University of Kent

School of Economics Discussion Papers

What Determines Post-Compulsory Educational Choice? Evidence from the Longitudinal Survey of Young People in England

William Collier, Javier Valbuena and Yu Zhu

May 2011

KDPE 1112

What Determines Post-Compulsory Educational Choice? Evidence from the Longitudinal Survey of Young People in England

William Collier[†], Javier Valbuena[‡] and Yu Zhu[§]

May 2011

ABSTRACT

Using a unique dataset which is rich in both family background and attainment in education, we find that educational attainments at the end of the compulsory schooling stage are powerful predictors for post-compulsory educational choices in England. In particular, the single academic success indicator of achieving the Government's *gold standard in GCSE*, is able to explain around 30% of the variation in the proportion of young people studying for academic qualifications. Instrumental-variables estimation which exploits variations in birth weight and school starting age suggest that over half of the least-squares effect of achieving the *gold standard in GCSEs* on studying for academic qualifications is due to individual heterogeneity (ability bias) or simultaneity bias (reverse causation). Nonetheless, conditional on the young person working towards a higher-level qualification, we find strong evidence of a highly significant causal effect of achieving the *gold standard* when choosing between the academic or vocational pathway.

Keywords: Educational choice, instrumental-variable estimation

JEL Classification: I21, J24, P36

Acknowledgements: We wish to thank Antonio Di Paolo, Colm Harmon, Jagjit Chadha and participants at the 2010 WPEG annual conference, the First Lisbon Research Workshop on Economics and Econometrics of Education, XIX Meeting of the Economics of Education Association and XXXV Simposio de la Asociación Española de Economía for their helpful comments. All errors and omissions remain the responsibility of the authors.

[†] **Corresponding author**: School of Economics, University of Kent, Canterbury, CT2 7NP. Email: W.J.Collier@kent.ac.uk

[‡] School of Economics, University of Kent, Canterbury, CT2 7NP. Email: jv51@kent.ac.uk

[§] School of Economics, University of Kent, Canterbury, CT2 7NP. Email: Y.Zhu-5@kent.ac.uk

1. Introduction

Most of the empirical literature on the analysis of returns to education is concerned with the differential returns at different levels of qualifications. More recently, researchers have started to examine the rate of return associated with different types of qualifications. In this regard, there is agreement, for the UK at least, that returns to academic qualifications are higher than those to vocational qualifications at the same level (see e.g. Robinson 1997, Conlon 2001, Dearden *et al*. 2002, McIntosh 2006 and Dickerson 2008). This pattern appears to be remarkably robust with respect to the method of estimation, the data source used, and the specification of qualifications in regression models (e.g. focusing on the highest qualification or using all qualifications as the preferred measure).

Understanding the causes of a persistent gap along the academic-vocational lines is not only of academic interest, but also of enormous policy relevance. Indeed, many of the recent educational reforms in the UK such as the introduction of GCSEs, AS Levels, and the new range of diplomas for 14 to 19 year olds, have an explicit aim of '*breaking down the artificial barriers between academic and vocational education*' (DfES 2005).[1]

This paper focuses on the determinants of post-compulsory educational choice in the UK, including the choice between the academic-vocational route, using a unique dataset which is rich in both family background and attainment in education, as well as post-16 plans.

We find that educational attainments at the end of the compulsory schooling stage are powerful predictors for post-compulsory educational choices in England. In particular, the single indicator of achieving the Government's *gold standard in GCSE*, which emphasizes the core subjects of Maths and English, is able to explain around 30% of the variation in the proportion of young people studying for academic qualifications.[2]

We also investigate the extent to which the impact of initial academic success on post-compulsory educational choices reflects a causal relationship. Instrumental-variables estimates which exploit variations in birth weight and school starting age by month of birth induced by school entry rules suggest that over half of the least-squares effect of achieving the *gold standard in GCSEs* on studying for academic qualifications is due to individual heterogeneity (ability bias) or simultaneity bias (reverse causation). Nonetheless, conditional on the young person working towards a higher-level qualification, we find strong evidence of

[1] GCSE stands for General Certificate of Secondary Education, while AS stands for Advanced Subsidiary.
[2] Students achieve the "gold standard" in GCSE if they achieve five or more A* to C passes, including English and Mathematics.

a highly significant causal effect of achieving the *gold standard* when choosing between the academic or vocational pathway.

The remainder of the paper proceeds as follows. Section 2 presents a brief overview of the existing literature. Section 3 outlines relevant features of the English education system. Section 4 discusses the data and their relative merits. Empirical results are reported in Section 5. Section 6 concludes.

2. Literature Review

Compared to the vast literature within the economics of education that is concerned with the differential returns at different levels of qualifications, empirical research into the rate of return associated with different types of qualifications is fairly sparse. Indeed, there have been only a handful of recent empirical studies for the UK which attempt to distinguish between various forms of academic and vocational qualifications.

A key contribution of the human capital theory is the distinction between *general* and *specific* human capital (Becker 1964). General human capital (such as literacy or numeracy) is useful to all employers, while specific human capital refers to skills or knowledge that are useful only to a single occupation or industrial sector. Broadly speaking, one can equate academic qualifications with general human capital and vocational qualifications with specific human capital.

To the best of our knowledge, Robinson (1997) provides the first UK study in the parity of returns between academic and vocational qualifications in the labour market. Using the Quarterly Labour Force Survey (QLFS), he concludes that men and women with academic qualifications at one level in the National Qualifications Framework (NQF) earn on average as much as those with vocational qualifications set notionally one level higher. Even controlling for occupations, academic qualifications are still found to be associated with higher earnings.

Using the QLFS and the National Child Development Study (NCDS), Conlon (2001) finds a statistically significant gap in hourly wage in favour of academic qualifications for working age males in the UK, at every level of qualification within the NQF. Moreover, this wage gap is also rising in the level of the qualification hierarchy.

Dearden *et al*. (2002) also reports higher returns to academic qualifications relative to those to vocational qualifications at the same level, using the QLFS and the NCDS data, as well as the British data from the 1995 International Adult Literacy Survey (IALS). They

show that while returns to academic qualifications are homogenous across the distribution of ability, as measured by scores of reading and maths tests taken at age 7, returns to vocational qualifications are significantly higher for low ability individuals.

Unlike earlier studies which focus on the level and type of the **highest** qualifications, both McIntosh (2006) and Dickerson (2008) use the 'all qualifications' specifications. Methodologically, the former approach focuses on the **marginal** return, while the latter is concerned with the 'average' rate of return, i.e. the return measured across all individuals who hold that particular qualification, holding all other qualifications constant.[3] Despite the differences in model specification, their findings, which are based on the QLFS data, are consistent with earlier studies which report that academic qualifications yield higher returns than vocational ones at the same NQF level. Moreover, both studies also find that lower level vocational qualifications fare particularly badly, with either zero or even negative rates of return.

Potential causes for a persistent differential return by type of qualifications can be usefully divided into demand side and supply factors. A leading candidate of the demand side explanation is *skill-biased technological change* (Berman *et al*., 1994), which increases demand towards more educated labour which exhibits a high level of general human capital and which is quick to learn and to adapt in a fast-changing working environment. The leading supply side explanation is possible self-selection into the academic or vocational qualification pathways on the basis of ability (Conlon 2005).

In this paper, we focus on supply side factors since we consider them to be more relevant for young people at this important stage of their educational development.

3. Relevant features of the English education system

The school education system in England can be divided into three stages: primary education (Reception Year and Year 1 to Year 6), compulsory secondary education (Year 7 to Year 11) and post-compulsory secondary education (Year 12 to Year 13). By law, all children of compulsory school age (between 5 and 16 years old) must receive a full-time education. The current school leaving age of 16 in England and Wales has been in force since

[3] Estimation of the 'marginal' and 'average' rate of return is now common in the empirical literature. Jenkins et al (2007), for example, presents estimates of both average and marginal returns when considering the returns to low level vocational qualifications.

September 1973, as a result of the **Raising of School Leaving Age (RoSLA) Order** of 1972.[4]

The academic year in England runs from 1 September to 31 August with three terms starting in September, January, and April, respectively. Under the English education system, children must commence school at the beginning of the term after they turn 5. While many local education authorities (LEAs) operate a triple-entry-point system that admits children at the beginning of the term in which they turn 5, the system that is becoming increasingly popular in England and Wales is based on a single-entry-point.[5] Under this system, all children commence school in September of the academic year in which they turn 5, regardless of age.[6]

By law, a child in England is generally not allowed to leave school before their 16th birthday. A single school leaving date was introduced in 1997. This requires students to remain in school until the last Friday in June in the school year in which they turn 16 (usually the end of Year 11). This does of course mean that a small minority of students leave school aged 15.

At the end of five years of compulsory secondary education, students in England and Wales take examinations across a range of subjects at the level of General Certificate of Secondary Education (GCSE). The GCSE is a single subject examination introduced in 1988 and marked by independent exam boards.[7] Students usually take at least 5 (there is no upper limit) GCSE examinations in different subjects, including mathematics and English. Students are given a letter score of A-G where A is the top grade.[8] Although grades A-G are all pass grades officially, only grades A to C are given much credence by most employers, and regarded as equivalent to the 'pass' grades in the previous O-Level examinations. GCSEs are part of the *National Qualifications Framework* which is the official qualification

[4] Current policy reforms will see a further increase in the school leaving age with young people in England and Wales continuing in education or training to age 17 from 2013 and to age 18 from 2015.

[5] Under the triple-entry-point system, children born between May and August could receive two terms fewer education (in Reception Year) compared with classmates born in the autumn who start in September. The Labour Government decided to bring forward the starting date from the term before a child's fifth birthday to the September after their fourth, following the 2009 Independent Review of the Primary Curriculum by Sir Jim Rose (DCSF 2009, recommendation 14).

[6] Crawford *et al*. (2007) estimate that around one half of all children born between 1997 and 1999 started school in an LEA where a single-entry-point system was in operation.

[7] The introduction in 1988 of the GCSE marked a turning point in UK educational system, in removing streaming before the age of 16. Since the 1950s, secondary school students who were academically inclined took Ordinary Level (at age 16) and Advanced Level (at age 18) examinations, which were an essential requirement to enter higher education. Less academically oriented pupils could take the Certificate of Secondary Education (CSE) at 16 before they left school.

[8] In 1994, the A* grade was introduced to distinguish the very top end of achievement.

accreditation system for the whole UK except Scotland. A GCSE at grades D–G is a Level 1 qualification, while a GCSE at grades A*–C is a Level 2 qualification. Post-compulsory secondary-education qualifications are Level 3 while Higher Education (HE) qualifications are classified as Levels 4 to 8.[9]

After taking GCSEs students may leave secondary schooling, continue on to further education colleges (typically for vocational or technical courses), or take a higher level of secondary school examinations known as 'A-Levels' (typically in 2-4 subjects) after a further two years of study. A-Levels (short for Advanced level) are traditionally required for university entrance in the UK. Since the introduction of the GCSEs in 1988 has largely removed academic streaming before the age of 16, most young people will only have their first opportunity to choose between an academic or vocational pathway once they have completed compulsory education.

4. Data

This paper is based on the Longitudinal Study of Young People in England (LSYPE), also known as Next Steps, which is a major innovative panel study of young people which brings together data from a number of different sources The first wave of the LYSPE was conducted between March and October 2004 with a sample of young people aged between 13 and 14 who were studying in Year 9 in schools in England. Over 15,000 young people and their families were interviewed at Wave 1. Subsequent waves have been collected annually thereafter.

LSYPE was commissioned by the former Department for Education and Skills (DfES) and is now managed by the Department for Education (DfE) with an aim to improve understanding of the key factors affecting young people's progress in transition from the later years of compulsory education (which ends at the age of 16) through any subsequent education or training, to entry into the labour market. Accordingly, the LYSPE provides a rich source of socioeconomic information at the individual and household level including the young person's personal characteristics and family background, their attainment in education and post-16 plans, and information regarding the school(s) they attend or have attended.

The LYSPE study brings together data from a wide range of sources which can assist in identifying the variety of influences expected to impact on student learning and progression. In particular, the LYSPE data have been linked to administrative records such as

[9] For example, Certificate based awards are HE Level 4 qualifications; Level 8 qualifications include Doctorates.

the *National Pupil Database* (NPD) and other data sources such as geo-demographic data from the 2001 Census. For confidentiality reasons, these linked administrative data were not included in the early public-access release of the LSYPE data. However, a small set of variables extracted from the NPD, covering GCSEs gained and grades, alongside Key Stage test scores for LSYPE respondents, have since been added to recent releases of the public-access LSYPE file.

In order to exploit the rich information in LSYPE to study the choice of the academic or vocational pathway upon completion of compulsory education, we select a sample of UK born young people who provide full interviews, together with their mothers, across the first 4 waves of the LYSPE data. Young persons' born before the 1st September 1989 or after the 31st August 1990 are excluded from the sample on the basis that they violate the school entry rule for the Wave 1 cohort. We also exclude a small number of cases where the mother is aged 60 or over (the state retirement age) at Wave 4.[10] This results in a final sample of 9,190 young people, of which 4,570 (49.7%) are boys for our analysis.

At the time of the Wave 4 interview, 7,538 (82%) of our sample are reported as undertaking either school/college courses or apprenticeships/work-based training which will lead to a qualification. Of these, 5,196 (69%) are studying for academic qualifications such as A Levels (including its component units AS and A2 Levels), AVCEs (Advanced Vocational Certificate in Education),[11] or GCSEs.[12] More than 99% of young people who have chosen the academic pathway are studying full-time. This compares with 73% of those who have chosen the vocational route.

Table 1 presents sample summary statistics by gender for all variables used in our empirical analyses. The first thing to note is that there is a very significant gender gap in our primary outcome variable, namely studying for an academic qualification. Upon completion of compulsory education, 61% of all girls are studying for academic qualifications, as opposed to only 52% of all boys. This gender gap is statistically significant at the 5% level.

Around one quarter of young people in our sample are non-white. While there are no significant differences across gender in the probability of being born prematurely (around

[10] It is conventional to exclude pensioners in empirical studies, to minimize the problem of unobserved heterogeneity (in say health and preferences).

[11] Despite its name, we have decided to treat AVCEs (formerly known as Vocational A Levels) as academic qualifications because they are full-time education based at schools or colleges, unlike traditional vocational routes such as apprenticeships. In our sample, there are only 51 young people taking AVCEs.

[12] Of the 351 young people studying for GCSEs post-compulsory schooling, only 35% have achieved the NQF Level 2 threshold and 8% achieved the *gold standard*, suggesting many of them are retaking subject (or retaking examinations) to improve their grades.

28%) or by single parents (around 18%), boys tend to have higher birth weight on average than girls. Boys are also more likely to self-report any disability or long-term health problems, which may or may not affect their schooling, in Wave 1. One in eight of young people receive free school meals, due to low family income.

Approximately one quarter of young people live with a lone mother at age 16 (Wave 4) in our sample. In our econometric analysis, we control for a full list of mother's characteristics at Wave 4, including race, qualifications, partnership status, number of other children (i.e. siblings to the young person) and employment status. We also control for household income reported in Wave 1. Given the 50% non-response rates of household income, we include a dummy for missing income variables rather than omitting the information or indeed dropping one half of the sample.

There are no notable differences in mother's characteristics across gender lines, except that boys appear to be less likely to be living with a lone mother.[13] Around one quarter of mothers have post-secondary qualification (NQF4 or above), of which nearly half have degrees. Around 13% of mothers have upper-secondary qualifications (NQF3) while almost 30% have NQF2 which is awarded upon successful completion of compulsory education. One in 5 of all mothers have no formal qualifications while another 11% have only Level 1. About 30% of mothers report a vocational qualification as her highest qualification.

It is clear that there is significant gender gap in educational attainment at around age 16 in favour of girls: 57% of girls achieve the critical benchmark of 5 or more GCSEs at Grades A*-C including Maths and English, only 49% of boys manage to reach the same standard. Girls are also more likely to achieve the Level 2 threshold which simply requires ***any*** 5 GCSEs with grade C or above. Interestingly, boys do almost as well as girls in GCSE Maths. The contextual value added KS2-KS4 scores used by the government, which is supposed to measure progress between the time one finishes primary education (Key Stage 2) and the time one completes compulsory secondary education (Key Stage 4), suggest that the gender-gap in academic attainment actually widens during these stages.

Boys are more likely than girls to attend private schools in Wave 1. However, the overall proportion is small, at 4% or below. Around 44% of mothers think the overall quality of the school is very good, and only 10% or less think it is poor (omitted category being

[13] Dahl & Moretti (2008) report that child gender affects both marital status and fertility. They find that fathers of boys are substantially more likely to be living with their children compared to fathers of girls. Parents of girls are also significantly more likely to be divorced than those of boys.

good). It turns out that the gender gap in actual educational attainment is well reflected by differences in parental aspirations and subjective assessment of the likelihood that the young person will continue in full-time education at 16 and go to higher education.

Finally, given the distinct differences in educational attainment across gender described above, the last column of Table 1 highlights those variables for which the equality of means across gender is rejected at the 5% significance level. It is clear from these descriptive statistics that a pooled specification would be hard to justify given the wide differences in own and mother's characteristics, parental aspirations, as well as educational attainment at 16. Therefore, we estimate regressions for boys and girls separately to allow for gender-specific effects, while maintaining a common model specification.

5. Findings

5.1 Determinant of post-compulsory educational choices

There has been a heated debate in the economics of education literature on the (relative) roles of family background, school environment (peer effects etc) and ability in determining individual's educational attainment. In this paper, we contribute to this debate by exploiting the unusually rich information in LSYPE and investigate two key choices affecting young people as they approach completion of compulsory education. Firstly, we consider the young person's choice between the pursuit of an academic qualification versus other options in the education system or entry into the labour market. Secondly, we examine the choice between the academic and vocational pathways conditional on working towards a qualification. Both issues are of enormous policy relevance, but are poorly understood so far.

Our empirical approach starts by attempting to quantify the relative importance of the different factors emphasized by different researchers in the literature, sometimes due to data availability problems. We proceed by successively adding new sets of control variables in a Linear Probability Model (LPM) of whether to study for academic qualification immediately after the completion of the compulsory education stage.[14]

Our baseline model (Model 1) controls for a comprehensive list of own characteristics of the young person and those of the mother which includes race, educational attainment, marital (partnership) status, indicator for step-families and number of siblings,

[14] Angrist & Pischke (2009) show that whilst a non-linear model may fit the conditional expectation function for Limited Dependent Variable (LDV) models more closely than a Linear Probability Model (LPM), this matters little when evaluating marginal effects. Furthermore, a linear model offers clear advantages when one extends the regression framework to consider instrumental variable estimation.

labour market status and family income when the young person was 13. These variables are widely available in labour force or household surveys and have been used extensively in empirical labour economics.

A second model (Model 2) augments the baseline specification to include the NPD records which capture educational attainments at the end of compulsory schooling stage such as indicators for achieving the Government's *gold standard* in GCSE, that is attaining five or more GCSEs at grades A*-C, including Maths and English.[15] In Model 3, we extend these additional controls to include parental aspirations measures and school type from Wave 1, when the young person was aged 13. Finally, we reduce the most comprehensive specification to a parsimonious model.

Table 2a and 2b present LPM estimates for models M1 through M3 as well as the parsimonious model for boys and girls respectively. Our baseline model contains most of the family background variables found in the literature. These results appear both meaningful and appropriate: Most of the variables are individually significant, and have the expected signs. For instance, any disability or long-term health problems reported by the young person decreases the chance of pursuing academic qualifications post 16. Conversely, higher educational qualification of the mother and higher family income are positively related to studying for academic qualification. However, all these family background variables as a whole can only explain no more than 14% of the variation in the proportion studying for academic qualifications for either gender.

It is apparent from Model 2 that passing any qualification threshold in the NQF classification has a positive effect on the chance of studying for academic qualifications. Comparing to someone who leaves school without any qualifications, achieving NQF Level 1 and Level 2 thresholds increases the chance of studying for academic qualification by 11 and 33 percentage points respectively for boys, or 4 and 29 percentage points for girls. What is most striking is that merely including the core subjects of Maths and English in the 5 GCSE subjects at grades A*-C required to achieve a Level 2 NQF qualification has an ***additional*** 26 and 22 percentage point effect for boys and girls respectively. While the adjusted-R^2 measures have tripled for both boys and girls, many of the family background variables, most notably young person's disability and family income, lose statistical significance in this extended model. This implies that many of the family background variables impact on the

[15] Indeed, the children in our sample are the first school cohort to face the *gold standard*, introduced by the Department for Children, Schools and Families (DCSF) in 2006.

outcome only through their effect on prior educational attainment. This result is consistent with Heckman (2008) who finds that family environments of young children are major predictors of cognitive and socio-emotional abilities.

Model 3 reports our most comprehensive results having additionally accounted for school type and parental aspirations at age 13. It is unsurprising to see school quality as assessed by the main parent does not matter, given that we have already controlled for actual educational attainment at age 16. On the other hand, parental aspirations as regards the young person's educational attainment are all statistically significant. Some of these variables reflect parental preferences which are likely to differ by parental education, while others could be thought of as proxies for the ability of the young person. It is worth noting that the goodness-of-fit of the regression as measured by the adjusted-R^2 only improves modestly, while the size of effects of age 16 attainments and mother's qualifications are reduced markedly.

Given that only one-third of the regressors in Model 3 are statistically significant at the conventional 5% level, we use the naive stepwise regression technique to arrive at a parsimonious model. This final model is presented in the last columns of Table 2a and 2b only contains variables which are statistically significant at the 5% level for either boys or girls, but maintains almost the same explanatory power as in Model 3.

5.2 The importance of achieving the *gold standard in GCSEs*

It is clear from the parsimonious specification reported in Table 2 that the single most important predictor for studying for academic qualifications at age 16 in England is educational attainment at the end of the compulsory education stage, represented by whether having achieved the NQF Level 2 threshold, and in particular whether having achieved the *gold standard of GCSEs*. Simply passing any 5 GCSEs at grades A*-C would increase the probability of academic studies by around 20 percentage points, while achieving good grades in the core subjects of Maths and English among the 5 subjects, which emphasizes key skills of numeracy and literacy, will add a further 22 percentage points for both boys and girls.

This finding is in line with the conventional wisdom (see e.g. Nuffield Foundation 2009) and the Government's views as summarized in a White Paper by the DfES:

> *"By far the best-known and best-understood qualifications for young people in this country are the GCSE and the A level. The overwhelming majority of young people who do well at GCSE level go on to take A level" – DfES 2005 White Paper, p19*

Given the relative importance of prior educational attainment in pursuing an academic qualification, Table 3a and 3b assess the importance of achieving the *gold standard* on the probability of studying for academic qualifications for boys and girls separately. In this instance, we commence with a model with a single regressor for achieving the *gold standard.* We then successively add regressors which capture other age 16 educational attainments and age 13 school type and parental aspirations. Note that all of the regressors reported in Table 3 are subsets of the parsimonious specification reported in Table 2.The parsimonious specification is reproduced in the last column in Table 3 to facilitate comparison.

It is really striking that the variable for achieving the *gold standard* alone accounts for 32% of the variation in the proportion studying towards an academic qualification for boys and 28% for girls. Those who passed this critical threshold are 52 to57 percentage points more likely to pursue an academic qualification than those who failed. A comparison of Model 2 and Model 3 to Model 1 reveals that adding other educational attainment measures and parental aspirations increases the explanatory power by only 11 and 9 percentage points, for boys and girls respectively. Finally, including all the family background controls in the parsimonious specification adds a mere 0.5 percentage points for boys and 1.4 percentage points for girls to the explanatory power of Model 3, a model consisting of educational variables only.

We interpret the above findings as compelling evidence that prior educational attainment as represented by having achieved the *gold standard in GCSEs*, which emphasizes key skills in numeracy and literacy by including the core subjects, is the overriding determinant for pursing academic qualifications post 16.

5.3 Determinant of post-compulsory educational choices

Interesting as it might be, we can only interpret this strong relationship we find as a correlation, because of potential ability bias and simultaneity issues (e.g. those who intended to drop out were also less likely to pass GCSEs). For policy interventions, one would be interested in identifying the causal relationship.

Since all Wave 4 educational attainments are effectively jointly determined, we keep only the indicator for having achieved the *gold standard* in our empirical specifications. We also omit the type of school indicator and parental aspirations from Wave 1, for fear of

endogeneity problems.[16] The mother's employment status and family income variables are excluded for similar reasons. However we do condition on a full set of dummy variables for mother's educational qualifications and partnership status.

The causal effect of achieving the *gold standard* is identified through two instrumental variables. The first one exploits the exogenous variation in the relative school starting age (SSA) by month of birth within the same school cohort group, induced by the English school entry policy. Under a single-entry-point system, a child born on the 1 September 1989 will be the oldest in this school cohort while another child born on the 31st August 1990 will be the youngest.

Drawing on 18 research studies published between 2000 and 2008 for a range of countries, the survey by Sharp *et al*. (2009) concludes that there is overwhelming evidence of statistically significant effects for relative age, i.e. comparing the youngest to the oldest in the academic year group. Pupils who are younger in the year group fare less well in attainment tests, commonly measured by test scores in maths, reading and writing. For recent UK evidence, see e.g. Crawford *et al*. 2007, and Walker and Zhu 2009.

We expect some noise in the actual SSA (which we do not observe in our data), due to the fact that different school entry rules are in operation in different LEAs.[17] However, what matters for identification purposes is whether month of birth is *statistically* correlated with the probability of achieving *gold standard in GCSEs* while having no direct impact on the post-compulsory educational choice. Figure 1 shows that a September-born boy is 7 percentage points more likely than his August-born counterpart to pass the threshold. The corresponding gap for girls is a striking 15 percentage points. This implies that on average, the chance of reaching the *gold standard* for both boys and girls is increasing in the predicted SSA (using the single-entry-point rule).

Our second instrument relies on birth weight. There has been compelling evidence of an adverse effect of low birth weight on school outcomes in the literature. Indeed, birth weight has been used routinely as an instrument for schooling differences in within-twins analysis of wages (see e.g. Behrman *et al*. 1994, Neumark 1999, Behrman and Rosenzweig

[16] For example, rich parents who are worried about their child's educational performance are more likely to send the child to private schools, which produce superior academic results on average, not least because of better resources (see e.g. Green *et al*. 2008).

[17] Admittedly, children exposed to multi-entry-points systems will receive different length of education (up to 2 terms) at the end of the compulsory education stage (see Footnote 2). This idea has been exploited by Del Bono and Galindo-Rueda (2004) for the UK and Black *et al*. (2008) for Norway.

2004, Miller *et al* 2005 and Black *et al* 2007).[18] More recently, Oreopoulos *et al* (2008) and Fletcher (2010) provide evidence of an impact of birth weight on medium term educational outcomes.

Figure 2 suggests a strong positive relationship between birth weight and the probability of achieving the *gold standard*. An underweight (less than 2.5 kg) boy is 9 percentage points less likely than a normal weight boy (between 2.5 and 4.5 kg) to achieve the gold standard. The corresponding gap for girls is 6 percentage points. The results for overweight births (over 4.5 kg) are somewhat surprising but may be due to small cell sizes, as only 2.6% of boys and 1.1% of girls fall into this category in our data.[19] In our empirical specification, we use the log of birth weight to proxy the effect of birth weight.[20]

Whilst we are confident that both of the above instruments can capture the causal effect of achieving the gold standard in pursing an academic qualification beyond age 16, we are aware of the debate in the literature regarding the use of birth weight and month of birth as valid instruments when determining educational outcomes and the bias that may arise when such instruments are only very weakly correlated with the endogenous variable (see for example, Angrist and Krueger, 1991; Bound *et* al 1995). Accordingly, we report a range of diagnostic tests to support the validity of our instruments [21]

Table 4 presents Instrumental Variable (IV) estimates with the corresponding LPM results for both boys and girls.[22] The first-stage estimates of the IV model are also shown in the bottom panel of the table, together with the relevant diagnostics tests for validity of instruments. By and large, the family background variables maintain their signs and statistical significances when we endogenize the *gold standard* indicator, although the sizes of the effects appear to be larger under the IV specification.

In contrast, the IV estimates for achieving the *gold standard* are 60% lower than their LPM counterparts for boys and 45% lower for girls respectively. Moreover, only for girls is the IV estimate marginally statistically significant (p=0.09). This implies that over half of the effect of achieving the *gold standard* on choosing the academic pathway is probably driven by individual heterogeneity (ability) or reverse causation.

[18] The use of twins data enables the researcher to capture the genetic influences on children and much of their exposure to a shared environment.
[19] Cesur and Rashad (2008) find a negative association between high birth weight (>4.5 kg) and low test scores for children in the US.
[20] A quadratic term is dropped from the final specification due to lack of statistical significance.
[21] See Stock and Yogo (2005) for a discussion of weak instruments and how to test for them.
[22] The corresponding Probit models produce very similar marginal effects to the LPM estimates, and are only shown in the Appendix.

The diagnostic tests are strongly supportive of the validity of our instruments. All instruments are at least individually significant at the 5% level for both genders. Furthermore, the Cragg-Donald Wald F-statistics for the excluded restrictions are well above the recommended threshold of 10 in both models, meaning we do not have a weak-instrument problem. Indeed, the F-statistics are above the critical value for 20% relative bias for boys and that for 15% relative bias for girls. This implies that our IV estimates have been successful in removing most of the bias in the LPM estimates. Finally, we are unable to reject the null of exogeneity of instruments according to the Sargan test.[23]

In Table 5, we repeat this exercise using a sample which conditions on the young person pursuing either school or college courses, or apprenticeships and work-based trainings, which lead to a qualification (N=7538). In other words, we drop the 18% of young people who are not working towards any qualifications from the reference category. Once again, our instruments easily pass the diagnostic tests. Moreover, the IV estimates are statistically significant for both boys and girls. Although the size of the IV estimate for boys is still around 30% lower than the OLS counterpart, the size of the causal effect for girls is virtually identical to the OLS estimate. We interpret this as evidence of a strong causal effect of academic success at the compulsory education stage, as represented by achieving the *gold standard*, on choosing the academic as opposed to the vocational pathway, especially for girls, conditional on the young person is working towards a qualification.

6. Conclusions

This paper is concerned with the determinants of educational choices, including the choice between the academic and the vocational pathway, immediately after the completion of compulsory education in England. While earlier studies have demonstrated convincingly that returns to academic qualifications are significantly higher than those to vocational qualifications at notionally equivalent levels for the UK, there are hardly any empirical studies that have assessed the relative contributions of family background, prior educational attainment and attributes of schools.

[23] The diagnostic tests reported in Table 4 are consistent and efficient under the null hypothesis of homoskedasticity. Angrist and Pischke (2009) show that whilst LPM residuals are heteroskedastic by construction, heteroskedasticity may matter very little. To check the validity of our findings, we re-estimated the results in Table 4 with a robust VCE. The robust standard errors and associated diagnostic tests are qualitatively identical to those reported in Table 4 indicating that our conclusions remain valid. Indeed, Hansen's J statistic of overidentifying restrictions yields the same interpretation as the Sargan test reported previously.

Using a unique dataset which is rich in both family background and attainment in education, we find that all family background variables combined explain no more than 14% of the variation in the decision to pursue an academic qualification upon completion of compulsory education at age 16. By contrast, educational attainments at the end of the compulsory schooling stage are much powerful predictors for post-compulsory educational choices. In particular, the single academic success indicator of achieving the Government's *gold standard in GCSE*, i.e. attaining 5 or more GCSEs including Maths and English at grades A*-C, which emphasizes key skills in numeracy and literacy, is able to explain around 30% of the variation in the proportion of young people studying for academic qualifications. Moreover, many family background variables, most notably family income and child's disability indicators, are no longer statistically significant once educational attainment at age 16 is included. This implies that family background factors impact on post-compulsory educational choices mainly through their contribution to educational attainment achieved prior to the end of compulsory schooling (i.e. attainment at age 11, 13 etc). This finding is consistent with Chowdry et al (2010) who report that around *two fifths* of the gap in educational attainment between young people from rich and poor backgrounds at the end of compulsory schooling (age 16) can be accounted for by prior educational attainment.

We also investigate the extent to which the impact of initial academic success on post-compulsory educational choices reflects a causal relationship, not least because of policy considerations. Instrumental-variable estimates which exploit variations in birth weight and school starting age by month of births induced by school entry rules suggest that the IV estimates for achieving the *gold standard* are 60% lower than their LPM counterparts for boys, and 45% lower for girls respectively. Only in the latter case is the IV estimate marginally statistically significant (p=0.09). This implies that much (over half) of the effect of achieving the *gold standard* on pursuing the academic pathway is probably driven by individual heterogeneity (ability) or reverse causality.

Nevertheless, if we exclude the 18% or so young people who are not working towards any qualifications from our sample, then the IV estimates for studying for academic qualifications (as opposed to vocational ones) are statistically significant for both boys and girls. While the size of the IV estimate for boys remains around 30% lower than the OLS counterpart, the size of the causal effect for girls is virtually identical to the corresponding OLS estimate. Thus, conditional on the young person working towards a qualification, there appears to be a strong causal effect of academic success at the compulsory education stage, as

represented by achieving the *gold standard*, on choosing the academic as opposed to the vocational pathway. This effect is most evident for girls.

Our results are consistent with a substantial and persistent earnings gap between the academic and vocational qualifications at the same level in the National Qualifications Framework, and clearly at odds with a notional parity of esteem of the two tracks when young people make their educational choices upon completion of compulsory schooling.

A further cause for concern, especially from an equity perspective, is the realisation that the chance of academic success in the UK appears to be heavily affected by birth weight and month of birth. While the former is knowingly related to socio-economic factors which might need expensive long-term solutions, the latter is a pure artefact created by the school-entry rules operating in the country and hence warrants policy interventions to counterbalance the apparent disadvantage of an early school starting age.

Table 1: Summary Statistics for Family Characteristics

Variable Name	Boys	Girls	Equality at 5%?
Studying for any academic qualifications (dep. var.)	0.518 (0.500)	0.611 (0.487)	No
Young Person (YP)'s Own Characteristics:			
Non-white	0.243 (0.429)	0.274 (0.446)	No
Premature Birth (by 1+ week)	0.280 (0.449)	0.279 (0.449)	
Log birth weight (kilograms)	1.196 (0.208)	1.152 (0.205)	No
Any disability (health problem) in Wave 1	0.153 (0.360)	0.119 (0.324)	No
Any disability (health problem) affecting schooling in Wave 1	0.065 (0.246)	0.051(0.221)	No
Single-parent family at birth	0.188 (0.390)	0.176 (0.380)	
Mother's Characteristics (measured at age 16, or Wave 4):			
Mother non-white	0.223 (0.416)	0.251 (0.434)	No
Mother's highest qualification is degree or above	0.122 (0.327)	0.118 (0.323)	
Mother's highest qualification is NQF4 but below degree	0.138 (0.345)	0.131 (0.338)	
Mother's highest qualification is NQF3	0.135 (0.342)	0.135 (0.342)	
Mother's highest qualification is NQF2	0.285 (0.452)	0.294 (0.456)	
Mother's highest qualification is NQF1	0.115 (0.319)	0.113 (0.317)	
Mother has no qualification (reference category)	0.205 (0.404)	0.211 (0.408)	
Mother's highest qualification is vocational	0.309 (0.462)	0.305 (0.460)	
Mother Married (reference category)	0.704 (0.457)	0.687 (0.464)	
Mother cohabiting	0.059 (0.235)	0.054 (0.227)	
Mother is lone parent	0.238 (0.426)	0.259 (0.438)	No
Indicator for step-family	0.102 (0.302)	0.097 (0.296)	
Number of siblings in the household (to YP)	1.503 (1.167)	1.530 (1.181)	
Any non-resident siblings	0.270 (0.444)	0.278 (0.448)	
Mother works full-time	0.411 (0.492)	0.414 (0.493)	
Mother works part-time	0.314 (0.464)	0.305 (0.460)	
Family incomes (measured at age 13, or Wave 1):			
Log gross annual HH income in Wave 1	5.000 (5.070)	5.044 (5.075)	
Log gross annual HH income in Wave 1 missing	0.503 (0.500)	0.499 (0.500)	
Educational Attainment at age 16 (Wave 4)			
Gold standard in GCSEs	0.487 (0.500)	0.566 (0.496)	No
Achieving NQF Level 2 threshold	0.617 (0.486)	0.705 (0.456)	No
Achieving NQF Level 1 threshold	0.310 (0.462)	0.248 (0.432)	No
Maths A*-C	0.594 (0.491)	0.614 (0.487)	
Maths D-G	0.350 (0.477)	0.352 (0.478)	
Contextual value added KS2-KS4	5.783 (58.066)	7.368 (52.531)	
Contextual value added KS2-KS4 missing	0.069 (0.254)	0.063 (0.242)	
Receiving free school meals	0.122 (0.327)	0.132 (0.338)	
Parental Aspirations at age 13 (Wave 1)			
Private school	0.041 (0.198)	0.034 (0.182)	
Parent think overall quality of school very good	0.432 (0.495)	0.446 (0.497)	
Parent think overall quality of school poor	0.107 (0.309)	0.109 (0.312)	
Parent think YP will continue in full-time education at 16	0.687 (0.464)	0.808 (0.394)	No
Parent would like YP to continue in f-t education at 16	0.776 (0.417)	0.881 (0.325)	No
Parent think YP unlikely to go into Higher Education (HE)	0.333 (0.471)	0.239 (0.427)	No
Parent think YP unlikely to go into HE missing	0.052 (0.222)	0.066 (0.248)	No
Number of Observations	4570	4620	

Notes: Standard errors in parentheses.

Table 2a: Linear Probability Model, Boys

Studying for any academic qualifications	**Model 1**	**Model 2**	**Model 3**	**Parsi-monious**
Young Person(YP)'s Own Characteristics:				
Non-white	0.060*	0.049	0.017	
Premature Birth (by 1+ week)	-0.004	-0.016	-0.018	
Log birth weight (kilograms)	0.013	-0.050	-0.057*	
Any disability (health problem) in Wave 1	-0.009	0.001	-0.003	
Any disability (health problem) affecting schooling in Wave 1	-0.190***	-0.029	-0.011	
Single-parent family at birth	-0.029	0.011	0.016	
Mother's Characteristics (measured at age 16, or Wave 4, unless stated otherwise :				
Mother non-white	0.155***	0.108***	0.066**	0.078***
Mother's highest qualification is degree+	0.366***	0.110***	0.060***	0.038***
Mother's highest qualification is NQF4	0.340***	0.106***	0.073***	
Mother's highest qualification is NQF3	0.227***	0.048*	0.027	
Mother's highest qualification is NQF2	0.115***	0.009	0.014	
Mother's highest qualification is NQF1	0.008	-0.038*	-0.027	
Mother's highest qualification is vocational	-0.080***	-0.024	-0.018	
Mother cohabiting	-0.111***	-0.046*	-0.041	-0.046*
Mother is lone parent	-0.121***	-0.042**	-0.038**	-0.044***
Indicator for step-family	-0.064**	-0.039*	-0.028	-0.023
No of siblings in the HH (to YP)	-0.039***	-0.014**	-0.014***	-0.013**
Any non-resident siblings	-0.062***	-0.031**	-0.027	
Mother works full-time	-0.018	-0.037**	-0.025	
Mother works part-time	0.014	-0.026	-0.005	
Log gross annual HH income in Wave 1	0.037***	0.003	-0.007	
Log gross annual HH income in Wave 1 missing	0.394***	0.050	-0.048	
Educational Attainment at age 16 (Wave 4)				
Gold standard in GCSEs		0.264***	0.213***	0.221***
Achieving NQF Level 2 threshold		0.334***	0.217***	0.188***
Achieving NQF Level 1 threshold		0.106***	0.041	
Maths A*-C		0.002	-0.011	
Maths D-G		-0.091***	-0.078	-0.061***
Contextual value added KS2-KS4		0.000***	0.001***	0.001***
Contextual value added KS2-KS4 missing		0.072***	-0.023	
Receiving free school meals		-0.027	-0.036*	
Parental Aspirations at age 13 (Wave 1)				
Private school			0.129***	0.121***
Parent think overall quality of school very good			0.016	
Parent think overall quality of school poor			-0.013	
Parent think YP will continue in full-time education at 16			0.084***	0.090***
Parent would like YP to continue in f-t education at 16			0.072***	0.070***
Parent think YP unlikely to go into Higher Education (HE)			-0.161***	-0.168***
Parent think YP unlikely to go into HE missing			-0.121***	-0.124***
Adj-R^2	0.139	0.397	0.437	0.435

Notes: N=4570. * p<0.1; ** p<0.05; *** p<0.01.

Table 2b: Linear Probability Model, Girls

Studying for any academic qualifications	**Model 1**	**Model 2**	**Model 3**	**Parsi-monious**
Young Person(YP)'s Own Characteristics:				
Non-white	0.106***	0.057*	0.038	
Premature Birth (by 1+ week)	0.040**	0.023	0.026*	
Log birth weight (kilograms)	0.104***	0.065**	0.062**	
Any disability (health problem) in Wave 1	0.023	0.035	0.036	
Any disability (health problem) affecting schooling in Wave 1	-0.172***	0.048	-0.034	
Single-parent family at birth	-0.067***	-0.027	-0.029*	
Mother's Characteristics (measured at age 16, or Wave 4, unless stated otherwise :				
Mother non-white	0.140***	0.110***	0.083***	0.101***
Mother's highest qualification is degree+	0.336***	0.110***	0.074***	0.064***
Mother's highest qualification is NQF4	0.299***	0.090***	0.062**	
Mother's highest qualification is NQF3	0.195***	0.035	0.017	
Mother's highest qualification is NQF2	0.135***	0.030	0.022	
Mother's highest qualification is NQF1	0.028	-0.015	-0.020	
Mother's highest qualification is vocational	-0.074***	-0.027	-0.027	
Mother cohabiting	-0.103***	-0.078***	-0.072***	-0.081***
Mother is lone parent	-0.122***	-0.054***	-0.053***	-0.070***
Indicator for step-family	-0.086***	-0.055**	-0.050**	-0.062***
No of siblings in the HH (to YP)	-0.025***	-0.009	-0.009	-0.009*
Any non-resident siblings	-0.071***	-0.021	-0.017	
Mother works full-time	-0.011	-0.013	0.010	
Mother works part-time	0.032*	0.003	0.006	
Log gross annual HH income in Wave 1	0.040***	0.010	0.004	
Log gross annual HH income in Wave 1 missing	0.368***	0.098	0.044	
Educational Attainment at age 16 (Wave 4)				
Gold standard in GCSEs		0.216***	0.196***	0.220***
Achieving NQF Level 2 threshold		0.289***	0.195***	0.207***
Achieving NQF Level 1 threshold		0.044	-0.002	
Maths A*-C		0.071	0.050	
Maths D-G		-0.004	-0.001	-0.035
Contextual value added KS2-KS4		0.001***	0.001***	0.001***
Contextual value added KS2-KS4 missing		0.069***	0.069***	
Receiving free school meals		-0.016	-0.029	
Parental Aspirations at age 13 (Wave 1)				
Private school			0.068	0.088***
Parent think overall quality of school very good			0.007	
Parent think overall quality of school poor			-0.033*	
Parent think YP will continue in full-time education at 16			0.068***	0.073***
Parent would like YP to continue in f-t education at 16			0.077***	0.073***
Parent think YP unlikely to go into Higher Education (HE)			-0.145***	-0.151***
Parent think YP unlikely to go into HE missing			-0.086***	-0.092***
Adj-R^2	0.134	0.362	0.389	0.387

Notes: N=4620. * p<0.1; ** p<0.05; *** p<0.01.

Table 3a: The Importance of achieving the *gold standard in GCSEs*, Boys

Studying for any academic qualifications	**Model 1**	**Model 2**	**Model 3**	**Parsi-monious**
Educational Attainment at age 16 (Wave 4)				
Gold standard in GCSEs	0.567***	0.299***	0.222***	0.221***
Achieving NQF Level 2 threshold		0.268***	0.192***	0.188***
Contextual value added KS2-KS4		0.000***	0.001***	0.001***
Maths D-G		-0.073***	-0.058***	-0.061***
Parental Aspirations at age 13 (Wave 1)				
Private school			0.121***	0.121***
Parent think YP will continue in FTED at 16			0.090***	0.090***
Parent would like YP to continue in FTED at 16			0.078***	0.070***
Parent think YP unlikely to go into HE			-0.195***	-0.168***
Parent think YP unlikely to go into HE missing			-0.137***	-0.124***
Mother's Characteristics (measured at age 16, or Wave 4):				
Mother non-white				0.078***
Mother's highest qualification is degree+				0.038**
Mother cohabiting				-0.046*
Mother is lone parent				-0.044***
Indicator for step-family				-0.023
No of siblings in the HH (to YP)				-0.013**
Adj-R^2	0.321	0.366	0.430	0.435

Notes: N=4570. * p<0.1; ** p<0.05; *** p<0.01.

Table 3b: The Importance of achieving the *gold standard in GCSEs*, Girls

Studying for any academic qualifications	**Model 1**	**Model 2**	**Model 3**	**Parsi-monious**
Educational Attainment at age 16 (Wave 4)				
Gold standard in GCSEs	0.523***	0.277***	0.230***	0.220***
Achieving NQF Level 2 threshold		0.284***	0.216***	0.207***
Contextual value added KS2-KS4		0.001***	0.001***	0.001***
Maths D-G		-0.0400*	-0.032	-0.035
Parental Aspirations at age 13 (Wave 1)				
Private school			0.092***	0.088***
Parent think YP will continue in FTED at 16			0.077***	0.073***
Parent would like YP to continue in FTED at 16			0.078***	0.073***
Parent think YP unlikely to go into HE			-0.181***	-0.151***
Parent think YP unlikely to go into HE missing			-0.094***	-0.092***
Mother's Characteristics (measured at age 16, or Wave 4):				
Mother non-white				0.101***
Mother's highest qualification is degree+				0.064***
Mother cohabiting				-0.081***
Mother is lone parent				-0.070***
Indicator for step-family				-0.062***
No of siblings in the HH (to YP)				-0.009*
Adj-R^2	0.283	0.331	0.373	0.387

Notes: N=4620. * p<0.1; ** p<0.05; *** p<0.01.

Table 4: Comparing LPM with IVs, All young people aged 16/17

Studying for any academic qualifications	**BOYS**		**GIRLS**	
	OLS	**IV**	**OLS**	**IV**
(Second Stage) Results:				
Gold standard in GCSEs	**0.517**	0.208	**0.475**	*0.265*
	(0.013)	(0.194)	(0.013)	(0.155)
Mother's Characteristics (measured at age 16 or Wave 4):				
Mother non-white	**0.170**	**0.192**	**0.170**	**0.193**
	(0.016)	(0.022)	(0.015)	(0.023)
Mother's highest qualification is degree+	**0.146**	**0.293**	**0.146**	**0.251**
	(0.023)	(0.095)	(0.023)	(0.081)
Mother's highest qualification is NQF4	**0.137**	**0.275**	**0.129**	**0.221**
	(0.029)	(0.092)	(0.029)	(0.074)
Mother's highest qualification is NQF3	**0.078**	**0.181**	**0.064**	**0.137**
	(0.024)	(0.069)	(0.025)	(0.060)
Mother's highest qualification is NQF2	0.029	**0.089**	**0.051**	**0.099**
	(0.019)	(0.043)	(0.019)	(0.040)
Mother's highest qualification is NQF1	-0.030	-0.002	-0.002	0.017
	(0.023)	(0.030)	(0.023)	(0.028)
Mother's highest qualification is vocational	-0.029	**-0.065**	*-0.032*	**-0.054**
	(0.018)	(0.029)	(0.019)	(0.025)
Mother cohabiting	*-0.046*	**-0.093**	**-0.090**	**-0.105**
	(0.028)	(0.042)	(0.028)	(0.031)
Mother is lone parent	**-0.077**	**-0.127**	**-0.087**	**-0.124**
	(0.014)	(0.035)	(0.014)	(0.031)
Indicator for step-family	**-0.048**	**-0.073**	**-0.066**	**-0.089**
	(0.021)	(0.027)	(0.022)	(0.028)
No of siblings in the HH (to YP)	**-0.014**	**-0.026**	-0.007	*-0.013*
	(0.005)	(0.009)	(0.005)	(0.007)
Adj-R^2	0.354		0.317	
Root MSE	0.402	0.426	0.403	0.414
Observations	4570		4620	
First Stage Results:				
Log birth weight (kilograms)		**0.113**		**0.066**
		(0.038)		(0.033)
Relative (school starting) age in months		**-0.007**		**-0.011**
		(0.002)		(0.002)
Diagnostic Tests:				
Cragg-Donald Wald F-stat for excluded Restrictions		**11.21**		**16.86**
(p-value)		(0.000)		(0.000)
Critical Value for 10% relative bias		19.93		19.93
Critical Value for 15% relative bias		11.59		11.59
Critical Value for 20% relative bias		8.75		8.75
Sargan (Anderson-Rubin /Hansen) $\chi^2_{(1)}$		0.029		3.750
(P-value)		(0.866)		(0.053)
Hansen's J Statistic $\chi^2_{(1)}$		0.029		3.241
(P-value)		(0.864)		(0.072)

Notes: Standard errors in parentheses. **Bold** and *italic* cases indicate statistical significance at the 5% and the 10% levels respectively.

Table 5: Comparing LPM with IVs, Conditional on Working towards a Qualification

Studying for any academic qualifications	**BOYS**		**GIRLS**	
	OLS	**IV**	**OLS**	**IV**
(Second Stage) Results:				
Gold standard in GCSEs	**0.476**	**0.350**	**0.432**	**0.435**
	(0.014)	(0.177)	(0.013)	(0.128)
Mother's Characteristics (measured at age 16 or Wave 4):				
Mother non-white	**0.129**	**0.133**	**0.146**	**0.146**
	(0.017)	(0.018)	(0.016)	(0.017)
Mother's highest qualification is degree+	**0.101**	*0.155*	**0.121**	*0.120*
	(0.025)	(0.080)	(0.024)	(0.062)
Mother's highest qualification is NQF4	**0.102**	*0.156*	**0.115**	*0.114*
	(0.033)	(0.082)	(0.030)	(0.059)
Mother's highest qualification is NQF3	*0.047*	0.084	**0.056**	0.055
	(0.027)	(0.059)	(0.026)	(0.048)
Mother's highest qualification is NQF2	0.020	0.045	**0.046**	0.046
	(0.022)	(0.042)	(0.020)	(0.033)
Mother's highest qualification is NQF1	-0.036	-0.022	-0.002	-0.002
	(0.028)	(0.034)	(0.025)	(0.026)
Mother's highest qualification is vocational	-0.022	-0.035	**-0.047**	**-0.047**
	(0.021)	(0.028)	(0.019)	(0.023)
Mother cohabiting	-0.036	-0.054	**-0.076**	**-0.075**
	(0.033)	(0.042)	(0.030)	(0.033)
Mother is lone parent	**-0.069**	**-0.087**	**-0.084**	**-0.084**
	(0.017)	(0.030)	(0.014)	(0.027)
Indicator for step-family	*-0.046*	*-0.051*	**-0.047**	*-0.046*
	(0.025)	(0.027)	(0.023)	(0.027)
No of siblings in the HH (to YP)	**-0.016**	**-0.020**	-0.008	-0.008
	(0.006)	(0.009)	(0.006)	(0.007)
Adj-R^2	0.298		0.287	
Root MSE	0.398	0.401	0.379	0.379
Observations	3612		3926	
First Stage Results:				
Log birth weight (kilograms)		**0.132**		**0.079**
		(0.038)		(0.036)
Relative (school starting) age in months		**-0.008**		**-0.013**
		(0.002)		(0.002)
Diagnostic Tests:				
Cragg-Donald Wald F-stat for excluded Restrictions		**11.88**		**21.44**
(p-value)		(0.000)		(0.000)
Critical Value for 10% relative bias		19.93		19.93
Critical Value for 15% relative bias		11.59		11.59
Critical Value for 20% relative bias		8.75		8.75
Sargan (Anderson-Rubin /Hansen) $\chi^2_{(1)}$		0.000		3.521
(P-value)		(0.991)		(0.061)
Hansen's J Statistic $\chi^2_{(1)}$		0.000		3.038
(P-value)		(0.991)		(0.081)

Notes: Standard errors in parentheses. **Bold** and *italic* cases indicate statistical significance at the 5% and the 10% levels respectively.

Figure 1: Effect of Month of Birth on Attaining the *Gold standard in GCSEs*

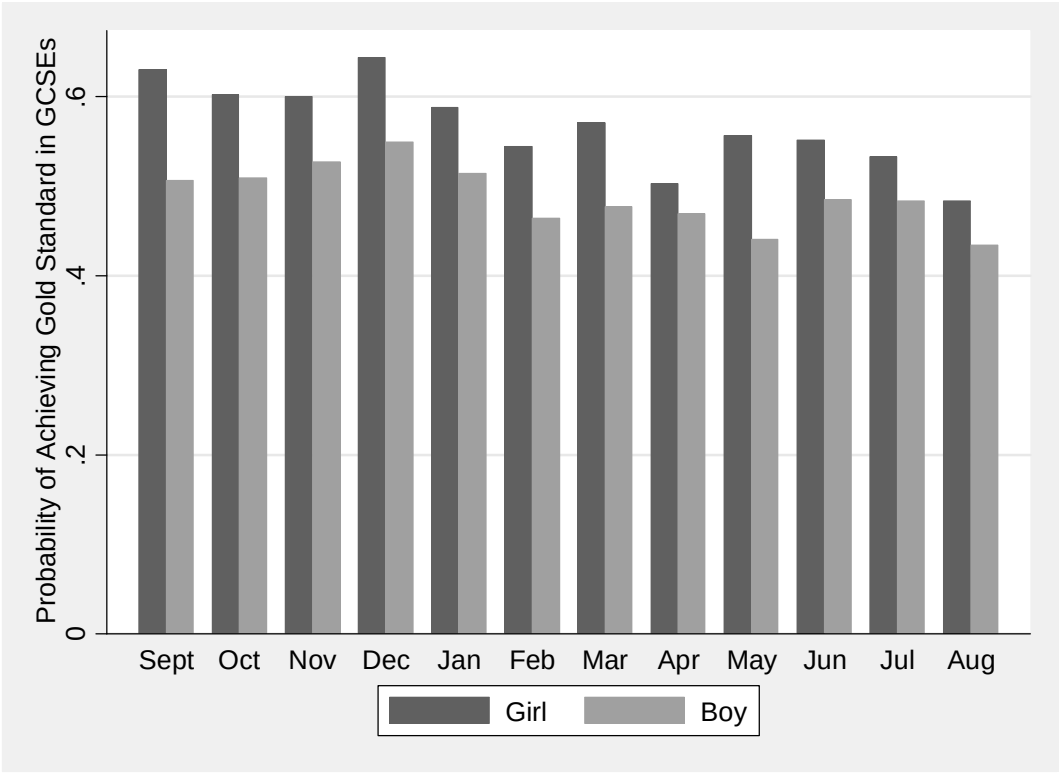
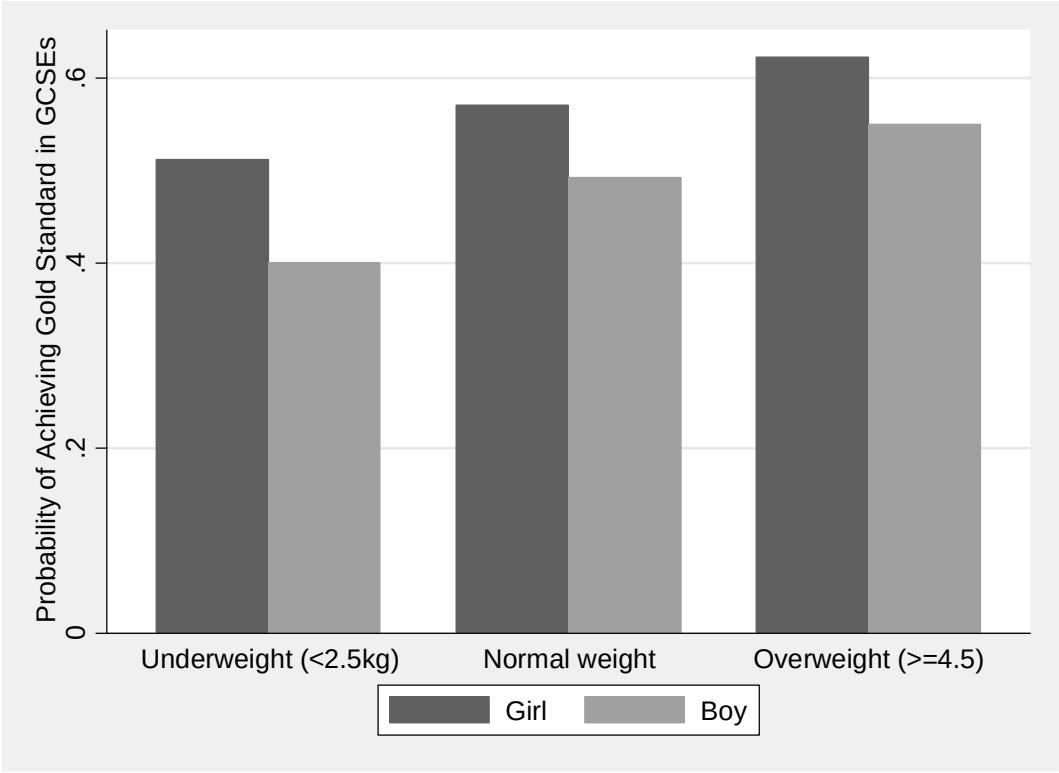


Figure 2: Effect of Birth Weight on Attaining the *Gold standard in GCSEs*



References

Angrist, J. and A. Krueger (1991), 'Does Compulsory School Attendance affect Schooling and Earnings?', *Quarterly Journal of Economics*, 106, 4, 979-1014.

Angrist, J. and J-S. Pischke (2009), *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton University Press.

Becker, G. (1964), *Human Capital*, Chicago University Press.

Behrman, J., Rosenzweig, M. And P. Taubman (1994), 'Endowments and the Allocation of Schooling in the Family and in the Marriage Market: the Twins Experiment', *Journal of Political Economy*, 52, 1131-1174.

Behrman, J. and M. Rosenzweig (2004), 'The Returns to Birthweight', *Review of Economics and Statistics*, 86, 586-601.

Berman, E., J. Bound and Z. Griliches (1994), 'Changes in the Demand for Skilled Labor Within US Manufacturing: Evidence from the Annual Survey of Manufactures', *The Quarterly Journal of Economics*, 109, 367-397.

Black, S., P. Devereux, and K.G. Salvanes (2008), 'Too Young to Leave the Nest? The Effects of School Starting Age', *NBER Working Paper*, No. 13969.

J. Bound, D. Jaeger and R. Baker (1995), 'Problems with Instrumental Variable Estimation when the Correlation between the Instruments and the Endogenous Explanatory Variable is Weak', *Journal of the American Statistical Association*, 90, 430, 443-450.

Cesur, R. And I. Rashad (2008), 'High Birth Weight and Cognitive Outcomes', *NBER Working Paper*, No. 14524.

Chowdry, H., Crawford, C. and A. Goodman (2010), 'The Role of Attitudes and Behaviours in Explaining Socio-Economic Differences in Attainment at Age 16, *IFS Working Paper 10/15*.

Conlon, G. (2001), 'The Differential in Earnings Premia Between Academically and Vocationally Trained Males in the United Kingdom', *CEEDP 0011*, London School of Economics.

Conlon, G. (2005), 'The Determinants of Undertaking Academic and Vocational Qualifications in the United Kingdom', *Education Economics*, 13, 299 – 313.

Crawford, C., Dearden, L. and C. Meghir (2007), 'When You Are Born Matters: the Impact of Date of Birth on Child Cognitive Outcomes in England', *IFS Commentary*, www.ifs.org.uk/docs/born_matters_report.pdf.

Dahl, G. and E. Moretti (2008), 'The Demand for Sons', *The Review of Economic Studies,* 75, 1085-1120.

DCSF (Department for Children, Schools and Families) (2009), *Independent Review of the Primary Curriculum*, the Stationery Office.

Dearden, L., McIntosh, S., Myck, M. and A. Vignoles (2002), 'The Returns to Academic and Vocational Qualifications in Britain', *Bulletin of Economic Research*, 54(3), 249-274.

Dearden, L., McGranahan, L. and B. Sianesi (2004), 'An In-depth Analysis of the Returns to National Vocational Qualifications Obtained at Level 2', *CEE DP46*, London School of Economics.

Del Bono, D. and F. Galindo-Rueda (2004), 'Do a Few Months of Compulsory Schooling Matter? The Education and Labour Market Impact of School Leaving Rules', *IZA DP* 1233.

DfES (Department for Education and Skills) (2005), *14-19 Education and Skills*, White Paper, the Stationery Office.

Dickerson, A. (2008), 'The Distribution and Returns to Qualifications in the Four Countries of the UK', *Sector Skills Development Agency Research Report*, 21A.

Fletcher, J. (2011), 'The Medium Term Schooling and Health Effects of Low Birth Weight: Evidence from Siblings', *Economics of Education Review*, 30, 3, 517-527.

Green, F., Machin, S., Murphy, R. and Y. Zhu (2008), 'Competition for Private and State School Teachers'. *Journal of Education and Work***,** 21(5), 383-404.

Heckman, J.J. (2008), 'Schools, Skills and Synapses', *IZA DP* 3515.

Jenkins, A., Greenwood, C. and A. Vignoles (2007), "The Returns to Qualifications in England: Updating the Evidence Base on Level 2 and Level 3 Vocational Qualifications", CEE DP, September 2007.

McIntosh, S. (2006), 'Further Analysis of the Returns to Academic and Vocational Qualifications', *Oxford Bulletin of Economics and Statistics*, *68*, 225–251.

Miller, P., Mulvey, C. and N. Martin (2005), 'Birth weight and schooling and earnings: estimates from a sample of twins', *Economics Letters*, 86, 387–392.

Nuffield Foundation, (2009), *The Nuffield Review of 14-19 Education*. http://www.nuffield14-19review.org.uk/.

Oreopoulos, P., Stabile, M., Walld, R. and L. Roos (2008), 'Short-, Medium-, and Long-Term Consequences of Poor Infant Health: An Analysis Using Siblings and Twins', *Journal of Human Resources*, 43, 88-138.

Robinson, P. (1997), 'The Myth of Parity of Esteem: Earnings and Qualifications', *Centre for Economic Performance Discussion Paper*, No. 354, London School of Economics.

Sharp, C., George, N., Sargent, C., O'Donnell, S. and M. Heron (2009), *International Thematic Probe: The Influence of Relative Age on Learner Attainment and Development*, National Foundation for Educational Research.

Stock, J. H., and M. Yogo. 2005, 'Testing for weak instruments in linear IV regression', in Andrews, D. and Stock, J. (eds), *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, Cambridge University Press, New York.

Walker, I. and Y. Zhu (2009), 'The Causal Effect of Teen Motherhood on Worklessness', *University of Kent Studies in Economics*, 09/17.

Appendix

Table A1: Probit Estimates, Marginal Effects

Studying for any academic qualifications	**All Young People**		**Young People Working towards a Qualification**	
	BOYS	**GIRLS**	**BOYS**	**GIRLS**
Gold standard in GCSEs	**0.538**	**0.490**	**0.485**	**0.436**
	(0.014)	(0.014)	(0.015)	(0.016)
Mother's Characteristics (measured at age 16 or Wave 4):				
Mother non-white	**0.234**	**0.207**	**0.156**	**0.157**
	(0.021)	(0.018)	(0.020)	(0.015)
Mother's highest qualification is degree+	**0.202**	**0.189**	**0.129**	**0.150**
	(0.030)	(0.026)	(0.027)	(0.020)
Mother's highest qualification is NQF4	**0.187**	**0.153**	**0.122**	**0.115**
	(0.037)	(0.032)	(0.034)	(0.026)
Mother's highest qualification is NQF3	**0.106**	**0.077**	0.052	**0.060**
	(0.033)	(0.031)	(0.032)	(0.027)
Mother's highest qualification is NQF2	*0.045*	**0.068**	0.024	**0.056**
	(0.027)	(0.024)	(0.027)	(0.021)
Mother's highest qualification is NQF1	-0.037	0.002	-0.035	0.008
	(0.034)	(0.030)	(0.035)	(0.027)
Mother's highest qualification is vocational	-0.042	*-0.040*	-0.026	**-0.054**
	(0.026)	(0.024)	(0.026)	(0.023)
Mother cohabiting	-0.064	**-0.121**	-0.044	**-0.086**
	(0.040)	(0.039)	(0.043)	(0.039)
Mother is lone parent	**-0.107**	**-0.113**	**-0.087**	**-0.099**
	(0.020)	(0.019)	(0.022)	(0.018)
Indicator for step-family	**-0.065**	**-0.088**	-0.053	**-0.057**
	(0.031)	(0.029)	(0.033)	(0.029)
No of siblings in the HH (to YP)	**-0.022**	-0.010	**-0.021**	-0.009
	(0.008)	(0.007)	(0.008)	(0.007)
Pseudo R^2	0.284	0.283	0.287	0.287
Observations	4570	4620	3612	3926

Notes: Standard errors in parentheses. **Bold** and *italic* cases indicate statistical significance at the 5% and the 10% levels respectively.