THE CENTRE FOR MARKET AND PUBLIC ORGANISATION

The Causal Effect of Education on Wages Revisited

Matt Dickson

September 2009
Working Paper No. 09/220

Published in the Oxford Bulletin of Economics and Statistics, DOI: 10.1111/j.1468-0084.2012.00708.x

Centre for Market and Public Organisation
Bristol Institute of Public Affairs
University of Bristol
2 Priory Road
Bristol BS8 1TX
http://www.bristol.ac.uk/cmpo/

Tel: (0117) 33 10799
Fax: (0117) 33 10705
E-mail: cmpo-office@bristol.ac.uk

The Centre for Market and Public Organisation (CMPO) is a leading research centre, combining expertise in economics, geography and law. Our objective is to study the intersection between the public and private sectors of the economy, and in particular to understand the right way to organise and deliver public services. The Centre aims to develop research, contribute to the public debate and inform policy-making.

CMPO, now an ESRC Research Centre was established in 1998 with two large grants from The Leverhulme Trust. In 2004 we were awarded ESRC Research Centre status, and CMPO now combines core funding from both the ESRC and the Trust.

ISSN 1473-625X

CMPO Working Paper Series No. 09/220

The Causal Effect of Education on Wages Revisited

Matt Dickson

CMPO, University of Bristol and IZA

September 2009

Abstract

This study estimates the return to education in Britain using two instrumental variable (IV) estimators: one exploits variation in schooling associated with early smoking, the other uses the raising of the school leaving age; both affect a sizeable proportion of the sample. Early smoking is found to be a strong and valid IV and unlike previous IV strategies uses variations in education at numerous points across the distributions of (i) education, and (ii) ability. Thus whilst still a 'local average treatment effect' the estimate is closer to the average effect of additional education, akin to least squares but corrected for endogeneity.

Keywords: human capital, endogeneity, local average treatment effect

JEL Classification: I20, J30

Electronic version: www.bristol.ac.uk/cmpo/publications/papers/2009/wp220.pdf

Acknowledgements

Many thanks for useful comments to Ian Walker, Mark Stewart, Simon Burgess, Sarah Smith, Colm Harmon, Jennifer Smith, Paul Gregg and seminar participants at University of Warwick and the Centre for Market and Public Organisation

Address for Correspondence
CMPO, Bristol Institute of Public Affairs
University of Bristol
2 Priory Road
Bristol
BS8 1TX
m.r.dickson@bristol.ac.uk
www.bristol.ac.uk/cmpo/people/researchers/dickson
www.bristol.ac.uk/cmpo/

CMPO is jointly funded by the Leverhulme Trust and the ESRC

1 Introduction

The causal effect of education on wages has long been a parameter of interest not only to labour economists but also to governments and individuals themselves. Much ink has been spent and many regressions run in pursuit of an answer yet consensus remains elusive. Restricting focus to estimates of returns in the UK, the range varies from the very low (Devereux and Hart, 2010) to the perhaps surprisingly high (Harmon and Walker, 1995). Choice of econometric technique – and more specifically for instrumental variables studies, the choice of instrument – is non-trivial and has implications both for the size of the estimate and its external validity.

This paper estimates the causal effect of education on wages using two alternative methods of instrumentation. Estimates derived using variations in schooling associated with early smoking behaviour are compared with estimates derived by exploiting the impact on schooling of the 1972 raising of the minimum school leaving age (RoSLA) in England and Wales. The RoSLA instrument follows in the tradition of Card (1995) and similar papers[1], which use institutional factors or elements of the budget constraint to create instruments. This earlier research using IV methods covers a wide range and it is well established that when the effect of treatment is heterogeneous, IV estimates a 'local average treatment effect' (LATE)[2]. The LATE will be instrument dependent and captures the effect of the treatment only for those induced to change treatment status by the particular instrument chosen. These IV estimates of the return to education may only apply to quite a specific and unrepresentative group, for example RoSLA affected only those who had wanted to leave school early therefore in this case IV estimates the effect of additional schooling for those at the bottom of the schooling distribution who were forced to stay longer.

In contrast, early smoking is correlated with the schooling decisions of individuals across

[1]The first notable paper to use instrumental variables to estimate the return to education was Angrist and Krueger (1991) while Harmon and Walker (1995) were the first to exploit changes in the minimum school leaving age in the UK to form IVs.

[2]See Imbens and Angrist (1994) for the theoretical exposition of LATE.

the distributions of both (i) education, and (ii) ability. We interpret the estimates from this IV as closer to an average effect of additional schooling akin to least squares but corrected for endogeneity. Our contribution is therefore two-fold: first, we present the case that the early smoking IV LATE is more appropriate than the RoSLA LATE for drawing inference concerning the average return to education in the population, implementing each IV using that same data from the British Household Panel Survey. Secondly, as we have two genuinely different mechanisms generating variation in education and providing two instruments, we are able to test the validity of the exclusion restrictions, something that is rarely possible to do. We find the OLS estimate to be considerably downward biased at 4.6% compared with the IV estimates of 12.9% using early smoking, 10.2% using RoSLA and 12.5% using both instruments simultaneously. This suggests that schooling in Britain has a high wage return not only at the lower part of the distribution but across a range of education levels.

The next section considers the problem of estimating the return to education and IV solutions in the literature. Section 3 establishes the case for early smoking as an instrument for education. Section 4 describes the data and estimation approach, while section 5 presents the results before evaluating various hypotheses regarding the instrument. Section 6 then compares the early smoking instrument estimates with RoSLA IV results, before section 7 exploits the presence of two instruments to formally test the validity of these instruments. Section 8 offers some concluding remarks.

2 Literature

The foundation of the education returns literature has been Mincer's (1974) human capital earnings function:

$$\ln w_i = X_i'\varphi + \beta S_i + \epsilon_i \tag{1}$$

in which w_i is the wage, X_i is a vector of the individual's characteristics, including experience

and experience-squared, and S_i is the number of years of schooling, determined by:

$$S_i = X_i'\gamma + u_i \tag{2}$$

This human capital earnings function tells us the expected (log) wage that an individual will earn given his/her observable characteristics and years of education. It is well known that if this relationship in equation (1) is estimated by least squares, a causal interpretation of the estimated parameter $\hat{\beta}$ requires $\mathrm{E}(X_i\ \epsilon_i) = 0$ and $\mathrm{E}(S_i\ \epsilon_i) = 0$, the latter of which entails $\mathrm{E}(u_i\ \epsilon_i) = 0$.

The potential for the unobserved characteristics that determine schooling choice to also be correlated with wage has always been a concern to labour economists. Early research concentrated on the issue of 'ability bias' which suggested that $\mathrm{E}(u_i\ \epsilon_i) > 0$ since the residuals pick up ability which is positively correlated with both wages and schooling, resulting in OLS estimates being unambiguously biased upwards. In contrast, Griliches (1977) concluded that 'ability bias' was in reality small and was overwhelmed by the attenuation bias introduced by measurement error in the schooling variable, with the result that OLS under-estimated the actual return to education.

At the start of the 1990s, a number of economists suggested that OLS estimates of the return to education may suffer from a further bias – 'discount rate bias' (see Lang, 1993; Card, 1994). In Becker's (1964) model of human capital formation, with standard assumptions[3], an individual will accumulate human capital to the point where the marginal rate of return on the last unit of education is equal to his/her discount rate. An individual's discount rate reflects both his/her access to finance to fund current investment in education whilst deferring earnings and also his/her rate of time preference. If individuals differ in their preferences and in their financial resources, this will result in different discount rates and lead to variation in the point at which they stop acquiring education – a higher discount

[3]i) workers maximize the discounted present value of lifetime wealth; ii) time in school is independent of time in work, or alternatively lifetimes are infinitely lived; iii) there are no direct costs of education; iv) the effect of experience on earnings is multiplicative.

rate resulting in a lower optimal level of education. Therefore schooling choice may differ amongst individuals of the *same* ability because of differences in individual discount rates. If the unobserved discount rate that affects education also independently affects wages then OLS estimates will incur bias from this source.

It could be the case that individuals who have a higher discount rate have more ambition or determination to get into the labour market and earn money. This drive is rewarded in higher wages and also these individuals are more likely to choose career paths with steep wage curves. Consequently a higher discount rate is associated with lower education but also a higher wage conditional on education, thus $\mathrm{E}(u_i\ \epsilon_i) < 0$ and OLS estimates of the return to education will be negatively biased. However, it may be that the opposite is true: Munasinghe and Sicherman (2000) present evidence from the NLSY[4] that smoking can proxy for rate of time preference, and that after controlling for a rich set of covariates, smokers (high discount rate individuals) experience lower initial wages and lower wage growth than non-smokers, which would suggest that high discount rate individuals are not selecting into steep wage growth occupations. If the wages of high discount rate individuals are lower (conditional on education) and grow more slowly then OLS estimates will be upward biased.

Discount rate and ability are both sources of variation in levels of schooling, moreover they interact. Though (arguably) unconditionally uncorrelated, Lang (1993) points out that once we condition on a level of education, discount rate and ability *are* correlated: those with a higher discount rate will have higher ability. Recalling Becker's model this makes sense: if two individuals have chosen the same level of schooling it means that for each, at that point, the marginal return to schooling is equal to their discount rate. Thus the individual with the higher discount rate has a higher marginal return at that level of education, indicating that they have higher ability. A higher value of discount rate will reduce schooling, but conditional on schooling in the wage equation, a higher discount rate will mean higher ability and a higher wage: thus $\mathrm{E}(u_i\ \epsilon_i) < 0$. This potential mechanism through which discount

[4]National Longitudinal Survey of Youth, US data

rate affects the joint process of education and earnings again suggests a negative bias in the OLS estimates. If both ability bias and discount rate bias affect the OLS estimate of the return to education but work in opposite directions, then *a priori* the net bias in the coefficient is indeterminate.

Much of the literature has focused on the IV solution to this endogeneity problem: identifying a variable which affects schooling but does not independently enter into the earnings equation, and is uncorrelated with its error term. Numerous instruments have been employed, with many studies reviewed in Card (2001).

While IV has the advantage that we can potentially derive estimates purged of the biases discussed above, it also has some shortcomings. Weak instruments (that is, those that although uncorrelated with wages are hardly correlated with schooling) and invalid instruments (those that although correlated with schooling, may also be correlated with wages) may be worse than no instruments at all – as Bound, Jaeger and Baker (1993) put it "the cure can be worse than the disease".

Much attention has been given to the weak instruments issue in the econometrics literature of the last 15 years and it is now well established (see *inter alia*, Baum, Schaffer and Stillman (2007), Murray (2006)) that two-stage least squares performs very poorly in the presence of weak instruments: not only are point estimates biased, the estimated standard errors of parameters are too small, leading to inaccurate inference.

Further, Bound, Jaeger and Baker (1995) show that even a small correlation between the instrument and the error term in the earnings equation can result in a large bias in the IV estimates even in large samples – a problem that is compounded if the instrument is weak. Even when an instrument is significant at conventional levels, it may still be weak and lead to the problems of bias and unreliable inference outlined above; hence the need to econometrically verify the strength of any candidate instrument.

It is not routinely possible however to test an instrument for correlation with the error term in the earnings equation as to do that we would first need to estimate the wage equation

to give us a valid error term which requires a consistent estimator for φ and β, which we can only find if we have an alternative instrument that we know is valid and strong in the first place. The advantage in having multiple instruments is that this allows us to determine the validity of the preferred instrument (early smoking), exploiting the validity of the other instrument available (RoSLA). In addition to the formal econometric test, further supportive evidence for the validity of the early smoking instrument is provided by the reduced forms, by intuition and by the consistency of results estimated with different instruments. While it is rarely possible to completely dismiss the fear of invalidity, there is firm evidence in favour of the validity of early smoking as an instrument.

As outlined above, an additional problem with IV in the presence of heterogeneous treatment response is that it captures a LATE: the IV estimate of the schooling coefficient β is a weighted average of the marginal returns to education for those whose schooling choice is influenced by the instrument, conditional on X_i[5]. While contributing useful information regarding the return for specific groups, previous LATE estimates have provided little indication of what the return to schooling may be for a more general population.

For example, increasing the minimum school leaving age only affects those who want to leave at the minimum age, and so the LATE is the return for this specific group. Figure 1 shows the education leaving age density when the minimum school leaving age is 15 compared with when it is 16. It is clear that in the upper ranges the densities are very similar, and that the increase in minimum school leaving age affects only the lower part of the distribution of leaving ages. This concurs with the evidence of Chevalier *et al.* (2004) who use a large sample of data from the General Households Survey (GHS) and find – using a number of tests of the equality of distributions – that RoSLA only affected the attainment of those at the bottom of the schooling distribution: there was not a ripple effect further up. Whether the individuals affected are predominantly high discount rate or predominantly low ability will determine whether we expect the IV estimate from RoSLA to be higher or lower than

[5]In order to give this LATE interpretation, there is a monotonicity requirement that all individuals have the same signed response to the instrument.

OLS. In either case it may not be a good guide to a more general return to education.

Therefore it is important that an instrument avoids these three prominent problems: being correlated with the structural equation error term, being only weakly correlated with the endogenous regressor or capturing a LATE that is not informative when it comes to answering the question we want to ask.

3 Instrumenting Education Using Early Smoking

Evans and Montgomery (1994) proposed using whether or not an individual smoked when they were young as an instrument for schooling[6]. The intuition behind the instrument starts with the observation that just as schooling is not randomly assigned across the population, neither is the decision to engage in (un)healthy habits. Evans and Montgomery note that "one of the most persistent relationships in health economics is that more educated people have better health and better health habits" (1994, p1). This view is supported by a number of reviews of the empirical evidence on the link between health and education by Grossman (see Grossman, 2006). After extensively reviewing the evidence Grossman concludes that completed years of formal schooling is the most important correlate of good health, and this statement applies whether health is being measured by mortality rates, morbidity rates, self-evaluated health status or psychological well being. In the UK, Oreopoulos (2006) uses data from the GHS which asks individuals to self-report their health status, and finds that an additional year of schooling increases the chance that an individual will report good health by 6.0 percentage points, and reduces the chance of reporting poor health by 3.2 pp. There remains a debate as to whether or not this education-health relationship is causal, with Evans and Montgomery citing a quite different explanation for the relationship, proposed by Victor Fuchs (1982). Fuchs argues that unobserved differences in the rate of time preference determine both the number of years schooling that an individual attains and

[6]This IV strategy has also been pursued by Chevalier and Walker (2002) using GHS and National Child Development Study (NCDS) data, and by Fersterer and Winter-Ebmer (2003) for Austrian data.

their investments in health, as both decisions involve a trade off between current costs and the discounted value of future benefits.

As with Becker's model of human capital accumulation, in a health accumulation model individuals invest in health until the marginal return to health investment equals their discount rate. If an individual has a higher discount rate because of her rate of time preference, she cares less about the future and more about the present and will therefore *ceteris paribus* quit formal education at a younger age and be less (more) likely to engage in good (bad) health habits. If the correlation between a health habit, such as smoking, and education is driven by a common unobserved factor (time-preference) then this health habit could potentially be used as an instrument for education.

In their work on rational addiction Becker and Murphy (1988) posit that the decision to smoke reflects discount rate in that it indicates an individual's rate of time preference; we follow Evans and Montgomery in arguing that, as a result, smoking as a teenager can be used as a valid instrument for education. One way in which Fuchs supported his hypothesis was to show that education at age 24, when education levels vary considerably, is as important a predictor of smoking at 17 – when most individuals have the same level of education – as it is a predictor of smoking at 24 (see Farrell and Fuchs, 1982). Using a larger dataset than the main estimation sample, we implement a probit of current smoking using completed years of schooling amongst the explanatory variables, and repeat the probit for smoking at age 16. The marginal effects estimated at the mean of the explanatory variables suggest that for each additional year of schooling the probability of being a current smoker falls by 2.7% (significant at the 1% level). In the probit for smoking at 16, it is estimated that each additional year of completed education reduces the probability of having smoked at age 16 by 3.8% (significant at the 1% level, see Table B-1). Thus completed education is a significant determinant of early smoking – suggesting that it is not greater education that determines the decision not to smoke: education predicts early smoking as well as later smoking, suggesting that another underlying factor (time preference) is determining both.

Intuitively, the decision that an individual makes at age 16 as to whether or not to continue in education is likely to be significantly affected by her discount rate – whether that is because of access to financial resources or because of the individual's rate of time preference. In the UK this is the first point at which individuals can choose to leave education, moreover it remains the case that staying in school post-16 and taking A-levels is the major route into university. Moreover, whether an individual chooses to smoke at 16 is also likely to be determined in large part by their rate of time preference. Of the sample individuals who have ever smoked, 60.5% were smoking when age 16 and 81.1% were smoking when age 18[7]. Therefore it is clear that the majority of individuals who ever smoke, first take that decision at around the same time that they are making decisions over the continuation of their education. This concurrence in the timing of the smoking and school leaving decisions generates a statistically precise and quantitatively large correlation between years of education and early smoking. Thus smoking at 16 satisfies the first criterion for an instrument: it is relevant as it is strongly correlated with completed education. Moreover, the effect of early smoking on years of schooling is sizeable (just under one year less education is completed on average by those who smoke when 16 *ceteris paribus*), therefore the instrument works through a substantial variation in education.

In addition to looking at the first stage equation for years of schooling, looking at the reduced form for the dependent variable of interest (log hourly wage), supports the argument that early smoking can be used to instrument for education. Appendix Table B-2 shows that the smoker-at-16 indicator has a significant negative coefficient in the reduced form regression, which is in line with intuition: those who smoked when 16 have lower wages than those who did not, with the argument being that this is driven wholly by the difference in average years of schooling between the two groups.

The second criterion is validity: the instrument must not be correlated with wage. As it is a past health habit that is instrumenting for education in the equation for *current* wage,

[7]These figures refer to the estimation sample, for the largest sample available in the BHPS the figures are almost identical: 61.0% and 79.7%.

there should not be a correlation via an income effect: the contemporary wage can have no impact on the disposable income of 16 year old deciding whether or not to smoke. Moreover, theoretically whether one smoked at 16 should have no independent direct effect on *current* wage. It is by no means certain that current smoking affects current wage via a productivity effect – in fact, using BHPS data Brune (2007) shows that there is only weak evidence that current smoking causally reduces current wages. Thus a link between smoking at 16 and current wage would be even more speculative.

Further support for this contention comes from evidence in the literature that there are no harmful effects of early drug and alcohol use on later economic prospects. Burgess and Propper (1998) show this for adolescent men in the US whose economic outcomes 10 years later are unaffected by soft drug and alcohol use, while MacDonald and Pudney (2000) show similar results for young British men: there are no negative effects on employment or occupational attainment of teenage use of soft drugs. Taking all of this into consideration, there is no reason to think that smoking at 16 would affect current wage – and as individuals age and move further away from being 16 this is even more so the case. Moreover, there is a good degree of movement between smoking and non-smoking amongst the BHPS sample of men, with 41.0% of men who *did* smoke when they were 16 having stopped by the time they are first observed in the data, while 37.5% of the men who are smokers when first observed in the data *were not* smokers at age 16. In fact, of the 585 men who change smoking status between 16 and their first panel observation, 46.3% are taking up smoking. In light of these arguments, we believe that smoking at 16 can *itself* be legitimately excluded from the wage equation.

The next important question is whether early smoking is correlated with other unobservable characteristics that *do* affect current wage. If the rate of time preference that characterises early smokers leads to other unobserved human capital investments throughout life that independently affect wage then this would invalidate the instrument. Alternatively it may be the case that time-preference at age 16 affects investment in education but that

conditional on this there is not an additional effect on earnings. Evidence from the psychology and psychiatry literatures suggests that for men in particular, time horizons during the teenage years are often shorter than is the case later in life and that time-preferences do not settle until men are in their early 20s (see Steinberg (2007), Silverman (2003), Read and Read (2004), Romer (2010)). This would lower the correlation between teenage and adult time-preferences for an individual, thus if there is an independent (of education) effect of time-preference on wages, this may not be correlated with education decisions made when 16. Furthermore, recent work by Grafova and Stafford (2009), examining the wage effects of smoking history in the US, suggest that conditional on education there is no difference in wages between those who never smoke and those who smoke earlier in life but then give up. Anger and Kvasnicka (2010) for Germany and Brune (2007) for the UK find analogous results. This also suggests that – at least for those who later quit – early smoking does not indicate unobservable characteristics and investments that will influence later wages. The substantial degree of movement between smoking status at 16 and smoking status when an individual is observed in the data also implies that the correlation between early smoking and later unobserved characteristics is not clear cut.

Finally, given that we have more than one instrument the system is over identified and we can exploit this to econometrically test for a correlation between early smoking and the unobservables in the wage equation. Section 7 details the results of numerous approaches to investigating such a correlation, each of which support the conclusion that early smoking ***is*** a valid instrument for education[8].

If we accept that early smoking satisfies these two criteria of relevance (and non-weakness) and validity then an indicator for early smoking can be used as an instrument: influencing

[8]Bonjour *et al* (2003) use data from St. Thomas' Hospital's Adult Twins Registry to estimate returns to education, and also exploit information on early smoking behaviour to consider whether this can be used as a valid instrument. Bonjour *et al* conclude that early smoking is picking up family background rather than time-preference, however their sample is comprised almost exclusively of female twins. Though it may be the case that for young females – where historically smoking prevalence has been lower – early smoking *is* more of a reflection of background than time-preference, this is less likely to be the case for young men, moreover the background/ability arguments are tackled below.

schooling through changing the marginal costs of schooling in a way which is conditionally uncorrelated with the unobservables.

Furthermore, though it captures a LATE – the return for a group of individuals who have lower education because of a higher than average discount rate – it is shown below that this is a group comprised of individuals across the education distribution and of all abilities, thus the estimate is arguably more representative of the return for the population as a whole than estimates identified by alternative IV estimation strategies.

4 Data and Estimation

The data comes from the British Household Panel Survey (BHPS) which is a nationally representative survey of the population which began in 1991 and follows the sample individuals each year. In 1999 in addition to the core survey there was a supplementary component in which questions were asked regarding previous health habits. This allows the creation of a 15-wave pooled-panel dataset containing variables describing individuals' characteristics, a dummy to indicate whether the individual smoked when 16, education, and current hourly wage rate. Since the previous health habits question was only asked in wave 9, the constructed panel only contains observations from individuals present in that wave, but for these individuals all of their available observations are included. The sample includes males who are in full-time employment (30+hours per week), are not self-employed and are in the age range 19 to 65 inclusive[9].

There are issues of measurement error when using number of years of schooling as the measure of education, however in order to make the results comparable with the majority of the literature the number of years of schooling is used as the education variable[10]. The BHPS does not ask how many years education an individual has nor when the individual

[9]This age range captures 'prime-age' males and ensures that smoking at 18 is not the same as current smoking for any individuals, as smoking at 18 will be used as an instrument in a robustness check.

[10]Formally: Years-of-schooling = (age left education - 5); thus we assume a school start age of 5, which is the compulsory school start age in the UK.

first left full-time education, rather it asks the age at which the individual left school and age at which he/she left further education. This creates a potential problem when people return to full-time education after a number of years away. If an individual completes GCSEs, A-levels, a standard 3-year degree, then a Masters degree and then a PhD (3 years) this would equate to 21 years of education, therefore any individual with more than 21 years recorded education is excluded. This removes observations from just 84 individuals (3.6% of those with years of schooling calculated)[11]. With respect to earnings, it is standard to use the log of hourly earnings and so again for comparability this is what is used – the log of real wage (using 2006 pounds as the base)[12].

The pooled-panel dataset constructed contains 21,256 observations from 2,266 males with each individual having between 1 and 15 observations; the mean number of observations per individual is 9.38, median 10. In all of the regressions standard errors are clustered at the level of the individual[13]. Appendix A Table A-1 contains summary statistics for the estimation sample, with the breakdown by early smoking status in Table A-2.

We begin by estimating the conventional human capital earnings function where the dependent variable is the natural log of real hourly wage, and the explanatory variables are age, age-squared, and years-of-schooling[14]. Also included are controls for ethnicity, for region (using the 10 standard regions) in order to pick up regional real wage differentials, year-of-birth[15] and its square to pick up cohort effects[16] and dummies for parental characteristics.

[11]The results are robust to an alternative assumption of recoding such that anyone with education greater than 21 years education is recorded as having 21 years of education.

[12]The log wage distribution is trimmed such that the top and bottom 1% within each year are excluded.

[13]In order to avoid issues around differential attrition, we have re-estimated the models using both inverse probability weighting and also including in the regressions a variable indicating the number of observations that each individual has, and in each case the results remain. As the first stage involves regression of years-of-schooling – which is time-invariant – on characteristics, we re-estimate the model using just one observation (their first) for each member of the sample but then all of the observations in the second stage, bootstrapping to get the correct standard errors in each stage. There is no substantive change in the conclusions. Similarly the models can be estimated on any single wave and the nature of the results does not change. See Appendix D, Tables D-1 to D-5.

[14]Mincer's specification included experience and experience-squared however in the absence of information on labour market experience, age and age-squared are routinely used instead hence this is the approach taken.

[15]Year-of-birth is rescaled such that 1897=1,..., 1989=93, since in the range 1897-1989 the birth years in the data, year-of-birth and year-of-birth-squared are perfectly collinear.

[16]Including a higher order polynomial in a suitably rescaled year-of-birth does not alter the results nor

The parental characteristics included are the standard occupational classification of the job of both the individual's father and mother when the individual is 14 years of age, and a dummy to indicate that the individual lived with both natural parents from birth up to the age of 16. These parental characteristics are included since in their absence the smoking at 16 variable could be picking up background characteristics correlated with education and smoking at 16. Including year dummies in the model would be problematic since both age and year-of-birth are included, however controls are included for whether it was the early-, mid-, late-1990s or post-2000 to allow for business cycle effects[17].

5 Results and Analysis

The first column of Table 1 reports the OLS estimate of the human capital earnings function, the third column reports the IV results using smoking at 16 as the instrument, with the fifth column reporting the results from the reduced form equation for years of schooling (the second, fourth and sixth columns contain the standard errors). Looking at the fifth column, we can see that individuals who smoke when they are 16 have on average 0.88 fewer years of schooling than those who do not. The standard error is 0.108 giving an absolute value of the t-statistic of 8.13. Therefore smoking when 16 is strongly significant for education, and the parameter precisely estimated. This is encouraging given the concerns raised by *inter alia* Staiger and Stock (1997) and Bound *et al.* (1995) concerning the precision of first stage estimates. The R^2 of 0.246 is higher than the R^2 for first stage regressions in some other IV studies[18] and the partial-R^2 of the effect of the instrument on years-of-schooling having partialled out the effect of the other covariates is 0.0289 which is high relative to the guidelines given by Bound *et al.* (1995).

Applying Stock and Yogo's (2005) formal test for weak instruments, the Kleibergen-Paap

add to precision in the estimates and so in the interests of parsimony only a quadratic is used.

[17]These dummies are significant in the wage equation, though their inclusion/exclusion does not alter the coefficient on the instrument (1st stage) or $\hat{S}_i$ in the second stage.

[18]Harmon and Walker (1995) for example have a first stage R^2 of 0.147.

rk Wald F-statistic of 66.167 suggests a very strong instrument: the critical value is 16.38, hence the overwhelmingly indication is that there is not a problem of weak identification introducing bias to the coefficient on years of schooling[19].

Turning to columns 1 and 3, the OLS estimate suggests that an additional year of schooling increases wage by 4.6% whereas the IV estimate of the return is 12.9%. We expect that the IV results will be less precisely estimated than the OLS, and while the standard error on years of schooling in the instrumented regression does increase, the coefficient remains precisely estimated and significant at all conventional levels. The large difference in the estimated coefficients suggests that years of schooling is endogenous, and this conclusion is strengthened if the residual from the first stage reduced form equation is included as a regressor in the OLS regression, providing a Hausman test of the endogeneity of schooling. The absolute value of the t-statistic on this residual is 4.78. Estimation of the IV using the Fuller-LIML estimator rather than standard 2SLS-IV, in order to be as robust as possible to any potential bias in the IV estimates, does not result in any substantive change to the estimated coefficients or standard errors: the return to schooling in the IV estimation remains 12.9, st. err. of 0.020 (see Appendix C, Table C-1).[20].

The estimates of a 4.6% return by OLS rising to 12.9% by IV are in line with those found in majority of other studies using UK data, particularly Harmon and Walker (1995, 1999):

[19]Critical values relevant when standard errors are clustered have not (at time of writing) been tabulated, however Baum *et al.* (2007) suggest applying with caution the critical values for the *i.i.d.* errors case, or alternatively falling back on the original Staiger and Stock (1997) rule-of-thumb that the F-statistic should be 10 or more. In either case the null here is soundly rejected.

[20]The modified LIML estimator introduced by Fuller, with the Fuller parameter, a, set to 1 is regarded as most robust to any potential weakness of the instrument.

	OLS	IV	Instrument
Harmon and Walker:			
(1995) using FES data	6.1%	15.3%	RoSLA (1947 and 1972)
(1999) using GHS data	4.9%	14.0%	RoSLA (1972) and educational reforms
(2000) using NCDS data	5.0%	9.9%	peer effects and education system level effects
Chevalier and Walker:			
(2002) BHPS (6 waves)	6.4%	20.5%	RoSLA (1972)
(2002) NCDS	6.1%	8.0%	smoking at 16
(2002) GHS	6.4%	9.5%	smoking at 14/16/18
Devereux and Hart:			
(2010) GHS	7.0%	4.9%	RoSLA (1947) RDD estimate
(2010) NESPD	N/A	3.7%	RoSLA (1947) TSTSLS design
Grenet:			
(2009) LFS	9.3%	7.7%	RoSLA (1972) RDD estimate

More recently Devereux and Hart (2010) have used both the GHS and the New Earnings Survey Panel Dataset to re-assess returns to education estimates using the 1947 raising of the school leaving age in England from 14 to 15. They implement a Two Sample Two Stage Least Squares approach and contrary to Oreopoulos's original (2006) GHS results they find very low returns for men, in the order of 3 to 4%[21]. Devereux and Hart suggest that the low return may be due to the leaving age of 15 not coinciding with any exit exams; moreover this reform affected a cohort of individuals a generation older than those in our data. Grenet (2009) exploits the 1972 RoSLA and using Labour Force Survey data finds a return for men in the range of 7-8%. The 1972 RoSLA IV estimate we find in the BHPS is 10.2%, which is slightly above Grenet's estimate, with the early smoking estimate higher still, however this is in line with the early smoking IV capturing the return for a wider range of individuals – not just those at the lower end of the schooling distribution.

[21]Oreopoulos has himself now revised his original IV and RDD estimates – which ranged from 9.4% to 15.8% depending on the sample rules – downwards to 6–7%, consistent with the findings of Devereux and Hart, see Oreopoulos (2008).

5.1 Testing for a spurious relationship

5.1.1 Is it a background effect?

This observed relationship between smoking at age 16 and educational attainment could be driven by something other than rate of time preference, something that also affects wages and therefore invalidates the instrument. It could be argued for example, that poorer socio-economic background lowers education and increases the likelihood of smoking. Clearly socio-economic background may influence the decision to smoke at 16, hence the need to control as much as possible for individuals' socio-economic characteristics from the time when they were a teenager and were making their decisions over education and smoking. The main specification of the model includes such variables and thus should take this effect out of the coefficient on the early smoking indicator. As a robustness check we also estimate the model without the parental characteristic variables, see Table 5. The effect of removing these background characteristics is that the magnitude of the coefficient on the smoker at 16 indicator in the schooling demand equation increases slightly to -1.08 (s.e. 0.113, t =-9.61). The F-statistic on the exclusion of the instrument is 92.39, with the partial-R^2 of the instrument of 0.0400, again both suggesting a strong instrument[22]. The estimated return to education in the wage equation is 12.1% (s.e. 0.016, t =7.56). Thus with the nature of the estimate remaining unchanged, it is clear that the inclusion/exclusion of parental characteristics is not driving the results. Moreover, if we are worried that the early smoking variable is picking up background, the implication would be that the estimate absent the background controls is upward biased. Without the controls the estimate goes down rather than up which implies that early smoking is not simply picking up the effect of these background factors.

A further concern may be that teenage smoking is influenced by peer group behaviour and that peer group effects may also translate into choices that affect later wages, imparting

[22]Clearly with the F-statistic even higher than before, the Stock-Yogo formal tests of weak identification continue to overwhelmingly suggest a strong instrument.

a positive bias to our estimate. Evidence from the peer effects literature suggests that peers' behaviour influences teenage smoking decisions (see Gaviria and Raphael, 2001, Nakajima, 2006, Clark and Lohéac, 2007), however research using UK data also provides support to an alternative view: that individuals choose to smoke and then choose like-minded peers (see De Vries et al, 2003). Ideally information on teenage peer group would be available in order to consider this channel of influence – though even then the endogeneity of peer group formation and smoking initiation would remain a problem – but unfortunately such variables are not available in our data. However, to the extent that children from similar family backgrounds form peer groups, some peer effects will be controlled for by the inclusion of parental background characteristics. Again, the fact that the point estimate of the return to education increases when parental characteristics are included indicates that absent these controls the estimate was not upward biased, which is reassuring.

5.1.2 Is it an ability effect?

The correlation between smoking and education is also consistent with an alternative hypothesis: that those with lower unobserved ability will acquire less education and are more likely to smoke. If it is the case that we are primarily picking up some measure of ability then we would expect that – by definition – smoking at 16 only affects the education of individuals at the lower end of the ability distribution. However if we assume that the residual from the OLS log wage regression is a reasonable proxy for ability, we can divide this residual distribution into quintiles and examine whether smoking at 16 is a feature only of low ability (low residual) individuals or if it is something that individuals of all abilities engage in.

Table 4 shows the numbers who smoke at age 16 in each quintile of this residual log wage distribution. The left-side panel of the table shows that in the lowest quintile approximately 44% of the males smoked at 16. This figure falls to approximately 39% in the next quintile up and the next after that (30%) before rising again in the fourth quintile (34%). Despite a fall in the last quintile, the figure for the percentage of individuals who smoked at age 16

is still as high as 23% in the highest quintile of the residual log wage distribution. There are fewer smokers at 16 in the higher quintiles of the distribution, nevertheless there remain substantial numbers of smokers at 16 in the highest quintiles which indicate the highest ability individuals.

In addition, Figure 2 plots the density of education leaving age for smokers at 16 and non-smokers at 16. If it was only low educated, low ability individuals who smoke at 16 then we would expect the densities to look very different with very little mass in the upper ranges for the early smokers. However, while the density for non-smokers at 16 does have a greater mass around 21 and less around 15/16 (suggesting more non-smokers go to university), it is quite close to being a general right-ward shift of the distribution compared with the smokers at 16. This is consistent with the idea that A-levels are the main route into university – we would expect more lower discount rate individuals to remain in school at 16 and the result of this is the lower percentage leaving at 16 and the resulting higher percentage leaving at around 21. Elsewhere the picture is very similar but with the smokers at 16 distribution to the left of the non-smokers. This all fits with the discount rate hypothesis which says that there are smokers and non-smokers at 16 of all abilities and that smoking at 16 has an effect to reduce education at all points of the distribution[23].

If we continue to use the log wage residual distribution as a proxy for ability and, again dividing it into five quintiles, look at the first stage reduced form schooling equations, we can see that the effect of smoking at 16 is actually increasing as we move up the distribution. The left side of Table 8 shows that in the lowest quintile, schooling is reduced by 0.77 years, this is equivalent to a reduction of 6.21% of the mean number of years of education in this group. In the second and third quintiles the reduction in education associated with early smoking is even greater both in absolute terms and relative to mean education in these quintiles. The fourth quintile is affected the least by early smoking but still it is associated with three-

[23]It is true that younger cohorts have been acquiring more education and smoking less which would also lead to a shift of the curve to the right for non-smokers at 16, however Figure 3 repeats Figure 2 by cohort, revealing the same pattern in all cohorts.

quarters of a year less education, and in the highest quintile the estimated reduction is 0.88 years, 6.9% of mean education in this quintile. Table 4 shows that there are significant numbers of individuals who smoke at 16 in all of the quintiles thus these results are not due to small numbers of early smokers, and the coefficient on smoking at 16 is significant at the 1% level in all quintiles. Far from only affecting the low ability individuals, this evidence indicates that smoking at 16 has a greater absolute and relative effect on the highest ability individuals.

To further pursue the hypothesis that individuals who have lower ability are likely to get less education and are more likely to smoke, the results are replicated using smoking at age 18 rather than smoking at age 16. Age 18 is the point at which individuals in the UK have to decide whether to remain in education and go to university, and this decision is likely to be affected by their rate of time preference. Moreover, it is more difficult to argue that smokers at 18 are more likely to be lower ability than higher ability individuals. The right panel of Table 4 shows the numbers who smoke at age 18 in the quintiles of the log wage residual distribution. The table illustrates that in the lowest quintile the smokers at 18 out number non-smokers (54% v 46%), and this remains the case in the next quintile up (52% smokers v 48% non). As with smoking at 16, the numbers who did smoke are generally lower as we move up the quintiles yet in the highest quintile, still as much as 35% of the individuals smoked at 18. There are a higher number of individuals who smoked at 18 in the upper quintiles than in the corresponding table for smoking at 16, indeed in each quintile there are more smokers at 18 than there were at 16, at least a 10 percentage point swing to smokers from non-smokers compared with the age 16 measure. This supports the idea that teenage smoking is a habit that high discount rate individuals of all abilities engage in.

Results using smoking at 18 as the instrument are displayed in Table 9. Looking at the fifth column, the reduced form equation for schooling, smoking at 18 reduces education by 0.75 years. This is lower than the corresponding reduction associated with smoking at 16 but this is consistent with the time preference story: smokers at 18 have a higher discount rate

than *non-smokers* at 18 but, *ceteris paribus*, *smokers at 16* will have a higher discount rate than *smokers at 18*. If smokers at 18 have a lower discount rate relative to those who smoke at 16, they will remain in education longer thus we expect that the reduction in education for smokers at 18 is not as much as it is for smokers at 16. The robust standard error on smoking at 18 is 0.108, giving a t-statistic with an absolute value of 6.93, therefore the parameter remains precisely estimated. The first stage regression is very similar to the first stage regression when using smoking at 16. The R^2 for this first stage regression is 0.242 so again high relative to other studies' findings and the Kleibergen-Paap rk Wald F-statistic of 48.025 again rejects even a hint of weak identification.

Turning to column 3, the estimated return to schooling when we instrument with smoking at 18 is slightly higher at 13.5% than the corresponding figure using smoking at 16 (12.9%). The parameter remains precisely estimated, with a standard error of 0.023 (t =5.76).

As the results using smoking at 18 are very similar to those using smoking at 16, and given the distribution of smokers at 16 and 18 throughout the wage distribution, this is evidence to support the hypothesis that early smoking is picking up the time preference of the individual rather than being a proxy for ability. Moreover, if smoking at 16 was capturing ability then the IV estimate would be biased up and we would expect smoking at 18 to give a lower IV estimate. While the estimates using smoking at 16 and 18 are not statistically different to each other, the point estimate using smoking at 18 is in fact higher. Estimates using smoking at 17 again give similar results (see Table 10), with the size of the first stage effect of smoking at 17 consistent with the time preference explanation of the effect of smoking at different ages on education choice, and the IV estimate of the return to education again suggesting that using smoking at 16 as an instrument does not upwardly bias the estimate.

6 Instrumenting Using RoSLA

We now turn to examining how this broader based IV estimate compares with an estimate derived using RoSLA. The minimum school leaving age was raised in England and Wales from 15 to 16 in 1972 such that if an individual was 16 by the end of August 1973 he/she was allowed to leave school in the June of 1973, while if the individual was only 15 at the end of August 1973 he/she would have to remain another year at school. Thus those born after August 1957 face a minimum school leaving age of 16. In Scotland this reform took place in August 1976 therefore individuals born after August 1960 face a minimum school leaving age of 16[24].

The sample is split almost exactly in half with 51.5% of individuals facing a minimum leaving age of 16. Including a quadratic in year-of-birth means that the smooth changes in schooling resulting from younger cohorts generally gaining more education are controlled for, while the identification derives from the discontinuity in years of education induced by the RoSLA – a well established IV strategy. For those born in 1958, the first year in which all individuals are affected by the law change, the proportion of individuals leaving school at age 15 or earlier is 1.9%, a fall from 17.4% the year before. The reduced form for log wage, including an indicator for the individual facing a minimum school leaving age of 16, shows that the raising of the school leaving age is associated with a statistically significant increase in log wage, see Appendix B, Table B-2.

Table 2 displays the RoSLA estimates: column 5 contains the first stage regression results, column 3 the returns to schooling estimate instrumenting using RoSLA (while column 1 contains the OLS results for comparison).

Looking first at column 5, the raising of the school leaving age is associated with an increase in education of 0.564 years and the coefficient is precisely estimated. Again, it is notable that the R^2 (0.227) is higher than has been found in similar studies. The partial-R^2

[24]The minimum school leaving age was raised from 14 to 15, in 1947 for England and Wales, 1946 for Scotland, however, in the estimation sample there are only 73 individuals (3.22%) who face a minimum school leaving age of 14 hence we have concentrated on the later change to create an instrument.

for the instrument in the first stage is 0.0044 which is smaller than for the early smoking instrument but compares well with Harmon and Walker (1995) for their first stage, and with Bound *et al.* (1995). The F-statistic on the exclusion of the instrument from the first stage is 7.49. While this is below Staiger and Stock's (1997) rule-of-thumb guide of 10, taken with the partial R^2, the overall picture is not of a weak instrument[25]. The size of the average increase in education, controlling for other covariates in the first stage, is comparable with that found in other RoSLA studies (see Harmon and Walker (1995), Oreopoulos (2006)).

Turning to column 3, the estimated return to schooling is 10.2% when we instrument using RoSLA. This is more than double the size of the OLS return though below the other IV estimate. However it is not as precisely estimated, the standard error is 0.051 giving a t-statistic of 1.99, significant at the 5% level.

Table 6 displays the results for the specification excluding parental characteristics. The effect on the estimated return to education is minor – reducing from 10.2% to 10.0% (s.e. 0.042, t=2.41), signficant at the 5% level. More importantly, in this specification the F-statistic on the exclusion of the instrument from the first stage increases to 9.98 thus almost exactly attaining Staiger and Stock's threshold for a non-weak instrument. As the estimated coefficient on years of schooling is almost identical to the main specification case, when the F-statistic was only 7.49, this suggests that there is no weak instrument bias in the estimated coefficient on years of schooling in the main specification.

The question is whether the group whose return to education is captured when instrumenting using RoSLA is comprised mainly of those with low ability or is it mainly those with high discount rate particularly because of financial constraints? *If* the group whose return is identified by the RoSLA instrument (which is by definition a low education group) is in the main comprised of low ability individuals then we would expect that the return for this group would be lower than the return we find with the smoker at 16 instrument – as individuals of all abilities are in the early smokers group. The imprecision of the estimate

[25]Using the Fuller(1) estimator – which is the most robust to the presence of a potentially weak instrument – the result is almost identical, see Appendix C, Table C-2.

using RoSLA does not allow us to conclude that it is definitely smaller than the smoking at 16 IV estimate, however one test of the extent to which RoSLA affects individuals of different abilities is to repeat the first stage regressions by quintile of the log wage residual ("ability") distribution. The results from these regressions are in right hand section of Table 8.

The table shows that the raising of the minimum school leaving age increases the number of years of schooling by 1.04 years in the lowest quintile, which is 8.4% of the mean number of years schooling for this group. Being almost exactly one year extra education this suggests that in this lower quintile of the (proxy) ability distribution, all the individuals wished to leave school at the minimum age. In the second lowest quintile RoSLA increases the number of years of schooling by 0.84 years which is 6.9% of the mean for this group. In the three quintiles above this the increase in education associated with RoSLA is much smaller in absolute and relative terms than in both of the lowest two quintiles but in none of these higher quintiles is the dummy for minimum school leaving age of 16 close to being statistically significant. This is consistent with the hypothesis that the low education group affected by RoSLA are generally lower ability – if they were mainly high discount rate then we would expect to see a similar effect across the log wage residual distribution.

Though the imprecision of the RoSLA IV estimate prevents a concrete conclusion that it is indeed lower than the early smoking IV estimate, comparison of the two suggests that the RoSLA group are lower ability. This, and the results from looking at where in the proxy ability distribution each instrument is working, supports the contention that it is more appropriate to generalise from the early smoking IV estimate to the rest of the population: as unlike RoSLA, the estimate is not capturing a LATE that is only a lower ability group.

7 Testing the Instruments

Having more than one instrument allows a more formal test of whether the exclusion restrictions are valid. RoSLA and early smoking provide two independent sources of exogenous variation in education and allow us to genuinely test the validity of the exclusion restrictions. The Hansen J-test is more compelling when one of the instruments is definitely thought to be valid, and there are strong arguments that the RoSLA instrument is valid (it was a policy change exogenous to the individual).

Instrumenting using both early smoking and RoSLA (see Table 3) and then performing the Hansen J-test results in a test statistic of 0.202, p-value 0.6529, which is a comprehensive failure to reject the null hypothesis that the instruments are valid. The first stage R^2 is high at 0.250 and the F-statistic on the exclusion of the instruments is 36.83 with a partial R^2 on the instruments of 0.0332, all of which suggests that the instruments are strong as well as valid[26].

The Hansen J-test provides formal support for the validity of the early smoking instrument, which earlier evidence has shown to be a strong instrument. Furthermore, in order to re-enforce the evidence of the Hansen test, it can be decomposed to illustrate directly the validity of the early smoker instrument specifically: by using the RoSLA instrument to just identify the system of equations and then taking the valid estimates of the error from the structural equation and regressing them on the early smoker instrument. The results of such an exercise are contained in Appendix Table B-3. As can be seen, there is no relationship between the residuals from the structural equation and the early smoker indicator.

Finally, an alternative IV regression can be run in which RoSLA is used as the identifying instrument, while the early smoker indicator is included as one of the X variables. Appendix Table B-4 shows the coefficient estimates when this exercise is carried out. While early smoking affects education in the first stage (with a coefficient almost identical to the specifications

[26]The Kleibergen-Paap *rk* Wald F-statistic indicates that the Stock-Yogo tests of weak identification are easily passed; moreover, using the weak-instrument robust Fuller(1) LIML estimator, the results are almost identical, see Appendix C, Table C-3.

in which it is used as an instrument) it is completely insignificant in the structural equation. This evidence supports the contention that teenage discount rate – as captured by early smoking – affects choice of education, conditional on the other variables in X, but then has no further independent effect on log wage. All of these results suggest that early smoking is both a strong ***and*** valid instrument.

Returning to the estimation results when using both RoSLA and early smoking as instruments, Table 3 shows that the coefficient on each instrument in the first stage is almost identical to the case when the instruments are used separately (*cf* Tables 1 and 2), and the estimated return to education using both instruments together is 12.5%, significant at the 1% level. The standard error (0.019) is lower than is the case when either of the instruments are used singly, so the extra variation in schooling that comes with using both instruments results in a more precise estimate of the IV return to education, as we would expect[27]. The problem with this strategy is that using both instruments makes the interpretation less straightforward: exactly whose return the resulting LATE estimator captures is unclear. Given that the effects of each instrument in the first stage are similar to their impacts when used separately, and that the early smoking instrument is the stronger, the resulting IV estimate is very close to the return for individuals with high discount rates because of their rate of time preference.

8 Conclusions

This paper presents three IV estimates of the return to education: the early smoking estimate of 12.9%, the RoSLA estimate of 10.2% and the combined estimate of 12.5%, all of which are statistically significant and considerably higher than the OLS estimate of 4.6%. The analysis looking at the effects on different quintiles of the proxy ability distribution, suggests that the RoSLA IV captures the return for lower-ability individuals who wanted to leave at the

[27]For completeness, Table 7 displays the estimates when instrumenting using both instruments and just the basic set of controls – as with the other specifications, excluding background characteristics results in no substantive change to the results.

minimum leaving age but were forced to stay longer. In contrast, early smoking is a behaviour engaged in by individuals of all abilities and education levels, who have high discount rates because they have a rate of time preference that reflects a preference for the present. This is important as, unlike the RoSLA IV, it implies that the smoking IV captures the average marginal return to education for a group comprised of individuals (a) at different points in the education distribution, and (b) with different levels of ability. Hence the IV estimate derived is closer to an average marginal return to education, purged of the bias of OLS, and is thus more appropriate for drawing inference about the return to education in the population as a whole. Various tests indicate that early smoking is picking up time-preference as opposed to ability or background and importantly, exploiting the over-identification, further tests suggest that early smoking is uncorrelated with the wage equation error and is therefore a valid instrument for education.

There remains the question of why the OLS estimates are consistently found to be below IV estimates – irrespective of the instrument chosen – when measurement error in standard micro surveys could only sensibly account for a relatively small attenuation in the OLS coefficient. Moreover it appears from this study that a negative 'discount rate bias' is not a major factor biasing the OLS estimates downwards. The 'discount rate bias' story suggests that a high discount rate will reduce education and might also independently increase wages. However, testing for a correlation between the discount rate (as captured by early smoking) and the wage equation error term suggests no relationship, hence it does not appear that 'discount rate bias' is the factor biasing the OLS estimates downwards.

We believe that the results that this and other IV studies find can be reconciled when we consider the assumptions imposed by Mincer's human capital earnings function: specifically, that each additional year of schooling has the same proportional effect on earnings. If we consider that rather than being linear, the education–log earnings profile is actually concave, then it seems that the instruments that are commonly used isolate a 'local average treatment effect' for groups of individuals who are primarily located at points in the education

distribution at which there is a higher average return than the global average estimated by OLS. The individuals affected by RoSLA may be of lower ability, however if all individuals have a higher marginal return to schooling at lower levels of schooling then this is consistent with the estimate from the RoSLA IV being higher than the OLS estimate.

Similarly, the early smoking IV estimate is a LATE and while it uses variations across a broader range of the education distribution, Figure 2 indicates that the identified treatment effect is weighted towards individuals located at points in the mid-to-lower education range, which are regions where the return is higher on average than the global average estimated by OLS. Compared with the RoSLA return, the early smoking IV estimate derives from individuals with a greater range of abilities, hence we would expect it to be a higher return. However, it also derives from a range of education margins, some of which will have a lower return than the return at the minimum schooling margin identified by RoSLA, hence it is perhaps unsurprising that the aggregate effect is an estimate close to that from the RoSLA IV. It needs also to be noted that while significant, the imprecision of the RoSLA IV coefficient in particular means that the conclusion that the RoSLA estimate is indeed below the early smoking IV estimate must remain tentative.

Given the consistent evidence from these and other IV estimates of the return to education, it seems that the linearity in returns assumption of Mincer's human capital earnings function is the principal reason why we consistently find lower OLS estimates.

Support for this conclusion comes from implementing the OLS regression *only* for those who left school at their contemporary minimum leaving age. The estimated return to education when doing this is 19.7%. Though the endogeneity of years of schooling in this regression is not dealt with, the much greater coefficient on years of schooling does suggest that when estimating over the entire range of education levels, the linearity in returns assumption contributes significantly to the lowering of the OLS coefficient.

From a policy perspective, the evidence here implies that education continues to be a worthwhile investment, which supports the UK Government's policy of raising of the

education leaving age to 17 (in 2013) and then up to 18 (by 2015) in England and Wales. Moreover it is not only those located at the high school margin who receive a high return – the average over a broader range of the distribution is also higher than OLS would suggest. Even allowing for IV's imprecision we can conclude the return for males in Britain is likely to be closer to 10% than 5%.

References

Angrist, J. D. and Krueger, A. B. (1991). 'Does compulsory schooling attendance affect schooling and earnings?', *Quarterly Journal of Economics* Vol 106, pp. 979-1014.

Anger, S. and Kvasnicka, M. (2010). 'Does smoking really harm your earnings *so* much? Biases in current estimates of the smoking wage penalty', *Applied Economic Letters*, Vol. 17, pp. 561-564.

Baum, C. F., Schaffer, M. E. and Stillman, S. (2007). *Enhanced routines for instrumental variables/GMM estimation and testing*, Working Paper no. 667, Boston College Economics, Boston, MA.

Becker, G. S. (1964). *Human capital: A theoretical & empirical analysis with special reference to education*, Columbia University Press, New York.

Becker, G. S. and Murphy, K. M. (1988). 'A theory of rational addiction', *Journal of Political Economy*, Vol. 96, pp. 675-700.

Bonjour, D., Cherkas, L. F., Haskel, J. E., Hawkes, D. D. and Spector, T. D. (2003). 'Returns to Education: Evidence from UK Twins', *American Economic Review*, Vol. 93, pp. 1799-1812.

Bound, J., Jaeger, D. A. and Baker, R. (1993). *The cure can be worse than the disease: A cautionary tale regarding instrumental variables*, Technical Working Paper no. 0137, National Bureau of Economic Research, Cambridge, MA.

Bound, J., Jaeger, D. A. and Baker, R. (1995). 'Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak', *Journal of the American Statistical Association*, Vol. 90, pp. 443-450.

Brune, L. (2007). 'The smoker's wage penalty puzzle – evidence from Britain', Working paper no. 2007-31, ISER, Univeristy of Essex, Colchester, UK.

Burgess, S. and Propper, C. (1998). 'Early health-related behaviours and their impact on later life chances: evidence from the US', *Health Economics*, Vol. 7, pp. 381-399.

Card, D. (1994). *Earnings, schooling and ability revisited*, Working Paper no. 4832, National Bureau of Economic Research, Cambridge, MA.

Card, D. (1995). 'Using geographic variation in college proximity to estimate the return to schooling', in Christofides, L. N., Grant, E. K. and Swidinsky, R. (eds), *Aspects of Labour Market Behaviour: essays in honour of John Vanderkamp*, University of Toronto Press, Toronto.

Card, D. (1999). 'The causal effect of education on earnings', in Ashenfleter, O. C. and Card, D. (eds), *Handbook of Labor Economics vol. 3A*, Elsevier, Amsterdam.

Card, D. (2001). 'Estimating the return to schooling: Progress on some persistent econometric problems', *Econometrica*, Vol. 69, pp. 1127-1160.

Chevalier, A., Harmon, C., Walker, I. and Zhu, Y. (2004). 'Does education raise productivity, or just reflect it?', *Economic Journal*, Vol. 114, pp. F499-F517.

Chevalier, A. and Walker, I. (2002). 'Further estimates of the returns to education in the UK', in Harmon, C., Walker, I. and Westergard-Nielsen, W. (eds) *The Returns to Education Across Europe*, Edward Elgar.

Clark, A. and Lohéac, Y. (2007). '"It wasn't me, it was them!" Social influence in risky behaviour by adolescents', *Journal of Health Economics*, Vol. 26, pp. 763-784.

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

De Vries, H., Engels, R., Kremers, S., Wetzels, J. and Mudde, A. (2003). 'Parents' and friends' smoking status as predictors of smoking onset: findings from six European countries', *Health Education Research*, Vol. 18, pp. 627-636.

Evans, W. N. and Montgomery, E. (1994). *Education and health: where there's smoke there's an instrument*, Working Paper no. 4949, National Bureau of Economic Research, Cambridge, MA.

Farrell, P. and Fuchs, V. R. (1982). 'Schooling and health: the cigarette connection', *Journal of Health Economics*, Vol. 1, pp. 217-30.

Fersterer, J. and Winter-Ebmer, R. (2003). 'Smoking, discount rates and the return to education', *Economics of Education Review*, Vol. 22, pp. 561-566.

Fuchs, V. R. (1982). 'Time preferences and health: An exploratory study', in Fuchs, V. R. (ed) *Economic Aspects of Health*, University of Chicago Press, Chicago.

Fuller, W. A. (1977). 'Some properties of a modification of the limited information estimator', *Econometrica*, Vol. 45, pp. 939-954.

Gaviria, A. and Raphael, S. (2001). 'School-based peer effects and juvenile behavior', *The Review of Economics and Statistics*, Vol. 83, pp. 257-268.

Griliches, Z. (1977). 'Estimating the returns to schooling: some econometric problems', *Econometrica*, Vol. 45, pp. 1-22.

Grafova, I. and Stafford, F. P. (2009). 'The Wage Effects of Personal Smoking History', *Industrial & Labor Relations Review*, Vol. 62, pp. 381-393.

Grossman, M. (2006). 'Education and nonmarket outcomes', in Hanushek, E. and Welch, F. (eds) *Handbook of the economics of education, Volume 1*, Elsevier, Amsterdam.

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

Harmon, C. and Walker, I. (1999). 'The marginal and average return to schooling in the UK', *European Economic Review*, Vol. 43, pp. 879-887.

Harmon, C. and Walker, I. (2000). 'The returns to the quantity and quality of education: Evidence for men in England and Wales', *Economica*, Vol. 67, pp. 19-35.

Imbens, G. W. and Angrist, J. D. (1994). 1Identification and estimation of local average treatment effects', *Econometrica*, Vol. 62, pp. 467-475.

Lang, K. (1993). *Ability bias, discount rate bias and the return to education*, Unpublished manuscript, Department of Economics, Boston University, Boston, MA.

MacDonald, Z. and Pudney, S. (2000). 'The wages of sin? Illegal drug use and the labour market', *Labour*, Vol. 14, pp. 657-674.

Mincer, J. A. (1974). *Schooling, Experience and Earnings*, Columbia University Press, New York.

Munasinghe, L. and Sicherman, N. (2000). *Why do dancers smoke? Time preference, occupational choice and wage growth*, Working Paper no. 7542, National Bureau of Economic Research, Cambridge, MA.

Murray, M. P. (2006). 'Avoiding invalid instruments and coping with weak instruments', *Journal of Economic Perspectives*, Vol. 20, pp. 111-132.

Nakajima, R. (2007). 'Measuring peer effects on youth smoking behaviour', *The Review of Economics and Statistics*, Vol. 74, pp. 897-935.

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

Oreopoulos, P. (2008). *Estimating average and local average treatment effects of education when compulsory schooling laws really matter: Corrigendum*, Unpublished manuscript, Department of Economics, University of Toronto, Canada.

Read, D. and Read, L. (2004). 'Time discounting over the lifespan', *Organizational Behaviour and Human Decision Processes*, Vol. 94, pp. 22-32.

Romer, D. (2010). 'Adolescent risk taking, impulsivity, and brain development: implications for prevention', *Developmental Psychobiology*, Vol. 52, pp. 263-276.

Silverman, I. (2003). 'Gender difference in delay of gratification: a meta-analysis', *Sex Roles*, Vol. 49, pp. 451-463.

Staiger, D. and Stock, J. H. (1997). 'Instrumental variables regression with weak instruments', *Econometrica*, Vol. 65,pp. 557-586.

Steinberg, L. (2007). 'Risk taking in adolescence. New perspectives from brain and behavioural science', *Current Directions in Psychological Science*, Vol. 16, pp. 55-59.

Stock, J. H. and Yogo, M. (2005). 'Testing for weak instruments in IV regression', in Andrews, D. W. K. and Stock, J. H. (eds) *Identification and inference for econometric models: A festschrift in honor of Thomas Rothenberg*, Cambridge University Press, Cambridge.

Tables

Table 1: OLS and Early Smoking IV Returns to Education Estimates for Log Hourly Wage: full set of controls

	OLS		IV: smoker at 16		IV: first stage	
	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.
years of schooling	0.046***	0.003	0.129***	0.020	— —	— —
smoker at 16 indicator	— —	— —	— —	— —	−0.876***	0.108
age	0.099***	0.004	0.094***	0.005	0.056***	0.022
age^2	−0.001***	0.000	−0.001***	0.000	−0.001***	0.000
year-of-birth	−0.016**	0.007	−0.052***	0.011	0.398***	0.041
year-of-birth^2	0.000***	0.000	0.001***	0.000	−0.003***	0.000
region: North	0.047	0.038	0.054	0.044	−0.103	0.272
region: Yorkshire	0.003	0.033	−0.022	0.041	0.331	0.253
region: North West	0.054*	0.032	0.023	0.040	0.402	0.253
region: East Midlands	−0.010	0.032	−0.005	0.038	−0.034	0.235
region: East Anglia	0.015	0.039	−0.009	0.048	0.366	0.324
region: South East	0.142***	0.028	0.082**	0.037	0.757***	0.206
region: South West	0.023	0.034	0.015	0.041	0.175	0.237
region: Wales	−0.012	0.040	−0.019	0.045	0.081	0.285
region: Scotland	0.028	0.036	−0.021	0.044	0.643**	0.262
ethnicity: Black	0.114	0.105	0.115	0.117	−0.164	0.779
ethnicity: Asian	−0.136*	0.071	−0.312***	0.105	1.965***	0.485
ethnicity: Other	−0.048	0.103	−0.234**	0.119	2.067*	1.111
father's occ class: 1	0.116***	0.028	0.020	0.041	1.122***	0.214
father's occ class: 2	0.121***	0.038	−0.077	0.065	2.268***	0.291
father's occ class: 3	0.089**	0.043	−0.043	0.058	1.499***	0.321
father's occ class: 4	0.065*	0.036	−0.053	0.051	1.320***	0.305
father's occ class: 5	0.038*	0.023	0.011	0.028	0.335**	0.170
father's occ class: 6	0.014	0.035	−0.074	0.048	0.991***	0.305
father's occ class: 7	0.103***	0.040	0.066	0.049	0.467	0.330
father's occ class: 9	−0.021	0.029	0.028	0.035	−0.551***	0.197
father's occ class: 10	0.029	0.027	0.027	0.030	−0.012	0.186
mother's occ class: 1	0.047	0.049	0.035	0.061	0.112	0.411
mother's occ class: 2	0.015	0.054	−0.103	0.070	1.433***	0.439
mother's occ class: 3	0.056	0.048	0.053	0.057	0.046	0.387
mother's occ class: 4	0.055	0.040	0.014	0.048	0.485	0.307
mother's occ class: 5	0.010	0.049	0.031	0.058	−0.117	0.417
mother's occ class: 6	0.025	0.040	0.029	0.045	0.054	0.311
mother's occ class: 7	0.055	0.041	0.057	0.048	−0.083	0.312
mother's occ class: 9	−0.004	0.038	0.034	0.044	−0.461	0.284
mother's occ class: 10	0.004	0.032	−0.006	0.036	0.115	0.253
'nuclear family' to 16	0.028	0.019	0.001	0.022	0.247*	0.136
mid 1990s	−0.045***	0.009	−0.050***	0.010	0.067	0.046
late 1990s	−0.065***	0.014	−0.070***	0.016	0.080	0.081
post 2000	−0.033	0.021	−0.040*	0.023	0.108	0.126
constant	−0.754***	0.250	−0.607**	0.287	−0.471	1.664
# observations	21256		21256		21256	
# individuals	2266		2266		2266	
R^2	0.265		0.072		0.246	
F-test on exclusion of smoking at 16 from first stage: 66.17; Partial R^2 of instrument = 0.0289						

Notes: *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level
Standard errors are clustered at the level of the individual and robust.
'nuclear family' to 16 means lived with both natural parents from birth to age 16.
Reference categories: West Midlands, white, did not live with both natural parents to 16, father/mother occupational class 'plant/machine operative'.
Occupational Class dummies: (1) management, (2) professional, (3) associate professional/technical, (4) clerical/secretarial, (5) craft and related, (6) personal/protective services, (7) sales, (9) other, (10) self-emp/unemp.

Table 2: OLS and RoSLA IV Returns to Education Estimates for Log Hourly Wage: full set of controls

	OLS		IV: RoSLA		IV: first stage	
	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.
years of schooling	0.046***	0.003	0.102**	0.051	— —	— —
min. school LA=16	— —	— —	— —	— —	0.564***	0.206
age	0.099***	0.004	0.095***	0.005	0.056**	0.022
age^2	−0.001***	0.000	−0.001***	0.000	−0.001***	0.000
year-of-birth	−0.016**	0.007	−0.040*	0.023	0.427***	0.041
year-of-birth^2	0.000***	0.000	0.000**	0.000	−0.004***	0.000
region: North	0.047	0.038	0.051	0.041	−0.080	0.272
region: Yorkshire	0.003	0.033	−0.014	0.040	0.320	0.256
region: North West	0.054*	0.032	0.033	0.041	0.386	0.255
region: East Midlands	−0.010	0.032	−0.007	0.035	−0.035	0.234
region: East Anglia	0.015	0.039	−0.001	0.047	0.324	0.327
region: South East	0.142***	0.028	0.101**	0.051	0.741	0.208
region: South West	0.023	0.034	0.017	0.038	0.114***	0.240
region: Wales	−0.012	0.040	−0.017	0.043	0.093	0.290
region: Scotland	0.028	0.036	−0.005	0.050	0.658**	0.266
ethnicity: Black	0.114	0.105	0.114	0.110	0.037	0.746
ethnicity: Asian	−0.136*	0.071	−0.255*	0.139	2.146***	0.515
ethnicity: Other	−0.048	0.103	−0.174	0.152	2.214**	1.074
father's occ class: 1	0.116***	0.028	0.051	0.069	1.162***	0.216
father's occ class: 2	0.121***	0.038	−0.013	0.128	2.404***	0.298
father's occ class: 3	0.089**	0.043	0.000	0.093	1.585***	0.333
father's occ class: 4	0.065*	0.036	−0.015	0.083	1.440***	0.308
father's occ class: 5	0.038*	0.023	0.020	0.029	0.322*	0.172
father's occ class: 6	0.014	0.035	−0.046	0.064	1.046***	0.313
father's occ class: 7	0.103***	0.040	0.078	0.049	0.484	0.339
father's occ class: 9	−0.021	0.029	0.012	0.044	−0.592***	0.196
father's occ class: 10	0.029	0.027	0.028	0.028	0.043	0.186
mother's occ class: 1	0.047	0.049	0.039	0.056	0.107	0.426
mother's occ class: 2	0.015	0.054	−0.065	0.094	1.378***	0.454
mother's occ class: 3	0.056	0.048	0.054	0.053	0.007	0.395
mother's occ class: 4	0.055	0.040	0.027	0.050	0.453	0.317
mother's occ class: 5	0.010	0.049	0.025	0.054	−0.240	0.430
mother's occ class: 6	0.025	0.040	0.027	0.042	−0.070	0.322
mother's occ class: 7	0.055	0.041	0.057	0.044	−0.053	0.324
mother's occ class: 9	−0.004	0.038	0.022	0.047	−0.491*	0.293
mother's occ class: 10	0.004	0.032	−0.003	0.034	0.103	0.264
'nuclear family' to 16	0.028	0.019	0.010	0.026	0.330**	0.137
mid 1990s	−0.045***	0.009	−0.048***	0.010	0.063	0.047
late 1990s	−0.065***	0.014	−0.068***	0.015	0.075	0.083
post 2000	−0.033	0.021	−0.038*	0.023	0.094	0.129
constant	−0.754***	0.250	−0.655**	0.280	−1.459	1.681
# observations	21256		21256		21256	
# individuals	2266		2266		2266	
R^2	0.265		0.177		0.227	
F-test on exclusion of min. school LA=16 from first stage: 7.49; Partial R^2 of the instrument = 0.0044						

Notes: *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level
Standard errors are clustered at the level of the individual and robust.
'nuclear family' to 16 means lived with both natural parents from birth to age 16.
Reference categories: West Midlands, white, did not live with both natural parents to 16, father/mother occupational class 'plant/machine operative'.
Occupational Class dummies: (1) management, (2) professional, (3) associate professional/technical, (4) clerical/secretarial, (5) craft and related, (6) personal/protective services, (7) sales, (9) other, (10) self-emp/unemp.

Table 3: OLS and Both instruments IV Returns to Education Estimates for Log Hourly Wage: full set of controls

	OLS		IV: both		IV: first stage	
	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.
years of schooling	0.046***	0.003	0.125***	0.019	— —	— —
smoker at 16 indicator	— —	— —	— —	— —	−0.874***	0.107
min. school LA=16	— —	— —	— —	— —	0.556***	0.202
age	0.099***	0.004	0.094***	0.005	0.054**	0.022
age^2	−0.001***	0.000	−0.001***	0.000	−0.001***	0.000
year-of-birth	−0.016**	0.007	−0.050***	0.011	0.399***	0.041
year-of-birth^2	0.000***	0.000	0.001***	0.000	−0.004***	0.000
region: North	0.047	0.038	0.053	0.044	−0.097	0.272
region: Yorkshire	0.003	0.033	−0.021	0.040	0.347	0.253
region: North West	0.054*	0.032	0.024	0.039	0.409	0.253
region: East Midlands	−0.010	0.032	−0.006	0.037	−0.014	0.235
region: East Anglia	0.015	0.039	−0.008	0.048	0.398	0.325
region: South East	0.142***	0.028	0.084**	0.036	0.767***	0.207
region: South West	0.023	0.034	0.015	0.040	0.192	0.236
region: Wales	−0.012	0.040	−0.019	0.045	0.082	0.286
region: Scotland	0.028	0.036	−0.019	0.043	0.705***	0.263
ethnicity: Black	0.114	0.105	0.115	0.116	−0.114	0.788
ethnicity: Asian	−0.136*	0.071	−0.305***	0.103	1.975***	0.493
ethnicity: Other	−0.048	0.103	−0.226*	0.116	2.021*	1.080
father's occ class: 1	0.116***	0.028	0.024	0.040	1.118***	0.213
father's occ class: 2	0.121***	0.038	−0.068	0.062	2.271***	0.290
father's occ class: 3	0.089**	0.043	−0.038	0.056	1.485***	0.319
father's occ class: 4	0.065*	0.036	−0.048	0.050	1.324***	0.303
father's occ class: 5	0.038*	0.023	0.012	0.027	0.322*	0.170
father's occ class: 6	0.014	0.035	−0.070	0.046	0.968***	0.303
father's occ class: 7	0.103***	0.040	0.067	0.048	0.501	0.330
father's occ class: 9	−0.021	0.029	0.026	0.034	−0.542***	0.194
father's occ class: 10	0.029	0.027	0.027	0.030	0.000	0.185
mother's occ class: 1	0.047	0.049	0.036	0.060	0.079	0.414
mother's occ class: 2	0.015	0.054	−0.098	0.069	1.379***	0.442
mother's occ class: 3	0.056	0.048	0.053	0.057	0.018	0.388
mother's occ class: 4	0.055	0.040	0.016	0.048	0.451	0.310
mother's occ class: 5	0.010	0.049	0.030	0.057	−0.104	0.414
mother's occ class: 6	0.025	0.040	0.028	0.045	0.030	0.313
mother's occ class: 7	0.055	0.041	0.057	0.047	−0.111	0.316
mother's occ class: 9	−0.004	0.038	0.033	0.044	−0.488*	0.285
mother's occ class: 10	0.004	0.032	−0.005	0.036	0.099	0.256
'nuclear family' to 16	0.028	0.019	0.002	0.022	0.251*	0.136
mid 1990s	−0.045***	0.009	−0.049***	0.010	0.073	0.046
late 1990s	−0.065***	0.014	−0.070***	0.016	0.092	0.081
post 2000	−0.033	0.021	−0.039*	0.023	0.120	0.126
constant	−0.754***	0.250	−0.613**	0.283	−0.157	1.663
# observations	21256		21256		21256	
# individuals	2266		2266		2266	
R^2	0.265		0.088		0.250	

F-test on exclusion of instruments from first stage: 36.83; Partial R^2 of the instrument = 0.0332
Hansen's J-test of overidentification = 0.202, p-value = 0.6529

Notes: *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level
Standard errors are clustered at the level of the individual and robust.
'nuclear family' to 16 means lived with both natural parents from birth to age 16.
Reference categories: West Midlands, white, did not live with both natural parents to 16, father/mother occupational class 'plant/machine operative'.
Occupational Class dummies: (1) management, (2) professional, (3) associate professional/technical, (4) clerical/secretarial, (5) craft and related, (6) personal/protective services, (7) sales, (9) other, (10) self-emp/unemp.

Table 4: Smokers at 16/18 by quintile of the mean log wage residual distribution

quintile	Non-smoker at 16	Smoker at 16	Total	Non-smoker at 18	Smoker at 18	Total
1	256	198	454	209	245	454
	56.39%	43.61%	100.00%	46.04%	53.96%	100.00%
2	278	175	453	216	237	453
	61.37%	38.63%	100.00%	47.68%	52.32%	100.00%
3	319	134	453	265	188	453
	70.42%	29.58%	100.00%	58.50%	41.50%	100.00%
4	299	154	453	255	198	453
	66.00%	34.00%	100.00%	56.29%	43.71%	100.00%
5	349	104	453	295	158	453
	77.04%	22.96%	100.00%	65.12%	34.88%	100.00%
Total	1501	765	2266	1240	1026	2266
	66.24%	33.76%	100.00%	54.72%	45.28%	100.00%

Notes: OLS log wage regression (Table 1 column 1) run on pooled panel dataset, residuals are taken and the mean residual for each individual is calculated. These are then ranked into 5 quintiles as a measure of unobserved ability.

Table 5: OLS and Early Smoking IV Returns to Education Estimates for Log Hourly Wage: basic set of controls

	OLS		IV: smoker at 16		IV: first stage	
	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.
years of schooling	0.052***	0.003	0.121***	0.016	— —	— —
smoker at 16 indicator	— —	— —	— —	— —	−1.087***	0.113
age	0.098***	0.004	0.094***	0.005	0.067***	0.023
age^2	−0.001***	0.000	−0.001***	0.000	−0.001***	0.000
year-of-birth	−0.014**	0.007	−0.048***	0.011	0.466***	0.042
year-of-birth^2	0.000***	0.000	0.001***	0.000	−0.004	0.000
region: North	0.041	0.038	0.048	0.044	−0.124	0.286
region: Yorkshire	−0.003	0.033	−0.018	0.039	0.269	0.265
region: North West	0.050	0.033	0.022	0.039	0.432	0.270
region: East Midlands	−0.016	0.032	−0.006	0.037	−0.122	0.258
region: East Anglia	0.010	0.040	−0.006	0.047	0.318	0.338
region: South East	0.143***	0.028	0.080***	0.036	0.946***	0.219
region: South West	0.023	0.034	0.010	0.040	0.271	0.251
region: Wales	−0.018	0.040	−0.019	0.044	0.003	0.305
region: Scotland	0.020	0.036	−0.023	0.043	0.688**	0.283
ethnicity: Black	0.117	0.093	0.104	0.113	−0.015	0.751
ethnicity: Asian	−0.150**	0.070	−0.290***	0.098	1.844***	0.519
ethnicity: Other	−0.042	0.095	−0.221*	0.119	2.406**	0.996
mid 1990s	−0.047***	0.009	−0.049***	0.010	0.040	0.050
late 1990s	−0.068***	0.014	−0.068***	0.016	0.019	0.086
post 2000	−0.038*	0.021	−0.037	0.023	0.014	0.133
constant	−0.849***	0.247	−0.590**	0.280	−2.204	1.697
# observations	21256		21256		21256	
# individuals	2266		2266		2266	
R^2	0.251		0.098		0.143	
F-test on exclusion of instrument from first stage: 92.39; Partial R^2 of the instrument = 0.0400						

Notes: *** significant at 1% level, ** significant at 5% level, * significant at 10% level. Standard errors are clustered at the level of the individual and robust.

Table 6: OLS and RoSLA IV Returns to Education Estimates for Log Hourly Wage: basic set of controls

	OLS		IV: RoSLA		IV: first stage	
	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.
years of schooling	0.052***	0.003	0.100**	0.042	— —	— —
min. school LA=16	— —	— —	— —	— —	0.691***	0.219
age	0.098***	0.004	0.095***	0.005	0.068***	0.023
age^2	−0.001***	0.000	−0.001***	0.000	−0.001***	0.000
year-of-birth	−0.014**	0.007	−0.038*	0.022	0.501***	0.042
year-of-birth^2	0.000***	0.000	0.000**	0.000	−0.004***	0.000
region: North	0.041	0.038	0.046	0.041	−0.089	0.286
region: Yorkshire	−0.003	0.033	−0.014	0.038	0.246	0.268
region: North West	0.050	0.033	0.030	0.041	0.418	0.273
region: East Midlands	−0.016	0.032	−0.009	0.035	−0.129	0.258
region: East Anglia	0.010	0.040	−0.001	0.045	0.277	0.343
region: South East	0.143***	0.028	0.099*	0.052	0.936***	0.223
region: South West	0.023	0.034	0.014	0.038	0.202	0.256
region: Wales	−0.018	0.040	−0.019	0.042	0.022	0.314
region: Scotland	0.020	0.036	−0.010	0.048	0.710**	0.288
ethnicity: Black	0.117	0.093	0.108	0.105	0.251	0.700
ethnicity: Asian	−0.150**	0.070	−0.248**	0.120	2.075***	0.560
ethnicity: Other	−0.042	0.095	−0.167	0.149	2.566***	0.939
mid 1990s	−0.047***	0.009	−0.048***	0.010	0.036	0.050
late 1990s	−0.068***	0.014	−0.068***	0.015	0.013	0.089
post 2000	−0.038*	0.021	−0.037*	0.022	−0.004	0.137
constant	−0.849***	0.247	−0.668**	0.303	−3.375*	1.727
# observations	21256		21256		21256	
# individuals	2266		2266		2266	
R^2	0.251		0.176		0.113	
F-test on exclusion of instrument from first stage: 9.98; Partial R^2 of the instrument = 0.0058						

Notes: *** significant at 1% level, ** significant at 5% level, * significant at 10% level. Standard errors are clustered at the level of the individual and robust.

Table 7: OLS and Both Instruments IV Returns to Education Estimates for Log Hourly Wage: basic set of controls

	OLS		IV: both		IV: first stage	
	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.
years of schooling	0.052***	0.003	0.118***	0.015	— —	— —
smoker at 16 indicator	— —	— —	— —	— —	−1.083***	0.113
min. school LA=16	— —	— —	— —	— —	0.673***	0.213
age	0.098***	0.004	0.094***	0.005	0.064***	0.023
age^2	−0.001***	0.000	−0.001***	0.000	−0.001***	0.000
year-of-birth	−0.014**	0.007	−0.047***	0.010	0.466***	0.042
year-of-birth^2	0.000***	0.000	0.001***	0.000	−0.004***	0.000
region: North	0.041	0.038	0.048	0.043	−0.117	0.285
region: Yorkshire	−0.003	0.033	−0.018	0.039	0.292	0.264
region: North West	0.050	0.033	0.023	0.039	0.444*	0.268
region: East Midlands	−0.016	0.032	−0.006	0.037	−0.095	0.257
region: East Anglia	0.010	0.040	−0.005	0.046	0.360	0.340
region: South East	0.143***	0.028	0.083**	0.036	0.958***	0.219
region: South West	0.023	0.034	0.011	0.039	0.292	0.250
region: Wales	−0.018	0.040	−0.019	0.044	0.007	0.306
region: Scotland	0.020	0.036	−0.021	0.043	0.764***	0.283
ethnicity: Black	0.117	0.093	0.104	0.112	0.041	0.763
ethnicity: Asian	−0.150**	0.070	−0.285***	0.097	1.868***	0.528
ethnicity: Other	−0.042	0.095	−0.215*	0.117	2.351**	0.958
mid 1990s	−0.047***	0.009	−0.048***	0.010	0.049	0.049
late 1990s	−0.068***	0.014	−0.068***	0.016	0.035	0.086
post 2000	−0.038*	0.021	−0.037	0.023	0.030	0.133
constant	−0.849***	0.247	−0.600**	0.277	−1.811	1.697
# observations	21256		21256		21256	
# individuals	2266		2266		2266	
R^2	0.251		0.109		0.148	
F-test on exclusion of instruments from first stage: 51.63; Partial R^2 of the instrument = 0.00454						
Hansen J-test of overidentification: 0.190, p-value = 0.6630						

Notes: *** significant at 1% level, ** significant at 5% level, * significant at 10% level. Standard errors are clustered at the level of the individual and robust.

Table 8: First Stage IV Regression Coefficients on Smoker at 16 indicator and Minimum School Leaving Age of 16 indicator, by quintile of the mean log wage residual distribution

	IV (first stage): Smoking		IV (first stage): RoSLA	
quintile	smoker at 16	R^2	MSLA=16	R^2
1	−0.773***	0.268	1.044**	0.262
#obs = 3684	(0.265)		(0.510)	
mean years of schooling 12.41				
2	−1.044***	0.317	0.837*	0.292
#obs = 4285	(0.227)		(0.458)	
mean years of schooling 12.09				
3	−0.950***	0.329	0.315	0.309
#obs = 4461	(0.249)		(0.496)	
mean years of schooling 12.30				
4	−0.747***	0.257	0.398	0.240
#obs = 4496	(0.213)		(0.388)	
mean years of schooling 12.28				
5	−0.879***	0.341	0.080	0.321
#obs = 4330	(0.241)		(0.435)	
mean years of schooling 12.65				

Notes: *** signficant at 1% level, ** significant at 5% level, * significant at 10% level
Robust standard errors clustered at individual level in parenthesis.
Other covariates included in regressions are as Table 1.

Table 9: OLS and Smoker at 18 IV Returns to Education Estimates for Log Hourly Wage: full set of controls

	OLS		IV: smoker at 18		IV: first stage	
	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.
years of schooling	0.046***	0.003	0.135***	0.023	— —	— —
smoker at 18 indicator	— —	— —	— —	— —	−0.745***	0.108
age	0.099***	0.004	0.093***	0.005	0.054**	0.022
age^2	−0.001***	0.000	−0.001***	0.000	−0.001***	0.000
year-of-birth	−0.016**	0.007	−0.054***	0.012	0.399***	0.041
year-of-birth^2	0.000***	0.000	0.001***	0.000	−0.003***	0.000
region: North	0.047	0.038	0.054	0.045	−0.121	0.274
region: Yorkshire	0.003	0.033	−0.024	0.042	0.319	0.254
region: North West	0.054*	0.032	0.020	0.041	0.414	0.253
region: East Midlands	−0.010	0.032	−0.005	0.038	−0.041	0.236
region: East Anglia	0.015	0.039	−0.011	0.050	0.392	0.324
region: South East	0.142***	0.028	0.077**	0.039	0.760***	0.209
region: South West	0.023	0.034	0.014	0.042	0.138	0.238
region: Wales	−0.012	0.040	−0.020	0.046	0.063	0.287
region: Scotland	0.028	0.036	−0.025	0.046	0.600**	0.264
ethnicity: Black	0.114	0.105	0.115	0.119	−0.212	0.774
ethnicity: Asian	−0.136*	0.071	−0.325***	0.112	2.081***	0.511
ethnicity: Other	−0.048	0.103	−0.248**	0.124	2.112*	1.091
father's occ class: 1	0.116***	0.028	0.013	0.044	1.160***	0.213
father's occ class: 2	0.121***	0.038	−0.092	0.071	2.327***	0.292
father's occ class: 3	0.089**	0.043	−0.053	0.063	1.514***	0.326
father's occ class: 4	0.065*	0.036	−0.062	0.054	1.362***	0.309
father's occ class: 5	0.038*	0.023	0.009	0.029	0.340**	0.170
father's occ class: 6	0.014	0.035	−0.081	0.051	0.983***	0.308
father's occ class: 7	0.103***	0.040	0.063	0.050	0.493	0.329
father's occ class: 9	−0.021	0.029	0.032	0.036	−0.551***	0.196
father's occ class: 10	0.029	0.027	0.026	0.031	−0.008	0.186
mother's occ class: 1	0.047	0.049	0.034	0.063	0.062	0.411
mother's occ class: 2	0.015	0.054	−0.112	0.074	1.352***	0.443
mother's occ class: 3	0.056	0.048	0.053	0.059	0.074	0.388
mother's occ class: 4	0.055	0.040	0.011	0.050	0.471	0.310
mother's occ class: 5	0.010	0.049	0.033	0.060	−0.199	0.420
mother's occ class: 6	0.025	0.040	0.029	0.046	0.021	0.313
mother's occ class: 7	0.055	0.041	0.058	0.049	−0.113	0.313
mother's occ class: 9	−0.004	0.038	0.037	0.046	−0.489*	0.287
mother's occ class: 10	0.004	0.032	−0.007	0.037	0.088	0.256
'nuclear family' to 16	0.028	0.019	−0.001	0.022	0.258*	0.135
mid 1990s	−0.045***	0.009	−0.050***	0.010	0.066	0.046
late 1990s	−0.065***	0.014	−0.070***	0.016	0.079	0.082
post 2000	−0.033	0.021	−0.040*	0.024	0.100	0.127
constant	−0.754***	0.250	−0.596**	0.293	−0.399	1.675
# observations	21256		21256		21256	
# individuals	2266		2266		2266	
R^2	0.265		0.042		0.242	
F-test on exclusion of smoking at 18 from first stage: 48.02; Partial R^2 of instrument = 0.0236						

Notes: *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level
Standard errors are clustered at the level of the individual and robust.
'nuclear family' to 16 means lived with both natural parents from birth to age 16.
Reference categories: West Midlands, white, did not live with both natural parents to 16, father/mother occupational class 'plant/machine operative'.
Occupational Class dummies: (1) management, (2) professional, (3) associate professional/technical, (4) clerical/secretarial, (5) craft and related, (6) personal/protective services, (7) sales, (9) other, (10) self-emp/unemp.

Table 10: OLS and Smoker at 17 IV Returns to Education Estimates for Log Hourly Wage: full set of controls

	OLS		IV: smoker at 17		IV: first stage	
	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.
years of schooling	0.046***	0.003	0.130***	0.021	— —	— —
smoker at 17 indicator	— —	— —	— —	— —	−0.816***	0.108
age	0.099***	0.004	0.094***	0.005	0.055**	0.022
age^2	−0.001***	0.000	−0.001***	0.000	−0.001***	0.000
year-of-birth	−0.016**	0.007	−0.052***	0.012	0.398***	0.041
year-of-birth^2	0.000***	0.000	0.001***	0.000	−0.003***	0.000
region: North	0.047	0.038	0.054	0.045	−0.123	0.274
region: Yorkshire	0.003	0.033	−0.022	0.041	0.302	0.254
region: North West	0.054*	0.032	0.022	0.040	0.383	0.253
region: East Midlands	−0.010	0.032	−0.005	0.038	−0.065	0.235
region: East Anglia	0.015	0.039	−0.009	0.049	0.351	0.323
region: South East	0.142***	0.028	0.081**	0.037	0.739***	0.208
region: South West	0.023	0.034	0.015	0.041	0.157	0.238
region: Wales	−0.012	0.040	−0.019	0.045	0.062	0.285
region: Scotland	0.028	0.036	−0.022	0.044	0.594**	0.263
ethnicity: Black	0.114	0.105	0.115	0.117	−0.172	0.772
ethnicity: Asian	−0.136*	0.071	−0.314***	0.106	1.972***	0.492
ethnicity: Other	−0.048	0.103	−0.236**	0.119	2.147**	1.087
father's occ class: 1	0.116***	0.028	0.019	0.042	1.132***	0.213
father's occ class: 2	0.121***	0.038	−0.079	0.067	2.282***	0.293
father's occ class: 3	0.089**	0.043	−0.045	0.060	1.508***	0.325
father's occ class: 4	0.065*	0.036	−0.055	0.052	1.319***	0.303
father's occ class: 5	0.038*	0.023	0.010	0.028	0.335**	0.170
father's occ class: 6	0.014	0.035	−0.075	0.048	1.007***	0.308
father's occ class: 7	0.103***	0.040	0.065	0.049	0.510	0.331
father's occ class: 9	−0.021	0.029	0.029	0.035	−0.550***	0.196
father's occ class: 10	0.029	0.027	0.027	0.030	−0.026	0.186
mother's occ class: 1	0.047	0.049	0.035	0.061	0.081	0.411
mother's occ class: 2	0.015	0.054	−0.104	0.071	1.368***	0.441
mother's occ class: 3	0.056	0.048	0.053	0.058	0.043	0.393
mother's occ class: 4	0.055	0.040	0.014	0.049	0.463	0.308
mother's occ class: 5	0.010	0.049	0.032	0.059	−0.139	0.418
mother's occ class: 6	0.025	0.040	0.029	0.046	0.034	0.313
mother's occ class: 7	0.055	0.041	0.057	0.048	−0.087	0.312
mother's occ class: 9	−0.004	0.038	0.035	0.045	−0.492*	0.285
mother's occ class: 10	0.004	0.032	−0.006	0.037	0.082	0.254
'nuclear family' to 16	0.028	0.019	0.001	0.022	0.252*	0.136
mid 1990s	−0.045***	0.009	−0.050***	0.010	0.063	0.046
late 1990s	−0.065***	0.014	−0.070***	0.016	0.074	0.081
post 2000	−0.033	0.021	−0.040*	0.023	0.097	0.126
constant	−0.754***	0.250	−0.605**	0.288	−0.469	1.671
# observations	21256		21256		21256	
# individuals	2266		2266		2266	
R^2	0.265		0.067		0.245	
F-test on exclusion of smoking at 18 from first stage: 57.14; Partial R^2 of instrument = 0.0267						

Notes: *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level
Standard errors are clustered at the level of the individual and robust.
'nuclear family' to 16 means lived with both natural parents from birth to age 16.
Reference categories: West Midlands, white, did not live with both natural parents to 16, father/mother occupational class 'plant/machine operative'.
Occupational Class dummies: (1) management, (2) professional, (3) associate professional/technical, (4) clerical/secretarial, (5) craft and related, (6) personal/protective services, (7) sales, (9) other, (10) self-emp/unemp.

Figures

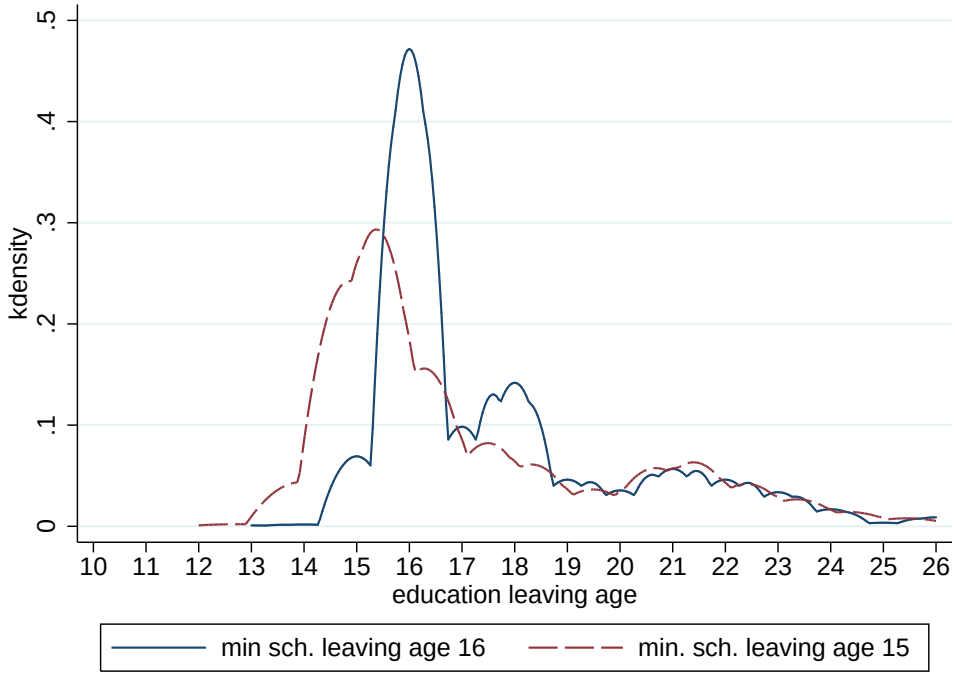

Figure 1: Education Leaving Age Density, by Minimum School Leaving Age

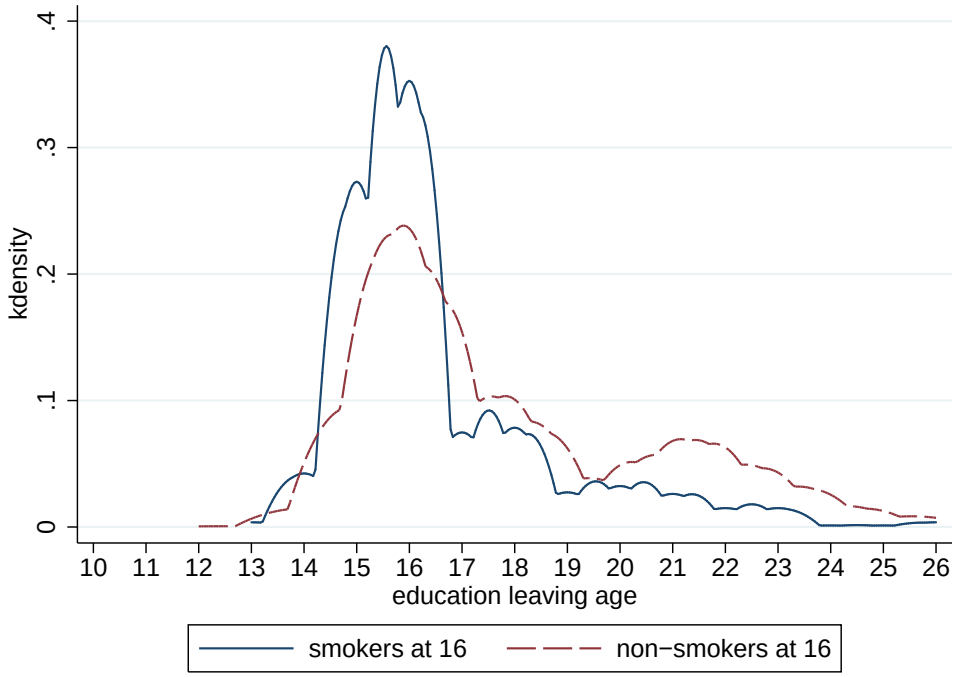

Figure 2: Education Leaving Age Density, by Smoker at 16 Status

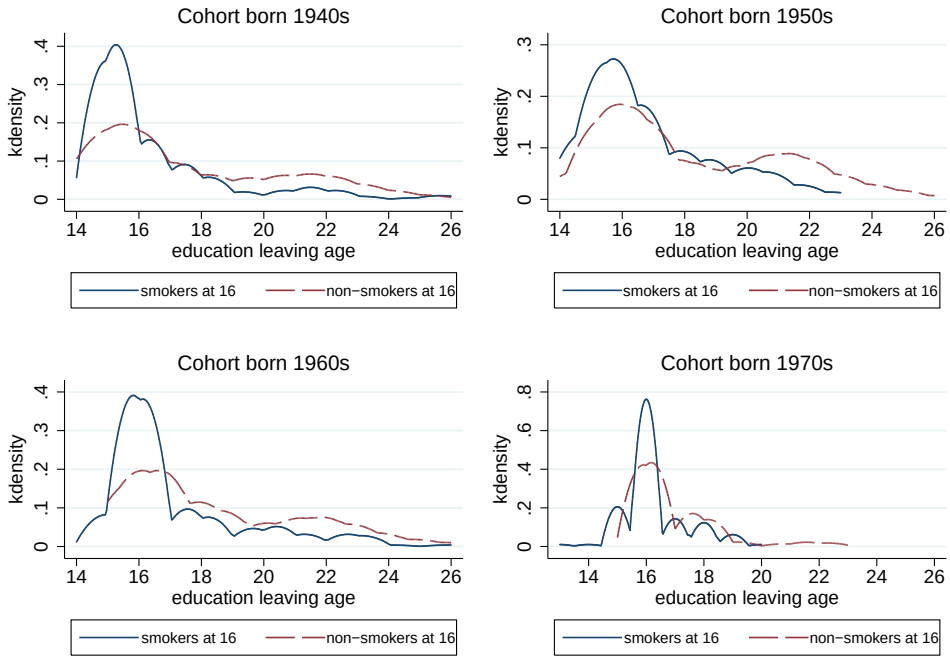

Figure 3: Education Leaving Age Density, by Smoker at 16 Status and cohort

A Sample Summary Statistics

Table A-1: Sample Summary Statistics, total

Variable	Mean	Std. Dev.	Min	Max	Variable	Mean	Std. Dev.	Min	Max
number of observations per person	9.380	4.516	1	15	ethnicity: Other	0.007	0.083	0	1
log wage	2.214	0.437	0.848	3.407	father's occ class: 1	0.139	0.346	0	1
years of schooling	12.507	2.646	7	21	father's occ class: 2	0.064	0.244	0	1
smoker at age 16	0.315	0.465	0	1	father's occ class: 3	0.038	0.191	0	1
minimum school leaving age was 16	0.537	0.499	0	1	father's occ class: 4	0.049	0.216	0	1
age	39.991	10.757	19	65	father's occ class: 5	0.234	0.423	0	1
cohort: born in the 1920s	0.003	0.052	0	1	father's occ class: 6	0.044	0.205	0	1
cohort: born in the 1930s	0.050	0.219	0	1	father's occ class: 7	0.032	0.177	0	1
cohort: born in the 1940s	0.203	0.402	0	1	father's occ class: 8	0.171	0.377	0	1
cohort: born in the 1950s	0.274	0.446	0	1	father's occ class: 9	0.086	0.280	0	1
cohort: born in the 1960s	0.319	0.466	0	1	father's occ class: 10	0.143	0.350	0	1
cohort: born in the 1970s	0.146	0.354	0	1	mother's occ class: 1	0.037	0.188	0	1
cohort: born in the 1980s	0.005	0.069	0	1	mother's occ class: 2	0.026	0.159	0	1
region: North	0.068	0.252	0	1	mother's occ class: 3	0.032	0.175	0	1
region: Yorkshire	0.097	0.296	0	1	mother's occ class: 4	0.098	0.297	0	1
region: North West	0.106	0.308	0	1	mother's occ class: 5	0.029	0.168	0	1
region: East Midlands	0.092	0.290	0	1	mother's occ class: 6	0.073	0.260	0	1
region: West Midlands	0.086	0.280	0	1	mother's occ class: 7	0.066	0.248	0	1
region: East Anglia	0.043	0.202	0	1	mother's occ class: 8	0.051	0.220	0	1
region: South East	0.280	0.449	0	1	mother's occ class: 9	0.084	0.277	0	1
region: South West	0.100	0.300	0	1	mother's occ class: 10	0.505	0.500	0	1
region: Wales	0.051	0.221	0	1	'nuclear family' to 16	0.831	0.375	0	1
region: Scotland	0.076	0.265	0	1	early 1990s	0.195	0.396	0	1
ethnicity: White	0.974	0.160	0	1	mid 1990s	0.213	0.409	0	1
ethnicity: Black	0.004	0.062	0	1	late 1990s	0.221	0.415	0	1
ethnicity: Asian	0.016	0.124	0	1	post 2000	0.371	0.483	0	1
total number of observations		21265							

Notes: 'nuclear family' to 16 means lived with both natural parents from birth to age 16.
Occupational class dummies: (1) management, (2) professional, (3) associate professional/technical, (4) clerical/secretarial, (5) craft and related, (6) personal/protective services, (7) sales, (8) plant/machine operative, (9) other, (10) self-emp/unemp.

Table A-2: Sample Summary Statistics, by Early Smoking Status

	Smoker at 16				Non-Smoker at 16			
Variable	Mean	Std. Dev.	Min	Max	Mean	Std. Dev.	Min	Max
log wage	2.120	0.423	0.878	3.395	2.257	0.437	0.848	3.407
years of schooling	11.646	2.081	8	21	12.902	2.781	7	21
smoker at age 16	1.000	0.000	1	1	0.000	0.000	0	0
min. sch leaving age 16	0.486	0.500	0	1	0.561	0.496	0	1
age	41.230	11.337	19	65	39.421	10.431	19	65
cohort: born in the 1920s	0.004	0.063	0	1	0.002	0.046	0	1
cohort: born in the 1930s	0.063	0.244	0	1	0.044	0.206	0	1
cohort: born in the 1940s	0.258	0.438	0	1	0.177	0.382	0	1
cohort: born in the 1950s	0.250	0.433	0	1	0.284	0.451	0	1
cohort: born in the 1960s	0.266	0.442	0	1	0.344	0.475	0	1
cohort: born in the 1970s	0.150	0.357	0	1	0.145	0.352	0	1
cohort: born in the 1980s	0.008	0.091	0	1	0.003	0.057	0	1
region: North	0.057	0.231	0	1	0.073	0.260	0	1
region: Yorkshire	0.103	0.304	0	1	0.095	0.293	0	1
region: North West	0.103	0.303	0	1	0.107	0.309	0	1
region: East Midlands	0.095	0.293	0	1	0.091	0.288	0	1
region: West Midlands	0.076	0.266	0	1	0.090	0.287	0	1
region: East Anglia	0.050	0.219	0	1	0.039	0.193	0	1
region: South East	0.272	0.445	0	1	0.284	0.451	0	1
region: South West	0.119	0.324	0	1	0.091	0.288	0	1
region: Wales	0.044	0.204	0	1	0.055	0.228	0	1
region: Scotland	0.082	0.274	0	1	0.074	0.261	0	1
ethnicity: White	0.991	0.097	0	1	0.966	0.181	0	1
ethnicity: Black	0.001	0.037	0	1	0.005	0.070	0	1
ethnicity: Asian	0.006	0.075	0	1	0.020	0.140	0	1
ethnicity: Other	0.002	0.049	0	1	0.009	0.094	0	1
father's occ class: 1	0.123	0.329	0	1	0.146	0.353	0	1
father's occ class: 2	0.036	0.185	0	1	0.077	0.266	0	1
father's occ class: 3	0.027	0.161	0	1	0.043	0.203	0	1
father's occ class: 4	0.032	0.177	0	1	0.057	0.231	0	1
father's occ class: 5	0.257	0.437	0	1	0.223	0.416	0	1
father's occ class: 6	0.036	0.187	0	1	0.048	0.213	0	1
father's occ class: 7	0.036	0.187	0	1	0.031	0.172	0	1
father's occ class: 8	0.194	0.396	0	1	0.161	0.367	0	1
father's occ class: 9	0.113	0.316	0	1	0.074	0.261	0	1
father's occ class: 10	0.146	0.353	0	1	0.142	0.349	0	1
mother's occ class: 1	0.031	0.173	0	1	0.039	0.195	0	1
mother's occ class: 2	0.018	0.134	0	1	0.030	0.169	0	1
mother's occ class: 3	0.028	0.166	0	1	0.033	0.179	0	1
mother's occ class: 4	0.085	0.279	0	1	0.104	0.305	0	1
mother's occ class: 5	0.043	0.204	0	1	0.023	0.148	0	1
mother's occ class: 6	0.098	0.297	0	1	0.062	0.240	0	1
mother's occ class: 7	0.051	0.221	0	1	0.072	0.259	0	1
mother's occ class: 8	0.053	0.223	0	1	0.050	0.218	0	1
mother's occ class: 9	0.092	0.288	0	1	0.080	0.272	0	1
mother's occ class: 10	0.501	0.500	0	1	0.507	0.500	0	1
'nuclear family' to 16	0.795	0.404	0	1	0.848	0.359	0	1
early 1990s	0.196	0.397	0	1	0.194	0.396	0	1
mid 1990s	0.212	0.409	0	1	0.214	0.410	0	1
late 1990s	0.219	0.414	0	1	0.222	0.415	0	1
post 2000	0.373	0.484	0	1	0.371	0.483	0	1
# observations per person	8.753	4.650	1	15	9.700	4.414	1	15
# individuals	765				1501			

Notes: See Table A-1 for variable descriptions.

B In Support of Early Smoking as a Valid Instrument

Table B-1: The Effect of Schooling on Probability of Current and Early Smoking

	Current Smoker		Smoker at 16		
	marginal fx		marginal fx	z	x-bar
years of schooling	−0.027***	−7.24	−0.038***	−8.38	12.306
age	0.005	1.35	0.002	0.50	42.374
age^2	0.000***	−2.99	0.000	−0.36	1939.210
year-of-birth	−0.016***	−2.66	−0.002	−0.30	59.190 (=1955)
year-of-birth^2	0.000***	2.84	0.000	0.25	3647.300
region: North	−0.032	−0.76	−0.033	−0.66	0.066
region: Yorkshire	0.037	0.99	0.031	0.69	0.098
region: North West	0.037	0.97	0.031	0.69	0.104
region: East Midlands	0.056	1.46	0.060	1.30	0.094
region: East Anglia	0.047	1.05	0.112*	1.94	0.043
region: South East	0.055*	1.72	0.047	1.23	0.285
region: South West	−0.005	−0.12	0.059	1.29	0.097
region: Wales	0.062	1.36	0.006	0.11	0.053
region: Scotland	0.065	1.57	0.063	1.28	0.078
ethnicity: Black	−0.121	−1.27	−0.181	−1.39	0.006
ethnicity: Asian	0.225***	3.18	−0.176**	−2.29	0.016
ethnicity: Other	−0.016	−0.16	−0.174	−1.55	0.008
father's occ class: 1	−0.022	−0.74	−0.041	−1.17	0.141
father's occ class: 2	−0.094**	−2.39	−0.094*	−1.88	0.058
father's occ class: 3	0.013	0.28	−0.099*	−1.81	0.035
father's occ class: 4	−0.055	−1.28	−0.104**	−2.10	0.047
father's occ class: 5	0.012	0.46	−0.010	−0.33	0.236
father's occ class: 6	0.022	0.49	−0.089*	−1.79	0.042
father's occ class: 7	0.006	0.13	−0.002	−0.03	0.032
father's occ class: 9	0.009	0.29	0.016	0.43	0.094
father's occ class: 10	−0.012	−0.39	−0.056*	−1.67	0.151
mother's occ class: 1	−0.017	−0.33	−0.026	−0.39	0.037
mother's occ class: 2	0.070	1.07	0.003	0.04	0.027
mother's occ class: 3	−0.050	−0.87	−0.036	−0.51	0.030
mother's occ class: 4	−0.036	−0.84	−0.023	−0.42	0.089
mother's occ class: 5	−0.023	−0.41	0.048	0.67	0.028
mother's occ class: 6	0.025	0.51	0.036	0.62	0.068
mother's occ class: 7	−0.054	−1.16	−0.083	−1.48	0.060
mother's occ class: 9	−0.017	−0.39	−0.057	−1.08	0.083
mother's occ class: 10	−0.012	−0.33	−0.060	−1.29	0.532
'nuclear family' to 16	−0.062***	−2.86	−0.099***	−3.73	0.820
mid 1990s	0.002	0.22	0.000	0.06	0.223
late 1990s	0.037***	2.90	0.035***	2.85	0.200
post 2000	−0.006	−0.30	0.002	0.08	0.371
# individuals	2805		2805		
# observations	33298		33298		
obs. prob.	0.287		0.344		
pred. prob. (at x-bar)	0.276		0.331		

Notes: Reference categories: West Midlands, white, did not live with both natural parents to 16, father/mother occupational class 'plant/machine operative'. Occupational Class dummies: (1) management, (2) professional, (3) associate professional/technical, (4) clerical/secretarial, (5) craft and related, (6) personal/protective services, (7) sales, (9) other, (10) self-emp/unemp.

Table B-2: Reduced Form for Log Hourly Wage: Smoker at 16 instrument, RoSLA instrument and Both instruments

	Smoker at 16		Min. School LA=16		Both	
Dep. Var: log hourly wage	Coeff.	Robust Std. Err.	Coeff.	Robust Std. Err.	Coeff.	Robust Std. Err.
smoker at 16 indicator	−0.113***	0.016	— —	— —	−0.113***	0.016
min. school LA=16	— —	— —	0.058**	0.028	0.057**	0.027
age	0.101***	0.004	0.101***	0.004	0.101***	0.004
age^2	−0.001***	0.000	−0.001***	0.000	−0.001***	0.000
year-of-birth	−0.000	0.007	0.003	0.007	−0.000	0.007
year-of-birth^2	0.000**	0.000	0.000	0.000	0.000*	0.000
region: North	0.040	0.040	0.043	0.040	0.041	0.039
region: Yorkshire	0.021	0.035	0.019	0.035	0.023	0.035
region: North West	0.075**	0.034	0.072**	0.034	0.075**	0.034
region: East Midlands	−0.010	0.033	−0.010	0.034	−0.008	0.033
region: East Anglia	0.038	0.040	0.032	0.041	0.041	0.040
region: South East	0.179***	0.029	0.177***	0.029	0.180***	0.029
region: South West	0.037	0.034	0.029	0.034	0.039	0.034
region: Wales	−0.009	0.042	−0.007	0.043	−0.009	0.042
region: Scotland	0.062*	0.037	0.062*	0.037	0.068*	0.037
ethnicity: Black	0.094	0.115	0.118	0.113	0.099	0.115
ethnicity: Asian	−0.059	0.066	−0.036	0.066	−0.058	0.066
ethnicity: Other	0.032	0.132	0.052	0.129	0.028	0.130
father's occ class: 1	0.165***	0.029	0.170***	0.029	0.164***	0.029
father's occ class: 2	0.215***	0.040	0.233***	0.041	0.216***	0.041
father's occ class: 3	0.150***	0.046	0.162***	0.047	0.149***	0.047
father's occ class: 4	0.117***	0.040	0.132***	0.041	0.117***	0.040
father's occ class: 5	0.054**	0.024	0.053**	0.024	0.053**	0.024
father's occ class: 6	0.054	0.039	0.061	0.038	0.051	0.038
father's occ class: 7	0.126***	0.042	0.127***	0.043	0.129***	0.042
father's occ class: 9	−0.043	0.031	−0.048	0.031	−0.042	0.031
father's occ class: 10	0.025	0.028	0.032	0.028	0.026	0.028
mother's occ class: 1	0.050	0.052	0.050	0.053	0.046	0.052
mother's occ class: 2	0.082	0.060	0.076	0.060	0.076	0.061
mother's occ class: 3	0.059	0.051	0.055	0.052	0.056	0.051
mother's occ class: 4	0.077*	0.042	0.074*	0.043	0.073*	0.042
mother's occ class: 5	0.016	0.052	−0.000	0.054	0.018	0.052
mother's occ class: 6	0.035	0.043	0.020	0.044	0.033	0.043
mother's occ class: 7	0.047	0.043	0.051	0.044	0.044	0.043
mother's occ class: 9	−0.025	0.040	−0.028	0.042	−0.028	0.040
mother's occ class: 10	0.009	0.034	0.008	0.036	0.007	0.034
'nuclear family' to 16	0.033	0.021	0.044**	0.021	0.033	0.021
mid 1990s	−0.041***	0.009	−0.042***	0.009	−0.040***	0.009
late 1990s	−0.060***	0.015	−0.061***	0.015	−0.058***	0.015
post 2000	−0.026	0.021	−0.028	0.022	−0.025	0.021
constant	−0.668**	0.261	−0.803***	0.263	−0.636**	0.261
# observations	21256		21256		21256	
# individuals	2266		2266		2266	
R^2	0.217		0.205		0.218	

Notes: *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level
Standard errors are clustered at the level of the individual and robust.
'nuclear family' to 16 means lived with both natural parents from birth to age 16.
Reference categories: West Midlands, white, did not live with both natural parents to 16, father/mother occupational class 'plant/machine operative'. Occupational Class dummies: (1) management, (2) professional, (3) associate professional/technical, (4) clerical/secretarial, (5) craft and related, (6) personal/protective services, (7) sales, (9) other, (10) self-emp/unemp.

Table B-3: Regression of Residuals from Structural Equation (when using the RoSLA IV) on the Smoker at 16 indicator

	OLS
smoker at 16 indicator	−0.022
	(0.016)
constant	0.007
	(0.009)
# observations	21256
# individuals	2266
R^2	0.001

Notes: *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level
Robust standard errors clustered at individual level in parenthesis.

Construction of the dependent variable $\hat{\epsilon_i}$:
Log hourly wage estimated by IV regression, first stage equation estimated:
$\hat{S}_i = X_i\tilde{\gamma} + Z_i\tilde{\pi} + u_i$ where Z_i is the min. school LA=16 indicator.
Second stage equation (1) estimated:
$log(w_i) = X_i\varphi + \hat{S}_i\beta_i + \epsilon_i$.
The residuals are recovered and these $\hat{\epsilon_i}$ are the dependent variable.

Table B-4: IV Estimates of The Return to Education using RoSLA, including Smoker at 16 status as an Explanatory Variable

	IV: RoSLA		IV: first stage	
Dep. Var: log hourly wage	Coeff.	Robust Std. Err.	Coeff.	Robust Std. Err.
years of schooling	0.102∗	0.052	— —	— —
min. school LA=16	— —	— —	0.556∗∗∗	0.202
smoker at 16 indicator	−0.024	0.049	−0.874∗∗∗	0.107
age	0.095∗∗∗	0.005	0.054∗∗	0.022
age^2	−0.001∗∗∗	0.000	−0.001∗∗∗	0.000
year-of-birth	−0.041∗	0.022	0.399∗∗∗	0.041
year-of-birth^2	0.000∗∗	0.000	−0.004∗∗∗	0.000
region: North	0.051	0.041	−0.097	0.272
region: Yorkshire	−0.013	0.041	0.347	0.253
region: North West	0.034	0.042	0.409	0.253
region: East Midlands	−0.006	0.035	−0.014	0.235
region: East Anglia	0.001	0.048	0.398	0.325
region: South East	0.102∗	0.053	0.767∗∗∗	0.207
region: South West	0.019	0.038	0.192	0.236
region: Wales	−0.017	0.042	0.082	0.286
region: Scotland	−0.004	0.052	0.705∗∗∗	0.263
ethnicity: Black	0.110	0.109	−0.114	0.788
ethnicity: Asian	−0.259∗	0.133	1.975∗∗∗	0.493
ethnicity: Other	−0.178	0.146	2.021∗	1.080
father's occ class: 1	0.051	0.068	1.118∗∗∗	0.213
father's occ class: 2	−0.015	0.123	2.271∗∗∗	0.290
father's occ class: 3	−0.002	0.089	1.485∗∗∗	0.319
father's occ class: 4	−0.018	0.078	1.324∗∗∗	0.303
father's occ class: 5	0.020	0.029	0.322∗	0.170
father's occ class: 6	−0.047	0.061	0.968∗∗∗	0.303
father's occ class: 7	0.078	0.050	0.501	0.330
father's occ class: 9	0.013	0.042	−0.542∗∗∗	0.194
father's occ class: 10	0.026	0.028	0.000	0.185
mother's occ class: 1	0.038	0.056	0.079	0.414
mother's occ class: 2	−0.064	0.095	1.379∗∗∗	0.442
mother's occ class: 3	0.054	0.052	0.018	0.388
mother's occ class: 4	0.028	0.050	0.451	0.310
mother's occ class: 5	0.028	0.053	−0.104	0.414
mother's occ class: 6	0.030	0.042	0.030	0.313
mother's occ class: 7	0.055	0.044	−0.111	0.316
mother's occ class: 9	0.022	0.047	−0.488∗	0.285
mother's occ class: 10	−0.003	0.034	0.099	0.256
'nuclear family' to 16	0.008	0.023	0.251∗	0.136
mid 1990s	−0.048∗∗∗	0.010	0.073	0.046
late 1990s	−0.068∗∗∗	0.016	0.092	0.081
post 2000	−0.037	0.023	0.120	0.126
constant	−0.620∗∗	0.267	−0.157	1.663
# observations	21256		21256	
# individuals	2266		2266	
R^2	0.178		0.250	

Notes: *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level. Standard errors are clustered at the level of the individual and robust. 'nuclear family' to 16 means lived with both natural parents from birth to age 16. Ref. categories: West Midlands, white, did not live with both natural parents to 16, father/mother occupational class 'plant/machine operative'. Occupational Class dummies: (1) management, (2) professional, (3) associate professional/technical, (4) clerical/secretarial, (5) craft and related, (6) personal/protective services, (7) sales, (9) other, (10) self-emp/unemp.

C Robustness to using Fuller(1) LIML Estimator

Table C-1: OLS and Early Smoking IV Returns to Education Estimates for Log Hourly Wage: full set of controls

	OLS		IV: smoker at 16		IV: first stage	
		Robust		Robust		Robust
Dep. Var: log hourly wage	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.
years of schooling	0.046***	0.003	0.129***	0.020	— —	— —
smoker at 16 indicator	— —	— —	— —	— —	−0.876***	0.108
age	0.099***	0.004	0.094***	0.005	0.056***	0.022
age^2	−0.001***	0.000	−0.001***	0.000	−0.001***	0.000
year-of-birth	−0.016**	0.007	−0.052***	0.011	0.398***	0.041
year-of-birth^2	0.000***	0.000	0.001***	0.000	−0.003***	0.000
region: North	0.047	0.038	0.054	0.044	−0.103	0.272
region: Yorkshire	0.003	0.033	−0.022	0.041	0.331	0.253
region: North West	0.054*	0.032	0.023	0.040	0.402	0.253
region: East Midlands	−0.010	0.032	−0.005	0.038	−0.034	0.235
region: East Anglia	0.015	0.039	−0.009	0.048	0.366	0.324
region: South East	0.142***	0.028	0.082**	0.037	0.757***	0.206
region: South West	0.023	0.034	0.015	0.041	0.175	0.237
region: Wales	−0.012	0.040	−0.019	0.045	0.081	0.285
region: Scotland	0.028	0.036	−0.021	0.044	0.643**	0.262
ethnicity: Black	0.114	0.105	0.115	0.117	−0.164	0.779
ethnicity: Asian	−0.136*	0.071	−0.312***	0.105	1.965***	0.485
ethnicity: Other	−0.048	0.103	−0.234**	0.119	2.067*	1.111
father's occ class: 1	0.116***	0.028	0.020	0.041	1.122***	0.214
father's occ class: 2	0.121***	0.038	−0.076	0.065	2.268***	0.291
father's occ class: 3	0.089**	0.043	−0.043	0.058	1.499***	0.321
father's occ class: 4	0.065*	0.036	−0.053	0.051	1.320***	0.305
father's occ class: 5	0.038*	0.023	0.011	0.028	0.335**	0.170
father's occ class: 6	0.014	0.035	−0.074	0.048	0.991***	0.305
father's occ class: 7	0.103***	0.040	0.066	0.049	0.467	0.330
father's occ class: 9	−0.021	0.029	0.028	0.035	−0.551***	0.197
father's occ class: 10	0.029	0.027	0.027	0.030	−0.012	0.186
mother's occ class: 1	0.047	0.049	0.035	0.061	0.112	0.411
mother's occ class: 2	0.015	0.054	−0.103	0.070	1.433***	0.439
mother's occ class: 3	0.056	0.048	0.053	0.057	0.046	0.387
mother's occ class: 4	0.055	0.040	0.015	0.048	0.485	0.307
mother's occ class: 5	0.010	0.049	0.031	0.058	−0.117	0.417
mother's occ class: 6	0.025	0.040	0.029	0.045	0.054	0.311
mother's occ class: 7	0.055	0.041	0.057	0.048	−0.083	0.312
mother's occ class: 9	−0.004	0.038	0.034	0.044	−0.461	0.284
mother's occ class: 10	0.004	0.032	−0.006	0.036	0.115	0.253
'nuclear family' to 16	0.028	0.019	0.001	0.022	0.247*	0.136
mid 1990s	−0.045***	0.009	−0.050***	0.010	0.067	0.046
late 1990s	−0.065***	0.014	−0.070***	0.016	0.080	0.081
post 2000	−0.033	0.021	−0.040*	0.023	0.108	0.126
constant	−0.754***	0.250	−0.607**	0.287	−0.471	1.664
# observations	21256		21256		21256	
# individuals	2266		2266		2266	
R^2	0.265		0.073		0.246	
F-test on exclusion of smoking at 16 from first stage: 66.17; Partial R^2 of instrument = 0.0289						

Notes: *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level
Standard errors are clustered at the level of the individual and robust.
'nuclear family' to 16 means lived with both natural parents from birth to age 16.
Reference categories: West Midlands, white, did not live with both natural parents to 16, father/mother occupational class 'plant/machine operative'. Occupational Class dummies: (1) management, (2) professional, (3) associate professional/technical, (4) clerical/secretarial, (5) craft and related, (6) personal/protective services, (7) sales, (9) other, (10) self-emp/unemp.

Table C-2: OLS and RoSLA IV Returns to Education Estimates for Log Hourly Wage: full set of controls

	OLS		IV: RoSLA		IV: first stage	
		Robust		Robust		Robust
Dep. Var: log hourly wage	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.
years of schooling	0.046***	0.003	0.101**	0.051	— —	— —
min. school LA=16	— —	— —	— —	— —	0.564***	0.206
age	0.099***	0.004	0.095***	0.005	0.056**	0.022
age^2	−0.001***	0.000	−0.001***	0.000	−0.001***	0.000
year-of-birth	−0.016**	0.007	−0.040*	0.023	0.427***	0.041
year-of-birth^2	0.000***	0.000	0.000**	0.000	−0.004***	0.000
region: North	0.047	0.038	0.051	0.041	−0.080	0.272
region: Yorkshire	0.003	0.033	−0.013	0.040	0.320	0.256
region: North West	0.054*	0.032	0.033	0.041	0.386	0.255
region: East Midlands	−0.010	0.032	−0.007	0.035	−0.035	0.234
region: East Anglia	0.015	0.039	−0.001	0.046	0.324	0.327
region: South East	0.142***	0.028	0.102**	0.051	0.741***	0.208
region: South West	0.023	0.034	0.017	0.038	0.114	0.240
region: Wales	−0.012	0.040	−0.017	0.042	0.093	0.290
region: Scotland	0.028	0.036	−0.005	0.050	0.658**	0.266
ethnicity: Black	0.114	0.105	0.114	0.110	0.037	0.746
ethnicity: Asian	−0.136*	0.071	−0.254*	0.138	2.146***	0.515
ethnicity: Other	−0.048	0.103	−0.172	0.151	2.214**	1.074
father's occ class: 1	0.116***	0.028	0.052	0.068	1.162***	0.216
father's occ class: 2	0.121***	0.038	−0.011	0.127	2.404***	0.298
father's occ class: 3	0.089**	0.043	0.001	0.092	1.585***	0.333
father's occ class: 4	0.065*	0.036	−0.014	0.082	1.440***	0.308
father's occ class: 5	0.038*	0.023	0.020	0.029	0.322*	0.172
father's occ class: 6	0.014	0.035	−0.045	0.063	1.046***	0.313
father's occ class: 7	0.103***	0.040	0.078	0.049	0.484	0.339
father's occ class: 9	−0.021	0.029	0.012	0.044	−0.592***	0.196
father's occ class: 10	0.029	0.027	0.028	0.028	0.043	0.186
mother's occ class: 1	0.047	0.049	0.039	0.056	0.107	0.426
mother's occ class: 2	0.015	0.054	−0.064	0.093	1.378***	0.454
mother's occ class: 3	0.056	0.048	0.054	0.052	0.007	0.395
mother's occ class: 4	0.055	0.040	0.028	0.050	0.453	0.317
mother's occ class: 5	0.010	0.049	0.024	0.054	−0.240	0.430
mother's occ class: 6	0.025	0.040	0.027	0.042	−0.070	0.322
mother's occ class: 7	0.055	0.041	0.057	0.044	−0.053	0.324
mother's occ class: 9	−0.004	0.038	0.022	0.047	−0.491*	0.293
mother's occ class: 10	0.004	0.032	−0.003	0.034	0.103	0.264
'nuclear family' to 16	0.028	0.019	0.010	0.026	0.330**	0.137
mid 1990s	−0.045***	0.009	−0.048***	0.010	0.063	0.047
late 1990s	−0.065***	0.014	−0.068***	0.015	0.075	0.083
post 2000	−0.033	0.021	−0.038*	0.022	0.094	0.129
constant	−0.754***	0.250	−0.656**	0.279	−1.459	1.681
# observations	21256		21256		21256	
# individuals	2266		2266		2266	
R^2	0.265		0.178		0.227	
F-test on exclusion of min. sch. LA=16 from first stage: 7.49; Partial R^2 of instrument = 0.0044						

Notes: *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level
Standard errors are clustered at the level of the individual and robust.
'nuclear family' to 16 means lived with both natural parents from birth to age 16.
Reference categories: West Midlands, white, did not live with both natural parents to 16, father/mother occupational class 'plant/machine operative'. Occupational Class dummies: (1) management, (2) professional, (3) associate professional/technical, (4) clerical/secretarial, (5) craft and related, (6) personal/protective services, (7) sales, (9) other, (10) self-emp/unemp.

Table C-3: OLS and Both Instruments IV Returns to Education Estimates for Log Hourly Wage: full set of controls

	OLS		IV: both		IV: first stage	
		Robust		Robust		Robust
Dep. Var: log hourly wage	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.
years of schooling	0.046***	0.003	0.125***	0.019	— —	— —
smoker at age 16	— —	— —	— —	— —	−0.874***	0.107
min. school LA=16	— —	— —	— —	— —	0.556***	0.202
age	0.099***	0.004	0.094***	0.005	0.054**	0.022
age^2	−0.001***	0.000	−0.001***	0.000	−0.001***	0.000
year-of-birth	−0.016**	0.007	−0.050***	0.011	0.399***	0.041
year-of-birth^2	0.000***	0.000	0.001***	0.000	−0.004***	0.000
region: North	0.047	0.038	0.053	0.044	−0.097	0.272
region: Yorkshire	0.003	0.033	−0.021	0.040	0.347	0.253
region: North West	0.054*	0.032	0.024	0.039	0.409	0.253
region: East Midlands	−0.010	0.032	−0.006	0.037	−0.014	0.235
region: East Anglia	0.015	0.039	−0.008	0.048	0.398	0.325
region: South East	0.142***	0.028	0.084**	0.036	0.767***	0.207
region: South West	0.023	0.034	0.015	0.040	0.192	0.236
region: Wales	−0.012	0.040	−0.019	0.045	0.082	0.286
region: Scotland	0.028	0.036	−0.019	0.043	0.705***	0.263
ethnicity: Black	0.114	0.105	0.115	0.116	−0.114	0.788
ethnicity: Asian	−0.136*	0.071	−0.305***	0.103	1.975***	0.493
ethnicity: Other	−0.048	0.103	−0.226*	0.116	2.021*	1.080
father's occ class: 1	0.116***	0.028	0.024	0.040	1.118***	0.213
father's occ class: 2	0.121***	0.038	−0.069	0.062	2.271***	0.290
father's occ class: 3	0.089**	0.043	−0.038	0.056	1.485***	0.319
father's occ class: 4	0.065*	0.036	−0.049	0.050	1.324***	0.303
father's occ class: 5	0.038*	0.023	0.012	0.027	0.322*	0.170
father's occ class: 6	0.014	0.035	−0.071	0.046	0.968***	0.303
father's occ class: 7	0.103***	0.040	0.067	0.048	0.501	0.330
father's occ class: 9	−0.021	0.029	0.026	0.034	−0.542***	0.194
father's occ class: 10	0.029	0.027	0.027	0.030	0.000	0.185
mother's occ class: 1	0.047	0.049	0.036	0.060	0.079	0.414
mother's occ class: 2	0.015	0.054	−0.098	0.069	1.379***	0.442
mother's occ class: 3	0.056	0.048	0.053	0.057	0.018	0.388
mother's occ class: 4	0.055	0.040	0.016	0.048	0.451	0.310
mother's occ class: 5	0.010	0.049	0.030	0.057	−0.104	0.414
mother's occ class: 6	0.025	0.040	0.028	0.045	0.030	0.313
mother's occ class: 7	0.055	0.041	0.057	0.047	−0.111	0.316
mother's occ class: 9	−0.004	0.038	0.033	0.044	−0.488*	0.285
mother's occ class: 10	0.004	0.032	−0.005	0.036	0.099	0.256
'nuclear family' to 16	0.028	0.019	0.002	0.022	0.251*	0.136
mid 1990s	−0.045***	0.009	−0.049***	0.010	0.073	0.046
late 1990s	−0.065***	0.014	−0.070***	0.016	0.092	0.081
post 2000	−0.033	0.021	−0.039*	0.023	0.120	0.126
constant	−0.754***	0.250	−0.613**	0.283	−0.157	1.663
# observations	21256		21256		21256	
# individuals	2266		2266		2266	
R^2	0.265		0.087		0.250	
F-test on exclusion of both instruments from first stage: 36.83; Partial R^2 of instrument = 0.0332						

Notes: *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level
Standard errors are clustered at the level of the individual and robust.
'nuclear family' to 16 means lived with both natural parents from birth to age 16.
Reference categories: West Midlands, white, did not live with both natural parents to 16, father/mother occupational class 'plant/machine operative'. Occupational Class dummies: (1) management, (2) professional, (3) associate professional/technical, (4) clerical/secretarial, (5) craft and related, (6) personal/protective services, (7) sales, (9) other, (10) self-emp/unemp.

D Robustness Regarding Use of Panel Dataset

Table D-1: Inverse Probability Weighting: OLS and IV Returns to Education Estimates for Log Hourly Wage

		smoker at 16		RoSLA		both	
	OLS	IV	(first stage)	IV	(first stage)	IV	(first stage)
years of schooling	0.046***	0.131***	— —	0.106**	— —	0.128***	— —
	(0.003)	(0.021)		(0.052)		(0.020)	
smoker at 16 indicator	— —	— —	−0.852***	— —	— —	— —	−0.848***
			(0.107)				(0.107)
min. school LA=16	— —	— —	— —	— —	0.561***	— —	0.540***
					(0.205)		(0.201)
R^2	0.279	0.079	0.246	0.180	0.228	0.094	0.249
F-test on exclusion of instrument(s)	— —	— —	62.97	— —	7.53	— —	34.70
Partial R^2 of the instrument(s)	— —	— —	0.0278	— —	0.0040	— —	0.0316
			Hansen J-test of overidentification			0.183, p-value = 0.6690	

Notes: All regressions: 21256 observations from 2266 individuals. Robust standard errors clustered at individual level in parenthesis.
*** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level
Controls included: age, age^2, year-of-birth, year-of-birth^2, region dummies (10 regions), ethnicity dummies (white, black, asian, other), survey period dummies (early 1990s, mid 1990s, late 1990s, post 2000), dummy for lived with both natural parents from birth to age 16, each parents' occupational class dummies: (1) management, (2) professional, (3) associate professional/technical, (4) clerical/secretarial, (5) craft and related, (6) personal/protective services, (7) sales, (8) plant/machine operative, (9) other, (10) self-emp/unemp.
All regressions include a constant.

Table D-2: Estimates including the no. obs. an individual has as a regressor: OLS and IV Returns to Education Estimates for Log Hourly Wage

		smoker at 16		RoSLA		both	
	OLS	IV	(first stage)	IV	(first stage)	IV	(first stage)
years of schooling	0.047***	0.126***	— —	0.100**	— —	0.123***	— —
	(0.003)	(0.020)		(0.051)		(0.018)	
individual's # obs.	0.009***	0.011***	−0.026*	0.010**	−0.022	0.011***	−0.026*
	(0.002)	(0.002)	(0.015)	(0.002)	(0.015)	(0.002)	(0.015)
smoker at 16 indicator	— —	— —	−0.881***	— —	— —	— —	−0.880***
			(0.108)				(0.107)
min. school LA=16	— —	— —	— —	— —	0.566***	— —	0.558***
					(0.206)		(0.201)
R^2	0.269	0.090	0.247	0.187	0.228	0.105	0.251
F-test on exclusion of instrument(s)	— —	— —	67.17	— —	7.56	— —	37.38
Partial R^2 of the instrument(s)	— —	— —	0.0293	— —	0.0044	— —	0.0336
			Hansen J-test of overidentification			0.189, p-value = 0.6635	

Notes: All regressions: 21256 observations from 2266 individuals. Robust standard errors clustered at individual level in parenthesis.
*** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level
Controls included: age, age^2, year-of-birth, year-of-birth^2, region dummies (10 regions), ethnicity dummies (white, black, asian, other), survey period dummies (early 1990s, mid 1990s, late 1990s, post 2000), dummy for lived with both natural parents from birth to age 16, each parents' occupational class dummies: (1) management, (2) professional, (3) associate professional/technical, (4) clerical/secretarial, (5) craft and related, (6) personal/protective services, (7) sales, (8) plant/machine operative, (9) other, (10) self-emp/unemp.
All regressions include a constant.

Table D-3: Estimates by wave: OLS and IV Returns to Education Estimates for Log Hourly Wage

Wave	OLS: log wage	IV: log wage	First stage: years of schooling	# obs.
1	0.048***	0.140***	−0.851***	1399
	(0.004)	(0.030)	(0.132)	
2	0.046***	0.123***	−1.050***	1377
	(0.004)	(0.023)	(0.130)	
3	0.048***	0.129***	−0.956***	1405
	(0.004)	(0.025)	(0.125)	
4	0.047***	0.127***	−0.863***	1447
	(0.004)	(0.028)	(0.126)	
5	0.048***	0.113***	−0.935***	1488
	(0.004)	(0.025)	(0.126)	
6	0.049***	0.155***	−0.854***	1580
	(0.004)	(0.029)	(0.120)	
7	0.044***	0.130***	−0.905***	1612
	(0.004)	(0.026)	(0.117)	
8	0.047***	0.112***	−0.885***	1595
	(0.004)	(0.026)	(0.121)	
9	0.041***	0.115***	−0.854***	1624
	(0.004)	(0.026)	(0.115)	
10	0.046***	0.135***	−0.889***	1523
	(0.004)	(0.026)	(0.118)	
11	0.048***	0.125***	−0.790***	1429
	(0.005)	(0.031)	(0.124)	
12	0.046***	0.132***	−0.756***	1328
	(0.005)	(0.033)	(0.130)	
13	0.049***	0.146***	−0.757***	1209
	(0.005)	(0.036)	(0.134)	
14	0.047***	0.140***	−0.788***	1143
	(0.005)	(0.033)	(0.136)	
15	0.052***	0.130***	−0.924***	1097
	(0.005)	(0.029)	(0.133)	

Notes: Robust standard errors clustered at individual level in parenthesis.
*** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level
Controls included: age, age^2, year-of-birth, year-of-birth^2, region dummies (10 regions), ethnicity dummies (white, black, asian, other), survey period dummies, dummy for lived with both natural parents from birth to age 16, each parents' occupational class dummies: (1) management, (2) professional, (3) associate professional/technical, (4) clerical/secretarial,(5) craft and related, (6) personal/protective services, (7) sales, (8) plant/machine operative, (9) other, (10) self-emp/unemp. All regressions include a constant.

Table D-4: Estimates using only one ob. per person in the first stage: OLS and IV using Early Smoking

	OLS		IV: smoker at 16		IV: first stage	
		Robust		Robust		Robust
Dep. Var: log hourly wage	Coeff.	Std. Err.	**Coeff.**	Std. Err.	Coeff.	Std. Err.
years of schooling	0.046***	0.003	0.133***	0.021	— —	— —
smoker at 16 indicator	— —	— —	— —	— —	−0.876***	0.097
age	0.099***	0.004	0.101***	0.006	−0.004	0.083
age^2	−0.001***	0.000	−0.001***	0.000	0.000	0.001
year-of-birth	−0.016***	0.007	−0.058***	0.012	0.411***	0.109
year-of-birth^2	0.000***	0.000	0.001***	0.000	−0.003***	0.001
region: North	0.047	0.038	0.061	0.044	−0.187	0.238
region: Yorkshire	0.003	0.033	0.006	0.039	0.136	0.223
region: North West	0.054	0.032	0.048	0.037	0.207	0.226
region: East Midlands	−0.010	0.032	−0.001	0.037	−0.104	0.224
region: East Anglia	0.015	0.039	0.010	0.049	0.180	0.302
region: South East	0.142***	0.028	0.102***	0.035	0.610***	0.186
region: South West	0.023	0.034	0.030	0.037	0.037	0.217
region: Wales	−0.012	0.040	−0.009	0.044	−0.042	0.270
region: Scotland	0.028	0.036	0.014	0.040	0.362	0.229
ethnicity: Black	0.114	0.105	0.096	0.166	−0.034	0.881
ethnicity: Asian	−0.136	0.071	−0.251***	0.092	1.385***	0.458
ethnicity: Other	−0.048	0.103	−0.180	0.140	1.615*	0.841
father's occ class: 1	0.116***	0.028	0.010	0.039	1.163***	0.146
father's occ class: 2	0.121***	0.038	−0.094	0.069	2.314***	0.209
father's occ class: 3	0.089**	0.043	−0.033	0.056	1.369***	0.236
father's occ class: 4	0.065*	0.036	−0.077	0.056	1.439***	0.219
father's occ class: 5	0.038*	0.023	0.007	0.027	0.359***	0.116
father's occ class: 6	0.014	0.035	−0.073	0.046	0.933***	0.189
father's occ class: 7	0.103***	0.04	0.057	0.048	0.490**	0.219
father's occ class: 9	−0.021	0.029	0.025	0.037	−0.494***	0.132
father's occ class: 10	0.029	0.027	0.007	0.030	0.152	0.126
mother's occ class: 1	0.047	0.049	0.025	0.064	0.159	0.296
mother's occ class: 2	0.015	0.054	−0.108	0.071	1.379***	0.313
mother's occ class: 3	0.056	0.048	0.032	0.063	0.179	0.312
mother's occ class: 4	0.055	0.04	0.008	0.049	0.491***	0.225
mother's occ class: 5	0.01	0.049	0.019	0.062	−0.044	0.263
mother's occ class: 6	0.025	0.04	0.030	0.049	0.031	0.228
mother's occ class: 7	0.055	0.041	0.054	0.049	−0.082	0.227
mother's occ class: 9	−0.004	0.038	0.030	0.049	−0.439***	0.201
mother's occ class: 10	0.004	0.032	−0.010	0.039	0.138	0.176
'nuclear family' to 16	0.028	0.019	−0.006	0.023	0.290	0.097
mid 1990s	−0.045***	0.009	−0.039	0.015	−0.130	0.211
late 1990s	−0.065***	0.014	−0.056	0.023	−0.144	0.347
post 2000	−0.033	0.021	−0.023	0.034	−0.055	0.533
constant	−0.754***	0.250	−0.541	0.351	−0.137	6.170
# observations	21256		13498		1432	
# individuals	2266		1398		1432	
R^2	0.265		0.220		0.250	
F-test on exclusion of instrument from first stage: 51.50; Partial R^2 of the instrument = 0.0302						

Notes: *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level
'nuclear family' to 16 means lived with both natural parents from birth to age 16.
Reference categories: West Midlands, white, did not live with both natural parents to 16, father/mother occupational class 'plant/machine operative'. Occupational Class dummies: (1) management, (2) professional, (3) associate professional/technical, (4) clerical/secretarial, (5) craft and related, (6) personal/protective services, (7) sales, (9) other, (10) self-emp/unemp.
IV second stage standard errors obtained by bootstrapping.

Table D-5: Estimates using only one ob. per person in the first stage: OLS and IV using RoSLA

	OLS		IV: RoSLA		IV: first stage	
		Robust		Robust		Robust
Dep. Var: log hourly wage	Coeff.	Std. Err.	**Coeff.**	Std. Err.	Coeff.	Std. Err.
years of schooling	0.046***	0.003	0.143**	0.058	— —	— —
min. school LA=16	— —	— —	— —	— —	0.487***	0.153
age	0.099***	0.004	0.101***	0.006	0.020	0.070
age^2	−0.001***	0.000	−0.001***	0.000	0.000	0.001
year-of-birth	−0.016***	0.007	−0.062**	0.028	0.408***	0.091
year-of-birth^2	0.000***	0.000	0.001***	0.000	−0.004***	0.001
region: North	0.047	0.038	0.058	0.048	−0.132	0.192
region: Yorkshire	0.003	0.033	0.000	0.045	0.142	0.178
region: North West	0.054	0.032	0.047	0.041	0.176	0.175
region: East Midlands	−0.010	0.032	0.004	0.039	−0.139	0.174
region: East Anglia	0.015	0.039	0.007	0.054	0.121	0.243
region: South East	0.142***	0.028	0.093*	0.049	0.609***	0.141
region: South West	0.023	0.034	0.026	0.039	−0.005	0.171
region: Wales	−0.012	0.040	−0.006	0.047	−0.023	0.211
region: Scotland	0.028	0.036	0.007	0.048	0.388**	0.185
ethnicity: Black	0.114	0.105	0.102	0.181	0.123	0.788
ethnicity: Asian	−0.136	0.071	−0.270**	0.134	1.571***	0.377
ethnicity: Other	−0.048	0.103	−0.202	0.173	1.756**	0.695
father's occ class: 1	0.116***	0.028	−0.004	0.078	1.217***	0.149
father's occ class: 2	0.121***	0.038	−0.120	0.155	2.443***	0.213
father's occ class: 3	0.089**	0.043	−0.052	0.101	1.474***	0.244
father's occ class: 4	0.065*	0.036	−0.093	0.102	1.566***	0.222
father's occ class: 5	0.038*	0.023	0.001	0.035	0.359***	0.118
father's occ class: 6	0.014	0.035	−0.084	0.074	0.994***	0.196
father's occ class: 7	0.103***	0.040	0.050	0.057	0.535**	0.223
father's occ class: 9	−0.021	0.029	0.029	0.048	−0.522***	0.133
father's occ class: 10	0.029	0.027	0.003	0.034	0.210	0.128
mother's occ class: 1	0.047	0.049	0.025	0.073	0.132	0.303
mother's occ class: 2	0.015	0.054	−0.130	0.108	1.373***	0.321
mother's occ class: 3	0.056	0.048	0.030	0.068	0.150	0.317
mother's occ class: 4	0.055	0.040	−0.001	0.058	0.485**	0.234
mother's occ class: 5	0.010	0.049	0.014	0.069	−0.112	0.274
mother's occ class: 6	0.025	0.040	0.027	0.052	−0.056	0.234
mother's occ class: 7	0.055	0.041	0.059	0.054	−0.080	0.233
mother's occ class: 9	−0.004	0.038	0.037	0.061	−0.471**	0.208
mother's occ class: 10	0.004	0.032	−0.013	0.042	0.135	0.184
'nuclear family' to 16	0.028	0.019	−0.011	0.033	0.376***	0.097
mid 1990s	−0.045***	0.009	−0.039**	0.015	−0.161	0.209
late 1990s	−0.065***	0.014	−0.056**	0.023	−0.137	0.349
post 2000	−0.033	0.021	−0.024	0.034	−0.030	0.539
constant	−0.754***	0.250	−0.533	0.377	−0.564	5.011
# observations	21256		13498		1398	
# individuals	2266		1398		1398	
R^2	0.265		0.160		0.229	
F-test on exclusion of instrument from first stage: 4.06; Partial R^2 of the instrument = 0.0029						

Notes: *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level
'nuclear family' to 16 means lived with both natural parents from birth to age 16.
Reference categories: West Midlands, white, did not live with both natural parents to 16, father/mother occupational class 'plant/machine operative'. Occupational Class dummies: (1) management, (2) professional, (3) associate professional/technical, (4) clerical/secretarial, (5) craft and related, (6) personal/protective services, (7) sales, (9) other, (10) self-emp/unemp.
IV second stage standard errors obtained by bootstrapping.