May 1957 and interviewed in April 1975, will be 17 at the interview date and so would have year of birth calculated as: $1975 - 17 = 1958$. This would be a severe measurement error problem for the key variable, as four birth months of *untreated* individuals would be included in the first cohort of the *treated* category.

Therefore, any attempt to assess the 1972 RoSLA using the NESPD data without the additional school cohort information that we obtained would be subject to a substantial degree of measurement error, which would be even more severe for individuals only observed in the panel post-2003.