Fernihough, Alan

Working Paper

Less is More? The child quantity-quality trade-off in early 20th century England and Wales

QUCEH Working Paper Series, No. 2017-07

Provided in Cooperation with:
Queen's University Centre for Economic History (QUCEH), Queen's University Belfast

QUCEH WORKING PAPER SERIES
http://www.quceh.org.uk/working-papers

LESS IS MORE? THE CHILD QUANTITY-QUALITY TRADE-OFF IN EARLY 20TH CENTURY ENGLAND AND WALES

Alan Fernihough (Queen's University Belfast)

Working Paper 2017-07

QUEEN'S UNIVERSITY CENTRE FOR ECONOMIC HISTORY
Queen's University Belfast
185 Stranmillis Road
Belfast BT9 5EE
September 2017

Less is More? The Child Quantity-Quality Trade-Off in Early 20th Century England and Wales*

Alan Fernihough†

Abstract

Whilst the child quantity-quality (QQ) model is theoretically well-established, the empirical literature offers only partial support. Motivated by the limited causal empirical evidence in both historic and contemporary societies, this study examines the relationship connecting fertility and child quality for individual families in England and Wales at the start of the 20th century. Using data from the 1911 census returns, I estimate whether reductions in family size reduce the probability of leaving school. To account for the endogenous nature of fertility decisions, I use the sex composition of the first two births in families with at least two children as an instrumental variable (IV) for family size. Overall, I find evidence in support of a child QQ effect, as children in the 13–15 age cohort born into smaller families were more likely remain in school. Whilst the IV results are very similar to the non-IV ones, one drawback is that the IV estimates are quite imprecise.

JEL-Classification: J10, N3, O10
Keywords: Quantity-Quality, Human Capital, Demographic Transition

*I am grateful to Cormac Ó Gráda, Neil Cummins, Alan de Bromhead, Zorina Khan, Sascha Becker, Omer Moav, Áine Doran and seminar participants at Bocconi University, Queen's University Belfast, London School of Economics, Stirling University, University College Dublin, and the University of Bayreuth for their helpful comments and advice. The usual disclaimer applies.

†Queen's Management School, Queen's University Belfast, Belfast BT9 5EE, United Kingdom. Email: a.fernihough@qub.ac.uk.

1 Introduction

The link between child quantity (a household's fertility) and quality (the allocation of resources promoting child welfare) has long been of interest to economists. Theoretical models that incorporate a child QQ relationship have the potential to make very powerful and long-reaching predictions that illuminate our understanding of demographic transitions, the rise in demand for human capital, and ultimately the path of economic development across both poor and rich economies (Galor and Weil, 2000; Galor and Moav, 2002). Furthermore, the child QQ trade-off has the potential to inform policy-makers, particularly in developing nations where high fertility and poverty persist (Moav, 2005).

The importance of the child QQ trade-off is underlined by the existence of a large empirical literature that seeks to measure this relationship. These empirical estimates rely on data samples that vary over time, space, and by observation-unit resolution.[1] However, the research findings in this body of work have been far from clear-cut, and a consensus on whether the child QQ model is relevant for economic research has yet to be achieved.

This paper offers an empirical assessment of the child QQ relationship in early 20th century England and Wales. The new evidence provided in this paper is superior to that offered in the existing literature for a number of reasons. Firstly, I employ a large data source from the complete population census and avoid inducing the ecological inference problems that may affect studies relying on aggregated data samples. Secondly, these data represent a rich source of family-level information. The detailed nature of this source allows me to not only model child quantity and quality, but also control for a wide array of confounding factors. These data also permit me to isolate exogenous variation in child quantity and thus estimate causal estimates of the QQ effect. Finally, the context in which these data were recorded is particularly germane. In 1911 British society was in the midst of a rapid and permanent economic and demographic shifts. If the child QQ trade-off underpins economic development, then the relationship should be unequivocally evident in this vital pre-Great War period. The parallels between early 20th century Britain

[1]For example, different resolutions might be aggregated cross-country macro-type data compared to disaggregated household-level data.

and contemporary least developed countries are also pertinent. Many low income-high fertility counties are similar across both dimensions to Britain over 100 years previously.[2]

This paper answers the following research question: does the number of siblings a child between the age of 13 and 15 has (quantity measure) influence the rate at which they leave school? The child QQ model predicts that children from smaller families should enter the labour force later and stay in school longer. This paper uses exogenous variation in fertility to account for the endogenous nature of family size. Angrist and Evans (1998) used the gender-composition of the first two births and examined whether fertility caused changes in female labour supply. This paper also uses a gender composition IV (a dummy variable indicating whether the first two births are the same gender), and finds a similar gender effects as in Angrist and Evans—families for whom the first two births are either both male or female, will be larger.

The same-sex IV operates as an exogenous shifter the the price of child quantity, that influences the demand for child quantity with a subsequent knock-on effect for child quality. It is generally recognized that the within-family sex ratio is an important determinant of child costs (Rosenzweig and Zhang, 2009). Families wherein the first two children are the same gender benefit from economies of scale which are less pronounced for non-same sex families. For example, two boys can more easily share clothes and other basic necessities. These economies of scale are less important for child quality. For example, the gender composition of the first two births should only matter for school attendance indirectly via the effect on child quantity.

A simple linear regression analysis of the conditional distribution of school attendance reveals a relationship consistent with child QQ theory. Children from larger families are more likely leave school at younger ages, according to a linear probability model (LPM) estimated using Ordinary Least Squares (OLS). The estimated QQ effect, an extra sibling reduces the probability of remaining in school

[2]For example, long-run estimates point to a British GDP per capita in 1911 of $4,709 (1990 Int. GK$), a figure that still exceeds the comparable figure for Africa in 2010: $2,034 (Bolt and van Zanden, 2014). Similarly at the turn of the 20th century, the United Kingdom had a total fertility rate of around 3.5 births per woman, whereas the comparable figure for Cameroon in 2014 was 4.7 (`http://www.gapminder.org/data/`).

by 2 per cent for boys and around 1.2 per cent for girls, is robust and consistent across an number of model specifications.

To account for the endogenous nature of fertility an IV analysis is performed. The results of this assessment indicate a substantial downward bias in the aforementioned QQ results. A marginal increase in one child is seen to reduce the likelihood of remaining in school by approximately 6.3 per cent for boys and 2.5 per cent for girls, although the point estimates in both analyses are measured with a degree of uncertainty.

One potential drawback of using IV to estimate our QQ regression model is that this methodology estimates the local average treatment effect (LATE) (Imbens and Angrist, 1994). In other words, the effect parameter we estimate only pertains to the subsample of families for whom the instrument was influential, the compliers. To examine the importance of this issue, I recast the empirical model in a way that allows me to measure both the size of the complier subsample and the demographic characteristics associated with it. The results in this section indicate that only around 1.2 per cent of the analysis sample can be defined as compliers although those to whom the IV QQ effects are relevant exhibit demographic characteristics similar to the wider population. These findings caution against any hasty conclusions surrounding the aforementioned larger IV QQ effects and call for nuanced interpretation.

The remainder of this paper proceeds as follows. Section 2 reviews the relevant literature on the child QQ model and its potential application for economic development. Section 3 discusses the data and contextualizes family-size decisions, childhood labour, and education in early 20th century England and Wales. The empirical analysis is reported in Section 4. Finally, Section 5 concludes.

2 Relevant Literature

Until the advent of the child QQ model by Becker and Lewis (1973), the negative income-fertility gradient puzzled economists. Conventional economic wisdom considers children to be a normal good. However, this convention appeared to be contradicted as income began to rise and fertility fall in Western economies from the mid-19th century on. In Becker and Lewis a coherent theoretical framework is presented wherein children are both a normal good and a rise in income can potentially reduces fertility. In their model parents care about both the number of children and the so-called “quality” of these children, and a rise in income can be associated with a fall in fertility because increases in income result in greater investment in child quality which in turn raise the shadow price of child quantity.

The relationships between income, fertility, and human capital underpin a number of theoretical models that explain economic growth in the very long run. Seminal papers on this topic by Galor and Weil (2000) and Galor and Moav (2002) both assume a child QQ mechanism and this mechanism is an important element that propelled global economic development in the 19th and 20th centuries. Generally, these papers all argue that the increase in income associated with the early phases of the industrial revolution stimulated a greater demand for child quality which, via the QQ trade-off, reduced fertility in Western societies. That the fertility transition happened at the same time as educational attainment levels were rising is consistent with this narrative. This mechanism and its relationship to economic growth is assumed to operate through the following channel: a greater emphasis on child quality leads to increased levels of human capital which in turn boost technological advancements and consequently economic development. As the boost in economic development will also lead to improvements in income which further impact upon child quantity and quality, a virtuous circle is said to exist connecting, fertility, human capital, technology, and income.

This virtuous circle tallies well with the historical pattern of development in most industrialised nations (Galor, 2012). However, the empirical relationship connecting child quantity and quality is somewhat less clear cut. The child QQ model implies that both an increase in child quantity should cause a decrease in quality and that increase in child quality should also lead to a decrease in fertility, and vice versa. Consequently, empirical research in this field has tended to focus on estimating the magnitude of the causal QQ effect, rather than the potentially endogenous conditional relationship. Rosenzweig and Wolpin (1980) were the first to examine the QQ trade-off from this perspective, as their paper used multiple (mainly twin) births as a source of exogenous variation in fertility to identify the effect of increases in quantity on schooling attainment (their measure of quality). Their data sample consisted of rural Indian families surveyed between 1969 and

1971. Rosenzweig and Wolpin results were the first to support the hypothesis that exogenous increases in fertility decrease child quality.

Black et al. (2005) used also multiple births to measure the effect of an exogenous shift of fertility on schooling attainment. Based on a large population return of individuals in late 20th century Norway, the results in Black et al. fail to support the QQ trade-off. Instead, their results stress the importance of within household allocation, in particular birth order effects, as higher birth order individuals appear to do worse than their lower birth order siblings. The exogenous variation created by the multiple birth phenomena has lent empirical support the QQ model based on population samples in 1990 China (Li et al., 2008) and the U.S. in 1980 (Cáceres-Delpiano, 2006). Using a dataset composed of Indonesian children surveyed in 2000 and exogenous variation caused by multiple births, Millimet and Wang (2011) found ambiguous support for the child QQ model when using "health" (measured as either height for age or BMI) as a child quality metric. In contrast, Angrist et al. (2010) found no evidence to support the child QQ model based on a large sample of Israeli's surveyed both in 1983 and 1995 population census.

The use of twin births as a source of exogenous variation in family size has been questioned. Rosenzweig and Zhang (2009) formulate a model wherein families who experience twin/multiple births reallocate greater child quality resources to their non-twin siblings. The empirical research of Bhalotra and Clarke (2016) points to additional source of bias. Using a large sample of births Bhalotra and Clarke show that twin births are not randomly distributed across all mothers. Mothers who are more likely to choose higher levels of child quality are also likely to give birth to twins thus biasing any econometric estimate of the QQ effect towards the null. Finally, Galor (2012) criticised the use of the twin instrument on the grounds that the IV imposed a non-optimal level of fertility on each families, and that this would not necessarily lead to a reduction in child quality but instead a reallocation of resources away from intergenerational transfers. All three criticisms in the above imply that methodologies that employ twin-based instruments lead to an underestimate of the child QQ relationship.

The sex-composition of the children born into a family has been proposed as an alternative to the twin-birth IV. This was first applied to the child QQ literature in Angrist et al. (2010) who found that a desire for a mixed gender composition meant that Israeli parents for whom the first two (and also three) births were the same gender were also more likely to have additional children. In the absence of sex selective abortion the gender of each birth is random and thus creates exogenous variation in fertility. The results presented in Angrist et al. are consistent with their aforementioned results obtained using the twin IV: no evidence of a child QQ trade-off. Conley and Glauber (2006) apply the sibling sex-composition instrument to a sample of data from the U.S. national census in 1990 and find some evidence in support of the child QQ theory, as children with a larger number of siblings are less likely to attend private school and more likely to be retained in the same year. Related studies have also used the son preference in Asian culture to identify the QQ relationship. Both Lee (2008) and Kugler and Kumar (2017) found evidence supporting the child QQ trade-off in late-20th century data samples collected in South Korea and India respectively.

Given the proposed importance of the child QQ model in long-run growth theory, a recent interest has emerged in examining the relationship between fertility and human capital in historic societies. Becker et al. (2010) found that a negative causal relationship connecting fertility and schooling existed on an aggregated level in 19th century Prussia. Their analysis shows that counties with lower fertility had higher levels of education, and furthermore, that variation in education in 1849 predicts the timing of fertility transitions at the end of the 19th century. The analysis performed by Murphy (2015) also found evidence in favour of the child QQ model at the département level in 19th century France. Klemp and Weisdorf (2016) use fecundity (which is measured as the number of days between the marriage date and the first birth) as instrument for family size and find that an increase in the number of siblings reduces the likelihood of literacy and of being employed in an occupation with greater prestige in later life. Hatton and Martin (2010) found that reductions in family size were an important determinant of the increase in stature observed amongst British children at the beginning of the 20th century.

The child QQ model predicts that within an economy that an increase in the preference for educated offspring decreases fertility and a number of researchers have empirically assessed the effect of education on fertility in historical samples.

Bleakley and Lange (2009) used the eradication of hookroom disease from the Southern United States to empirically test this models as the implementation this public health programme acted as shifter on the price of child quality. This programme was found to lower the cost of child quantity thus resulting in fertility reductions. Fernihough (2017) used the 1911 Irish census as a data source and found a strong relationship connecting school attendance to both net and gross fertility at the household. In the absence of any credible instrumental variables Fernihough followed the econometric approach outlined in Millimet and Tchernis (2013). This approach allows researchers to estimate causal effects in the absence of an exclusion restriction. Like Bleakley and Lange, the results presented in Fernihough point towards a substantial QQ effect.

Similar to Fernihough, this study uses census data collected in 1911. However, this study uses the more populous Census for England and Wales and benefits from the advantage of having a greater number of data points as well as a more industrialised economy. The existing evidence on fertility and education in late Victorian and Edwardian Britain suggests a role for the child QQ model. Pooley (2013) contextualises the massive demographic shift that occurred during this time period. The perception that children were economically useful assets for the household changed, and by the beginning of the Great War children are viewed as being more financially burdensome in the the short-run. Given the advantages afforded by education, and that by 1911 state education was both free and compulsory up until the age of 12, households began to focus more resources on educating smaller families.

The changing pattern discussed above is mirrored in the child labour literature. Humphries (2010) and Humphries (2013) links the rise of child labour to the economic model proposed in Basu and Pham (1998). This model proposes that child labour is caused by both a *luxury* axiom, children enter the labour force once the family budget falls below a certain threshold level, and a *substitution* axiom, as the relative productivity of children rises they will be more compelled to enter the labour force. The mechanisation of industry associated with the industrial revolution improved child productivity and this accompanied by static wage growth helps to illustrate why child labour rose during the initial phase of the industrial revolution. The introduction in compulsory schooling and child labour laws in the latter part of the 19th century helped to reduce the already falling incidence of child labour and subsequent increase in school attendance. Further historical context is provided in the following section.

3 Data Source and Historical Context

This paper uses data from the 1911 Census of England and Wales. The complete census returns are available online courtesy of the Integrated Census Microdata Project (I-CeM).[3] This source includes information for over 36 million people present on the 3rd of April 1911. As the 19th century progressed so did the depth of the census in Britain. Earlier censuses were simple headcounts. However, by the end of the 19th century the census enumerators were tasked not only with counting the number of people in the state, but also telling the state who these people were. The 1911 census differed from previous iterations as it was the first to survey fertility. The decision to survey married women about the number of children they have given birth to and the number who are still alive was driven by eugenicist concerns that marital fertility amongst the poorest was highest and this would lead to a deterioration in the genetic stock of the British population (Szreter, 1996, pp. 604–605).

The inclusion of the fertility questions in the 1911 Census of England and Wales means that we can measure household-level fertility accurately. Counting the number of children present in each household, a method used for measuring fertility in earlier censuses is potentially biased by children leaving home and child mortality. The advantage of these census returns is the abundance of data. The disadvantage of these data is that the census did not survey the population for retrospective information. As such, these data represent a snapshot of the population at one period of time. However, one can detect whether or not a child is in school based on their occupation response. Those who remained in school typically completed this field as "Scholar".

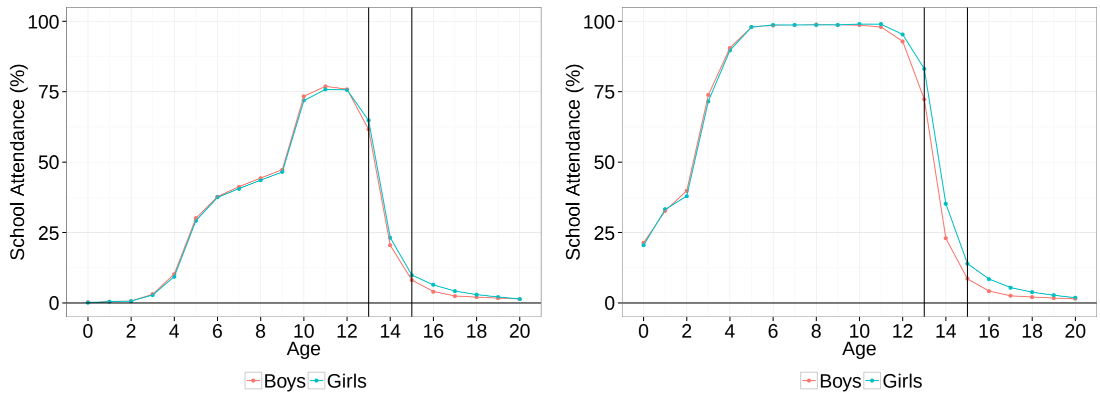
A child denoted as a "Scholar" provides a suitable measure of human capital and thus child-quality investment. However, this variable is sufficiently vague to

[3]These data can be accessed from the following url: `http://icem.data-archive.ac.uk/`.

warrant a further look at it's implications for our understanding of schooling in light of the British educational system at the dawn of the 20th century. The imperfect nature of using the "Scholar" response to infer schooling and the rate at which children were entering into the labour force was highlighted in the main census report published by the Registrar General.[4] Census enumerators were given instructions to ascertain whether or not children were "full time scholars" or "part time scholars" with a view to counting the number and circumstances of children in school. There was however some confusion surrounding this as some parents, completing the census on behalf of their child, left this field blank because they thought this field only related to full-time wage paying occupations.

Figure 1: School Attendance and Age

(a) Full Census (b) Excluding NA Occupation Respondents



Source: Full population data from 1911 census individual returns.

[4]Census of England and Wales, 1911, General Report with Appendices, pp. 147 & 163–165.

Figure 1 illustrates how some uncertainty surrounding the response to the occupation question amongst those completing the census manifested itself in the age-school attendance relationship. In the left-hand panel of Figure 1 shows the ratio of those entering "Scholar" or some variant (including misspellings) as their occupation by age. In this figure we see that this ratio increases until the age of 12 and declines thereafter. However, the proportion of those attending school never exceeds 75 per cent. This pattern is at odds with other historical evidence. The Elementary Education Act of 1880 established a school-leaving age of 10, which was subsequently raised to 11 in 1893 and 12 in 1899 (Trowler, 2003). Employed children under the age of 13 were also required to demonstrate that they had achieved a school-leavers certificate and employers who hired children without this evidence of basic education attainment were fined. Thus, one would expect to see near full attendance for those aged between five and twelve years of age. Furthermore, the plot shown in the left-hand panel of Figure 1 indicates that both boys and girls who are eight are less likely to attend school compared to their twelve year old counterparts. Again this is another pattern which does not stand up to scrutiny when one considers historical accounts of elementary schooling in England and Wales.

The plot in panel (b) of Figure 1 depicts a more accurate trajectory of the school attendance-age trajectory for both boys and girls. In this plot all of the children who left the occupational field blank have been removed from the sample. This plot shows near complete school attendance between the ages of five and eleven. The number in school begins to decline after the age of eleven. This paper uses data on children between the ages of thirteen and fifteen, marked with the vertical black bars. Firstly, as Figure 1 shows, the vast majority of children leave school during this stage of their life. Secondly, as we need to link sibship size at the individual level to the household mother's response, sibship size will be more closely aligned to completed fertility than with younger cohorts. The figures reported in panel (b) also tally well with those used in the Lewis report.[5] In this report Lewis, who used the 1911 census, estimated that around 95 per cent of those aged between 12 and 13 were in some engaged in some form of educational

[5]Great Britain. Board of Education. Committee on Juvenile Education in Relation to Employment After the War and Lewis, J.H., 1917.

activity.

The period between the mid-19th and early 20th centuries was one of dramatic change in the education system of England and Wales. Through a series of acts the British state sought to broaden the access to education and standardise educational material. Key in this was Foster's "Elementary Education Act" of 1870, which established over 3,000 school boards mostly funded by local ratepayers with the task of delivering a secular education to all aged between 5 and 12 (Gordon et al., 1991, p. 17). The 1902 Education Act further standardised the education system by replacing the school boards with 318 Local Education Authorities and making clearer distinctions between elementary, secondary, and technical institutions (Gordon et al., 1991, pp. 26 & 157). By this stage there existed public elementary schools which did not require school fees and in 1907 the government, under pressure from trade unions to do more for the working classes, required secondary schools to provide up 25 per cent of places to those non-fee payers graduating from the public elementary system (Gordon et al., 1991, p. 174).

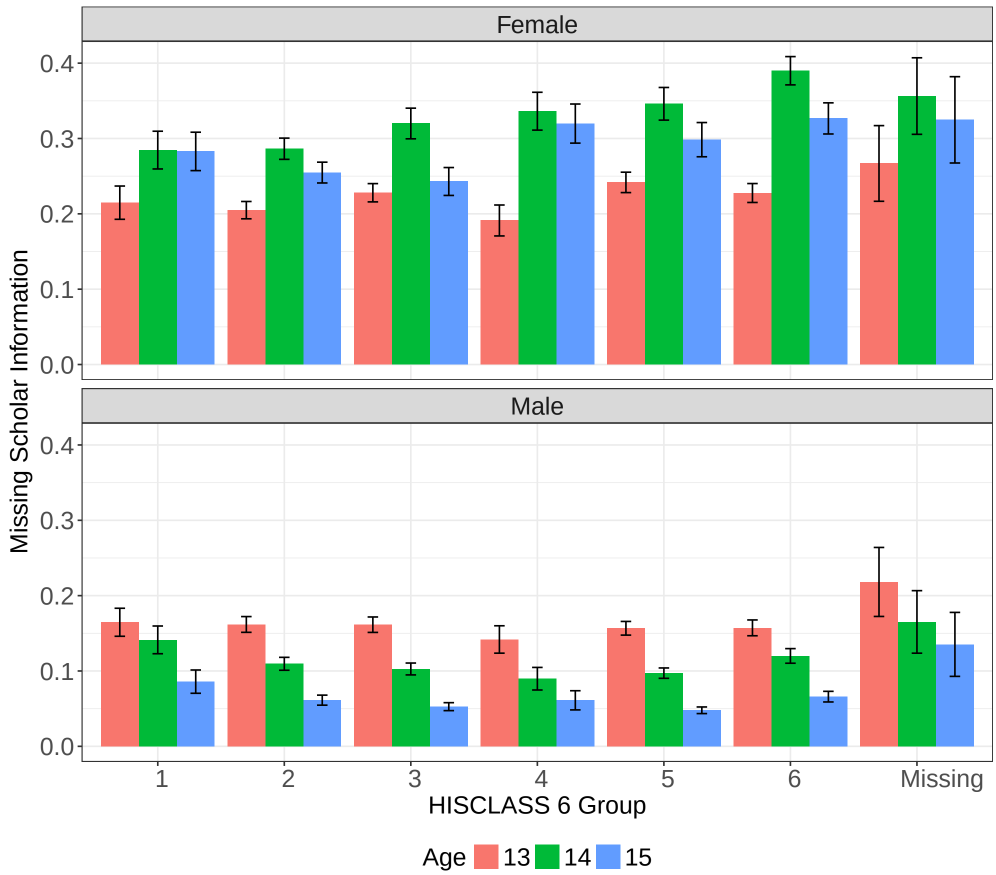
During the 1909–1910 academic year, 157,874 pupils attended secondary school, testament to the growing popularity of the secondary schools (Gordon et al., 1991, p. 175). That 46 per cent of those in secondary school were women highlights only a minor gender imbalance. Whilst this minor imbalance is also evident in Figure 1, it is worth considering what staying in school implied and how this differed amongst the sexes. Using unique data from personal biographies mostly recorded during the period 1750–1850, Humphries (2010, pp. 328–340) found a positive relationship connecting schooling to occupational prestige amongst a sample of men. Staying in school as well as attending Sunday school, and completing some form of adult education were all positively correlated with the CAMSIS occupational stratification score. If education could lead to socioeconomic advancement before 1911, then there is little to suggest that this had diminished by the beginning of the 20th century as schooling had largely replaced patronage as the main driver of recruitment to the civil service, army, and numerous other professions. However, the opportunities created by education were undoubtedly greater for men. Women who remained in school might have a greater chance of securing employment in retail, secretarial, and clerical occupational roles. Education also offered women the opportunity to prepare for life as a homemaker and the post-elementary school curricula for girls commonly featured dressmaking, cookery, and laundry whereas the curricula for boys focussed more on manual training (Gordon et al., 1991, p. 127).

Girls aged between 13 and 15 were more likely to be both finished schooling and outside the labour force. Many families removed their daughters from school to help with housework, effectively making then live-in servants (Pooley, 2013). This wasn't solely a pursuit of lower income working-class families, but also prevalent amongst the middle and upper classes (Gittins, 1982, pp. 69–70). The observations with missing occupational fields are thus a mixture of both children who were in school, working, and also both outside the labour force and education system. We would expect that the census microdata excludes more girls as they are more likely to leave the occupation cell of census form blank. This is confirmed in Figure 2 which shows the proportion of individuals with missing information stratified by age, gender, and father's occupational class. The father's occupations were first classified according to the HISCO scheme (van Leeuwen et al., 2002), which was then translated into the HISCLASS-6 categorisation ranking occupational prestige from 1 (the highest) to 6 (the lowest).[6] Additionally, this figure also includes a category for the small number of children (around 0.8 per cent) whose father's occupation return was missing.

There is an apparent and unsurprising disparity in missing occupational information between boys and girls. Girls are more likely to have left the occupation field blank. A number of other differences are also evident in this plot. The "missing" information declines with age amongst males but not by social class. This is consistent with the idea that these observations are "scholars", but did not enter this information. The pattern amongst females is more ambiguous. Here we see that the lowest prevalence of missing data is amongst the youngest (i.e. those aged 13). The increases that we see occurring after this age, in contrast to the variation amongst males, supports the historical accounts of that school leaving but not joining the labour force, either immediately or at all, was widespread. The female observations, in these instances will be a mix of scholars and non-scholars who remain outside the world of work.

[6] See Meier zu Selhausen et al. (2015) for another research paper that uses this scheme.

Figure 2: Empty Occupation Field by Father's HISCLASS Group Stratified by Age



Source: Analysis sample from 1911 census individual returns.

The timing of the transition from school to work clearly represents some form of quality investment in children. However, the transition from school to homemaker is somewhat less clear cut in this regard. Thus we omit all of those with missing occupation from the analysis sample. Later, I examine the potential bias that this omission may introduce via a method known as Multivariate Imputation by Chained Equations (MICE). These data are also, necessarily, trimmed across a number of other dimensions. Angrist and Evans (1998) applied similar sample restrictions. Firstly, we examine only "nuclear families" wherein both the mother and father are included in the same household form as the child. We remove a large number of observations with missing parents, although it is is certainly debatable how relevant the QQ model is in explaining fertility-schooling relationship amongst households where a parent is deceased or absent from the family residence for sustained periods of time. Our sample of children from nuclear families can measure the conditional correlation between sibship size and school attendance. The empirical strategy pursued here designed to measure causal effects relies on some intrahousehold variation, and thus some additional restrictions are required. All observations from households where the number of children present on the day of the census does not equal the number of children the mother has given birth to (and are alive) are removed. Finally, since this paper uses the sex composition of the first two births as an instrument for fertility we only examine only first-born singletons who have at least one other sibling. Using higher birth orders is problematic in this instance as schooling for the later born siblings will come from an endogenous sample if fertility is indeed endogenous (Angrist et al., 2010; Frühwirth-Schnatter et al., 2014).

Table 1 provides an illustration of how the composition of the 13 to 15 year olds changes as the sample restrictions outlined in the above are applied. We see that the number of observation falls from around 2.5 million to 121,775 individuals. The analysis sample is somewhat younger, and has younger parents, as one might expect when one compares these to all of those aged 13 to 15 in the general population. Consequently, the sibship size is less in the analysis sample.[7] One would expect for sibship size to fall as we move from the sample composed of nuclear families to children where all of the siblings are present in the household for the census. Here we can see that the average number of children present in the Analysis Sample (the final column) is four, which is smaller than the average of six in Nuclear Families, but still large by modern day standards in Western society. Comparisons of other demographic characteristics reveal only some relatively minor differences. Notably, there are few differences between both the proportion in school and observations with missing occupation information. Thus, the analysis sample provides a reliable sample of the population in terms of the school attendance measure.

[7]Note that there is no "Number of Siblings" average available for the Full Population as this information was taken from the indicated mother of each child.

Table 1: Internal Validity Of Sample

	Full Population	Nuclear Families	Alive= Present	First Borns, # Siblings > 0
Comparison of Means				
N	2,591,012	1,386,044	640,192	231,480
N (Occupation Incl.)	2,131,265	1,141,557	529,744	187,065
Male	0.495	0.515	0.512	0.515
Age	14.226	14.186	13.925	13.943
Servant Present	0.050	0.048	0.055	0.060
Number of Rooms in House	5.346	5.112	5.076	4.866
In Boarding School	0.004	0.000	0.000	0.000
Mother's Age		43.654	41.329	38.627
Father's Age		45.997	43.506	40.819
Number of Siblings		5.112	3.409	2.946
Proportion In School (Missing Observations Excluded)				
Age 13	0.774	0.774	0.766	0.774
Age 14	0.267	0.276	0.288	0.297
Age 15	0.092	0.096	0.111	0.112
Proportion Missing Occupation Observation				
Age 13	0.179	0.168	0.167	0.193
Age 14	0.216	0.214	0.199	0.217
Age 15	0.156	0.160	0.150	0.163

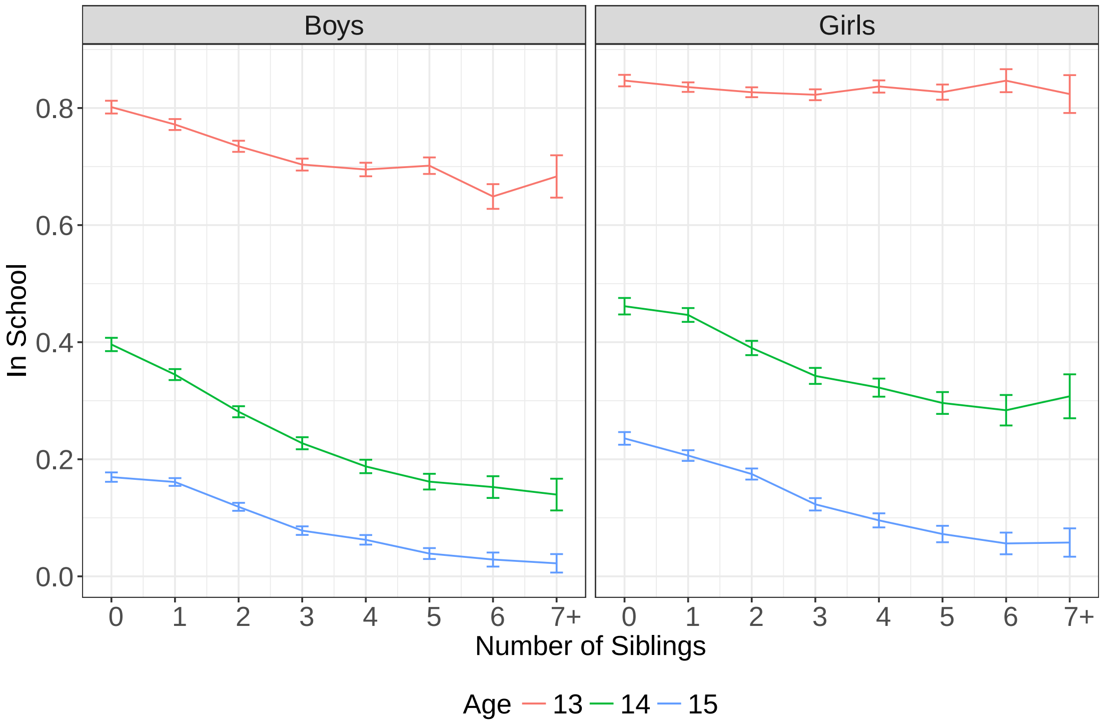
4 Empirical Analysis

The empirical analysis is based upon the following linear probability model (LPM):

$$\mathbf{s} = \alpha + \beta\mathbf{n} + \mathbf{X}\boldsymbol{\gamma} + \mathbf{u} \tag{1}$$

where $\mathbf{s}$ is a binary indicator variable that takes the value of 1 if the observation is a scholar and 0 otherwise, $\mathbf{n}$ is the number of siblings each child has, $\mathbf{X}$ contains various control variables for different model specifications, and finally $\mathbf{u}$ is the usual idiosyncratic error term.

Figure 3: School Attendance by Sibship Size Stratified by Age



Source: Analysis sample (but including only children) from 1911 census individual returns.

The parameter β is the slope coefficient and the child QQ theory predicts that this should be negative because increases in the number of siblings should reduce the likelihood that the child remains in school. Figure 3, plotting the trajectory of school attendance by sibship size for both boys and girls stratified by age, appears to be consistent with this prediction. This figure uses the analysis sample (but includes only children for completeness) and we see that girls are more likely to remain in school than their male counterparts, although this is more of an artefact of the removal of blank occupational returns observations from the analysis sample, as 13 to 15 year old girls were far more likely to be outside of both school and the labour force. In the left-hand panel there exists a negative sibship size-scholar gradient at all ages, but this is most pronounced for children aged 14 wherein a an only child is a scholar 40% of the time and a child with seven or more siblings would be expected to be in school in 15% of all instances. The relationship appears to follow a monotonic trajectory that is approximately linear. A broadly similar pattern is observed for girls, depicted in the right-hand panel of Figure 3. The major difference between the two panels is the flat gradient for girls aged 13.

4.1 OLS and 2SLS Analysis

Table 2 displays sibship size coefficient estimate from various OLS regression models. These results relate to the "Analysis Sample", so the aforementioned sample restrictions are in place. As noted in the literature review, the reasons for staying in school differed by gender and school curricula reflected this difference. By 1911, boys age 13–15 were being prepared for a life in work, whereas the female curriculum placed a greater emphasis on the home-making skills that would assist housewives. Therefore, we split all our analyses by gender. We also consider three separate model specifications, which differ on the basis of the control variables included in $\mathbf{X}$. The *Baseline* model only includes the following basic demographic controls: age, mother's age, and father's age. The *Standard* specification adds to this by including regional controls (11 NUTS1 regions[8]), a father's HISCLASS-6 fixed effect to account for socioeconomic differences between observations, two dummy variables that indicate whether the mother and father are in the labour force, and finally an indicator for observations living in an urban parish. Finally,

[8]These include: North East, North West, Yorkshire and The Humber, East Midlands, West Midlands, East, London, South East, South West, and Wales.

Table 2: Child is a Scholar: OLS Results

	(1)	(2)	(3)	(4)	(5)	(6)
Sibship Size	−0.023	−0.020	−0.018	−0.013	−0.013	−0.012
	(-0.025,-0.021)	(-0.022,-0.018)	(-0.020,-0.017)	(-0.015,-0.011)	(-0.015,-0.011)	(-0.014,-0.010)
Controls	Baseline	Standard	Full	Baseline	Standard	Full
Sample	Male	Male	Male	Female	Female	Female
Observations	106,083	106,083	106,083	80,975	80,975	80,975

Note: The baseline model contains controls for age, father's age and mother's age. The following fixed effects are added in the standard specification: region, father's HISCLASS-6 category, and indicators for whether the mother is in the labour force and urban residence. The full controls model uses the baseline controls specification alongside the following fixed effects: parish, number of rooms in house, number of servants in house, mother's occupation code, father's occupation code. The analysis sample consists of all first borns with at least one sibling with all siblings present. The 95% confidence intervals have been adjusted to account for clustering at parish level.

the most complete *Full* specification controls for the following in addition to the *Standard* specification: parish fixed effects, the number of rooms in the house, number of servants in the house, and finally mother and father occupation code fixed effects. The advantage of analysing three specifications is that it provides a rough assessment of the degree to which omitted variable bias is present.

Columns (1)–(3) of Table 2 list the child QQ coefficient for boys in the sample based on the three model specifications. The coefficient for sibship size in column (1) is -0.023 and the interpretation is simple: for every extra sibling a boy aged 13–15 is 2.3 per cent less likely to attend school. The sibship size coefficient falls to -2.0 per cent and -1.8 per cent in the more comprehensive *Standard* and *Full* specifications. This relatively small change suggests that omitted variable bias is a relatively minor concern. Note too the reasonably tight 95% confidence intervals (adjusted for clustering at the parish level) that show the precision to which these parameters are estimated. Indeed, the lower confidence interval in all cases is clearly far from the null. Columns (4)–(6) provide the equivalent QQ coefficient estimates for girls. Again we find evidence of a QQ effect of around -1.3 per cent. That this effect is robust across the three model specifications alleviates reasonable concerns of omitted variable bias being an important source of confounding variation. The gender difference in QQ effect found here is consistent with the idea that the context of the outcome variable, whether the child was in school, was not the same for girls as boys. We find a stronger QQ effect for boys, as one might expect if education was more of a parental "investment" than for girls, for whom education may have been more of a consumption-based decision.

The internal validity of the results shown in Table 2 is of some concern, as it restricts the analysis sample across a number of dimensions. This sample may not be a fair reflection of the population in England and Wales in 1911. Table 3 alleviates this potential concern, as it re-runs the regression models from Table 2 but on a larger sample drawn from all "nuclear families" (the summary statistics for which are in the third column of Table 1). The results in this table are consistent with the previous OLS results. However, these coefficients are slightly reduced in magnitude and the male-female dichotomy is less evident. The large sample size here means that the QQ effect parameters are estimated with even greater precision. This is reflected in the very narrow confidence intervals which, as before,

Table 3: Child is a Scholar: OLS Results from Nuclear-Family Sample

	(1)	(2)	(3)	(4)	(5)	(6)
Sibship Size	−0.017	−0.014	−0.012	−0.015	−0.013	−0.011
	(-0.018,-0.016)	(-0.015,-0.014)	(-0.012,-0.011)	(-0.016,-0.013)	(-0.013,-0.012)	(-0.012,-0.011)
Controls	Baseline	Standard	Full	Baseline	Standard	Full
Sample	Male	Male	Male	Female	Female	Female
Observations	652,463	652,463	652,463	487,464	487,464	487,464

Note: The baseline model contains controls for age, father's age and mother's age. The following fixed effects are added in the standard specification: region, father's HISCLASS-6 category, and indicators for whether the mother is in the labour force and urban residence. The full controls model uses the baseline controls specification alongside the following fixed effects: parish, number of rooms in house, number of servants in house, mother's occupation code, father's occupation code. The analysis sample consists of all children for whom both the mother and father are present in the census returns. The 95% confidence intervals have been adjusted to account for clustering at parish level.

do not come close to overlapping the null. The regression model results presented thus far point towards a substantial and precisely measured QQ effect. The difference between coefficients across different model specifications effectively downplay the importance of omitted variable bias. Whilst the *Full* specification contains a comprehensive set of control variables, they may not fully capture other sources of confounding variation. Furthermore, other research papers in this literature have highlighted simultaneity bias as additional factor which may bias the QQ coefficient estimated via OLS. To this end, we further the analysis using an instrumental variables (IV) approach and estimating the QQ effect via two-stage least squares (2SLS).

The prototypical 2SLS model of the child QQ effect is thus:

$$\mathbf{s} = \alpha + \beta\hat{\mathbf{n}} + \mathbf{X}\boldsymbol{\gamma} + \mathbf{u} \tag{2}$$

$$\mathbf{n} = \theta + \rho\mathbf{z} + \mathbf{X}\boldsymbol{\phi} + \mathbf{v} \tag{3}$$

where the endogenous regressor $\mathbf{n}$ is instrumented using the variable $\mathbf{z}$. Here, we follow Angrist et al. (2010) and Angrist and Evans (1998) and use the gender composition of the first two births as the instrumental variable. More formally, the variable $\mathbf{z}$ is a dummy variable that takes the value of one if the first two births are the same gender and zero otherwise. Since our analysis is split by gender and only contains first borns, this dummy variable can only take a value of one when both the eldest and second eldest sibling are male and vice-versa for females. The rationale for this instrument is that the sex-composition of the first number of births in a household is a significant causal factor determining whether or not parent's decide to have further births. In modern Western societies such as the United States or Israel, this instrument works because a significant proportion of the parents want at least one child of either both genders. How this IV works in a high-fertility society over one-hundred years ago is something this paper will investigate.

The validity of any IV approach typically relies on whether the instrumental variable's exclusion restriction holds. In other words we need to ask whether or not the gender composition of the first two births could influence whether or not those in our sample are in school in any ways that are unrelated to their sibship

Table 4: Child is a Scholar: 2SLS Results

	First Stage					
	(1)	(2)	(3)	(4)	(5)	(6)
Non-Mixed Gender Comp.	0.059	0.060	0.059	0.054	0.055	0.058
	(0.040,0.079)	(0.041,0.079)	(0.039,0.080)	(0.034,0.075)	(0.035,0.076)	(0.038,0.079)
	Second Stage					
	(1)	(2)	(3)	(4)	(5)	(6)
Sibship Size	−0.044	−0.057	−0.063	−0.070	−0.028	−0.025
	(-0.127,0.038)	(-0.134,0.021)	(-0.143,0.017)	(-0.173,0.032)	(-0.122,0.065)	(-0.109,0.059)
Controls	Baseline	Standard	Full	Baseline	Standard	Full
Sample	Male	Male	Male	Female	Female	Female
First Stage F-Stat	35.827	37.371	33.233	26.555	27.817	30.411
Observations	106,083	106,083	106,083	80,975	80,975	80,975

Note: The baseline model contains controls for age, father's age and mother's age. The following fixed effects are added in the standard specification: region, father's HISCLASS-6 category, and indicators for whether the mother is in the labour force and urban residence. The full controls model uses the baseline controls specification alongside the following fixed effects: parish, number of rooms in house, number of servants in house, mother's occupation code, father's occupation code. The analysis sample consists of all first borns with at least one sibling with all siblings present. The 95% confidence intervals have been adjusted to account for clustering at parish level.

size. Firstly, the gender composition variable is clearly exogenous, as parents cannot choose the sex of their children and sex-selective abortion was unavailable in early 20th century Britain. Secondly, the analysis sample removes all one-child families from these data, so there has to be at least one sibling for every observation. This means that parent's cannot observe the gender of the first born and then have no other births.

The relevance of the instrument is also important. In modern Western societies, families with more than three children are more the exception than the rule. In historic societies the opposite was true. Thus, one might be concerned that this IV does not have an effect because families are going to be large regardless of the gender of the first two births. The top section shown in Table 4 addresses this concern. This section displays the first-stage OLS coefficient estimates one obtains when the sibship size variable is regressed on the non-mixed gender composition variable in addition to the various control variables which vary from column to column based on the three model specifications. An interpretation of the first column is as follows: first born boys, who have a brother as the second born are predicted to have 0.059 extra siblings compared to first born boys with a female second born. The coefficient on the non-mixed gender composition are qualitatively similar. That these coefficients do not vary according to the number and extent of the control variables included is in keeping with the assumption that the IV is determined exogenously. This magnitude of the IV coefficient, which is approximately 0.06, is similar to the results reported by Angrist et al. (2010) and the relatively large first-stage F test statistics underline the instrument's relevance in this application. Unlike Angrist et al., little difference in these coefficients is seen across the genders. This, somewhat surprisingly, indicates the absence of any sex-preference in 1911 England and Wales as this was a patriarchal society. However, it is also well-documented that daughters, or at least one daughter, was also desirable as daughters could help with domestic duties, remove the need for any domestic service and thus "tip up wages" (Pooley, 2013).

The bottom half of Table 4 reveals the second stage IV results. The results relating to the male only sample are shown in the first three columns. Here, the estimated QQ effect varies from -0.044 to -0.063, so an extra sibling decreases the probability of school attendance by around around 5.5 per cent. These effects are larger in magnitude than the comparable OLS results shown in Table 2. However, these coefficient estimate are more uncertain and thus have much wider 95% confidence intervals than the comparable OLS estimates. These confidence intervals overlap both the previous OLS estimates and the null, such that one cannot rule out the possibility that the negative result is a result of sampling variation rather than the child QQ trade off. This pattern is repeated in the female only sample. Again we find a larger QQ effect estimate although the wider confidence intervals hasten against the conclusion that the original OLS estimates are biased downwards.

4.2 MICE Analysis

Section 3 notes the potential for missing data to impact upon this analysis. Without knowing whether or not around 20 per cent of observations with missing data are in school or have left school but are not working, we have excluded these from the analysis. However, one would like to know the potential impact that leaving these data observations has on the subsequent coefficient estimates. If excluding these observations creates a bias, in what direction does this bias go?

To assess the significance of the incomplete occupation field, I treat this problem as one of missing data and use a well-known method of multiple imputation: multivariate imputation by chained equations (MICE). We do not assume that these missing observations are missing at random, but instead use available covariates to perform logistic regression imputation (van Buuren and Groothuis-Oudshoorn, 2011). For example, in Figure 2 we observe that 14 year old girls with a father employed in an occupation that fits into the lowest HISCLASS grouping are less likely to enter information for the occupational field in the census. The predictive nature of the MICE method allows us to impute these missing observations based not completely at random but instead based on the characteristics of individuals most like the observations for whom these data are missing. As this procedure is computationally burdensome, I use the *Standard* model specification as this strikes a balance between having a reasonably rich set of predictive covariates to base the MICE sampling algorithm upon and being computationally feasible.

Table 5: Child is a Scholar: OLS and IV MICE Results, Pooled Coefficients

	(1)	(2)	(3)	(4)
Sibship Size	-0.021	-0.013	-0.060	-0.030
	(-0.022,-0.019)	(-0.015,-0.012)	(-0.142,0.023)	(-0.205,0.145)
Controls	Standard	Standard	Standard	Standard
Sample	Males	Females	Males	Females
Model	OLS	OLS	IV	IV
First-Stage F-Stat			39.728	11.73
Observations	119,133	112,337	119,133	112,337

The standard specification contains controls for age, gender, father's age and mother's age and the following fixed effects: region, father's HISCLASS-6 category, and indicators for whether the mother is in the labour force and urban residence. The analysis sample consists of all first borns with at least one sibling with all siblings present. The 95% confidence intervals have been adjusted to account for clustering at parish level. The pooled coefficients are calculated by setting the number of multiple imputations to 250.

Table 5 displays the OLS and IV results when multiple imputation methods are used. The MICE procedure accounts for uncertainty in the imputation methods by generating multiple imputation samples. In this instance, 250 samples were generated and analysed. The coefficients and other statistical information represent the pooled estimates from these multiple samples. As a point of comparison, one should look at columns (2) and (5) of Tables 2 and 4 respectively. The first feature of Table 5 is the increased sample size, increasing to 119,133 males and 112,337 females. The second feature is that the coefficients are very similar to those reported in the earlier tables. The OLS result for males goes from -0.020 to -0.021 and the IV estimate from -0.057 to -0.060. For females the OLS is the same once the missing data is incorporated (-0.013) and the IV method reports a slightly larger QQ effect of -0.030 compared to -0.028. Overall, these estimates do not indicate that the exclusion of these missing occupation data results in any bias to the QQ estimates. Finally, also of note are the measures of uncertainty, the 95% confidence intervals. It appears that including these imputed observations does little to improve the precision of these QQ effect estimates. One feature of the MICE method is that it accounts for the uncertainty of the imputed values, as these themselves are estimated from data. It appears that in this instance this extra uncertainty cancels out any boost to the precision that would have resulted by increasing the sample size.

4.3 LATE Interpretation

The IV results are somewhat inconclusive, although the most sensible interpretation of these results is that they fail to indicate that the OLS results are biased upwards. If anything, the OLS results are biased downwards, although a conventional Hausmann exogeneity test cannot reject the null that the sibship size variable is an exogenous regressor. Another feature of the IV results worth consideration is that they reveal the local average treatment effect (LATE) rather than the average treatment effect (ATE). Interpretation of the LATE is easiest when the endogenous regressor is a binary treatment variable and thus we change the sibship size variable **n** to a binary indicator variable that takes the value of one if the child in the sample has at least two siblings. The affect this change has on the results is shown in Table 6. The OLS results shown in the top panel are unchanged, having less siblings boosts the chance that you will remain in school. Again the IV results fail to indicate that the OLS results are biased away from the null but the caveat surrounding the large confidence intervals still applies. As these results are similar we can now use them to analyse how the LATE operates here.

Under the LATE framework, the IV coefficient estimate represents the effect of sibship size on schooling amongst the sample of compliers. Compliers represent children from families wherein the parents were influenced by the instrument. In other words, the children born to parent who were going to only have two children, but because their first two births were either both boys or girls they decided to have an extra child. The non-complying population includes "always-takers", those parents who will have more than two children regardless of the sex of the first two births, and "never-takers", those who will only have two children regardless of the circumstances.

Table 6: Child is a Scholar: OLS and 2SLS Results

	OLS					
	(1)	(2)	(3)	(4)	(5)	(6)
2 or More Siblings	−0.074	−0.064	−0.058	−0.040	−0.038	−0.034
	(-0.082,-0.066)	(-0.071,-0.057)	(-0.065,-0.052)	(-0.049,-0.032)	(-0.045,-0.031)	(-0.041,-0.028)
	2SLS					
	(1)	(2)	(3)	(4)	(5)	(6)
2 or More Siblings	−0.121	−0.156	−0.174	−0.271	−0.110	−0.099
	(-0.348,0.106)	(-0.371,0.058)	(-0.398,0.049)	(-0.668,0.126)	(-0.476,0.255)	(-0.437,0.239)
Controls	Baseline	Standard	Full	Baseline	Standard	Full
Sample	Male	Male	Male	Female	Female	Female
First-Stage F-Stat	78.369	79.664	67.535	25.776	26.034	26.964
Observations	106,083	106,083	106,083	80,975	80,975	80,975

Note: The baseline model contains controls for age, father's age and mother's age. The following fixed effects are added in the standard specification: region, father's HISCLASS-6 category, and indicators for whether the mother is in the labour force and urban residence. The full controls model uses the baseline controls specification alongside the following fixed effects: parish, number of rooms in house, number of servants in house, mother's occupation code, father's occupation code. The analysis sample consists of all first borns with at least one sibling with all siblings present. The 95% confidence intervals have been adjusted to account for clustering at parish level.

In these data the average child had 3 siblings so one might expect the "always-takers" to dominate, however we can use the routine proposed in Angrist and Pischke (2009) to figure out the size of the complier population without actually identifying the complying individuals. Under this the compliant subpopulation is:

$$P[d_{1i} > d_{0i} | d_i = 1] = \frac{P[z_i = 1](E[d_i | z_i = 1] - E[d_i | z_i = 0])}{P[d_i = 1]}$$

where d_{1i} or d_{0i} is the ith observation's treatment status, in this instance whether the observation has two or more siblings, given the IV takes a value of one or zero, $z_i = 1$ or $z_i = 0$ respectively. So the probability that individual i's treatment status changes because $z_i = 1$, $P[d_{1i} > d_{0i} | d_i = 1]$, is the product of the probability of having the first two births being the same gender ($P[z_i = 1]$) and effect that this IV has on fertility ($E[d_i | z_i = 1] - E[d_i | z_i = 0]$)) divided by the probability of having two or more siblings ($P[d_i = 1]$). Thus, the size of the subcomplier population for boys is:

$$\frac{0.505 \times 0.022}{0.767} \approx 0.014$$

and the equivalent subpopulation for girls is:

$$\frac{0.498 \times 0.015}{0.749} \approx 0.010.$$

So the large and uncertain effect found in the IV analysis relates only to a relatively small fraction of the population, although it is still identifying a causal effect. We can, as Angrist and Pischke show, further probe these data and see how different the compliant subsection of the population is different to the population generally. If large differences exist, this would call into question the external validity of the results.

Table 7 sheds some light into the sample differences between the compliers and overall sample. Some ambiguity exists surrounding the "urban" variable as it appears that complying male observations are less likely to reside in an urban parish compared to the male sample generally, however female compliers are one and a half times more likely to reside in urban parishes than the female sample overall.

Table 7: Complier Characteristics Ratios

Covariate	Mean (Everyone)	Mean (Compliers)	Ratio
	Male Sample		
Urban	0.099	0.048	0.486
Servants	0.058	0.139	2.400
Mother in Labour Force	0.123	0.087	0.711
Father Unskilled/Low Skilled	0.512	0.487	0.952
	Female Sample		
Urban	0.092	0.147	1.593
Servants	0.062	0.178	2.887
Mother in Labour Force	0.156	0.158	1.010
Father Unskilled/Low Skilled	0.487	0.414	0.850

Unlike the ambiguous urban variable, the other covariates tell a similar story in both the male and female samples. Compliers are between two and three times more likely to have a servant present in the house, a marker for wealth and a loftier socioeconomic position. Despite this, the the mother's labour force participation indicator and another a variable that indicates whether or not the father is in unskilled or low skilled employment do not appear to differ dramatically. Overall, Table 7 does not suggest that the complier subsample, whilst small, does not differ dramatically from the sample as whole. There is some evidence that compliers were more likely to reside in a household with a domestic servant, however the higher complier characteristic ratio merely could be a reflection of the relatively small proportion of households, around 6 per cent, with a domestic servant present in the full sample.

4.4 Nonlinear Analysis

Thus far we have assumed that the QQ effect is approximately linear. A one child increase leads to an approximate 2 per cent decline in the probability of school attendance across the sibship size distribution. In light of the LATE analysis performed in the above, this subsection probes this linearity assumption. To examine the potential for nonlinearity I use the control function probit IV method advocated in Rivers and Vuong (1988). This method is represented as follows:

$$\begin{aligned} \mathbf{s} =& 1(\alpha + f_1(\mathbf{n}) + \mathbf{X}\boldsymbol{\gamma} + \mathbf{u} + f_2(\hat{\mathbf{v}}) \geq 0) \qquad (4) \\ \mathbf{n} =& \theta + \rho\mathbf{z} + \mathbf{X}\boldsymbol{\phi} + \mathbf{v} \qquad (5) \end{aligned}$$

where the first stage residuals $\hat{\mathbf{v}}$ are included in the second stage to "control" for confounding endogeneity. Marra and Radice (2011), Rothe (2009), and Blundell and Powell (2004) have all documented how the control function approach can be extended to permit the nonlinearities in the endogenous regressor $\mathbf{n}$ and the first-stage residuals $\hat{\mathbf{v}}$ by using semiparametric or even less restrictive nonparametric estimators. For this application I adopt a more parsimonious approach and thus the functions f_1 and f_2 represent cubic polynomials.[9]

As highlighted in Blundell and Powell (2004) the estimated second stage model from the control function above can be used to trace out the Average Structural Function (ASF). The estimated ASF is also known as the policy response function and traces out the how the outcome $\Pr(\mathbf{s})$ changes in response to movements in the regressor of interest: $\mathbf{n}$. In this application the ASF can be calculated:

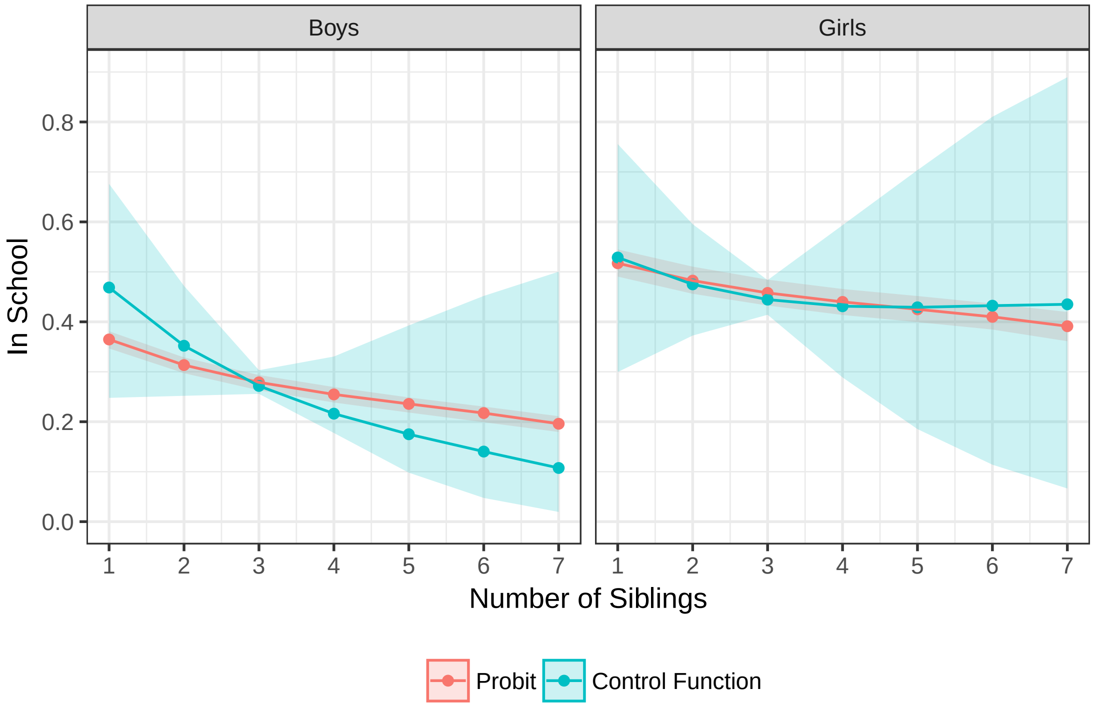
$$\widehat{ASF} = \Pr(\mathbf{s} = 1 | \mathbf{n} = k, \mathbf{v}, \mathbf{X}) = N^{-1} \sum_{i=1}^{N} \Phi \left(\hat{\alpha} + \hat{f}_1(k) + \bar{\mathbf{X}}_{\mathbf{i}} \hat{\boldsymbol{\gamma}} + \hat{f}_2(\hat{v}_i) \right) \qquad (6)$$

where $\widehat{ASF}$ is the probability that a child with a sibship size of k attends school.

Figure 4 displays the trajectory of both the Probit and IV Control Function methods ASFs stratified by gender. The left-hand panel contains the results for boys. Overall, we can see that the probability of school attendance falls in tandem with sibship size increases, as one might expect. The rate of decline appears to be constant, around -2 per cent, for boys in the conventional probit model. The IV

[9] More elaborate specifications such as higher order polynomials and crossproducts with covariates did not yield substantially different results. Similarly estimation of the second stage using a semiparametric estimator also failed to produce different results. In the interests of computational efficiency, the less elaborate estimator is favoured here as it eased the requirements to run the cluster bootstrap method of calculating uncertainty.

Figure 4: Probit and Control Function Average Structural Functions Stratifed by Gender



Note: 95% cluster bootstrap confidence interval shaded.

method appears to exhibit a degree of nonlinearity, wherein the QQ effect is larger at the lower end of the sibship size variable's distribution. However, once again the IV estimates are far more uncertain, as is evident in the 95% confidence intervals obtained via a cluster bootstrap. Such a large degree of uncertainty cautions against any conclusion supporting the presence of any significant nonlinear effect. The conventional probit ASF lies comfortably in the bounds of the control function estimates. A similar pattern is observed amongst the female only sample. For girls, shown in the right-hand panel, we see that the control function estimates appear to track the probit estimates which again are consistent with the existence of a child QQ trade-off. Again the imprecision of the IV-based methods are clearly evident.

5 Conclusion

Was sibship size an important determinant of educational attainment in early 20th century England and Wales? Considering the body of evidence presented in this paper, I can conclude that it was. The unconditional relationship between sibship size and school attendance indicates that school attendance amongst 13 to 15 year old boys falls by around 2 per cent for every extra sibling, with the corresponding effect for girls around 1 per cent, although remaining in education meant different things for males compared to females. This paper addressed the endogenous nature of the the sibship size variable in two ways. Firstly, I find that the reported unconditional QQ effects are robust to the inclusion of a large number of control variables and fixed effects. The conditional effects are almost identical to the unconditional effects, thus indicating the absence of omitted variable bias. Secondly, a variety of IV methods, in which the gender composition of the first two births is used as an instrument for the sibship size variable, also estimate a negative sibship size-school attendance relationship. The drawback with these IV methods is that the parameter estimates are far less precise when compared to their non-IV counterparts. Thus, whilst the IV estimates do not contradict the non-IV estimates, one would ideally like to have smaller standard errors. This remains a worthy goal for future research on this topic. These results also support hypotheses that claim the British demographic transition and subsequent expansion of post-primary education witnessed in the early 20th century did not occur independently. This suggests that policies promoting family planning and fertility control in LDCs could stimulate education demand as parents shift from child quantity to quality.

Conflict of Interest: The author declares that they have no conflict of interest.

References

Joshua D. Angrist and William N. Evans. Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size. *The American Economic Review*, 88(3):450–477, 1998.

Joshua D. Angrist and Jörn-Steffen Pischke. *Mostly Harmless Econometrics: An Empiricists Companion*. Princeton University Press, Princeton, NJ, 2009.

Joshua D. Angrist, Victor Lavy, and Analia Schlosser. Multiple Experiments for the Causal Link between the Quantity and Quality of Children. *Journal of Labor Economics*, 28(4):773–824, 2010.

Kaushik Basu and Van Pham. The Economics of Child Labor. *American Economic Review*, 88(3):412–427, 1998.

Gary S. Becker and H. Gregg Lewis. On the Interaction between the Quantity and Quality of Children. *Journal of Political Economy*, 81(2):279–288, 1973.

Sascha O. Becker, Francesco Cinnirella, and Ludger Woessmann. The Trade-Off between Fertility and Education: Evidence from Before the Demographic Transition. *Journal of Economic Growth*, 15(3):177–204, 2010.

Sonia R. Bhalotra and Damian Clarke. *The Twin Instrument*. IZA Discussion Paper No. 10405, 2016.

Sandra E. Black, Paul J. Devereux, and Kjell G. Salvanes. The More the Merrier? The Effect of Family Size and Birth Order on Children's Education. *The Quarterly Journal of Economics*, 120(2):669–700, 2005.

Hoyt Bleakley and Fabian Lange. Chronic Disease Burden and the Interaction of Education, Fertility, and Growth. *The Review of Economics and Statistics*, 91 (1):52–65, 2009.

Richard W Blundell and James L Powell. Endogeneity in Semiparametric Binary Response Models. *The Review of Economic Studies*, 71(3):655–679, 2004.

Jutta Bolt and Jan Luiten van Zanden. The Maddison Project: Collaborative Research on Historical National Accounts. *The Economic History Review*, 67 (3):627–651, 2014.

Julio Cáceres-Delpiano. The Impacts of Family Size on Investment in Child Quality. *Journal of Human Resources*, 41(4):738–754, 2006.

Dalton Conley and Rebecca Glauber. Parental Educational Investment and Children's Academic Risk Estimates of the Impact of Sibship Size and Birth Order from Exogenous Variation in Fertility. *Journal of Human Resources*, (41):722–737, 2006.

Alan Fernihough. Human Capital and the Quantity–Quality Trade-Off During the Demographic Transition. *Journal of Economic Growth*, 22(1):35–65, 2017.

Sylvia Frühwirth-Schnatter, Martin Halla, Alexandra Posekany, Gerald J. Pruckner, and Thomas Schober. The Quantity and Quality of Children: A Semi-Parametric Bayesian IV Approach. Economics Working Papers 2014–03, Department of Economics, Johannes Kepler University Linz, Austria, March 2014.

Oded Galor. The Demographic Transition: Causes and Consequences. *Cliometrica, Journal of Historical Economics and Econometric History*, 6(1):1–28, January 2012.

Oded Galor and Omer Moav. Natural Selection And The Origin Of Economic Growth. *The Quarterly Journal of Economics*, 117(4):1133–1191, 2002.

Oded Galor and David N. Weil. Population, Technology, and Growth: From Malthusian Stagnation to the Demographic Transition and Beyond. *The American Economic Review*, 90(4):806–828, 2000.

Diana Gittins. *Fair Sex: Family Size and Structure 1900–39*. Hutchinson, London, UK, 1982.

Peter Gordon, Richard Aldrich, and Dennis Dean. *Education and Policy in England in the Twentieth Century*. Woburn, London, UK, 1991.

Timothy J. Hatton and Richard M. Martin. The Effects on Stature of Poverty, Family Size, and Birth Order: British Children in the 1930s. *Oxford Economic Papers*, 62(1):157–184, 2010.

Jane Humphries. *Childhood and Child Labour in the British Industrial Revolution*. Cambridge University Press, Cambridge, UK, 2010.

Jane Humphries. Childhood and Child Labour in the British Industrial Revolution. *The Economic History Review*, 66(2):395–418, 2013.

Guido W. Imbens and Joshua D. Angrist. Identification and Estimation of Local Average Treatment Effects. *Econometrica*, 62(2):467–475, 1994.

Marc P B Klemp and Jacob Weisdorf. Fecundity, Fertility and the Formation of Human Capital. CAGE Online Working Paper Series 296, Competitive Advantage in the Global Economy (CAGE), 2016.

Adriana D Kugler and Santosh Kumar. Preference for Boys, Family Size, and Educational Attainment in India. *Demography*, 54(3):835–859, 2017.

Jungmin Lee. Sibling Size and Investment in Children's Education: An Asian Instrument. *Journal of Population Economics*, 21(4):855–875, 2008.

Hongbin Li, Junsen Zhang, and Yi Zhu. The Quantity-Quality Trade-Off of Children in a Developing Country: Identification using Chinese Twins. *Demography*, 45(1):223–243, 2008.

Giampiero Marra and Rosalba Radice. Estimation of a Semiparametric Recursive Bivariate Probit Model in the Presence of Endogeneity. *Canadian Journal of Statistics*, 39(2):259–279, 2011.

Felix P. Meier zu Selhausen, Marco H.D. van Leeuwen, and Jacob L. Weisdorf. Social Mobility among Christian Africans: Evidence from Ugandan Marriage Registers 1895-2011. CAGE Online Working Paper Series 239, Competitive Advantage in the Global Economy (CAGE), 2015.

Daniel L. Millimet and Rusty Tchernis. Estimation Of Treatment Effects Without An Exclusion Restriction: With An Application To The Analysis Of The School Breakfast Program. *Journal of Applied Econometrics*, 28(6):982–1017, 9 2013.

Daniel L. Millimet and Le Wang. Is the Quantity-Quality Trade-Off a Trade-Off for All, None, or Some? *Economic Development and Cultural Change*, 60(1): 155–195, 2011.

Omer Moav. Cheap Children and the Persistence of Poverty. *Economic Journal*, 115(500):88–110, 2005.

Tommy Murphy. Old Habits Die Hard (Sometimes). *Journal of Economic Growth*, 20(2):177–222, 2015.

Great Britain. Parliament. House of Commons. *Census of England and Wales, 1911, General Report with Appendices*. BPP 1917-18 XXXV [Cd.8491].

Great Britain. Board of Education. Committee on Juvenile Education in Relation to Employment After the War and J.H. Lewis. *Final Report of the Departmental Committee on Juvenile Education in Relation to Employnent After the War*. Number v. 1 in Cd. (Great Britain. Parliament). H.M. Stationery Office, 1917.

Siân Pooley. Parenthood, Child-Rearing and Fertility in England, 1850–1914. *The History of the Family*, 18(1):83–106, 2013.

Douglas Rivers and Quang H. Vuong. Limited Information Estimators and Exogeneity Tests for Simultaneous Probit Models. *Journal of Econometrics*, 39(3): 347–366, 1988.

Mark R. Rosenzweig and Kenneth I. Wolpin. Testing the Quantity-Quality Fertility Model: The Use of Twins as a Natural Experiment. *Econometrica*, 48(1):227–240, 1980.

Mark R. Rosenzweig and Junsen Zhang. Do Population Control Policies Induce More Human Capital Investment? Twins, Birth Weight and China's One-Child Policy. *Review of Economic Studies*, 76(3):1149–1174, 2009.

Christoph Rothe. Semiparametric Estimation of Binary Response Models with Endogenous Regressors. *Journal of Econometrics*, 153(1):51–64, 2009.

Simon Szreter. *Fertility, Class and Gender in Britain 1860–1940*. Cambridge University Press, Cambridge, UK, 1996.

Paul Trowler. *Education Policy*. Routledge, London, UK, 2003.

Stef van Buuren and Karin Groothuis-Oudshoorn. mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software, Articles*, 45(3):1–67, 2011.

Marco van Leeuwen, Ineke Maas, and Andrew Miles. *HISCO. Historical International Standard Classification of Occupations*. Leuven University Press,, Leuven, BE, 2002.