Devereux, Paul J.; Fan, Wen

Working Paper
Earnings returns to the British education expansion

Working Paper Series, UCD Centre for Economic Research, No. 11/11

Provided in Cooperation with:
UCD School of Economics, University College Dublin (UCD)

Suggested Citation: Devereux, Paul J.; Fan, Wen (2011) : Earnings returns to the British education expansion, Working Paper Series, UCD Centre for Economic Research, No. 11/11, UCD Centre for Economic Research, Dublin

This Version is available at:
http://hdl.handle.net/10419/72246

Terms of use:
Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.
Earnings Returns to the British Education Expansion
Paul J Devereux and Wen Fan, University College Dublin

WP11/11
June 2011
We study the effects of the large expansion in British educational attainment that took place for cohorts born between 1970 and 1975. Using the Quarterly Labour Force Survey, we find that the expansion caused men to increase education by about a year on average and gain about 8% higher wages; women obtained a slightly greater increase in education and a similar increase in wages. Clearly, there was a sizeable gain from being born late enough to take advantage of the greater educational opportunities offered by the expansion. Treating the expansion as an exogenous increase in educational attainment, we obtain instrumental variables estimates of returns to schooling of about 6% for both men and women.

Keywords: return to education; higher education expansion

* The authors acknowledge the Economic and Social Data Service (ESDS) for access to the Quarterly Labour Force Survey data. Devereux and Fan gratefully acknowledge financial support from the Irish Research Council for the Humanities and Social Sciences. We thank participants at the CMPO conference on the returns to education for useful insights. We also thank Colm Harmon and Damon Clark for helpful comments.
1. Introduction

Between 1989 and 1994, there was a large increase in higher educational participation in the U.K. The scale of the increase is quite impressive and it occurred subsequent to a long period of stagnation. From 1970 to 1989, the proportion of individuals who went to college was about 15% for men and 13% for women and these numbers were fairly constant from year to year. However, by the mid-1990s, these percentages had become about 30% and 35% respectively (Walker and Zhu 2008). Between 1988 and 1996, participation in higher education in the U.K. increased by 93%; the equivalent figure for the U.S. was 15% (OECD 2007). By any standards, this is a sharp increase in educational attainment over a short period of time.

When we look by birth cohort, the large expansion in educational attainment occurred for cohorts born between 1970 and 1975. Our focus in this study is twofold: First, we explore how educational attainment changed throughout the educational distribution for these cohorts. Interestingly, there were large increases in educational attainment even for persons who did not attain A-levels or acquire a college degree. Second, we examine how the wages of affected cohorts were impacted by the changes in educational opportunities.

For post-1975 cohorts, there was once again a fairly stagnant period with little increase in educational attainment. This has changed for more recent cohorts as the British government set targets for increased higher education and there has been much discussion about the merits of large increases in 3rd-level participation. By studying the impacts of the previous education expansion, we hope to provide some insight on the likely long-run affects of recent policy.
The results are also likely to be of interest for other reasons. Machin (2007) and Blanden and Machin (2004) show that the expansion of UK higher education since the 1960s predominantly benefited children from high-income families. They speculate that this is a major reason for the reduction in intergenerational mobility in the UK over this period. However, in the absence of direct estimates of the benefits of the education expansion to those who took advantage of it, it is difficult to be certain of this link.

Other researchers have worried that the large increase in the supply of educated individuals may have led to over-education in the labour market with college graduates doing jobs previously carried out by non-degree holders. Consistent with this idea, a pure signalling model suggests that increasing the supply of highly educated persons in certain cohorts would cause potential employers to revise downwards their estimate of the average ability of graduates in these cohorts. Both these factors imply that there may be a negligible benefit to being in a high-education cohort. On the other hand, standard human capital models suggest that the skills learned in school and university should increase the earnings of these cohorts.\(^1\)

In this paper, we investigate the effect of the education expansion on educational attainment and subsequent earnings of the affected cohorts. In particular, we use data from the Quarterly Labour Force Survey (QLFS) to study the outcomes of persons born between 1958 and 1982. Treating the expansion as an exogenous increase in educational attainment, we also report instrumental variable estimates of the implied return to education for both men and women.

\(^1\) Chevalier et al. (2004) show that the UK evidence suggests that human capital models are more relevant than signalling ones.
The structure of this paper is as follows: In the next section, we review some relevant literature. In section 3, we describe the Education Expansion (henceforth, EE). Section 4 describes the data and provides some descriptive statistics and section 5 discusses the empirical strategy. We report the results in section 6 and discuss the estimates in the context of the literature in section 7. Section 8 concludes.

2. Literature Review

There are several literatures that are relevant to this paper. First, there are a series of papers that have studied whether the return to a college degree (or the quality of job undertaken by college graduates) has fallen as a result of the large increase in the supply of university graduates resulting from the expansion in higher education (Walker and Zhu 2008 provide a literature review). Most of these papers compare the difference in outcomes of degree and non-degree holders in both pre- and post-expansion cohorts. Our reading of the literature is that evidence for a change in the degree/non-degree difference across cohorts is weak. A fundamental problem with these papers is that the increase in higher education (henceforth, HE) implies that the composition of graduates relative to non-graduates must have changed across cohorts — graduates are a much less elite group now than they were pre-expansion.2

Walker and Zhu (2008) improve on this methodology by restricting their sample to persons who have at least two A-level qualifications (see Section 3 for a description of the structure of the UK education system). This implies that they drop persons who would not be admitted to university and compare outcomes of those who chose to attend with those who chose not to. They find that this difference is not lower for the later-born

---

2 Sabadash (2010) links longer term changes in the UK skill premium to supply and demand factors.
cohorts who were able to take advantage of the education expansion. Their approach is problematic if the composition of persons who hold 2+ A-levels changed during the expansion. Later in the paper, we show that, during the education expansion, there was a sizeable increase in the proportion of persons who attained 2 or more A-levels. Therefore, it is likely that the underlying characteristics of this group have changed as well.

There are other related literatures that we now briefly discuss. Much research has looked at the impacts of changes in compulsory schooling laws (Harmon and Walker 1995; Oreopoulos 2006; Devereux and Hart 2010) and of school-building programmes (Duflo 2001; Berlinski and Galiani 2007). There is also some research on the effect of college openings on outcomes of women (Currie and Moretti 2003). Some of these papers have used cross-state variation while others have relied solely on cohort-level variation. As we discuss further below, Ichino and Winter-Ebmer (2004) use cohort-level variation in educational attainment in Germany and Austria that resulted from World War II to assess the effect of education on later outcomes. Our paper differs from the literature as it is the first attempt we are aware of to use cohort-level variation to study the effect of the UK education expansion on outcomes. It might reasonably be expected that the impact of the education expansion would differ from that of changes in compulsory schooling laws, school-building programmes, and from educational variation resulting from war. Also, as mentioned above, it is of particular policy relevance given the general trend of governments encouraging large increases in educational attainment.
3. Institutional Background

The UK Education System

In the UK, the academic year runs from 1st September to 31st August. Children generally start school at age 5. In England and Wales, the school-leaving age was raised from 15 to 16 on 1st September, 1973, so the cohorts we study here have to stay in school until they turn 16 years-old. Progression at school beyond the minimum leaving age of 16 is based on a series of nationally assessed examinations. Until 1986, students at 16 had to take either the lower-level Certificates of Secondary Education (CSE) exams or the more academically demanding Ordinary Level (O-Level) exams (the top grade (grade 1) achieved on a CSE was considered equivalent to O-Level grade C). While most CSE students tended to leave school at the minimum age, students who took O-Levels were much more likely to stay further in school. In 1986, CSEs and O-Levels were replaced by General Certificates of Secondary Education (GCSEs) with the first of the new exams taking place in 1988. Those staying on in school can then take Advanced Level (A-level) exams that are normally examined in three subjects at the end of secondary school at about age 18. A-levels are still the primary route into higher education and A-level grades are the main criteria for university entry. A minimum of two passing grades at A-level is a necessary but not sufficient condition for university entry.

The Education Expansion (EE)

Figure 1 (taken from Chowdry et al. 2010) shows the age participation index (API) by year as calculated by the UK Department for Education and Skills (DfES). Clearly, this index is flat from around 1970 to 1988 and then jumps from 1989 to 1994.

---

3 The DfES defines the Age Participation Index (API) as a composite measure based on the “the number of UK-domiciled young initial entrants to full-time and sandwich undergraduate courses of higher education in Great Britain, expressed as a proportion of the averaged Great Britain 18 to 19 year old population”.

As mentioned above, the increase is enormous, going from about 15% in 1988 to about 33% in 1994.\(^4\) Why was there such a large increase in educational attainment over this period? Different explanations are found in the literature. Walker and Zhu (2008) attribute it to a relaxation of government limits on student recruitment that occurred contemporaneously with a reduction in the government grant per student paid to the institution. Both factors provided a strong incentive to higher education institutions to increase enrolment.\(^5\) Another major change happened in the early 1990s with the establishment of similar funding arrangements for universities and polytechnics. The incorporation of many former polytechnics and colleges of higher education into the university sector in 1992 led to even greater expansion of capacity (particularly of degree-level courses).

There are also possible demand-side explanations for the expansion in higher educational attainment. Kogan and Hanney (2000) mention a shift in the economy away from manufacturing and towards services and a related increase in the perceived earnings return to education. Also over this period, there were changes in government support for students as the real value of grants fell during the late 1980s and 1990s. This was somewhat offset by an increased availability of student loans. Given loans are much less generous to students than grants, we suspect that these changes in the student funding environment were not an important factor in increasing HE enrollment.

Increases in educational attainment were not just in HE. Blanden and Machin (2004) show that the proportion of people who stayed in education beyond the

---

\(^4\) Greenaway and Haynes (2003) show that increases in governmental funding to HE did not keep pace with student numbers over this period so funding per student fell.

\(^5\) The then Education Secretary, Kenneth Baker, established the principle of university financing following the student providing obvious incentives for universities to increase their student numbers.
compulsory schooling age of 16 also increased rapidly in the late 1980s. They ascribe this acceleration to the introduction of a new examination system for 15/16 year olds, the General Certificate of Secondary Education (GCSE) with first exams in 1988, and the consequent improvement in exam results.\(^6\)

Thus, it appears that the large changes in educational attainment over the 1989 to 1994 period resulted mostly from two administrative changes -- a funding change for universities and polytechnics that encouraged greater HE enrolment and the introduction of the GCSE exam in secondary schools. It is unlikely that these can be considered as having independent effects. Once HE was expanded and made more accessible, it is likely that secondary school students responded by staying longer in school so that they were in a position to take advantage of these opportunities. Also, the GCSE introduction, by leading to increased grades, probably encouraged students to believe that they were good enough to attend HE. In studying the effects of increased educational attainment over this period, we are estimating the combined effect of both policy changes.\(^7\)

4. Data and Descriptives

The data we use are from the UK Quarterly Labour Force survey (QLFS). This is a quarterly survey that tracks particular households for 5 waves so each household is present for 5 consecutive surveys. We use waves from the first quarter of 1997 to the third quarter of 2009. We keep cohorts whose year of birth is between 1958 and 1982 and

---

\(^6\) Hodgson and Spours (2000) also suggest that the GCSE reforms led children to spend more years in education.

\(^7\) From the mid-1960s to the mid-1980s most local areas in England and Wales switched from an early tracking system which started at age 11 to a comprehensive secondary schooling system. The cohorts around the time of the educational expansion were predominantly in a comprehensive schooling system.
who are aged between 25 and 50. We also restrict the sample to persons born in the UK to ensure all persons were exposed to the UK education system.

We have several measures of educational attainment. First, we have age left full-time continuous education.\textsuperscript{8} We also have measures of whether specific qualifications were obtained. These include the number of A-levels or equivalent, and whether or not the person has a college degree.

Questions on earnings are asked in the first and fifth waves so there are a maximum of two earnings observations for each person.\textsuperscript{9} We set earnings to missing if total usual hours worked per week are less than 1 or greater than 80. In the analysis, we use both gross weekly pay in the person’s main job, and gross hourly pay which is a derived variable created in the QLFS. We deflate weekly and hourly earnings using the British quarterly retail price index with base period being the first quarter of 1997.

A critical variable in our analysis is birth cohort. The school starting rule in the U.K. is that children are supposed to start school in an academic year if they turn 5 by the August 31st immediately preceding that academic year. Given that the effect of changes in the education system is determined by the school cohort a child belongs to, it is natural for us to define cohort as a school cohort. The school cohort is defined by pooling together persons who should have started school in the same year, i.e. persons born between September 1st and the following August 31st.\textsuperscript{10} Further information on how

\textsuperscript{8} Education refers to continuous full time education, that is education without a break. Holiday jobs do not count as a break provided that the person intended to complete the course. In addition a gap of up to a year between going to school and going to college or university would not count as a break in continuous full time education. However, nursing training and similar vocational training undertaken while receiving a wage are not counted as part of the continuous education process.

\textsuperscript{9} We do not use observations from waves 2-4 in the analysis.

\textsuperscript{10} The compliance rate with the school starting age rule is very high in the UK, at above 95%. (Bedard and Dhuey 2006). Crawford \textit{et al.} (2010) provide a thorough analysis of the effects of school starting age in the UK.
exactly the variables we use are created from the survey information is available in the Data Appendix. Descriptive statistics for the sample are in Table 1.

**Changes across the Education Distribution**

Figures 2 to 6 show how the education distribution has changed by cohort. First, in figure 2, we plot average age left full time education. While some individuals have longer spells of education as a result of repeating years, repeating years is rare and we expect that most of the change in the cohort averages arise from changes in actual years of education achieved. Figure 2 shows that, in our data, this variable increases from about 17.5 for the 1969 cohort to 18.5 years for the 1976 cohort -- an increase of about 1 year during the expansion period. The increase was slightly larger for women than for men.

Figure 3 plots the proportion of the cohort that has at least an undergraduate degree. The increase for women is from about 12% in the 1969 cohort to 27% in the 1976 cohort. That is, the proportion of women who obtained a college degree more than doubled in 7 years. The increase for men is smaller going from about 15% to 24%. It is striking how flat the college degree series is before the expansion period and the series increases much more slowly after the expansion period. The proportion of women with a degree had barely changed between the 1958 and 1969 cohorts and, once the expansion was over, increased little between the 1976 and 1982 cohorts.

Figure 4 shows that the proportion that has at least 2 A-levels --- the level of achievement that is generally necessary to attend HE -- also increased during the expansion period. This is an important finding as it implies that the number with at least two A-levels is increasing with cohort and so the unobservables for this group are
probably changing across cohorts. Thus, we cannot use people without A-levels as a control group that would be assumed to be unaffected by the expansion.

While the increase in degree attainment is large, even the 12% to 27% increase for women only suggests an increase in years of education of 0.45 of a year on average (assuming a degree takes 3 years). Given an overall increase in education of about a year on average for women, it is clear that higher education expansion is not the dominant factor in the overall educational expansion. Indeed, years of education increased across the distribution. Figure 5 shows the proportion that stay in school beyond the minimum school leaving age of 16. Consistent with the other figures, it is clear that educational attainment also increased at this margin. Figure 6 shows that age left education also increased for persons who did not acquire A-levels.\textsuperscript{11}

Figures 7 and 8 plot average weekly earnings and hourly wages by school cohort and by gender. It is clear that there is a sizeable increase in earnings for female expansion cohorts despite being quite stable beforehand and afterwards. Although there is a decline in earnings for men over the expansion period, the downward trend is much slower over this period than either before or after. Given the presence of important age effects on earnings, these pictures are merely suggestive and motivate the analysis later when we explicitly control for age effects.

\textsuperscript{11} Age finished education also increased over the expansion period for persons with degrees and postgraduate qualifications.
5. Empirical Strategy

1. Cohort-Level Analysis of Education and Earnings

The empirical strategy is to use cohort-level variation in schooling and labour market outcomes. We assume that, in the absence of the educational expansion, the cohort-level changes could be described by a low-order cohort polynomial. The effects of the educational expansion are then identified as deviations from this flexible trend. The rationale behind this approach is that while outcomes such as earnings may differ across cohorts, there is no reason to believe that other factors that influence them do not change smoothly. A low-order polynomial relies on this assumption of smoothness but allows a lot of flexibility in terms of how exactly other cohort-level factors might influence earnings.

This strategy is closely related to the commonly used regression discontinuity (RD) methodology. However, in a classic RD design, the change in education would be discontinuous with a large jump for one particular cohort. Obviously, this is not the case here as the expansion occurs over about 6 cohorts. Therefore, we have to rely on stronger identification assumptions in this case.\(^\text{12}\)

We group cohorts as either being pre-expansion, post-expansion, or during-expansion. Our pre-expansion cohorts include persons born before 1970, the post-expansion cohorts are those born after 1975, and the during-expansion cohorts cover persons born 1970 to 1975.

Our basic specification is as follows:

\(^\text{12}\) In practice RD papers often use discrete variables such as year-of-birth and study changes that do not involve just a single continuous jump. For example, RD analyses of the 1947 change in the school-leaving age (Oreopoulos 2006; Devereux and Hart 2010) use year-of-birth as the running variable and face the problem that the jump in educational attainment takes place over two cohorts. Also see Clark (2010) for an RD design that is similar to that considered here.
\[ ED_{ic} = \alpha + \sum_{c=70}^{75} \beta_c COHORT_c + \delta AFTER_c + f(AGE_{ic}) + g(COHORT_c) + \varepsilon_{ic} \]  

(1)

where the \( i \) subscript denotes individuals and the \( c \) subscript denotes cohorts. \( ED \) is a measure of completed education and \( AFTER \) is a dummy variable for being in a post-expansion cohort. The omitted category is pre-expansion cohort. The coefficients on the \( COHORT \) dummies show the increase in education of each cohort relative to the average education level of the pre-expansion cohorts.\(^{13} \) Following Oreopoulos (2006) and Devereux and Hart (2010), we parameterise both the age and cohort polynomials using a quartic.\(^{14} \)

The specification for log earnings takes a similar form:

\[ \log y_{ic} = \phi + \sum_{c=70}^{75} g_c COHORT_c + \lambda AFTER_c + h(AGE_{ic}) + k(COHORT_c) + \nu_{ic} \]  

(2)

Here \( y \) is earnings. Note that we exclude education from equation (2) and this equation can be seen as a “reduced form” relationship between age, cohort, and earnings. Later we estimate the return to education by Two Stage Least Squares (2SLS) using equations (1) and (2). Monotonically increasing positive values of the cohort dummies in the earnings equation imply that the during-expansion cohorts have higher earnings than would be expected from secular cohort trends alone. If the increase in education resulting from the expansion leads directly to an increase in earnings, this will manifest itself as the cohort estimates from the education equation lining up with the cohort estimates from the earnings equation.

\(^{13} \) Note that we include a polynomial in cohort so the cohort dummies do not pick up any trend increases in education -- just that part of the increase in education that deviates from the underlying trend.

\(^{14} \) Oreopoulos (2006) and Devereux and Hart (2010) also report specifications without age controls and specifications with age dummies. We find estimates are very robust to how age is treated (see Table 4). The correlation between age and cohort is about .8 in our data.
In order to allow for the effects of labour market conditions on wages, we add the national unemployment rate at the time wages are reported into the earnings equation and the following 2SLS model.\footnote{Devereux and Hart (2006) show that wages in Britain are very sensitive to movements in the aggregate unemployment rate.} We also add a control for cohort size in our main education and earnings regressions as researchers have found a negative relationship between cohort size and earnings (Berger 1985; Brunello and Lauer, 2010) and even between cohort size and returns to education (Card and Lemieux, 2001). See the Data Appendix for more details on the national unemployment rate and cohort size measures.

The specification is parsimonious with control variables. We could include controls for variables such as marital status, number of children, and region of residence but we have chosen not to. While these variables correlate with wages, they are intermediate variables in the sense that they are all likely to be affected by the educational expansion. Thus, including them as controls would tend to bias the effects of the expansion on wages downwards. Race is pre-determined so we add a dummy variable for whether the person is white to all specifications.

Our method is related to that of other papers that use before, during, after designs. For example Ichino and Winter-Ebmer (2004) study the affect of the Second World War on educational attainment and subsequent earnings in Germany and Austria. They use pre-war (born before 1930), war-impacted (born 1930-39), and post-war (born 1940 onwards) groups in analysis. Like us, they allow for a low order polynomial in cohort in their analysis. However, our design differs in that they exploit a once-off fall in education for the war-impacted cohorts while we exploit a change in the trend of educational expansion.
attainment. Therefore, they study the impact of a war-impacted dummy while we are interested in all the individual cohort dummies for the during-expansion cohorts.

2. Estimating Returns to Education

If we are willing to add an additional assumption -- that the education expansion had no impact on any earnings-affecting factor other than education -- we can use the education expansion to estimate the return to education. The method is to use 2SLS where the education specification in equation (1) is the first stage and the reduced form for wages and earnings are described by equation (2). We include the cohort and age polynomials in the structural education equation so that the excluded instruments are the dummy variables for each of the 6 during-expansion cohorts and a dummy variable for being in a post-expansion cohort.16

6. Cohort-Level Results

Table 2 quantifies the findings of figures 2 and 3 by regressing the education variables on dummy variables for being in a cohort born in 1970, 71, 72, 73, 74, 75, and an additional dummy variable for being born 1976 or after. The regressions also include controls for a white dummy, a quartic function of cohort, a quartic function of age, and cohort size. We report robust standard errors that allow for arbitrary correlations among persons in the same cohort. Clearly the large changes in figures 2 and 3 are strongly statistically significant.

Changes in Average Weekly and Hourly Earnings

16 Note that when we use 2SLS to estimate the return to education, the national unemployment rate is included in the first stage so it is not used as an exclusion restriction.
Table 2 also includes estimates from regressing the wage variables on the same set of variables plus a control for the national unemployment rate at the time wages are reported. We can interpret the estimates on the cohort dummies as the increase in the log wage relative to that of pre-expansion cohorts, allowing for secular changes across cohort (through the quartic in year-of-birth). The estimates in Table 2 suggest that men born in 1975 gained about 8% higher hourly wages and 10% higher weekly earnings as a result of the expansion. For women, the equivalent gains are about 7% and 11%. This is a reasonably large gain in wages resulting solely from being born late enough to be able to take advantage of the educational expansion.

Another point to take from table 2 is that, consistent with the idea that the increased wages resulted from greater educational attainment, there is generally a monotonic relationship between cohort and wages – the upward trend in education between the 1969 and 1976 cohorts is matched by an upward trend in wages. This feature gives us some confidence that the wage gain is a result of the increases in educational attainment over this period. It also provides a motivation to use the expansion to estimate the return to an extra year of education and we carry out this exercise in the next section.

Returns to Education

Table 3 has estimates of the return to an extra year of education. Given most estimates in the literature have focussed on years of education, it is natural to use this variable for comparison purposes. Also, because educational attainment increased throughout the distribution over this period, it is very difficult to separate out the returns to different levels of education. The 2SLS estimates imply returns of about 6% for both
men and women. These are a little lower than the corresponding OLS estimates as would be consistent with more able individuals choosing to obtain more education. By using only cohort-level variation, the 2SLS estimates avoid this ability bias as it is reasonable to assume that average ability is very similar in neighbouring cohorts.

It is well-known that 2SLS is biased in finite samples when the model is over-identified and instruments are weak. However, Limited Information Maximum Likelihood (LIML) is approximately unbiased in over-identified models. Therefore, given our model is over-identified, we also implement LIML in Table 3. The estimates are almost exactly identical to 2SLS, suggesting that there is no problem with weak instruments in this application. This is not surprising given the large sample size and very precisely estimated effects of the cohort dummies on education.

Robustness Checks

Table 4 reports 2SLS estimates of the return to an additional year of education for a variety of specifications so as to assess the robustness of the estimates in Table 3. For brevity, we report results only for hourly wages – weekly earning estimates are similarly robust. Column (1) shows estimates from the basic specification used in Table 3. Because a quartic in age may not be appropriate, in column (2) we exclude age controls and, in column (3), we include a full set of age dummies (25 in all) instead of the age quartic. Estimates remain very similar.

For the remaining specifications we revert to the quartic in age. Our baseline 2SLS specification uses dummy variables for being in a cohort born in 1970, 71, 72, 73,

---

17 Ackerberg and Devereux (2009) show that LIML can be sensitive to heteroskedasticity and suggest alternative JIVE-type estimators. We have tried implementing their IJIVE estimator and found that it also provides estimates that are almost identical to 2SLS and LIML.
and an additional dummy variable for being born 1976 or after as excluded instruments. In column (4), we no longer exclude the post-expansion dummy from the regression and, in column (5) we remove the 1970 dummy from the instrument set and add a pre-expansion dummy to the regression (but not as an instrument). These changes are designed to force identification to come only from changes in education during the expansion years. In practice, the 2SLS estimates are very similar to those in column (1).

In columns (6) and (7), we reduce the number of instruments by replacing the dummy variables for being in a cohort born in 1970, 71, 72, 73, 74, 75 in the instrument set with two variables. These are a dummy variable for being born between 1970 and 1975 and an interaction between this variable and year-of-birth. This forces a linear relationship between year of birth and educational attainment over the expansion period. The estimates in Table 2 suggest that a linear functional form should fit quite well so it is not surprising that the 2SLS estimates in columns (6) and (7) are very close to those in the basic specification in column (1).

In Column (8) we add controls for regional dummies. Region dummies are potentially an intermediate variable as one way to get a return from higher education is to move to a higher-paying area. However, we feel this is a useful check as it may be that higher wages in some areas are correlated with higher costs of living. Thus, the region of residence controls can be seen as a control for the local cost of living. Their presence has no effect on the coefficient for men but reduces the estimate for women from about 5% to 4%.

Since Scotland has a different educational system from England and Wales, it would be useful to be able to exclude people who would never have considered attending
an educational institution outside of Scotland. The closest proxy for this would be to exclude persons born in Scotland but, unfortunately, this information is not available in many of the sample years. Instead, we use the usual region of residence variable and drop Scottish residents. The resulting estimates, in column (9), are similar to those from the full sample.

Wage and earnings data are missing for over 40% of the sample. This problem arises both because these variables are not available for certain groups (such as the self-employed) and also because respondents refuse or are unable to provide the relevant information. We have verified that there is no 2SLS relationship between education and whether wage and earnings data are missing so there is no indication that our estimates are biased due to missing data.

We have used a quartic cohort specification to be consistent with prior work. We have also examined the effects of using less flexible specifications, in particular quadratic and cubic functions of cohort. The estimates are somewhat sensitive to this but the general finding of moderate returns to education remains – for women the estimates are 4% with a quadratic and 5% with a cubic, the analogous estimates for men are 3% and 4% respectively. All of these estimates are statistically significant.

7. Discussion

It is natural to compare our returns to education estimates to other estimated returns in the literature. Our hourly wage estimates are 2SLS returns to schooling of about 6% for both men and women (Table 3). Compared to corresponding OLS estimates
which are about 8% for men and 10% for women, our findings suggest OLS is slightly biased upwards.

Comparisons between OLS and IV estimates are complicated by the possibility of heterogeneous returns to schooling. Imbens and Angrist (1994) show that, under a monotonicity assumption, the IV estimator provides a Local Average Treatment Effect (LATE). In other words, it calculates the average effect of the treatment for compliers (individuals whose behaviour is changed by the instrument) only. In our case, the monotonicity assumption implies that the increase in educational opportunities during the expansion does not cause anyone to choose less education than they would have if the expansion had never happened. The IV estimate provides no information about the returns to schooling for people whose education choice is not affected by the expansion. This means that, without extrapolation, the LATE is uninformative about the Average Treatment Effect (ATE). However, as pointed out by Oreopoulos (2006), it is reasonable to suppose that as the number of compliers becomes an increasingly large proportion of the sample, the LATE should converge towards the ATE. Therefore, a policy change that has effects throughout the educational distribution (such as the expansion we study) may cause the LATE to be closer to the ATE than otherwise.

Many studies in the schooling returns literature use compulsory schooling laws as instruments for education. Because these laws mostly impact the bottom tail of the education distribution, the LATE may be very different from the ATE. Moderate IV estimates of about 4-6% return to schooling have been found for some European countries (see Brunello and Miniaci (1999) for Italy, Levin and Plug (1999) for the Netherlands, Black et al. (2005) for Norway). However, Grenet (2009) for France, and
Pischke and von Wachter (2008) for Germany find IV estimates that are very low -- Pischke and von Wachter report estimates suggesting zero returns to schooling in Germany. Estimates from North America using compulsory schooling laws suggest larger returns to education of 10% or more (Oreopoulos 2006).

For Britain, Harmon and Walker (1995) find 15% returns to an extra year of education using both 1947 and 1973 changes to compulsory schooling laws. However, recent studies of the 1947 law change (Devereux and Hart, 2010) and of the 1973 law change (Grenet, 2009) find much smaller returns in each case: Devereux and Hart find zero estimates for women and about 4-7% for men, Grenet finds estimates of about 6-7% for both men and women.\textsuperscript{18} As such, our estimates of about 6% for both men and women are generally consistent with recent British estimates that use compulsory schooling laws.

In addition to compulsory schooling laws, researchers have looked at other sources of plausibly exogenous variation in higher education and attempted to investigate the relationship between schooling and earnings in terms of such policy changes. Pons and Gonzalo (2002) use availability of a college in the province to look at the returns to schooling for male workers in Spain. In earlier work, Card (1995) and Conneely and Uusitalo (1997) examine the schooling and earnings differentials associated with growing up near a college or university. Like us, these studies look at how people change their schooling decisions when provided with a greater opportunity of college entry. However, the availability of a local college may largely impact the type of people who would reduce their investment in schooling in the absence of a nearby college. These are likely

\textsuperscript{18} Dickson (2009) finds estimates of 10% for men using the 1973 change and Dickson and Smith (2011) show that most of this return appears to occur because of increased qualifications resulting from the law change. Grenet (2009) also stresses the role of qualifications.
to be people with potentially high returns to education. These three papers all find the resulting IV estimators are substantially above the corresponding OLS estimates.

Unlike the literature mentioned above, the British education expansion we study is a very different educational event that affected people’s education at both secondary school level and in higher education. As such, our estimates of the return to education may reflect the returns for a broader group of people than other studies in the literature. The (slightly lower than OLS) 6% returns we estimate are consistent with the notion that many previous studies have disproportionally identified the effect from high returns individuals.

8. Conclusions

The British education expansion increased people’s education by shifting the whole educational distribution of affected cohorts upwards. Using the Quarterly Labour Force Survey, we find that the expansion caused men to increase education by about a year on average and gain about 8% higher wages; women obtained a slightly greater increase in education and a similar increase in wages. Clearly, there was a sizeable gain from being born late enough to take advantage of the greater educational opportunities offered by the expansion.

Using cohort-level variation in education, our IV estimates of returns to schooling are about 6% for both men and women, which hold up against several robustness checks and are quite consistent with recent findings by Grenet (2009) who uses a difference-in-difference estimator to analyse returns to schooling in Britain. Our findings imply that the education expansion was effective in increasing educational levels and subsequent
earnings. It should be noted, however, that our estimation strategy relies on the strong assumption that any direct cohort effects on earnings are captured by a low-order polynomial. As such, further research on the educational expansion using different identification assumptions would be useful to corroborate our findings.
References


Figure 1 Long-Term Trends in Higher Education Participation (Age Participation Index)

Source: Chowdry et al. (2010).
Age Participation Index (API): A composite measure based on the “the number of UK-domiciled young initial entrants to full-time and sandwich undergraduate courses of higher education in Great Britain, expressed as a proportion of the averaged Great Britain 18 to 19 year old population”
Figure 2

Source: Labour Force Survey (authors’ calculations)

Figure 3

Source: Labour Force Survey (authors’ calculations)
Figure 4

Proportion of 2+ A-levels Achievement

Source: Labour Force Survey (authors’ calculations)

Figure 5

Proportion of People Leave School after Age 16

Source: Labour Force Survey (authors’ calculations)
Figure 6

Age Left Full Time Education for non-Alevels

Source: Labour Force Survey (authors’ calculations)

Figure 7

Mean Log(Hourly Wages)

Source: Labour Force Survey (authors’ calculations)
Figure 8

Source: Labour Force Survey (authors’ calculations)
### Table 1 Summary Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MEN</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>295412</td>
<td>35.784</td>
<td>6.251</td>
<td>25</td>
<td>50</td>
</tr>
<tr>
<td>White</td>
<td>295412</td>
<td>0.949</td>
<td>0.221</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Hours Worked per Week</td>
<td>251514</td>
<td>44.517</td>
<td>9.670</td>
<td>1</td>
<td>80</td>
</tr>
<tr>
<td>Unemployment Rate</td>
<td>295412</td>
<td>0.036</td>
<td>0.007</td>
<td>0.025</td>
<td>0.055</td>
</tr>
<tr>
<td>Cohort Size (in thousands)</td>
<td>295412</td>
<td>773.071</td>
<td>84.521</td>
<td>566.537</td>
<td>867.682</td>
</tr>
<tr>
<td>Log of Hourly Wage (real)</td>
<td>160503</td>
<td>2.161</td>
<td>0.517</td>
<td>0.006</td>
<td>4.604</td>
</tr>
<tr>
<td>Log of Weekly Wage (real)</td>
<td>160945</td>
<td>5.871</td>
<td>0.565</td>
<td>0.383</td>
<td>8.291</td>
</tr>
<tr>
<td>Region-of-Residence</td>
<td>295412</td>
<td>11.875</td>
<td>5.109</td>
<td>1</td>
<td>20</td>
</tr>
<tr>
<td>Age Left Full Time Continuous Education</td>
<td>288339</td>
<td>17.581</td>
<td>2.543</td>
<td>15</td>
<td>30</td>
</tr>
<tr>
<td>Proportion of First Degree Holders</td>
<td>295412</td>
<td>0.161</td>
<td>0.367</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>School Cohort</td>
<td>295412</td>
<td>1967.212</td>
<td>5.997</td>
<td>1958</td>
<td>1982</td>
</tr>
<tr>
<td><strong>WOMEN</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>326503</td>
<td>35.679</td>
<td>6.226</td>
<td>25</td>
<td>50</td>
</tr>
<tr>
<td>White</td>
<td>326503</td>
<td>0.946</td>
<td>0.226</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Hours Worked per Week</td>
<td>239601</td>
<td>31.822</td>
<td>12.736</td>
<td>1</td>
<td>80</td>
</tr>
<tr>
<td>Unemployment Rate</td>
<td>326503</td>
<td>0.036</td>
<td>0.007</td>
<td>0.025</td>
<td>0.055</td>
</tr>
<tr>
<td>Cohort Size (in thousands)</td>
<td>326503</td>
<td>772.304</td>
<td>84.950</td>
<td>566.537</td>
<td>867.682</td>
</tr>
<tr>
<td>Log of Hourly Wage (real)</td>
<td>174341</td>
<td>1.908</td>
<td>0.514</td>
<td>0.003</td>
<td>4.583</td>
</tr>
<tr>
<td>Log of Weekly Wage (real)</td>
<td>174788</td>
<td>5.219</td>
<td>0.803</td>
<td>0.390</td>
<td>8.280</td>
</tr>
<tr>
<td>Region-of-Residence</td>
<td>326503</td>
<td>11.867</td>
<td>5.150</td>
<td>1</td>
<td>20</td>
</tr>
<tr>
<td>Age Left Full Time Continuous Education</td>
<td>321315</td>
<td>17.578</td>
<td>2.378</td>
<td>15</td>
<td>30</td>
</tr>
<tr>
<td>Proportion of First Degree Holders</td>
<td>326503</td>
<td>0.155</td>
<td>0.362</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>School Cohort</td>
<td>326503</td>
<td>1967.333</td>
<td>6.007</td>
<td>1958</td>
<td>1982</td>
</tr>
<tr>
<td>Independent Variables</td>
<td>MEN</td>
<td>WOMEN</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------------</td>
<td>-----</td>
<td>-------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Education Degree Hourly Wages Weekly Earnings</td>
<td>Education Degree Hourly Wages Weekly Earnings</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cohort 70</td>
<td>0.019 0.006* 0.006 0.007</td>
<td>0.123*** 0.018*** -0.007** 0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.030] [0.003] [0.005] [0.006]</td>
<td>[0.003] [0.003] [0.006] [0.006]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cohort 71</td>
<td>0.209*** 0.015*** 0.013* 0.007</td>
<td>0.227*** 0.022*** 0.006 0.015**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.041] [0.004] [0.007] [0.009]</td>
<td>[0.034] [0.004] [0.005] [0.006]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cohort 72</td>
<td>0.342*** 0.027*** 0.024** 0.022</td>
<td>0.482*** 0.047*** 0.015* 0.023**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.050] [0.006] [0.011] [0.013]</td>
<td>[0.052] [0.006] [0.008] [0.010]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cohort 73</td>
<td>0.452*** 0.034*** 0.042** 0.052***</td>
<td>0.817*** 0.079*** 0.038*** 0.052***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.063] [0.009] [0.016] [0.018]</td>
<td>[0.073] [0.007] [0.012] [0.014]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cohort 74</td>
<td>0.711*** 0.050*** 0.060*** 0.067***</td>
<td>1.080*** 1.002*** 0.056*** 0.079***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.076] [0.012] [0.021] [0.023]</td>
<td>[0.096] [0.009] [0.015] [0.019]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cohort 75</td>
<td>0.775*** 0.053*** 0.077*** 0.090***</td>
<td>1.341*** 1.100*** 0.055*** 0.091***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.089] [0.015] [0.026] [0.027]</td>
<td>[0.119] [0.011] [0.019] [0.024]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post EE Cohort</td>
<td>0.937*** 0.059*** 0.080** 0.096***</td>
<td>1.548*** 1.300*** 0.071*** 0.108***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.107] [0.018] [0.031] [0.034]</td>
<td>[0.173] [0.016] [0.025] [0.031]</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: All specifications include a quartic function of year-of-birth, a quartic function of age, white, and cohort size; the quarterly unemployment rate enters the wages/earnings regressions only. Robust standard errors in brackets allow for clustering by year-of-birth. All wages/earnings regressions are weighted by the inverse of the number of observations on that individual in the regression.

*** p<0.01, ** p<0.05, * p<0.1
## Table 3 OLS, 2SLS, and LIML Estimates of Return to Education for British-born Persons Aged 25-50

<table>
<thead>
<tr>
<th></th>
<th>MEN</th>
<th>WOMEN</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Hourly Wages</td>
<td>Weekly Earnings</td>
</tr>
<tr>
<td><strong>OLS</strong></td>
<td>0.078***</td>
<td>0.069***</td>
</tr>
<tr>
<td></td>
<td>[0.003]</td>
<td>[0.003]</td>
</tr>
<tr>
<td><strong>2SLS</strong></td>
<td>0.062***</td>
<td>0.066***</td>
</tr>
<tr>
<td></td>
<td>[0.016]</td>
<td>[0.019]</td>
</tr>
<tr>
<td><strong>LIML</strong></td>
<td>0.061***</td>
<td>0.066***</td>
</tr>
<tr>
<td></td>
<td>[0.016]</td>
<td>[0.021]</td>
</tr>
<tr>
<td></td>
<td>[n=159,709]</td>
<td>[n=160,146]</td>
</tr>
</tbody>
</table>

**Notes:** All specifications include a quartic function of year-of-birth, a quartic function of age, white, the quarterly unemployment rate, and cohort size. Robust standard errors in brackets allow for clustering by year-of-birth. All regressions are weighted by the inverse of the number of observations on that individual in the regression. LIML is Limited Information Maximum Likelihood.

*** p<0.01, ** p<0.05, * p<0.1
Table 4 2SLS Estimates of Return to Education for British-born Persons Aged 25-50 (Hourly Wages)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>[n=159,709]</td>
<td>0.062***</td>
<td>0.071***</td>
<td>0.060***</td>
<td>0.060***</td>
<td>0.059***</td>
<td>0.069***</td>
<td>0.068***</td>
<td>0.062***</td>
<td>0.060***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[0.016]</td>
<td>[0.022]</td>
<td>[0.016]</td>
<td>[0.017]</td>
<td>[0.017]</td>
<td>[0.017]</td>
<td>[0.019]</td>
<td>[0.019]</td>
<td>[0.015]</td>
</tr>
<tr>
<td></td>
<td>[n=173,683]</td>
<td>0.053***</td>
<td>0.056***</td>
<td>0.050***</td>
<td>0.054***</td>
<td>0.050***</td>
<td>0.055***</td>
<td>0.056***</td>
<td>0.038***</td>
<td>0.056***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[0.014]</td>
<td>[0.017]</td>
<td>[0.014]</td>
<td>[0.014]</td>
<td>[0.011]</td>
<td>[0.017]</td>
<td>[0.017]</td>
<td>[0.014]</td>
<td>[0.016]</td>
</tr>
</tbody>
</table>

**Independent Variables**

- Age Controls: quartic, no dummies, quartic, quartic, quartic, quartic, quartic, quartic
- Additional Cohort Controls b: post-EE dummy, pre-EE dummy, no, no, no, no, no, no
- Region-of-Residence Dummies: no, no, no, no, no, no, no, yes, no
- Scottish Residents Included: yes, yes, yes, yes, yes, yes, yes, yes, no
- Excluded Instrumental Variables:
  - Cohorts 70-75, Post EE Dummy: √
  - Cohorts 70-75: √
  - Cohorts 71-75: √
  - Post-EE Dummy, EE Dummy, EE Linear Term: √
  - EE Dummy, EE Linear Term: √

Robust Standard Errors in Parentheses
*** p<0.01, ** p<0.05, * p<0.1

- a The sample size in column (9) is 140,032 for men and 151,040 for women.
- b. All specifications include a quartic function of school cohort.
Data Appendix

We use samples of the Quarterly Labour Force Survey (QLFS) from the first quarter of 1997 to the third quarter of 2009. We keep cohorts whose year of birth is between 1958 and 1982 and who are aged between 25 and 50.

Country of Birth

We identify individuals who were born in the UK using the CRY variable. (CRY equal to 1 up to 2000, between 1 and 5 from 2001 to 2006, and equal to 921, 922, 923, 924, or 926 from 2007 onwards). CRY is missing for about 15,000 observations in 2000 and we exclude these cases from the sample.

Education Variables

Our main education is age completed full time schooling (EDAGE). We recode EDAGE to missing if the reported age is less than 15 years or greater than 30.

From 1997 to 2009, there are three variables related to degree achievement: DEGREE (from 97Q1 to 03Q4), DEGREE4 (from 04SPRING to 06Q4), DEGREE71 (existing from 07Q1 to 09). While DEGREE records all degrees respondents have obtained, the other two only record the highest degree. We assume people who report having a higher degree are also first degree holders.

Earnings Measures

Our hours measure (TTUSHR) is total usual hours worked per week including overtime. We set this variable to missing if total usual hours are less than 1 or greater than 80.

The earnings variable we use is GRSSWK (gross weekly pay in main job) and the hourly wage measure we use is HOURPAY (gross hourly pay) which is a derived variable created in the QLFS. We deflate weekly and hourly earnings using the British quarterly retail price index with base period being the first quarter of 1997. We set to missing cases in which real hourly wages (in 1997 pounds) are less than £1 or greater than £100. We also set to missing cases in which real weekly earnings (in 1997 pounds) are less than £1 or greater than £4000. Additionally, we set wages and earnings to missing if usual weekly hours are not between 1 and 80.
**School Cohort**

We have information on month of birth (DOBM) and year of birth (DOBY) for survey years 1997 to 2002. For later years (up to the 3rd quarter of 2007), we have a variable AGEDFE that is age at the preceding August 31st. This allows us to assign people to school cohorts even if we don’t know month or year of birth. We define school cohort as being equal to the survey year minus AGEDFE if persons are interviewed in the 4th quarter. For persons interviewed in the first two quarters, school cohort is defined as survey year minus AGEDFE -1. This method uses the idea that all individuals who were the same age at the previous August 31st, are members of the same school cohort. For persons interviewed in the 3rd quarter, we use the interview week and AGEDFE to assign school cohort. For cases after the 3rd quarter of 2007, we estimate school cohort using age and interview week.

People born in the last 4 months of a calendar year are assigned to the next school cohort. For example, the 1958 cohort includes people who were born from September to December 1957 and people who were born from January to August 1958. Likewise, births in the last 4 months in 1982 are not included in 1982 school cohort.

**National Unemployment Rate Measures**

We measure unemployment rates using the variable on economic activity (ILODEFR). This is a self-reported variable where people report their economic activity in each survey quarter. We calculate the unemployment rate in each quarter as the proportion of people aged 25-50 who say that they are unemployed.

**Cohort Size Measures**

We calculate the cohort size using official national statistics on births by month and year from the Office for National Statistics website.


---

19 We use the WEEK variable to measure when exactly people are interviewed. The calendar week equals WEEK – 4 if WEEK is between 6 and 13, the calendar week equals WEEK + 9 if WEEK is between 1 and 4.