ISER Working Paper Series

The long-run effects of attending an elite school: evidence from the UK

Damon Clark
University of California, Irvine

Emilia Del Bono
Institute for Social and Economic Research
University of Essex

No. 2014-05
February 2014

INSTITUTE FOR SOCIAL & ECONOMIC RESEARCH

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

www.iser.essex.ac.uk

Non-technical summary

In several European countries and some US cities, students are tracked into different types of high school according to their ability; students perceived as academically able are allowed to attend 'elite' schools, while students perceived as less academically able are taught in 'non-elite' schools. An emerging body of evidence suggests that being tracked into the elite schools in these systems has, at best, small effects on test scores and college outcomes. This is surprising. As several of these studies document, parents have strong preferences for elite schools within these systems. One explanation is that parents overestimate the importance of being at school with other high ability students on educational outcomes. Another explanation is that parents are focused on outcomes other than educational attainment, such as crime for example. A third explanation is that parents are focused on longer-run outcomes, and that elite school attendance improves these outcomes despite apparently modest effects on test scores and college enrollment.

To date, there are few analyses of the long-run effects of elite school assignment. This paper begins to fill this gap by providing what we believe are the first estimates of the long-run impact of attending an elite school. These estimates make use of a large sample of students educated in a UK district that operated a selective high school system. Because assignment in this district was based on a strict formula, we can exploit our knowledge of this formula to generate credible estimates of the causal effects of attending an elite school within this system. The individuals in our sample attended school in the 1960s, and were followed and surveyed in 2001 (when they were in their late 40s). We can therefore estimate impacts on a range of long-run outcomes, including completed education, income, marriage, fertility and occupational success.

We find large impacts of elite school attendance on educational attainment of both men and women. For women, elite school attendance increased full-time education by almost one year and increased the probability of earning A-levels by 23 percentage points. For men, elite school attendance increased completed years of full-time education by more than one year and doubled the probability of degree receipt. These effects likely reflect the higher barriers (i.e., non-monetary costs) to further full-time education faced by students that attended non-elite schools. For women, we also find that elite school attendance significantly increased income and wages (by 20 percent and 10 percent respectively) and significantly decreased completed fertility (by around 0.5 children). However, for men elite school attendance had no effect on income or wages and no effect on fertility or marriage. We speculate this is because elite school attendance caused men to pursue further academic education at the expense of vocational training, such that the overall impact on human capital accumulation was ambiguous.

Our analysis shows that elite school attendance can have important long-run effects, including but not limited to effects on labor market outcomes. This suggests that selective school systems can generate a lottery in life chances, with important advantages accruing to students that perform well on the assignment tests. Our findings also suggest that the long-run impacts of school quality cannot be understood without reference to the wider education and labor market institutions facing students. For example, the small labor market impacts that we estimate for men may be driven by the vocational training options enjoyed by non-elite school students in this era. Men that attended non-elite schools in other settings may have enjoyed fewer such options. From a policy perspective, this research suggests that policy-makers would be advised to keep in mind the importance of related institutions when proposing changes to school resources and organization.

The Long-Run Effects of Attending an Elite School: Evidence from the UK*

Damon Clark
UC Irvine, NBER and IZA

Emilia Del Bono
University of Essex and IZA

6th February 2014

Abstract

This paper estimates the impact of elite school attendance on long-run outcomes including completed education, income and fertility. Our data consists of individuals born in the 1950s and educated in a UK district that assigned students to either elite or non-elite secondary schools. Using instrumental variables methods that exploit the school assignment formula, we find that elite school attendance had large impacts on completed education. For women, we find that elite school attendance generated large improvements in labor market outcomes and significant decreases in fertility; for men, we find no elite school impacts on any of these later-life outcomes.

JEL Classification: I2, J24, C31, C36

Keywords: Education, school quality, instrumental variables

*We would like to thank Heather Clark and David Leon for generous help with the Aberdeen data. We also thank the staff at the Aberdeen City Library for their hospitality and help with Education Committee minutes from the 1960s. We received many useful comments from Josh Angrist, Sonia Balothra, Kenneth Chay, Adeline Delavande, Yingying Dong, Steve Pudney and Luke Sibietta, as well as seminar participants at Cornell University, DIW Berlin, University of Essex, McGill University, SUNY Binghamton, MIT, Princeton University, the University of St. Andrews. This work contains statistical data from ONS which is crown copyright and reproduced with the permission of the controller HMSO and Queen's Printer for Scotland. The use of the ONS statistical data in this work does not imply the endorsement of the ONS in relation to the interpretation or analysis of the statistical data. We would like to thank Annarosa Pesole for her help in analyzing ONS data and Seok Min Moon for superb research assistance. Damon Clark acknowledges financial support through a National Educational Association/Spencer Foundation Post-Doctoral Research Fellowship. Emilia Del Bono acknowledges the support provided by the ESRC Centre on Micro-Social Change at ISER (grant RES-518-28-001).

1 Introduction

In many parts of the world, including several European countries and some US cities, students are tracked into different types of high school: students perceived as academically able into elite schools, students perceived as less academically able into non-elite schools. An emerging body of evidence suggests that being tracked into the elite schools in these systems has, at best, small effects on test scores and college outcomes (Angrist et al., 2011; Clark, 2010; Dobbie and Fryer, 2011; Pop-Eleches and Urquiola, 2013).[1] This is surprising. As several of these studies document, parents have strong preferences for elite schools within these systems. Indeed, much of the pressure to reform these systems stems from the perception that they represent a lottery in life chances, one in which the lucky winners assigned to the elite schools are given the prize of a better education and better later-life outcomes.[2]

One explanation for this combination of strong preferences and weak impacts is the possibility that parents do not understand the education production function: for example, they might overestimate the importance of peer effects. Another explanation is that parents are focused on other youth outcomes such as crime, which Deming (2010) shows can be improved when parents gain access to their preferred schools. A third explanation is that parents are focused on longer-run outcomes, and that elite school attendance improves these outcomes despite apparently modest effects on test scores and college enrollment. Unfortunately, there are few analyses of the long-run effects of elite school assignment, presumably a reflection of the difficulties associated with identifying exogenous variation in school assignments and then matching these to adult outcome data.[3] This is unfortunate because other education evalua-

[1] In a closely related study, Cullen et al. (2006) also find small test score effects of attending "better" schools in Chicago, in this case regular public schools that are high-achieving and popular with parents. An exception to this pattern of small effects is Jackson (2010), who finds larger effects of attending elite schools in Trinidad and Tobago.

[2] To analyze convincingly whether a selective or non-selective system is most effective we would require quasi-random assignment of students to different types of systems (such as that implemented by Duflo et al., 2011). We do not have access to this type of assignment hence make no claims as to which system is most effective.

[3] Dustmann et al. (2012) offer one such analyses, focusing on the impact of attending an

tions have revealed a disconnect between test score and long-run impacts (e.g., Garces et al. 2002; Krueger and Whitmore, 2001).

This paper begins to fill this gap by providing what we believe are the first estimates of the long-run impact of attending an elite school. These estimates make use of a large sample of students educated in a UK district that operated a selective high school system. Because assignment in this district was based on a strict formula, we can exploit our knowledge of this formula to generate credible estimates of the causal effects of attending an elite school within this system. The individuals in our sample attended school in the 1960s, and were followed and surveyed in 2001 (when they were in their late 40s). We can therefore estimate impacts on a range of long-run outcomes, including completed education, income, marriage, fertility and occupational success.

Our analysis produces three main findings. First, we find large impacts of elite school attendance on educational attainment. For women, we estimate that elite school attendance increased full-time education by almost one year and increased the probability of earning A-levels by 23 percentage points. Relative to average attainment among women with borderline scores that attended non-elite schools (which we refer to as the "control group" mean), this represents a 36 percent increase in completed years of post-compulsory education and a 60 percent increase in the likelihood of achieving A-level qualifications. For men, we estimate that elite school attendance increased completed years of full-time education by more than one year (approximately 60 percent of the control group mean) and doubled the probability of degree receipt. These

elite middle school in Germany. They instrument elite school attendance using date of birth relative to the school starting age, the idea being that older students will be more likely to be deemed suitable for the elite schools. The authors find that higher-track attendance in middle school has negligible effects on the type of secondary education received and on long-run outcomes such as wages and unemployment. An important caveat that could account for these findings is that, after being assigned, German students can move between tracks. There is much less scope for between-school mobility in the setting we consider. A related strand of literature considers the effect of changing the fraction of students assigned to the elite track. Duflo et al. (2011) use experimental variation to analyze tracking in Kenya; Guyon et al. (2012) argue convincingly that a Northern Ireland policy that resulted in an expansion of the elite track provides quasi-experimental variation in the size of the elite track. The relationship between average outcomes and the fraction of students tracked is an interesting and important one, but not one that we can address in this paper.

effects likely reflect the higher barriers (i.e., non-monetary costs) to further full-time education faced by students that attended non-elite schools. For example, as shown by Clark (2010), non-elite school students may have taken too narrow a range of courses to succeed in certain degree programs. We suspect these effects are larger than in those found in the previous literature because we suspect that these barriers are higher than those in other contexts (e.g., in the contemporary US context analyzed by Dobbie and Fryer, 2011, in which the SAT plays an important role in college admission and high school course-taking may be more similar across elite and non-elite schools).

Second, for women, we estimate that elite school attendance significantly increased income and wages (by 20 percent and 10 percent respectively) and significantly decreased completed fertility (by around 0.5 children). As such, these estimates are consistent with elite school effects on educational attainment leading to better labor market opportunities, and with better labor market opportunities leading to lower fertility. These large impacts on completed fertility are also consistent with Goldin's (2006) account of the importance of education in enabling a "quiet revolution" in women‘s economic and social roles in the 1970s.

Third, for men, we estimate that elite school attendance had no effect on income or wages and no effect on fertility or marriage. We speculate this is because elite school attendance caused men to pursue further academic education at the expense of vocational training, especially trade apprenticeships, such that the overall impact on human capital accumulation was ambiguous. We formalize this explanation using a school quality model similar to Card and Krueger (1996) but extended to include vocational training. We then show that several implications of this explanation are confirmed in the data. We also show that the data reject the implications of several alternative hypotheses.

We draw three conclusions from our analysis. First, elite school attendance can have important long-run effects, including but not limited to effects on labor market outcomes. Among other things, this suggests that selective school systems can generate a type of lottery in life chances, with important advantages accruing to students that perform well on the assignment tests. To the best of our knowledge, this is the first study to provide evidence in support of this point, one stressed by opponents of this system in the 1950s and 1960s.[4] This may also explain why parents exhibit strong preferences for elite-type schooling despite evidence that short-run effects can be small. Second, our findings suggest that the long-run impacts of school quality cannot be understood without reference to the wider education and labor market institutions facing students. For example, the large education impacts that we estimate likely reflect the barriers to further education faced by non-elite school students in this era. These barriers may be lower in other settings. Similarly, the small labor market impacts that we estimate for men may be driven by the vocational training options enjoyed by non-elite school students in this era. Men that attended non-elite schools in other settings may have enjoyed fewer such options.Third, from a policy perspective, it follows that policy-makers would be advised to keep in mind the importance of related institutions when proposing changes to school resources and organization. For example, in the contemporary US context, it seems plausible to suppose that elite school effects would be shaped by whether non-elite school students had ready access to Advanced Placement courses (Klopfenstein, 2004) and SAT-taking opportunities (Bulman, 2013; Goodman, 2012).

2 Institutions and data

2.1 The educational system in Aberdeen in the 1960s

Our data consists of a cohort of children born in the 1950s and educated in Aberdeen, Scotland. In the 1960s, the school system in Scotland was similar to

[4]The argument was that it was unfair and undesirable that life chances could hinge on the answers to a few questions on tests that children took at age elevent to enter these schools. These tests were thought to be decisive in part because there was limited scope for between-school transfers after age eleven (see the discussion in Galindo-Rueda and Vignoles, 2004).

that in the rest of the UK. Education was compulsory for all children aged 5 to 15. After 7 years of primary school, at about age 12, children were transferred to one of two types of secondary school: elite schools (known as "Senior Secondary Schools" in Scotland and "Grammar Schools" in England and Wales) and non-elite schools (known as "Junior Secondary Schools" in Scotland and "Secondary Modern Schools" in England and Wales).[5] In Aberdeen in the 1960s, there were three elite schools and 15 non-elite schools, three of which were private[6,7]

Secondary School Assignment

Secondary school assignment was determined by tests and assessments that took place during the last year of primary school. The tests comprised two intelligence tests (Verbal Reasoning Quotient (VRQ) tests), an English attainment test and an arithmetic attainment test, each standardized to have mean 100 and standard deviation 15. Two assessments (of ability in English and arithmetic) were provided by the student's primary school teacher. These were averaged and standardized to give a single teacher assessment with mean 100 and standard deviation 15. This was then added to the four test scores to give an overall assignment score with mean 500. The other assessment (of the student's suitability for an elite school) was provided by the primary school head teacher (categories were "suitable", "doubtful" or "unsuitable").

Students with assignment scores below 540 were allocated to a non-elite

[5]In the UK, elite high schools were established by the 1944 Education Act. Before the Act, these schools formed a class of private schools that offered scholarships in exchange for financial support from the local school district; after the Act, they received all of their funding from the district, were not allowed to charge fees and were required to admit students on the basis of academic potential assessed at the end of primary school. At its simplest, this involved all students in a district taking a test (the "11-plus"), with the elite school places going to the top-scoring students. The non-elite schools remained broadly unchanged after the Act, the important caveat being that while they previously educated all but those students that won scholarships to the elite schools, they now educated all students that failed the "11-plus".

[6]In addition to these, there were two special needs secondary schools and a convent.

[7]Appendix B Tables 1 and 2 use LFS data to show that the distribution of years of schooling and qualification attainment in our sample is broadly similar to that of comparison groups of individuals from the whole of the UK or the whole of Scotland.

school; students with assignment scores of 580 or more were assigned to an elite school unless assessed by their Head as "unsuitable"; students with scores between 560 and 579 and assessed by their Head as "suitable" were allocated to an elite school provided one of their intelligence test scores was at least 112. The names of the remaining students with assignment score between 540 and 579 were sent back to the Heads, who were required to arrange them in order of merit and write a report on each. The Appeals subcommittee used this information to allocate the remaining elite school places. At the end of this process, the allocation was publicly announced and parents could appeal.

As a result of this procedure, we expect that (i) no students with assignment scores of 539 or below would be assigned to an elite school; (ii) most students with assignment scores of 560 or above would be observed in an elite school; (iii) the fraction of students with assignment scores in the range 540-559 attending an elite school would be increasing in the assignment score. Our data are broadly consistent with this hypothesis. First, the distribution of assignment scores is as expected, with mean close to 500 (see Figure 1). Second, the relationship between school assignment and assignment scores has the expected pattern. This can be seen in Figure 2, which graphs elite school assignment against the assignment score.[8] The circles show the fraction of students attending an elite school for a 5-point interval of the score; the solid line is the probability of attending an elite school as predicted by a linear probability model of elite school attendance that includes a dummy variable for the borderline score range, a dummy for scores above this range and interactions between these dummies and the assignment score. The graph reveals that assignment probabilities are low for assignment scores less than 540, high for assignment scores greater than 560 and increasing in scores for

[8]What we term "elite school assignment" is actually "elite school attended" as reported by respondents. While the two could differ if respondents misreport the school actually attended, data from one of the four cohorts of students in our analysis (those observed in grade 7 in December 1962) suggests that any such differences are likely very small. In partiuclar, the number that attended a non-elite school in 1964 but report attending an elite school at the time of the postal survey is 7, while the number observed in an elite school in 1964 but who report having attended a non-elite school in 2001 is 3 (out of a total of 1,097 grade 7 survey respondents in 2001).

assignment scores in between.[9] The few students with scores below 540 that report attending an elite school may have won an appeal against an initial non-selective assignment.[10] The few students with scores above 560 that did not attend an elite school are likely those that primary Heads deemed "unsuitable" or "doubtful".

Curriculum and Exams

At this time, the minimum school leaving age was 15; hence all students could leave after three years in whichever secondary school they were assigned to. Students could stay in the elite schools for up to six years. In the third and fourth year, they could take courses leading to the Scottish Certificate of Education (SCE) "O grade" exams. In the fifth and sixth year, they could take courses leading to SCE "H grade" exams. In the sixth year, they could also take courses leading to a "Certificate of Sixth Year Studies". This was overseen by a different examinations board and was broadly equivalent to English "A-levels".

All of the non-elite schools allowed students to stay for four years and take courses leading to "O grade" exams. They also offered more vocational-type courses. To take more courses (e.g., leading to "H grade" exams), students had to transfer to an elite school. Elite school registers suggest that few students did this.[11]

Post-secondary options

As described by Findlay (1973), students could pursue degree courses at universities or teacher training courses at universities or teacher training col-

[9]Appendix A Figure 2 shows similar figures for students in different grades. These demonstrate that the rule was consistently applied for all four grades considered in our analysis.

[10]Grounds for appeal would likely have included the child being unwell on a test day or missing time at school through illness or family circumstances.

[11]For two of the three elite schools in Aberdeen, we gained access to school registers from the 1960s. These show that a small number of students entered the school for the first time at an age consistent with them having already spent four years in a non-elite school. We cannot match these students to our data, and our data do not contain information on whether a student transferred schools, but we view this as evidence that transfer opportunities were limited in our setting. We view such transfers as a mechanism that could decrease the cost of an initial non-elite school assignment.

leges. They could also pursue what Findlay describes as two main types of further education: technical and commercial. Technical education included higher-level type education leading to a Higher National Diploma (HND). This could be pursued at some universities and various "central institutions" (e.g., Colleges of Commerce, Agricultural Colleges, Nautical Colleges, Technical Colleges, Colleges and Schools of Art). In addition, it included lower-level education leading to lower-level qualifications (e.g., OND, HNC, ONC and City and Guilds qualifications). This could be pursued at colleges of further education serving the local area and would typically involve day release, block release, apprenticeship, sandwich or similar course arrangements. Commercial education was typically confined to further education colleges and included secretarial and business studies courses.

The apprenticeship system provided students with another alternative to the academic track. During the 1950s and the 1960s the system was based on a formal or informal agreement between a firm and an apprentice. This specified the length of the apprenticeship (between three and six years) and the classroom-based training component, typically day release to a technical college. The classroom-based component ensured that apprentices could acquire formal qualifications, such as City and Guilds or Business and Technology Education Council (BTEC) certificates (Steedman et al., 1998).

2.2 The Aberdeen Children of the 1950s

Our data come from the "Aberdeen Children of the 1950s" study. The study cohort consists of 12,150 children born in Aberdeen between 1950 and 1956 who participated in the Aberdeen Child Development Survey (Batty et al., 2004; Illsley and Wilson, 1981).[12] The target population consisted of all stu-

[12] Aberdeen is a coastal town in the North-East of Scotland. In the 1960s it was the third largest city in Scotland, its economy consisting of rapidly declining traditional industries, such as fishing and shipbuilding. Its fortunes changed dramatically with the discovery of the North Sea oil in 1971. The new oil industry offered more and well-paid high skilled jobs, and generated spillover effects on other sectors, including the state and service sector (Batty et al., 2004).

dents in primary school grades 3-7 in December 1962 (i.e., roughly aged 6-13). According to Illsley (2002), all students were covered by the study except for those attending three small private primary schools that did not take part (2.2 percent of targeted children).

In phase I of the study students were given a series of reading tests and asked to provide demographic information for themselves and their parents (address and date of birth). This information was used to link them to administrative records from the Aberdeen Maternity and Neonatal Databank (match rate 86 percent). These records included perinatal and social information collected throughout the course of their mother's pregnancy and their own birth. In phase II of the study (a year later, in July 1963) the students' medical records were extracted. In phase III of the study (March 1964) sociometric and behavioral data were collected from teachers, children and a 20 percent sample of parents. As a result of these data collection efforts, we know the father's occupation at the time of the child's birth, the premarital occupation of the child's mother, the father's occupation in 1962 (as described by the survey child) and the socio-economic status of the area in which the family lived at the time of the 1962 survey (based on dwelling age, ownership, building type and availability of domestic facilities).

District-held test score data were subsequently added to the dataset. These include all of the transfer tests and assessments discussed above (two IQ, one arithmetic, one English, one combined teacher estimate) and the scores of tests taken at ages 7 and 9.[13] The test at age 7 was called the "Moray House Picture Intelligence Test" and was used to screen students for a mental handicap; the test at age 9 was the "Schonell and Adams Essential Intelligence Test" used to screen for poor readers.

In 1998 a study team began to gather new information from the original participants using administrative records on pregnancies, hospital admissions and mortality, as well as administering a postal survey. Over 97 percent of

[13]This data is missing for the youngest students in the dataset, i.e. those attending grade 3 in December 1962. That is because the procedure used to assign these students had changed, and was based only on the two IQ tests.

the core population (N=11,727) were traced. Of these, 4 percent had died, 2.5 percent had emigrated and 0.6 percent were in the armed forces (Batty et al., 2004). The postal survey was conducted in 2001. Traced participants were sent a sex-specific questionnaire that obtained a response rate of 63.7 percent.

To construct the samples used in this paper, we start with the students matched to the Aberdeen Maternity Databank (N=12,150) then restrict the sample in several ways (see Appendix A Table 1a). First, we exclude individuals who moved outside Aberdeen during the period 1962-1964, as we do not have complete information on their test results and because the vast majority of them attended secondary schools outside Aberdeen. Second, we exclude some individuals on the basis of the primary school attended. In particular, we exclude: (i) individuals who attended either a private and/or faith primary school, as some of these did not take the assignment tests and others were much less likely to attend elite schools conditional on the assignment test score; (ii) individuals who attended elite secondary school during their primary school years, as these are observed to attend an elite secondary school irrespective of their assignment test scores; (iii) individuals who attended special needs schools; and (iv) individuals who attended primary schools outside Aberdeen.

Third, since we require information on school grade at the time of the first interview we exclude individuals for whom this is not available.[14] We also exclude individuals with missing assignment scores and missing age-7 and age-9 test scores, all of which are used in our analysis. Fourth, we exclude the roughly 40 percent of individuals who did not respond to the postal survey, which provides information on the type of secondary school attended and most of our outcome variables. Fifth, since we wish to compare individuals that attended elite and non-elite state schools, we exclude individuals that report attending a private secondary school (around 5 percent). Finally, since the

[14]Grade information was recorded on "Form A" (the one filled out by the children at the time of the first interview), but it was not added to the dataset until 1964, when it was collected as part of the sociometric data . Therefore, children with no sociometric data have no information on grades. Grade is recorded as a separate variable, and it is not based on date of birth, although the data suggest that there was not a lot of grade retention/promotion.

assignment procedure changed in 1966/67, we exclude the one cohort that was subject to this new procedure. Our final sample thus consists of 4,528 observations.

An obvious concern here is sample selection bias. First, we might worry that assignment to elite school affects survey response rates, with students assigned to elite schools being more or less likely to respond to the postal survey. Second, we might worry that assignment to elite school affects private school choices, with students assigned to elite school less likely to attend a private school. On the first point, the left panel of Appendix A Figure 1 reveals a positive relationship between assignment scores and survey response rates, but no evidence of a jump or a change in the slope within the borderline score range. Further tests performed on the cohort of children in grade 7 in December 1962, for whom we know secondary school assignment in March 1964 (independently of their postal survey response rate), confirm that there is no significant relationship between elite school status and the probability that individuals reply to the adult questionnaire. On the second point, the right panel of Appendix A Figure 1 suggests a negative correlation between private school attendance and test scores in the borderline score range (in contrast to the generally positive relationship), a possible indication that elite school attendance decreases the probability of attending private school.[15]

Although there is little we can do about possible biases caused by elite school impacts on private school attendance, two points are worth noting. First, any such bias would likely be upwards. In other words, for men, we would expect to estimate larger impacts on educational attainment and labor market outcomes in the presence of this bias than in the absence of this bias. With this in mind, it is interesting that we find no impact on male labor market outcomes and that our education estimates for boys are comparable to those for girls, for whom there can be no private school bias. Second, a comparison of the characteristics of the initial sample (left panel of Appendix A Table 1b) and the sample that excludes non-respondents and private school enrollees

[15]Only 0.4 percent girls in our sample attended a private high school. For boys this percentage was 7.6, comparable to the national figure at the time.

(central panel) does not reveal sharp differences, other than the expected difference in the fraction male. The two far right panels of Appendix A Table 1b report separate descriptive statistics for individuals that attended elite and non-elite schools. As expected, these reveal clear differences in ability and socio-economic characteristics. The difference in average ability (roughly two standard deviations as measured by the total assignment score) is particularly striking.

3 Empirical Strategy

3.1 Motivation

To motivate the IV strategy that we use to identify elite school effects, consider the following model for outcomes of individual i in the event that she attends an elite school (Y_{1i}) or a non-elite school (Y_{0i}):

$$\begin{aligned} Y_{0i} &= E[Y_{0i}|A_i] + u_{0i} \equiv g_0(A_i) + u_{0i} \\ Y_{1i} &= E[Y_{1i}|A_i] + u_{1i} \equiv g_1(A_i) + u_{1i} \end{aligned}$$

where A_i is the assignment score with S points of support, such that $A \in \{a_0, a_1 ..., a_S\}$ and $a_s - a_{s-1} > 0$, and the error terms u_{0i} and u_{1i} are mean-independent of A_i hence any functions of A_i including $g_0(A_i)$ and $g_1(A_i)$. The model for observed outcomes can then be written:

$$Y_i = g_0(A_i) + D_i \tau(A_i = a_s) + \{D_i(u_{1i} - u_{0i}) + u_{0i}\}$$

where $\tau(A_i = a_s) = E[Y_{1i} - Y_{0i}|A_i = a_s] = g_1(A_i = a_s) - g_0(A_i = a_s)$ and D_i is a dummy variable taking the value one if individual i attends an elite school and zero otherwise.

To begin, assume treatment effects are constant, such that:[16]

$$Y_i = g_0(A_i) + D_i\tau + u_{0i}$$

Even if we knew the form of $g_0(.)$, least squares estimates of this equation would be biased: students assigned to elite schools may have unobserved characteristics that would be associated with better outcomes even if they attended non-elite schools (i.e., $Cov(D_i, u_{0i}) > 0$).

As a first step to motivating our IV strategy, note that $D_i = E[D_i|A_i] + v_i = P(D_i = 1|A_i) + v_i$, where v_i is mean-independent of all functions of A_i. Substituting into the last equation gives:

$$Y_i = g_0(A_i) + P(D_i = 1|A_i)\tau + \{u_{0i} + v_i\}$$

where both components of the error term are mean-independent of $g_0(A_i)$ and $P(D_i = 1|A_i)$. It follows that if we knew the assignment score polynomial $g_0(A_i)$ and the assignment probability $P(D_i = 1|A_i)$, we could identify τ via a regression Y on $g_0(A_i)$ and $P(D_i = 1|A_i)$. The estimated τ would be positive (negative) if average outcomes exhibited a sharp increase (decrease) in the borderline range in which the probability of assignment increases from zero to one. If average outcomes did not change through this borderline range, then the estimated τ would be small.

Since we do not know $P(D_i = 1|A_i)$, we cannot use this approach. We could, however, instrument D_i with $P(D_i \widehat{=} 1|A_i)$ - the predicted value of D_i given A_i. Because the relationship between school assigment and assignment scores is highly nonlinear (c.f., Figure 2), the instrumental variable $P(D_i \widehat{=} 1|A_i)$ would have predictive power for D_i conditional on any smooth (e.g., low-order polynomial) functional chosen for $g_0(A_i)$. Moreover, because the instrumental variable $P(D_i \widehat{=} 1|A_i)$ is a function of A_i, it would be mean-independent of the error term. As noted by Wooldridge (2002), in a discussion

[16] More generally, the assumption is that treatment effects cannot be predicted at age 11, such that there is no correlation between the treatment and the gain from treatment (sometimes referred to as "selectivity bias").

of this type of strategy, if the model for $P(D_i = 1|A_i)$ is specified correctly, and if u_{0i} is homoscedastic, then this procedure is the efficient IV estimator (Procedure 18.1, discussed on p. 623). Even if the model for $P(D_i = 1|A_i)$ is specified incorrectly (e.g., if a linear probability model is used), then the instrument would still be valid and estimates would still be consistent.

In Appendix C, we argue that if treatment effects are heterogeneous, then this IV estimator would likely approximate the average effects among borderline students. There are two steps in the argument, which closely follows the argument developed by Angrist et al. (1996). First, we show that under some additional assumptions, the IV estimator would identify a weighted average of score-specific local average treatment effects (LATEs): $E[Y_i(1)-Y_i(0)|D_i(a_s)-D_i(a_{s-1}) = 1, A_i = a_s]$.[17] These LATEs capture the average effect among a particular subset of students: those that achieved score a_s and were assigned to an elite school but who would not have been assigned with score $a_s - 1$; only LATEs for borderline scores receive positive weight. Second, we argue that this weighted average of score-specific LATEs would likely approximate the average effect among borderline students.

3.2 IV strategy

The strategy described above used a fitted probability as a single instrument. The strategy that we actually implement employs as instruments various functions of the assignment score. Exactly the same arguments justify the validity of these instruments (they have predictive power for D and are functions of

[17] This result is based on two sets of assumptions. One imposes restrictions on the assignment probability, which is assumed to be zero to the left of the borderline range, one to the right of the borderline range and increasing within the borderline range. The other imposes restrictions on the $\tau(A_i)$ function, which is assumed to be constant within the borderline range (but not necessarily outside of it). The first set of assumptions generates a first-stage relationship that is a close approximation to the first-stage relationship that we actually work with (see Figure 1). The second set of assumptions is harder to assess, although the results of Monte Carlo simulations (available on request) suggest that even when these assumptions are violated, IV estimates will identify something close to the effect for the typical borderline student.

A hence mean-independent of the error term), although the single-instrument case is simpler to motivate. The main advantage of the multiple-instrument strategy, and the reason we pursue it, is that when treatment effects are constant, we can use the resulting over-identification restrictions to test the model specification.[18] To maximize efficiency, we implement this strategy using GMM and use J-tests of the overidentification restrictions.

In practice, we use as instruments the four variables required to fit the basic pattern seen in Figure 2: dummy variables indicating borderline scores and above-borderline scores and the interactions of these dummy variables and the assignment score. The associated first stage estimates are reported in Table 1. Note that these instruments do a good job of predicting elite school assignment, even when we include flexible functions (e.g., fourth-order polynomials) of the assignment score (i..e, the function $g_0(A_i)$) and covariates. This is revealed by tests of the hypothesis that the excluded instruments explain none of the variation in elite school assignment, which are easily rejected; the associated F-statistics are well in excess of the thresholds for instrument relevance suggested by the literature (e.g., Stock and Yogo, 2005).

3.3 Robustness checks

As with regression discontinuity approaches, a successful application of this IV approach requires that we can capture $g_0(A_i)$, the underlying relationship between outcomes and assignment scores. As with "global polynomial" regression discontinuity estimators, we proxy for this relationship using low-order polynomials. To ensure that our estimates are not biased by this choice, we implement five robustness tests.

Four of these tests are analogous to those used in many regression discontinuity applications. First, we check that our estimates are insensitive to the inclusion of covariates. Since we have an extensive set of covariates (in addition to the assignment scores), this first test should be quite powerful. Second,

[18] As recently illustrated in Parente and Santos Silva (2012), the validity of overidentifying restrictions does not ensure the validity of instruments. Instead, these tests are better interpreted as checking that the various instruments identify the same parameters.

we conduct falsification tests of the "effect" of elite school attendance on predetermined outcomes such as years of post-compulsory education predicted by covariates. Third, we check that the model provides an adequate fit to data outside of the borderline range. This is possible because for data outside of the borderline range, there is no scope for selection on unobservables conditional on the assignment score (i.e., the probability of elite school assignment is close to either zero or one). This is not true inside the borderline range; hence in this range we would not expect the model to fit the data well. Fourth, we check that our estimates are robust to alternative polynomial specifications. The remaining robustness check is the over-identification test discussed in the previous subsection.

3.4 Connection to other approaches

Regression Discontinuity Design (RDD)

At first glance, the "first stage" relationship seen in Figure 2 might appear to feature the "sawtooth" pattern documented by Angrist and Lavy (1999) in one of the first papers to use regression discontinuity methods. Upon closer inspection, it is clear that while the relationship seen in Figure 2 is highly non-linear, it is essentially continous. This implies that standard regression discontinuity methods cannot be applied. There is however a conceptual connection between the two approaches. Specifically, the regression discontinuity method exploits the idea that if the underlying relationship between outcomes and the running variable is smooth, then a positive treatment effect will be revealed as a discontinuity in the relationship between the outcome and the running variable at the point at which the treatment "switches on". As noted above, our method exploits the idea that if the underlying relationship between outcomes and the running variable (i.e., assignment score) is smooth, then a positive treatment effect will be revealed as a sharp increase (but not a jump) in outcomes across the borderline score range.

Although we cannot use standard regression discontinuity methods, we report estimates obtained using a non-standard implementation of the regression discontinuity approach. Specifically, we use data from the right of the borderline range to proxy $g_1(A_i)$ and predict the outcome given elite school attendance and assignment score 550, and we use data from the left of the borderline range to proxy $g_0(A_i)$ and predict the outcome given non-elite school attendance and assignment score 550. The difference between these predicted values provides an estimate of the treatment effect at assignment score 550. The advantage of this approach is that we do not need to assume that treatment effects are independent of assignment scores within the borderline range.[19] The key disadvantage is that extrapolation may generate non-robust estimates (as discussed by Angrist and Rokkanen, 2012). A second problem is that even outside of the borderline range, the treatment probability is not always zero or one, which complicates the estimation of $g_1(A_i)$ and $g_0(A_i)$. This is the main reason why we prefer our IV approach, although it is reassuring to find that our IV approach and this RD approach generate similar results, at least for the main outcomes.[20] Note that to implement this approach, we follow standard practice in the regression discontinuity literature (e.g., Lee and Lemieux, 2010) and approximate these functions using linear regression models (i.e., equivalent to estimating nonparametric regressions with uniform kernels). Rather than pick a single bandwidth, we present estimates and confidence intervals for a wide range of bandwidths.[21]

Kinked Regression Discontinuity Design (KRD)

[19]We could have used this approach to calculate treatment effects at any assignment score within the borderline range. We chose the 550 score because it is the midpoint of the range. This makes it a natural estimate to compare with the IV estimates (which we argued identified the average effect among borderline students). It also means that we extrapolate over the same distance from the left and right of the borderline range, which spans [540, 560]. If forecast errors are a convex function of the score range being extrapolated over, the sum of the forecast errors will be minimized at this score.

[20]If we wanted to push the RD approach further, we could generate estimates for every point in the borderline range (i.e., 540, 541,..559) and then generate a weighted average of these. Since we prefer the IV approach, we chose not to pursue this strategy.

[21]We conducted cross-validation analyses designed to select the optimal bandwidth (e.g., Imbens and Lemieux, 2008), adapted to account for the fact that we must extrapolate by 10 points. These generally pointed towards larger bandwidths (available upon request).

Because the first-stage relationship seen in Figure 2 is kinked, we might have considered the KRD approaches developed by Card et al. (2012) and Dong (2013). Those allows for "fuzzy kinks", as observed in our case, and the Card et al. (2012) application features two fuzzy kinks, as does ours. The main reason for not using the KRD approach is that in our case, the two kinks are relatively close. This has two implications. First, even with large numbers of observations it would be difficult to determine the shape of the relevant relationships around the kinks (i.e, the relationship between outcomes and assignment scores in the borderline range). Second, the dataset that we use is relatively small, such that each kink sample would likely feature fewer than 100 observations (the kink samples in Card et al. (2012) each feature almost 200,000 observations). Unfortunately, this means that the KRD approach is not feasible.

4 Long-term effects of elite school attendance

We now use the IV strategy described above to estimate the causal effects of attending an elite school on long term outcomes such as completed education, income, earnings, marriage and fertility. These estimates are obtained separately for men and women. We then pool men and women and report some of the main findings separately by low and high socio-economic status.

Educational attainment

Table 2 reports estimates of elite school effects on various measures of educational attainment. Our main measure, reported in Panel A, is the number of completed years of full-time education beyond the compulsory school leaving age (i.e., age left school plus years of full-time higher education less the compulsory schooling age facing these cohorts, i.e., 15). In using this measure we follow the labor economics literature and assume that there are constant returns to additional years of completed education (Card, 2001). We also consider specific qualification levels, and analyze separately the probability of achieving no qualifications (Panel B), CSEs and O-level equivalent qualifications obtained at age 15 (Panel C and D, respectively), A-level equivalent qualifications obtained at age 17 (Panel E), post-17 certificates and diplomas (Panel F), and degree or higher level qualifications (Panel G). For each of these outcomes we report least squares estimates (columns 1-3 and 5-8 for men and women, respectively) and GMM instrumental variable estimates (columns 4-5 and 9-10). In columns 1 and 6 our specifications include only a dummy for attending an elite school; all other specifications include a third-order polynomial in the assignment score; in columns 3, 5, 8 and 10 we also include a set of covariates that take into account individual differences in demographic characteristics, socio-economic characteristics, and previous attainment.

Panel A of Table 2 shows that, on average, relative to non-elite school male students, elite-school male students completed 3.2 additional years of full-time education (column 1). Once we control for a flexible function of the assignment score this effect reduces to 1.5 years; introducing covariates brings it down to 1.4 years. The GMM estimates are smaller than this, although statistically not different, and in the model including covariates (column 5) the estimated effect is 1.20 years. This is a very large effect. For example, if we consider that borderline male students that attended a non-elite school completed an average of 2 years of post-compulsory schooling (the "control mean" of 1.992), the estimate implies that elite school attendance increased years of post-compulsory education by 60 percent.[22]

The results for women are qualitatively similar. As reported in Table 2, on average, women that attended elite schools completed around 2.9 additional years of full-time education (column 6). After controlling for assignment scores and other covariates, the estimated effect is 0.70 years. The GMM estimates are consistent with these numbers, ranging from 0.83 (column 10) to 0.93 (column 9) years. Since the control mean for women is 2.3 years, the implied

[22]This is obviously not a control mean in the sense of a randomized trial, since we would expect some "negative selection" into the non-elite schools among the borderline students. Nevertheless, it seems like a reasonable counterfactual for the borderline students for whom we estimate elite school effects.

effect size is about 36 percent.

Our estimates pass four robustness checks analogous to those commonly employed in regression discontinuity analyses. First, our estimates are robust to the inclusion of covariates (see columns 4 and 5 and columns 9 and 10 of Table 2). Second, as seen in Figure 3, actual outcomes correspond closely to fitted values for students with assignment scores outside of the borderline range (i.e., less than 540 or greater than 560).[23] Third, our analysis passes a "falsification test". Specifically, our estimates suggest that elite school attendance has no "effect" on years of post-compulsory schooling predicted using the extensive set of background characteristics available in our data (see columns 3 and 6 of Appendix A Table 2, Panel A).[24] Fourth, our estimates are similar to those derived from models that use different polynomial specifications (see Appendix A Table 3, Panel A).[25]

Our estimates also pass two robustness checks specific to our GMM procedure. First, our estimates pass the over-identification test discussed above (see the last row in each panel). Second, our estimates are similar to those based on the regression discontinuity strategy described above. Recall that this uses local linear regression methods to extrapolate across the borderline range and estimate the effect of elite school attendance at assignment score 550. Appendix A Figure 3 reports these estimates for various bandwidths. For nearly all bandwidths, the confidence intervals contain the analogous GMM estimates (the one obtained in models without control variables). For larger bandwidths, the RD estimates are very close to the GMM estimate. This is reassuring, since the cross-validation procedures used to choose bandwidths (see footnote 18) suggest that larger bandwidths ought to be preferred.

[23]As noted above, for those with borderline scores, we expect that treatment assignment will be influenced by unobservables, and so we would not expect the model to fit the data well in this range.

[24]By contrast, OLS estimates suggest an impact of elite school attendance on the predicted outcome. This highlights the importance of dealing with omitted variables biase in the OLS estimates.

[25]The second-order polynomial estimate is generated using the subset of students with assignment scores of at least 460. It is clear from Figure 3 that a second-order polynomial would be too inflexible to fit to the entire range of data.

The remaining panels of Table 2 report education impacts as measured by qualifications earned. For men, note that while a large percentage have at least some qualifications (the control mean for those with no qualifications (Panel B) is 7.2 percent), the control mean for university degree (panel G) is low (13.6 percent). This suggests that while most borderline students in non-elite schools completed some qualifications, few proceeded to the end of the standard academic track. The main impacts of elite school attendance are to increase the probability of achieving A-level qualifications (by about 13 percentage points) and degree receipt (by 15 percentage points). The latter effect is more than double the control group mean. Since the academic track entails at least six years of full-time study (2 years for achieving A-level qualifications plus 4 years of university study), we would expect the effect on degree level qualifications alone to account for roughly 0.9 of the additional years of completed full-time education.

For women, elite school attendance is shown to reduce the probability of achieving lower-level qualifications such as CSEs (by about 20 percentage points) and O levels (by about 0.8 percentage points) and correspondingly increase the probability of earning higher-level qualifications, including A-levels (Panel E), HNC and teaching certificates (Panel F). The largest effects are found for A-level qualifications or equivalent. The effect size - more than 23 percentage points - is about 60 percent of the control group mean. Although the A-level estimates fail the overidentification tests and are somewhat sensitive to the order of the polynomial, they still pass all other robustness checks, including the introduction of covariates and falsification tests.

With our discussion of the relevant institutions in mind, these large effects on educational attainment are perhaps not surprising. As we noted, the path to a university degree was longer and harder for non-elite school students. For example, the non-elite schools were unlikely to offer many university-appropriate courses and transfer to elite schools was uncommon. In addition, since few non-elite school students attended university, default behavior and peer effects may have pushed students away from this path. We suspect that these institutional barriers facing students that attended non-elite schools in the

1960s hold the key to understanding why we estimate larger education effects than those found in other studies (e.g., Dobbie and Fryer, 2011; Dustmann et al., 2012).

Labor market outcomes

Table 3 reports estimates of elite school effects on gross annual income (Panel A), employment (Panel B), and imputed gross hourly wages (Panel C). Gross annual income and employment are measured at the time of the 2001 survey, when respondents were aged between 46 and 51. The income measure includes "personal current gross income from all sources", including interest from dividends and benefits. We impute gross hourly wages using occupation-specific means of hourly gross wages from the *New Earnings Survey* (NES).[26]

The consensus among labor economists is that an additional year of education increases hourly wages by around 8 percent (Card, 2001). As such, we might expect elite school attendance to generate significant impacts on income and wages. In fact, our estimates suggest that for men, elite school attendance had no significant effects on annual income, hourly wages or the probability of employment.[27] Remarkably, this is true for the least squares estimates (once assignment scores are controlled for) as well as the GMM estimates. In the next section we discuss what might account for this puzzling finding. For now, we note that the GMM estimates pass all of our robustness checks. In particular, models with covariates generate similar estimates (columns 4 and 5), the model appears to fit the data well (Figure 4), falsification checks do not reveal

[26]In order to compute occupation-specific earnings we take the period between 1997 and 2001 and restrict the sample to individuals working in Scotland and aged 45-55. The imputation of the earnings variables is based on 2-digit SOC 1990 classification. We would like to thank Annarosa Pesole for her help with the NES data.

[27]Income is recorded in 8 bands and a large group of individuals, especially men (about 20 percent), fall in the top interval. We produced several estimates of the impact of elite school attendance on the probability of being in the top interval, but none suggested any impacts. Similarly, we estimated elite school effects using several methods that account for the banded nature of the variable, including interval regression (Stewart, 1983), all of which generated results similar to those reported in the current analysis.

any impact of elite school attendance on measures of income predicted using control variables (Appendix A Table 2, Panel D), over-identification tests are passed, and the RD estimates tell a remarkably similar story across a wide range of bandwidths (Appendix A Figure 4). Although the estimates appear to be somewhat sensitive to the order of the polynomial in assignment scores (Appendix A Table 3), they remain statistically indistinguishable from zero.

By contrast, as reported in Table 3, we find that elite school attendance increased women's annual income by around 17-19 percent (Panel A). Although the estimate is not very precise, and is somewhat sensitive to the order of the polynomial, once again it passes most of our checks. Only the RD estimates are suggestive of potentially smaller effects (at larger bandwidths), but even in this case the effects are still above 10 percent. An obvious question is whether the effect for women is driven by increased hourly wages, increased labor supply or some combination of the two. Our results suggest that both effects could be at work. On the one hand, we estimate elite school effects on hourly wages of around 10 percent (Panel C), which is consistent with the effects typically associated with a year of full-time education. On the other hand, this alone cannot account for the almost 20 percent increase in annual income, so labor supply impacts seem likely. Although we find no elite school effects on whether women worked at all (Panel B), this obviously does not prelude effects on hours worked. Unfortunately, we do not have data on hours worked so we cannot examine this possibility.

Fertility and marriage

For men, the estimates reported in Table 4 suggest that there were no effects of elite school attendance on fertility and marriage outcomes.[28] Since we found that elite school attendance had no impact on labor market outcomes, this is

[28]One possible caveat to this summary is the suggestion of a positive elite school effect on the number of children (Panel B), although the estimate appears to be sensitive to the inclusion of covariates (compare columns 2 and 3, as well as 4 and 5) and is not statistically significant when these are excluded.

perhaps not surprising. For women, the estimates reported in Table 4 suggest that elite school attendance decreased the probability of having any children and decreased total fertility, although they do not reveal any significant effects on the timing of fertility or the probability of being married or getting divorced. Since we found that elite school attendance increased women's hourly wages and annual income, an effect on fertility is less surprising. As noted by Becker (1960) and Willis (1973), higher earnings power increases the opportunity costs of having children and will thereby decrease fertility.

Our estimates imply larger fertility effects of schooling than those found in the previous literature. For example, we find that elite school attendance (which translates into just less than one additional year of education) decreased the probability of having a child by 9 percentage points (about 10 percent of the control mean) and decreased the number of children by 0.47 (about 25 percent of the control mean).[29] Previous estimates of the causal effects of a year of education on completed fertility, all of which exploit changes to compulsory schooling laws, span positive effects of up to 0.2 children (Fort et al., 2011), insignificant effects (Black et al., 2008; Monstad et al., 2008; Silles, 2011), and negative effects of up to 0.3 children (Leon, 2004; Cygan-Rehm and Maeder, 2012).

The obvious starting point for trying to explain the difference between these findings is the difference between the "experiments" underlying them. Our analysis mimics that of an experiment in which high-ability girls are randomly assigned to an elite school environment and additional further academic education. In the experiments implicit in the related literature, low-ability girls are exposed to an additional period of compulsory education with broadly the same peers. With this difference in mind, there are several possible explanations for the larger effects that we find. For example, fertility may be influenced by women's perceptions of their role in society, which may be influenced by high-ability and ambitious high school peers. This is consistent with Goldin's

[29] Once again, our GMM estimates pass all of our robustness checks, including falsification checks (Appendix A Table 2), and are consistent with RD estimates (Appendix A Figure 5).

(2006) account of the key role of education in promoting a broader planning horizon and a new sense of identity for women of this generation.

Social Class

To this point, we have reported separate estimates by sex. This seems sensible, since the assignment rule was applied separately to girls and boys and, in some cases, they attended different elite schools. But it is also interesting to consider whether estimates differ by socio-economic status (SES), which would be the case if elite schools were especially helpful in taking low-SES students out of home and peer environments not conducive to educational success. To that end, we pool men and women and split our sample into low- and high-SES subsamples.[30]

The estimates, reported in Appendix A Table 4, reveal a surprising fact. Namely, that for education outcomes, the low-SES and high-SES control means are quite similar. This suggests that among borderline students that attended non-elite schools, the socio-economic gap in outcomes was small. With this in mind, it is perhaps not surprising that our estimates of elite school effects by SES are also quite similar. Since control means and effect sizes for other outcomes show no obvious or consistent differences, we conclude that among our sample elite school attendance did not have markedly different impacts on low- and high-SES individuals. While this may seem surprising at first glance - one might expect elite school effects to be larger for low-SES students (as found by Clark, 2010) - it is less surprising in light of the absence of any socio-economic gaps in the control means.

[30]Individuals are categorized as low SES if the father was in a semi-skilled or unskilled manual occupation or was not working; high socio-economic status is assigned to students whose father was in a skilled manual, non-manual, professional or managerial occupation.

5 Discussion and Interpretation

There is an obvious explanation for the women's results: elite school attendance makes further education more attractive (by reducing costs or increasing returns), such that elite school attendance increases completed education and thereby labor market productivity, wages and incomes. Increased wages also increase the opportunity costs of having children, such that elite school attendance reduces completed fertility.

The men's results are harder to explain, particularly our finding that elite school attendance increased completed education by more than one year but had no impact on incomes or wages. We speculate that the most likely explanation is that this generation of men enjoyed many vocational training options, especially trade apprenticeships, and that positive elite school effects on further education were offset by negative elite school effects on vocational training. Although we do not have the vocational training or apprenticeship data necessary to provide a direct test of this hypothesis, several facts are consistent with this story.

First, as seen in Appendix B Table 3, other data (Labour Force Survey (LFS) data) suggest that a large fraction of Scottish-born men of this generation completed an apprenticeship. The overall percentage is close to 40 percent and the percentage among men with some additional qualifications is even higher. Since our sample control means suggest that borderline men that attended non-elite schools typically had additional qualifications, it seems likely that at least one half of them completed a trade apprenticeship.

Second, simple economic reasoning suggests that elite school attendance likely decreased the probability of completing an apprenticeship. We set out a simple model to this effect in Appendix D; this extends the Card and Krueger (1996) model of school quality to include vocational training. We show that under some assumptions, elite school attendance could decrease the quantity of vocational training completed and could thereby decrease wages. The first key assumption is that elite school attendance increases the returns to academic education, decreases the costs of academic education and increases the cost of vocational training. The second key assumption is that while the returns to academic education are higher than the returns to vocational education, the relative return to vocational training is higher than the relative return to academic education. Both sets of assumptions seem reasonable, the second one because it generates a natural partition of students into those that leave school without further education or training, those that pursue vocational training and those that pursue academic education. This theoretical reasoning is supported by estimates of the elite school impact on years of part-time education completed (not reported but available on request). These suggest that elite school attendance decreased part-time years completed by roughly 0.7. Since vocational training typically includes a part-time education component, these estimates are consistent with elite school attendance reducing vocational training.

Third and most compelling, neither of these facts applies to women, for whom we did find positive and significant elite school effects on income and wages. The statistics reported in Appendix B Table 3 suggest that far fewer women completed apprenticeships, and for women we find little elite school impact on years of part-time education completed.

The fourth fact is derived from a testable prediction of the model that we present in Appendix D, namely, that the presence of vocational training will lower the returns to completed education. The intuition is that in addition to the usual (positive) "ability bias" component of the measured return (i.e., a positive correlation between academic education and ability), there will be an additional (negative) "vocational training bias" component (i.e., a negative correlation between academic education and vocational training). As seen in Appendix B Table 4 column 1, we do indeed find measured returns to be much larger for women than for men.

A fifth fact consistent with this explanation is that elite school attendance appears to increase the probability that men are found in occupations typically associated with academic education. In particular, Appendix A Table 5 shows that elite school attendance increased the probability that men worked as "professionals or associate professionals" by around 6 percentage points (or about 25 percent of the control mean), with an equivalent decrease in the probability that they worked in "managerial" occupations. These estimates are especially interesting when seen through the lens of Appendix A Table 6. Consistent with our education results, this shows that the fraction of degree holders in professional occupations is twice as large as the fraction in managerial occupations. Consistent with our income and wage results, this shows that on average, men working in professional occupations earn considerably less than men working in managerial occupations.

The comparison with women is again constructive. In Appendix A Table 5 we see that elite school attendance increases the probability that women work in "professional or associate professional" occupations, with an equivalent decrease in the probability that they work in "clerical and secretarial" occupations. Consistent with our education results, Appendix A Table 6 shows that the fraction of degree holders in professional occupations is roughly ten times as large as the fraction in clerical and secretarial occupations. Consistent with our income and wage results, the Table also shows that, on average, women working in professional occupations earn almost twice as much as women working in clerical and secretarial occupations.

There are several alternative explanations for our finding that elite school attendance did not increase income and wages for men. First, the result could be an artefact of the sample studied. For example, one can hypothesize that the wage distribution in Aberdeen has been compressed by the oil industry that emerged in the 1970s. To check that, we compared the returns to education estimated using the Aberdeen data with the returns estimated using LFS data (Appendix B Table 4). When we measure education using years of post-compulsory schooling, the estimates are remarkably similar: 0.072 versus 0.075. They are also fairly close when we disaggregate into several education categories. Importantly, there is no evidence to suggest that Aberdeen-UK differences are larger among men than among women. For this reason, we doubt this can explain the results.[31]

[31] The most direct way of assessing this hypothesis would be to measure what fraction of our sample work in the oil industry. Unfortunately, respondents were not asked about industry.

Second, one might wonder whether the returns are low because they are estimated for borderline men; perhaps the returns to elite school attendance are higher among higher-scoring men that were in a better position to benefit from the elite school experience. Against that, it is not obvious that elite school effects should be smallest for borderline students, and Angrist and Rokkanen (2012) find no evidence to suggest they are smaller for borderline students than for higher-scoring students (albeit in a different context). More persuasively in our view, it is hard to reconcile this explanation with our findings for women, which are also identified off the borderline group.

One could reconcile the borderline explanation and the women's results by noting that some elite schools were single-sex and by speculating that a single-sex experience benefited girls more than boys. As an indirect test of this hypothesis, we examined whether the returns to single-sex schools are larger for girls than boys. Least squares estimates based on models that are restricted to single-sex elite school students and that include the full set of covariates provide no support for this hypothesis.

In summary, while we cannot provide a decisive test of the vocational training explanation (without vocational training data), we find it to be plausible and we think that it fits the key facts. Since these facts undermine the most obvious alternative explanations, we suspect that the vocational training hypothesis is the most plausible explanation for the results that we find

Conclusion

What is the causal effect of being assigned to an elite secondary school in a selective school system? The balance of the existing evidence suggests small impacts on short- and medium-run outcomes such as high school test scores and college enrollment and attainment. In this paper we estimate the effects of elite school attendance on long-run outcomes such as completed education, income, wages, occupational attainment, marriage and fertility. We find that, on average, elite school attendance caused both men and women to complete almost one additional year of full-time education. For women, we find that elite school attendance also led to large increases in income and large decreases in fertility. For men, we find that elite school attendance had no impact on labor market outcomes.

These results support a claim made in the 1950s and 1960s by opponents of selective schooling, that important long-run outcomes could depend, via the elite school assignment decision, on how well a student performed on a single test taken at age eleven.[32] Proponents of selective schooling did not claim that elite school assignment had no impact, but rather that the assignment mechanism was generally reliable, and that selective schooling helped both high-ability students stretched by elite schooling and lower-ability students properly catered to by non-elite schooling. The debate surrounding selective versus non-selective schooling is interesting and important, but our study does not address it.[33] Instead, our study has tried to assess whether elite school attendance improved outcomes for borderline students, the only students for whom we can credibly identify effects.

In our view, it is difficult to explain our results without invoking various features of the relevant education and labor market institutions. For example, we argued that our education results likely reflect the barriers that faced non-elite school students wishing to pursue further academic education, while we speculated that our income results for men might reflect the choice these men faced between pursuing academic education and vocational training. The role that these institutions might have played provides an obvious explanation for why some of our estimates differ from those found in the previous literature (e.g., Dobbie and Fryer, 2011). These institutions also imply that some our results may be specific to the time period studied. For example, since the apprenticeship system became much less important after the mid-1970s, it is possible that labor market impacts may have been larger for men educated in

[32]For the students in our sample, assignment depended on four tests, teacher assessments and the Head's assessment. As discussed by Clark (2010), most districts in England used only two tests.

[33]See Galindo-Rueda and Vignoles (2004) for an account of the debate surrounding elite schools in the UK during the 1960s.

the late mid-1970s and beyond.

Our final point is that policy-makers might be advised to keep these institutions in mind when designing policies relating to school resources and organization, including policies relating to elite schools. For example, in the contemporary US context, there is compelling evidence to suggest that college outcomes are affected by whether or not students have access to Advanced Placement (AP) courses (Klopfenstein, 2004; Jackson, 2010) and SAT-taking opportunities (Bulman, 2013; Goodman, 2012) in high school. If non-elite school students in district A can take a wide range of AP courses and must sit the SAT by default, while non-elite school students in district B can take only a few AP courses and must travel to another high school to sit the SAT, we might expect elite school attendance to have a smaller impact in district A than district B. It follows that a district that wishes to expand elite schools might also want to ensure that these related policies and institutions are favorable to non-elite school students. To take an example relevant to the UK context, one effect of the recent increase in the compulsory schooling age (to 18) might be to reduce the educational advantages enjoyed by students assigned to elite schools in the few areas that still operate a selective system.

References

Angrist, J., Imbens, G. and D. Rubin (1996), Identification of Causal Effects Using Instrumental Variables, Journal of the American Statistical Association, 91(434): 444-455..

Angrist, J., and V. Lavy (1999), Using Maimonides‘ Rule to Estimate the Effect of Class Size on Scholastic Achievement, Quarterly Journal of Economics, 114(2): 533-575..

Angrist, J., Pathak, P. and C. Walters (2011), Explaining Charter School Effectiveness, NBER Working Paper No. 17332.

Angrist, J., and M. Rokkanen (2012), Wanna Get Away? RD Identification Away from the Cutoff, NBER Working Paper No. 18662.

Batty, G., Morton, S., Campbell, D., Clark, H., Smith, G., Hall, M., Macintyre, S., and D. Leon (2004), The Aberdeen Children of the 1950s Cohort Study: Background, Methods, and Follow-up Information on a New Resource for the Study of Life-Course and Intergenerational Effects on Health, Paediatric and Perinatal Epidemiology, 18: 221-239.

Becker, G. (1960), An Economic Analysis of Fertility, in Demographic and Economic change in Developed Countries, Universities-National Bureau of Economic Research Conference Series 11. NBER: Princeton, NJ, 209-231.

Black, S., Devereux, P. and K. Salvanes (2008), Staying in the Classroom and out of the Maternity Ward? The Effect of Compulsory Schooling Laws on Teenage Birth, Economic Journal, 118(530): 1025-1054.

Bulman, G. (2013), The Effect of Access to College Asseessments on Enrollment and Attainment, unpublished paper, Stanford University.

Card, D. (2001), Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems, Econometrica, 69: 1127-60.

Card, D. and A. Krueger (1996), Labour Market Effects of School Quality: Theory and Evidence, NBER Working Paper No. 5450.

Card, D., Lee, D., Pei, Zhuan, and A. Weber (2012), Nonlinear Policy Rules and the Identification and Estimation of Causal Effects in a Generalized Regression Kink Design, NBER Working Paper No. 18564.

Clark, D. (2010), Selective Schools and Academic Achievement, the B.E. Journal of Economic Analysis and Policy, 10(1): 1935-1682.

Cullen, J.B., Jacob, B. and S. Levitt (2006), The Effect of School Choice on Participants: Evidence from Randomized Lotteries, Econometrica, 74(5): 1191-1230.

Cygan-Rehm, C. and M. Maeder (2012), The Effect of Education on Fertility: Evidence from a Compulsory Schooling Reform, BGPE Discussion Paper No. 121.

Deming, D. (2010), Better Schools, Less Crime?, Quarterly Journal of Economics, 126(4): 2063-2115.

Dobbie, W. and R. Fryer (2011), Exam High Schools and Academic Achievement: Evidence from New York City, NBER Working Paper No. 17286.

Dong, Y. (2013), Regression Discontinuity without the Discontinuity, unpublished paper, University of California, Irvine.

Duflo, E., Dupas, P. and M. Kremer (2011), Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya, American Economic Review, 101(5): 1739-74.

Dustmann, C., Puhani, P. and U. Schönberg (2012), The Long-term Effects of School Quality on Labor Market Outcomes and Educational Attainment, CReAM Discussion Paper Series 1208, Department of Economics, University College London.

Findlay, I. (1973), Education in Scotland, World Series in Education. David and Charles: Newton Abbot.

Fort, M., Schneeweis, N. and R. Winter-Ebmer (2011), More Schooling, More Children: Compulsory Schooling Reforms and Fertility in Europe, CEPR Discussion Paper 8609.

Galindo-Rueda, F. and A. Vignoles (2004), The Heterogeneous Effect of Selection in Secondary Schools: Understanding the Changing Role of Ability, IZA Discussion Paper n. 1245.

Garces, E. and J. Currie, J. and D. Thomas (2002), Longer-term effects of Head Start, American Economic Review, 92(4): 999-1012.

Goldin, C. (2006), The quiet revolution that transformed women‘s employment, education, and family, American Economic Review, 96: 1-21.

Goodman, S. (2012), Learning from the Test: Raising Selective College Enrollment by Providing Information, unpublished paper, Columbia University.

Guyon, N., Maurin, E. and S. McNally (2012), The effect of tracking students by ability into different schools: a natural experiment, Journal of Human Resources, 47(3): 684-721.

Illsley, R. and F. Wilson (1981), Longitudinal studies in Aberdeen, Scotland. C. The Aberdeen child development survey, in Mednick S., Baert A., Bachmann B. (eds.) Prospective longitudinal research. An empirical basis for the primary prevention of psychosocial disorders. Oxford: Oxford Univerity Press.

Illsley, R. (2002), A City's Schools: from Inequality of Input to Inequality of Outcome, Oxford Review of Education, 28: 427-445.

Imbens, G. W. and J. D. Angrist (1994), Identification and Estimation of Local Average Treatment Effects, Econometrica, Econometric Society, 62(2):

467-75.

Imbens, G. and T. Lemieux (2008), Regression Discontinuity Design: A Guide to Practice, Journal of Econometrics, 142: 615-635.

Jackson, C. K. (2010), Do students benefit from attending better schools? Evidence from rule-based student assignments in Trinidad and Tobago, Economic Journal, 120(549): 1399-1429.

Klopfenstein, K. (2004), The Advanced Placement Expansion of the 1990s: How Did Traditionally Underserved Students Fare?, Education Policy Analysis Archives, 12(68).

Krueger, A. B. and D. Whitmore (2001), The effect of attending a small class in the early grades on college test taking and middle school test results: Evidence from Project STAR, Economic Journal, 111(468): 1-28.

Lee, D. and T. Lemieux (2010), Regression Discontinuity Designs in Economics, Journal of Economic Literature, American Economic Association, 48(2): 281-355.

Leon, A. (2004), The Effect of Education on Fertility: Evidence From Compulsory Schooling Laws, unpublished manuscript.

Monstad, K., Propper, C. and K. Salvanes (2008), Education and Fertility: Evidence from a Natural Experiment, Scandinavian Journal of Economics, 110(4): 827-85.

Parente, P. and J. Santos Silva, (2012), A Cautionary Note on Tests of Overidentifying Restrictions, Economics Letters, 115: 314-317.

Pop-Eleches, C. and M. Urquiola, (2013), Going to a Better Schools: Effects and Behavioral Responses, American Economic Review, 103(4): 1289-1324.

Silles, M. (2011), The Effect of Schooling on Teenage Childbearing: Evidence using Changes in Compulsory Education Laws, Journal of Population Economics, 24(2): 761-777

Steedman, H., Gospel, H. and P. Ryan (1998), Apprenticeship: A Strategy for Growth, A Special Report published by the Centre for Economic Performance, London School of Economics and Political Science.

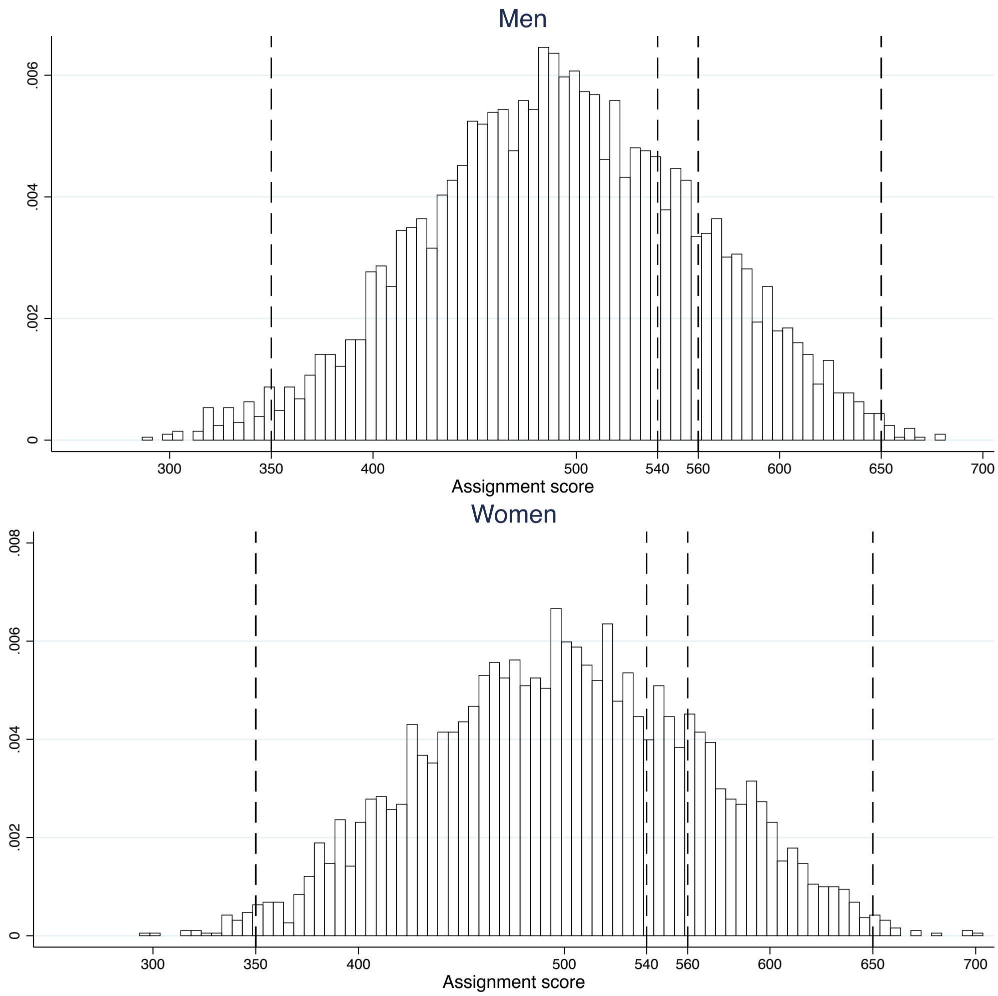
Stewart, M. (1983), On Least-Squares Estimation when the Dependent Variable is Grouped, Review of Economic Studies, 50: 737-753.

Stock, J. and M. Yogo (2005), Testing for Weak Instruments in Linear IV Regression, ch. 5 in J.H. Stock and D.W.K. Andrews (eds.), Identification and Inference for Econometric Models: Essays in Honor of Thomas J. Rothenberg, Cambridge University Press.

Willis, R. (1973), A New Approach to the Economic Theory of Fertility Behavior, Journal of Political Economy, 81: 514-64.

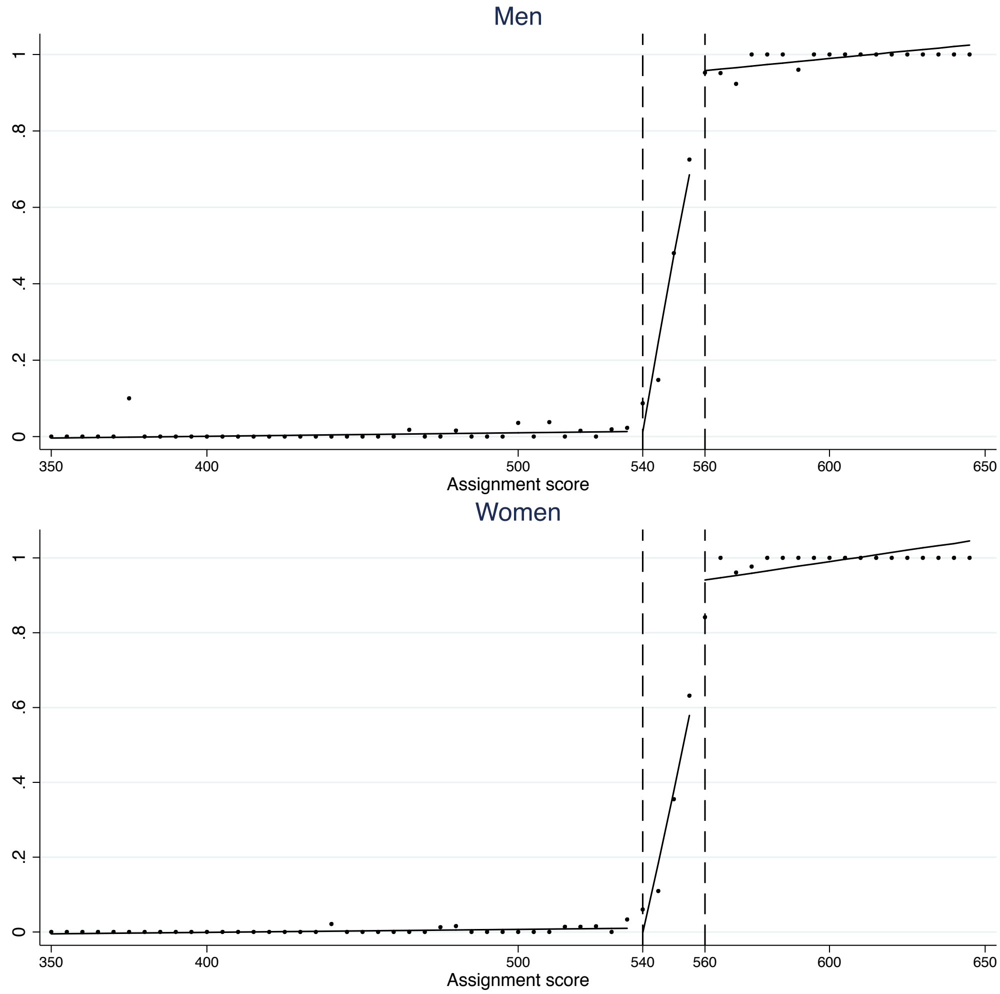
Wooldridge, J. (2002), Econometric Analysis of Cross Section and Panel Data. The MIT Press: Cambridge, Massachusetts.

Figure 1: Distribution of assignment scores



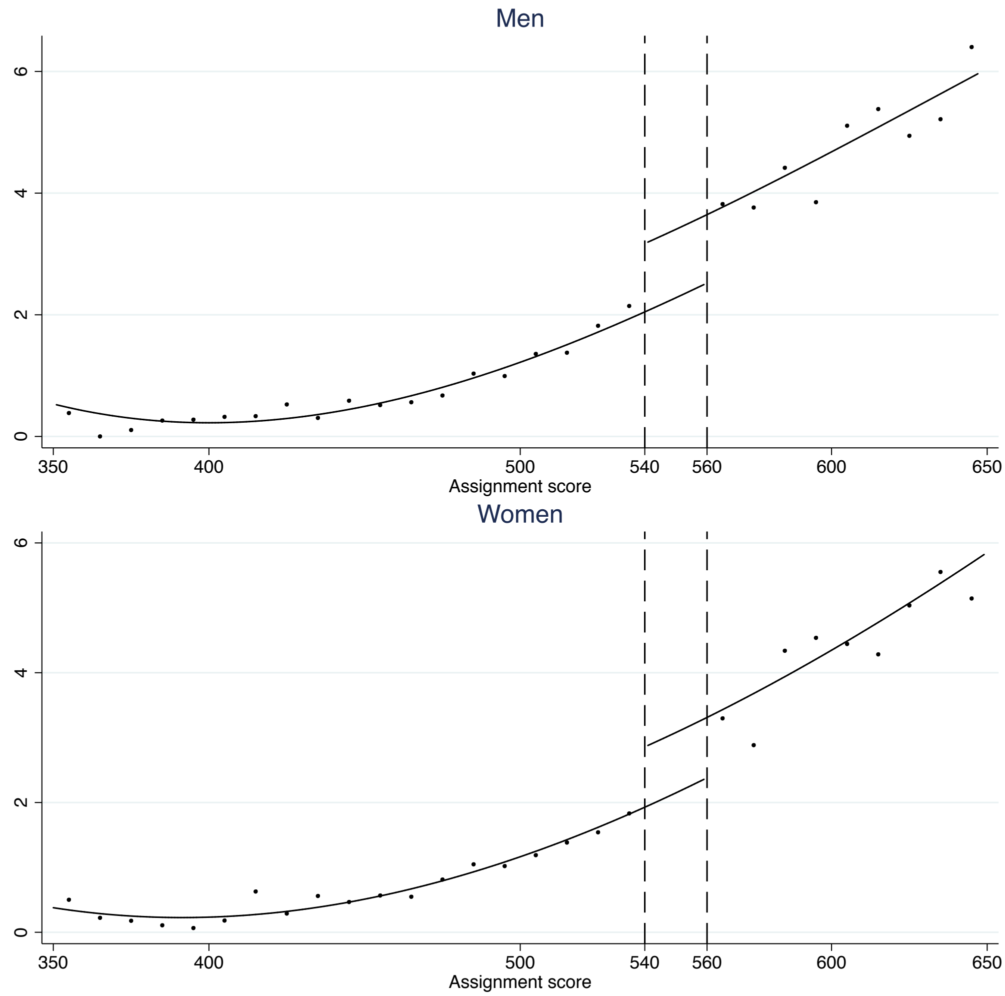
Note: Sample includes all children in grades 4-7 in 1962 with non-missing assignment scores. Each bar is drawn over an interval defined by 5 values of the score. Vertical lines plotted at the following score values: [350, 354], [540, 544], [560, 564], [650, 654].

Figure 2: Assignment score and the probability of attending an elite school



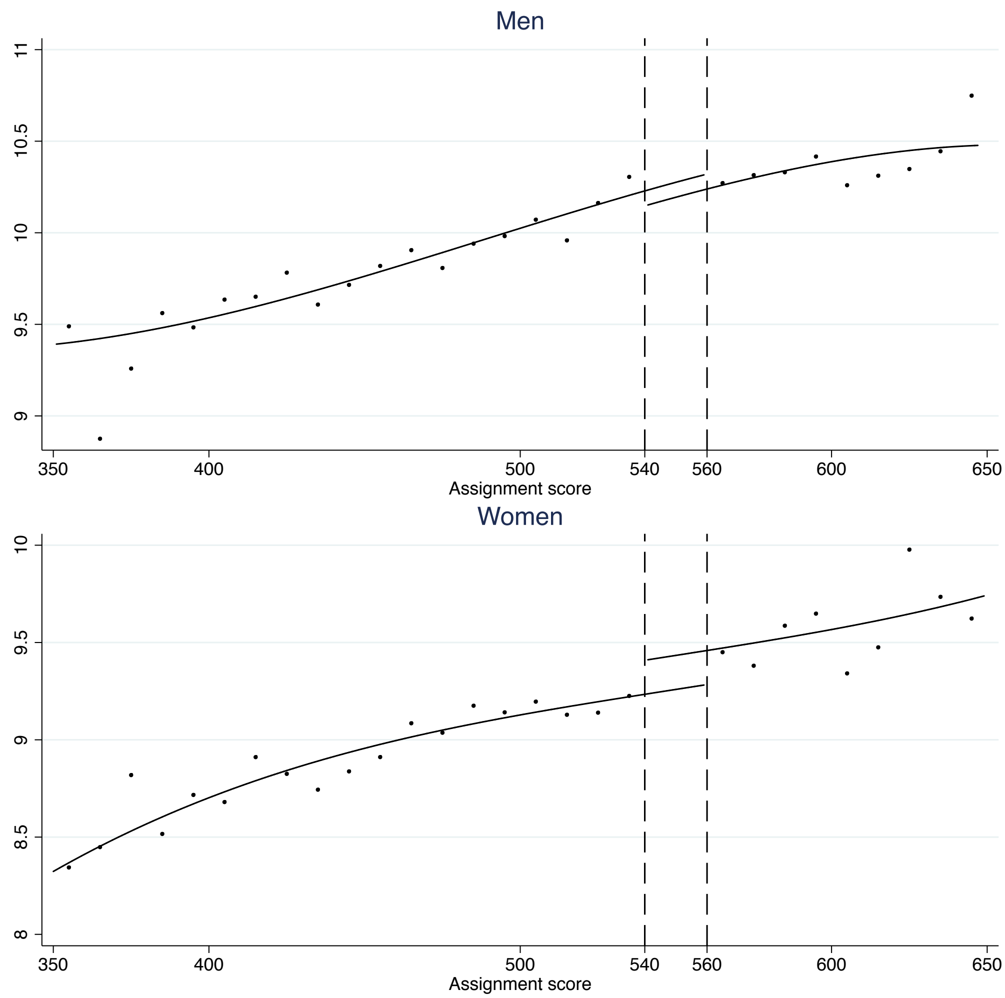
Note: Sample includes all men and women in grades 4-7 in 1962 who have non-missing assignment scores, responded to the 2001 survey and attended a non-private high school. Each circle represents the fraction of students in each cell that attended an elite school. Cells are defined over 5 assignment score values. The solid line represents the probability of elite school assignment predicted using the model described in the text.

Figure 3: Assignment score and years of post-compulsory schooling



Notes: Sample as in Figure 2. The solid lines are predicted outcomes from the model described in the text (estimated by GMM). The dots are outcome means corresponding to 10-score intervals. The first interval is [350, 360) and the last interval is [640, 650). Since the graph is designed to enable an assessment of fit in score ranges in which there is limited scope for selection on unobservables (i.e., outside of the marginal area [540,560)), the outcome means in this interval are not presented on the graph.

Figure 4: Assignment score and log annual income



Notes: See the notes to Figure 3. See text for a description of the construction of the hourly wage measure

Table 1: Elite school assignment

	MEN				
	(1)	(2)	(3)	(4)	(5)
Dummy [540-559]	0.822	0.810	0.826	0.821	0.823
	(0.054)	(0.055)	(0.056)	(0.056)	(0.058)
Dummy [540-559] * Score	0.045	0.045	0.045	0.045	0.045
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Dummy [560-649]	0.941	0.930	0.939	0.930	0.941
	(0.014)	(0.017)	(0.016)	(0.021)	(0.021)
Dummy [560-649] * Score	0.001	0.000	0.002	0.002	0.002
	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)
Score	Y	Y	Y	Y	Y
Score^2		Y	Y	Y	Y
Score^3			Y	Y	Y
Score^4				Y	Y
Xs					Y
Observations	2,072	2,072	2,072	2,072	2,072
R^2	0.852	0.852	0.852	0.852	0.858
F-test	2296	798.8	869.7	555.1	531.9
	WOMEN				
	(1)	(2)	(3)	(4)	(5)
Dummy [540-559]	0.678	0.677	0.709	0.684	0.681
	(0.056)	(0.057)	(0.059)	(0.059)	(0.058)
Dummy [540-559] * Score	0.038	0.038	0.039	0.039	0.038
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Dummy [560-649]	0.927	0.926	0.941	0.903	0.899
	(0.019)	(0.019)	(0.016)	(0.023)	(0.023)
Dummy [560-649] * Score	0.001	0.001	0.004	0.004	0.004
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
Score	Y	Y	Y	Y	Y
Score^2		Y	Y	Y	Y
Score^3			Y	Y	Y
Score^4				Y	Y
Xs					Y
Observations	2,397	2,397	2,397	2,397	2,397
R^2	0.872	0.872	0.872	0.872	0.876
F-test	4983	801.6	941.9	533.1	548.0

Notes: Sample includes all children in grades 4-7 in 1962 who reply to the 2001 survey and went to a non-private high school. Cells show coefficients from linear probability models of the probability of elite school assignment. "Xs" include school attended in 1962, grade in 1962, month of birth (within each grade), father's occupation, mother's socio-economic status and fourth-order polynomials of test scores at ages 7 and 9. The F-test statistics indicate the test statistics for the null hypothesis that the first 4 independent variables are jointly statistically insignificant.

Table 2: Impact of elite school attendance on educational attainment

	MEN					WOMEN				
	OLS estimates			GMM estimates		OLS estimates			GMM estimates	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Years of postcompulsory education										
Elite school	3.228	1.507	1.433	1.124	1.204	2.914	0.863	0.697	0.930	0.829
	(0.151)	(0.234)	(0.229)	(0.272)	(0.270)	(0.135)	(0.167)	(0.165)	(0.240)	(0.228)
F(Score)		Y	Y	Y	Y		Y	Y	Y	Y
Control variables			Y		Y			Y		Y
Control Mean	1.992	1.992	1.992	1.992	1.992	2.311	2.311	2.311	2.311	2.311
Observations	2,070	2,070	2,070	2,070	2,070	2396	2396	2396	2396	2396
Overid. test				0.454	0.581				0.204	0.102
Panel B: No qualifications										
Elite school	-0.274	-0.027	-0.013	0.028	0.019	-0.271	0.064	0.068	0.090	0.087
	(0.017)	(0.025)	(0.025)	(0.030)	(0.032)	(0.020)	(0.024)	(0.023)	(0.036)	(0.036)
F(Score)		Y	Y	Y	Y		Y	Y	Y	Y
Control variables			Y		Y			Y		Y
Control Mean	0.072	0.072	0.072	0.072	0.072	0.066	0.066	0.066	0.066	0.066
Observations	2,070	2,070	2,070	2,070	2,070	2396	2396	2396	2396	2396
Overid. test				0.676	0.708				0.132	0.176
Panel C: Ungraded CSE, school leaving certificate, other qualifications										
Elite school	0.057	0.008	0.024	0.026	0.055	-0.0275	-0.153	-0.155	-0.209	-0.196
	(0.024)	(0.040)	(0.041)	(0.055)	(0.055)	(0.023)	(0.046)	(0.048)	(0.055)	(0.056)
F(Score)		Y	Y	Y	Y		Y	Y	Y	Y
Control variables			Y		Y			Y		Y
Control Mean	0.176	0.176	0.176	0.176	0.176	0.359	0.359	0.359	0.359	0.359
Observations	2,070	2,070	2,070	2,070	2,070	2396	2396	2396	2396	2396
Overid. test				0.415	0.496				0.241	0.256
Panel D: O-level or equivalent										
Elite school	0.342	0.050	0.044	0.019	0.029	0.316	-0.024	-0.028	-0.084	-0.081
	(0.026)	(0.041)	(0.041)	(0.052)	(0.054)	(0.024)	(0.034)	(0.034)	(0.050)	(0.049)
F(Score)		Y	Y	Y	Y		Y	Y	Y	Y
Control variables			Y		Y			Y		Y
Control Mean	0.760	0.760	0.760	0.760	0.760	0.802	0.802	0.802	0.802	0.802
Observations	2,070	2,070	2,070	2,070	2,070	2396	2396	2396	2396	2396
Overid. test				0.0330	0.0440				0.413	0.488
Panel E: A-level or equivalent										
Elite school	0.522	0.210	0.201	0.104	0.128	0.601	0.274	0.243	0.261	0.234
	0.027	0.045	0.043	(0.059)	(0.055)	(0.025)	(0.039)	(0.038)	(0.050)	(0.049)
F(Score)		Y	Y	Y	Y		Y	Y	Y	Y
Control variables			Y		Y			Y		Y
Control Mean	0.464	0.464	0.464	0.464	0.464	0.401	0.401	0.401	0.401	0.401
Observations	2,070	2,070	2,070	2,070	2,070	2396	2396	2396	2396	2396
Overid. test				0.548	0.642				0.027	0.019
Panel F: HNC, teaching, etc.										
Elite school	0.076	0.021	0.021	-0.039	-0.015	0.268	0.075	0.048	0.113	0.099
	(0.027)	(0.048)	(0.048)	(0.064)	(0.066)	(0.024)	(0.037)	(0.038)	(0.051)	(0.050)
F(Score)		Y	Y	Y	Y		Y	Y	Y	Y
Control variables			Y		Y			Y		Y
Control Mean	0.344	0.344	0.344	0.344	0.344	0.293	0.293	0.293	0.293	0.293
Observations	2,070	2,070	2,070	2,070	2,070	2396	2396	2396	2396	2396
Overid. test				0.576	0.615				0.198	0.200
Panel G: Degree or higher										
Elite school	0.425	0.194	0.179	0.157	0.149	0.325	0.054	0.042	0.047	0.041
	(0.028)	(0.043)	(0.041)	(0.051)	(0.047)	(0.025)	(0.033)	(0.032)	(0.040)	(0.040)
F(Score)		Y	Y	Y	Y		Y	Y	Y	Y
Control variables			Y		Y			Y		Y
Control Mean	0.136	0.136	0.136	0.136	0.136	0.174	0.174	0.174	0.174	0.174
Observations	2,070	2,070	2,070	2,070	2,070	2396	2396	2396	2396	2396
Overid. test				0.386	0.399				0.373	0.336

Notes: Sample as in Table 1. "F(score)" refers to a third-order polynomial in assignement score; "Xs" are as in Table 1. GMM estimates use the instruments described in the text. Standard errors clustered by assignment score. "Control Mean" refers to the outcome mean among individuals with borderline scores (540,560) that attended non-elite schools.

Table 3: Impact of elite school attendance on labour market outcomes

	MEN					WOMEN				
	OLS estimates			GMM estimates		OLS estimates			GMM estimates	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Log annual income (all)										
Elite school	0.420	0.040	0.027	-0.082	-0.062	0.489	0.217	0.225	0.175	0.194
	(0.032)	(0.056)	(0.056)	(0.063)	(0.065)	(0.042)	(0.079)	(0.081)	(0.094)	(0.098)
F(Score)		Y	Y	Y	Y		Y	Y	Y	Y
Control variables			Y		Y			Y		Y
Control Mean	10.227	10.227	10.227	10.227	10.227	9.257	9.257	9.257	9.257	9.257
Observations	2,029	2,029	2,029	2,029	2,029	2,335	2,335	2,335	2,335	2,335
Overid. test				0.177	0.328				0.775	0.791
Panel B: Employment										
Elite school	0.055	-0.014	-0.013	-0.031	-0.023	0.053	-0.008	-0.008	-0.003	-0.006
	(0.014)	(0.025)	(0.024)	(0.032)	(0.032)	(0.016)	(0.029)	(0.030)	(0.037)	(0.039)
F(Score)		Y	Y	Y	Y		Y	Y	Y	Y
Control variables			Y		Y			Y		Y
Control Mean	0.944	0.944	0.944	0.944	0.944	0.862	0.862	0.862	0.862	0.862
Observations	2,066	2,066	2,066	2,066	2,066	2,387	2,387	2,387	2,387	2,387
Overid. test				0.809	0.556				0.830	0.940
Panel C: NES-imputed log hourly wage (working employees only)										
Elite school	0.356	0.062	0.067	-0.022	0.005	0.354	0.073	0.057	0.104	0.100
	(0.023)	(0.039)	(0.040)	(0.050)	(0.053)	(0.024)	(0.036)	(0.036)	(0.045)	(0.042)
F(Score)		Y	Y	Y	Y		Y	Y	Y	Y
Control variables			Y		Y			Y		Y
Control Mean	2.366	2.366	2.366	2.366	2.366	1.908	1.908	1.908	1.908	1.908
Observations	1,858	1,858	1,858	1,858	1,858	1,905	1,905	1,905	1,905	1,905
Overid. test				0.026	0.077				0.425	0.309

Notes: See the notes to Table 2.

Table 4: Impact of elite school attendance on fertility and marriage

	MEN					WOMEN				
	OLS estimates			GMM estimates		OLS estimates			GMM estimates	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Any child										
Elite school	-0.019	0.005	0.020	0.007	0.025	-0.069	-0.064	-0.054	-0.096	-0.090
	(0.021)	(0.039)	(0.038)	(0.052)	(0.053)	(0.017)	(0.032)	(0.032)	(0.041)	(0.040)
F(Score)		Y	Y	Y	Y		Y	Y	Y	Y
Control variables			Y		Y			Y		Y
Control Mean	0.864	0.864	0.864	0.864	0.864	0.898	0.898	0.898	0.898	0.898
Observations	2,068	2,068	2,068	2,068	2,068	2,396	2,396	2,396	2,396	2,396
Overid. test				0.565	0.549				0.374	0.404
Panel B: Number of children										
Elite school	-0.030	0.149	0.223	0.224	0.300	-0.214	-0.329	-0.296	-0.488	-0.473
	(0.060)	(0.107)	(0.104)	(0.150)	(0.148)	(0.050)	(0.103)	(0.104)	(0.114)	(0.108)
F(Score)		Y	Y	Y	Y		Y	Y	Y	Y
Control variables			Y		Y			Y		Y
Control Mean	1.864	1.864	1.864	1.864	1.864	2.066	2.066	2.066	2.066	2.066
Observations	2,068	2,068	2,068	2,068	2,068	2,396	2,396	2,396	2,396	2,396
Overid. test				0.812	0.577				0.287	0.360
Panel C: Age at first child										
Elite school	2.785	0.796	0.628	0.569	0.580	0.466	0.797	0.337	0.526	0.182
	(0.311)	(0.517)	(0.536)	(0.729)	(0.692)	(0.698)	(0.475)	(0.489)	(0.581)	(0.612)
F(Score)		Y	Y	Y	Y		Y	Y	Y	Y
Control variables			Y		Y			Y		Y
Control Mean	26.352	26.352	26.352	26.352	26.352	24.247	24.247	24.247	24.247	24.247
Observations	1,722	1,722	1,722	1,722	1,722	2,085	2,085	2,085	2,085	2,085
Overid. test				0.618	0.918				0.901	0.947
Panel D: First child before age 27										
Elite school	-0.229	-0.108	-0.082	-0.093	-0.062	-0.189	-0.049	-0.019	-0.035	-0.015
	(0.031)	(0.051)	(0.052)	(0.070)	(0.065)	(0.024)	(0.036)	(0.038)	(0.050)	(0.051)
F(Score)		Y	Y	Y	Y			Y	Y	Y
Control variables			Y		Y			Y		Y
Control Mean	0.648	0.648	0.648	0.648	0.648	0.780	0.780	0.780	0.780	0.780
Observations	1,722	1,722	1,722	1,722	1,722	2,085	2,085	2,085	2,085	2,085
Overid. test				0.324	0.588				0.996	0.971
Panel E: Currently married										
Elite school	0.012	-0.031	-0.021	-0.021	-0.019	0.008	0.012	0.002	0.043	0.040
	(0.027)	(0.047)	(0.047)	(0.061)	(0.062)	(0.019)	(0.037)	(0.038)	(0.046)	(0.048)
F(Score)		Y	Y	Y	Y		Y	Y	Y	Y
Control variables			Y		Y			Y		Y
Control Mean	0.864	0.864	0.864	0.864	0.864	0.766	0.766	0.766	0.766	0.766
Observations	2,063	2,063	2,063	2,063	2,063	2,389	2,389	2,389	2,389	2,389
Overid. test				0.140	0.134				0.652	0.720
Panel F: Ever divorced										
Elite school	-0.050	0.024	0.046	0.045	0.086	-0.031	-0.002	0.002	-0.019	-0.039
	(0.024)	(0.040)	(0.043)	(0.058)	(0.062)	(0.025)	(0.046)	(0.047)	(0.060)	(0.060)
F(Score)		Y	Y	Y	Y		Y	Y	Y	Y
Control variables			Y		Y			Y		Y
Control Mean	0.188	0.188	0.188	0.188	0.188	0.314	0.314	0.314	0.314	0.314
Observations	1,862	1,862	1,862	1,862	1,862	2,222	2,222	2,222	2,222	2,222
Overid. test				0.104	0.135				0.476	0.545

Notes: See the notes to Table 2.

Appendixes

Appendix A: Additional Figures and Tables

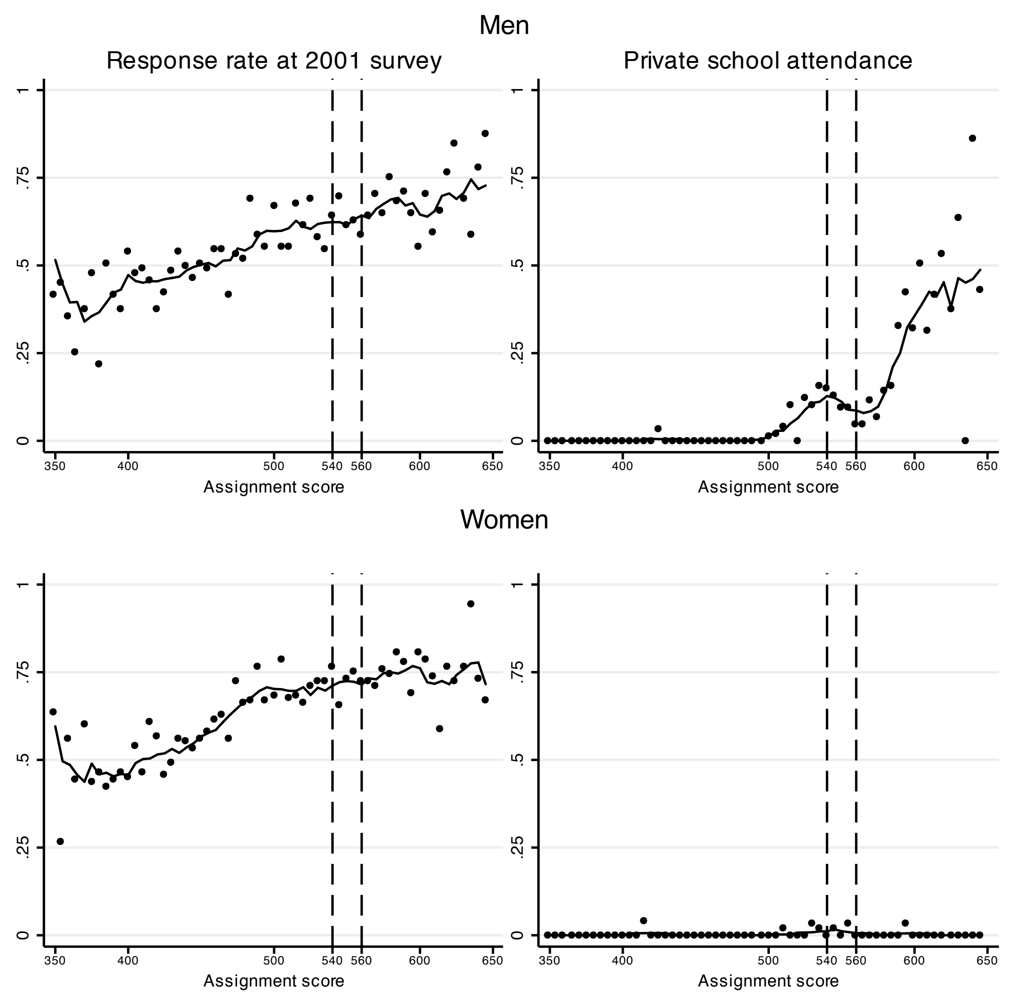
Appendix B: Aberdeen Cohort and Labour Force Survey Comparison

Appendix C: The Case of Heterogeneous Treatment Effects

Appendix D: A Model of School Quality with Vocational Training

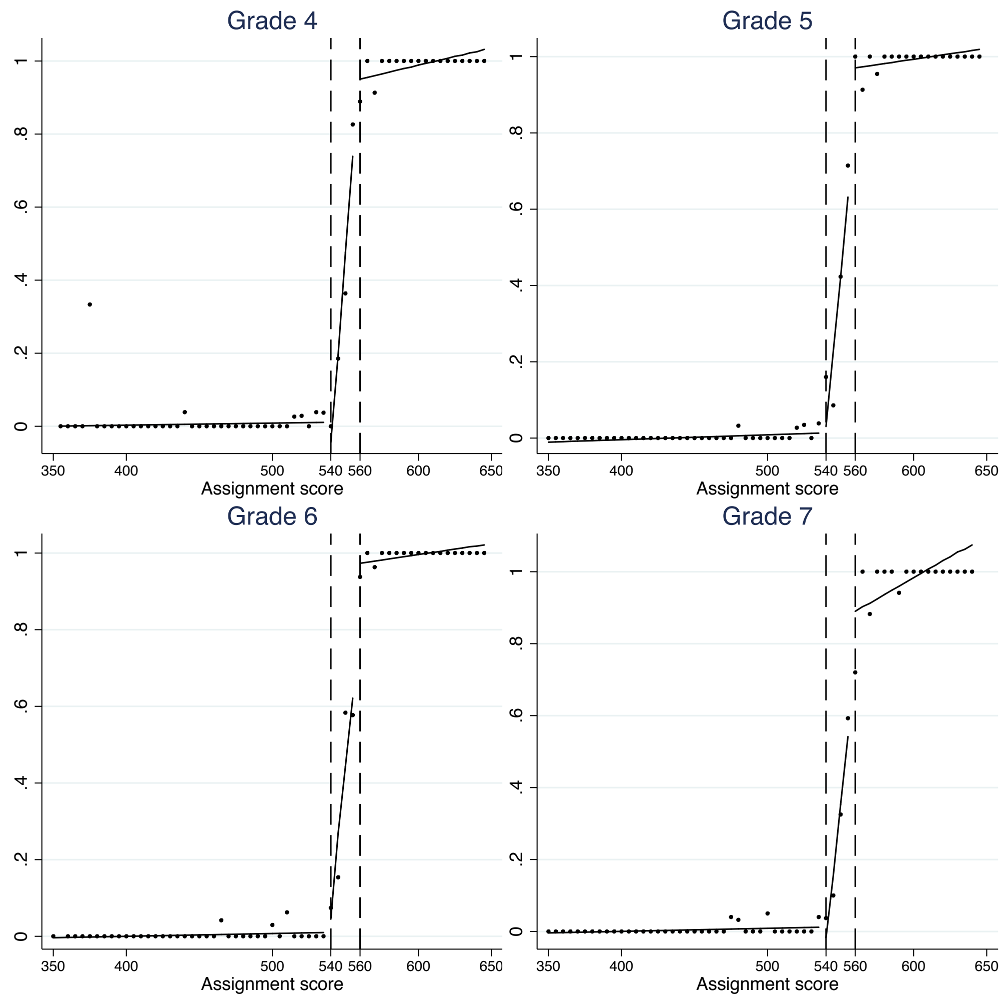
Appendix A: Additional Figures and Tables

Appendix A Figure 1: Assignment scores, survey response and private school attendance



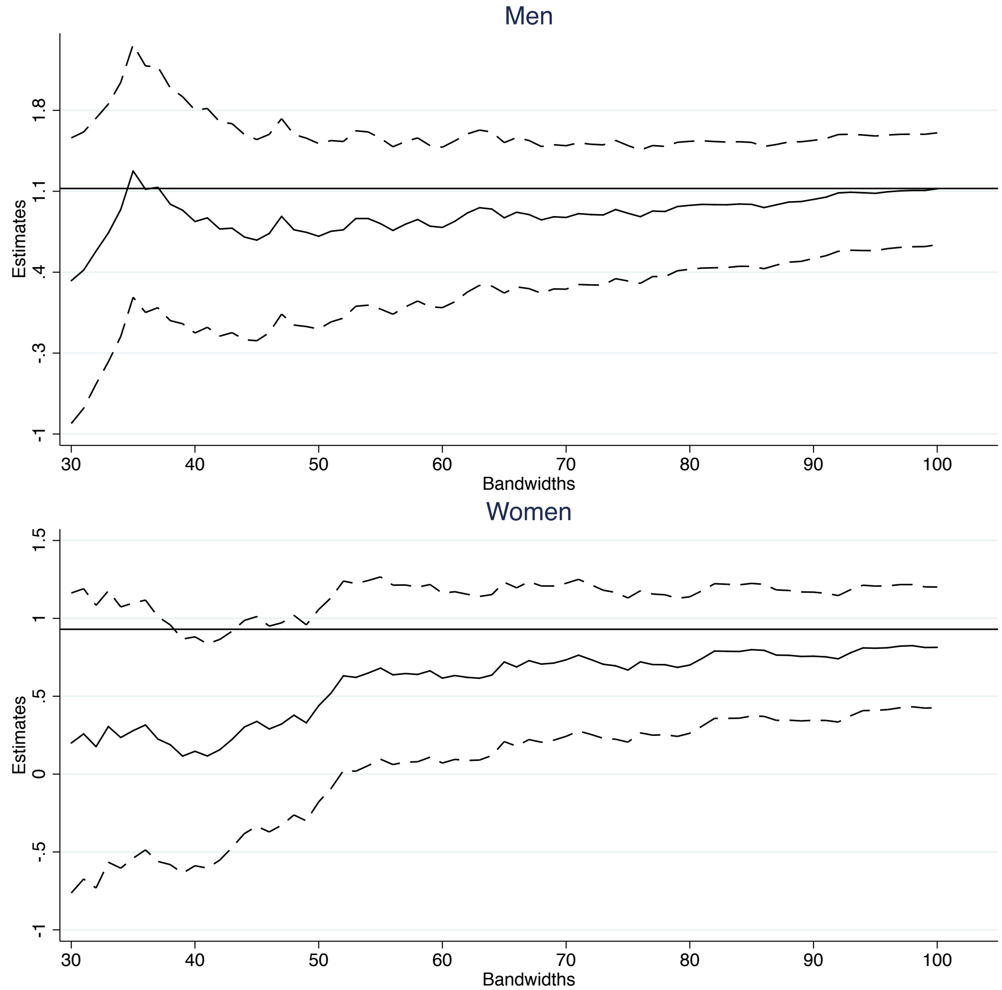
Notes: Sample includes all men and women in grades 4-7 in 1962. Dots represent the fraction of each cell who responded to the 2001 survey or who report attending a private school. Cells defined over 5-score intervals; vertical lines plotted at the [540, 544] and [560, 564] score values.

Appendix A Figure 2: Assignment score and elite school attendance by grade in 1962



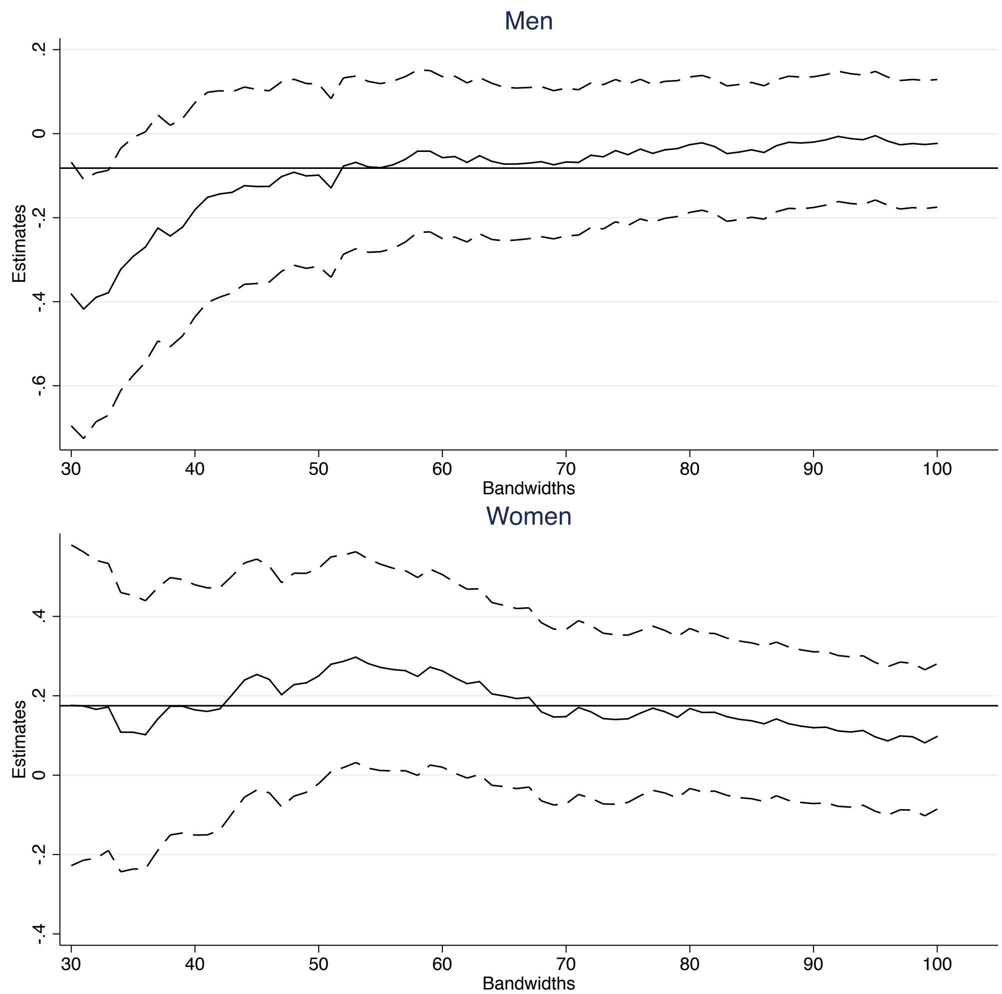
Notes: see notes to Figure 2.

Appendix A Figure 3: RD estimates of post-compulsory schooling impacts



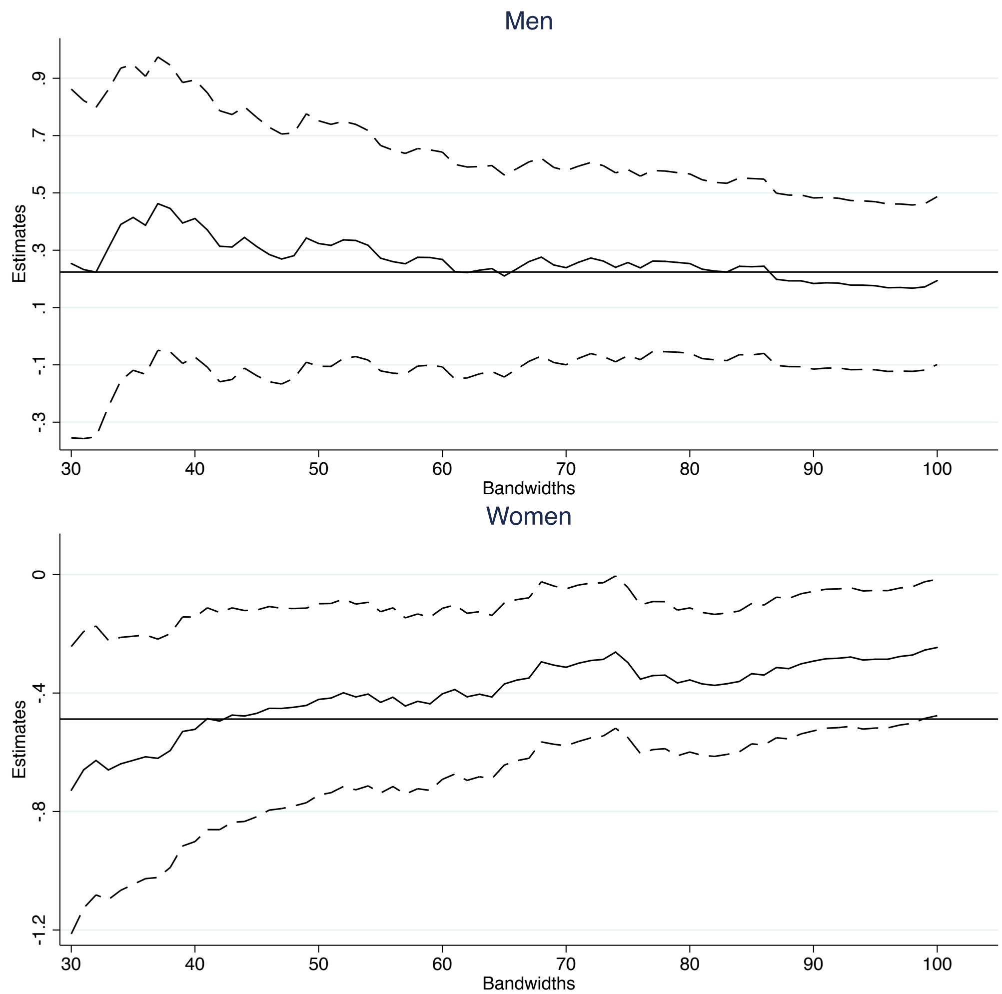
Notes: Sample includes all men and women in grades 4-7 in 1962 who reply to the 2001 survey and went to a non-private secondary school. RD estimation procedure described in the text. Solid line shows RD estimates corresponding to all bandwidths from 30 to 100; dashed lines show corresponding 95% confidence intervals. Solid horizontal lines represent IV estimates reported in columns (4) and (9) of Table 2, panel A.

Appendix A Figure 4: RD estimates of log income impacts



Notes: See the notes to Appendix A Figure 3. Solid horizontal lines represent IV estimates reported in columns (4) and (9) of Table 3, panel A.

Appendix A Figure 5: RD estimates of number of children impacts



Notes: See the notes to Appendix A Figure 3. Solid horizontal lines represent IV estimates reported in columns (4) and (9) of Table 4, panel B.

Appendix A Table 1a: Sample selection and definition of subsamples

	N	
Main sample	12150	Original sample
Moved 1962-64 (358)	11792	May have left Aberdeen
Not private primary school (255)		Some of these didn't take the test
Not RC primary school (229)		Most of them took the test, although they stay in a RC school
Not RC private primary school (81)		Same arguments as RC and private primary sc
Not primary elite school already (504)		These pupils stay in the school with probability 1 whatever the results of the test a
Not primary special school (167)		Some of these didn't take the test
Not primary outside Aberdeen school (12)	10544	Might go to secondary school outside Aberdeen
With grade info	10497	Because we will want to split by grade
With 11+ info	10219	Because we will need assignment scores
With test at 7, test at 9 info	9921	Because it will be important to regression-adjust for these variables
Replied to 2001 survey	5970	Because we will be looking at outcomes of respondents
Went to state secondary	5646	Because we will want to focus on elite vs. non-elite
Grade 4-7 in 1962	4528	Because assignment procedure changed

Notes: Description of sample selection procedure applied to original data

Appendix A Table 1b: Descriptive Statistics

	Sample with test at 7 and test at 9 information		Grade 4-7 in 1962		Grade 4-7 in 1962			
					Non-elite		Elite	
	mean	N	mean	N	mean	N	mean	N
Male	0.518	9921	0.464	4528	0.480	3489	0.412	1039
Age (Dec 1962)	9.727	9921	10.203	4528	10.216	3489	10.162	1039
	(1.436)		(1.158)		(1.160)		(1.151)	
Father's social class		9921		4528		3489		1039
Other (unemployed, disabled, etc.)	0.051		0.048		0.052		0.035	
Unskilled manual	0.170		0.161		0.184		0.085	
Semi-skilled manual	0.149		0.146		0.164		0.085	
Skilled manual, other	0.208		0.211		0.226		0.163	
Skilled manual, requiring apprent.	0.251		0.260		0.243		0.316	
Other non manual	0.109		0.115		0.093		0.189	
Intermediate/Technical	0.051		0.049		0.033		0.103	
Professional	0.011		0.010		0.005		0.026	
Grade in 1962		9921		4528		3489		1039
Grade 3	0.208		0.000		0.000		0.000	
Grade 4	0.205		0.260		0.261		0.253	
Grade 5	0.204		0.251		0.252		0.246	
Grade 6	0.195		0.247		0.242		0.266	
Grade 7	0.196		0.242		0.245		0.235	
Age 7 test score	106.735	9921	108.023	4528	104.073	3489	121.284	1039
	(15.172)		(14.552)		(13.068)		(11.011)	
Age 9 test score	110.612	9921	112.258	4528	107.305	3489	128.888	1039
	(16.214)		(15.358)		(12.868)		(10.632)	
Assignment score	496.046	7929	503.303	4528	479.078	3489	584.650	1039
	(66.433)		(62.825)		(48.085)		(29.123)	
Responded to 2001 survey	0.602	9921	1	4528	1	3489	1	1039
Went to private secondary school	0.040	5878	0	4528	0	3489	0	1039
Elite School	0.225	5646	0.230	4528	0	3489	1	1039

Notes: see Appendix A Table 1a for definition of each sample.

Appendix A Table 2: Falsification tests of elite school impacts

	MEN			WOMEN		
	OLS estimates		GMM estimates	OLS estimates		GMM estimates
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Years of post-compulsory education						
Elite school	2.038	0.282	-0.046	2.132	0.221	0.144
	(0.097)	(0.119)	(0.120)	(0.0934)	(0.098)	(0.117)
F(Score)		Y	Y		Y	Y
Control Mean	1.992	1.992	1.992	2.311	2.311	2.311
Observations	2,070	2,070	2,070	2,396	2,396	2,396
Overid. test			0.563			0.115
Panel B: A-level or equivalent						
Elite school	0.358	0.043	-0.028	0.415	0.035	0.033
	(0.016)	(0.019)	(0.020)	(0.017)	(0.018)	(0.022)
F(Score)		Y	Y		Y	Y
Control Mean	0.344	0.344	0.344	0.401	0.401	0.401
Observations	2,070	2,070	2,070	2,396	2,396	2,396
Overid. test			0.503			0.293
Panel C: Degree or higher						
Elite school	0.260	0.034	0.008	0.248	0.016	0.011
	(0.014)	(0.016)	(0.016)	(0.012)	(0.011)	(0.014)
F(Score)		Y	Y		Y	Y
Control Mean	0.136	0.136	0.136	0.174	0.174	0.174
Observations	2,070	2,070	2,070	2,396	2,396	2,396
Overid. test			0.951			0.294
Panel D: Log annual income (all)						
Elite school	0.333	0.038	-0.031	0.386	0.006	-0.021
	(0.016)	(0.020)	(0.018)	(0.019)	(0.018)	(0.023)
F(Score)					Y	Y
Control Mean	10.227	10.227	10.227	9,257	9,257	9,257
Observations	2.070	2.070	2.070	2,396	2,396	2,396
Overid. test			0.210			0.130
Panel E: Any child						
Elite school	-0.034	-0.012	-0.016	-0.041	-0.015	-0.008
	(0.004)	(0.007)	(0.009)	(0.003)	(0.006)	(0.007)
F(Score)		Y	Y		Y	Y
Control Mean	0.864	0.864	0.864	0.898	0.898	0.898
Observations	2,070	2,070	2,070	2,396	2,396	2,396
Overid. test			0.824			0.438
Panel F: Number of children						
Elite school	-0.097	-0.048	-0.056	-0.100	-0.058	-0.032
	(0.013)	(0.020)	(0.024)	(0.010)	(0.019)	(0.023)
F(Score)		Y	Y		Y	Y
Control Mean	1.864	1.864	1.864	2.066	2.066	2.066
Observations	2,070	2,070	2,070	2,396	2,396	2,396
Overid. test			0.452			0.678

Notes: see notes to Table 2. The dependent variables in these models are those predicted by a regression of the relevant outcome on: dummies for father's occupation, mother's socio-economic status, school and grade attended in 1962, continuous variables capturing relative age within the school-grade, and third-degree polynomials in test scores at age 7 and 9.

Appendix A Table 3: Impact elite school attendance: robustness checks

	MEN			WOMEN		
	GMM estimates			GMM estimates		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Years of post-compulsory education						
Elite school	1.195	1.204	0.981	0.761	0.829	0.760
	(0.286)	(0.270)	(0.304)	(0.244)	(0.228)	(0.259)
F(Score)	2nd degree	3rd degree	4th degree	2nd degree	3rd degree	4th degree
Control Mean	1.992	1.992	1.992	2.327	2.327	2.327
Observations	1,532	2,070	2,070	1879	2,396	2,396
Overid. test	0.754	0.581	0.587	0.11	0.102	0.103
Panel B: A-level or equivalent						
Elite school	0.167	0.128	0.0928	0.22	0.234	0.161
	(0.055)	(0.055)	(0.070)	(0.052)	(0.049)	(0.060)
F(Score)	2nd degree	3rd degree	4th degree	2nd degree	3rd degree	4th degree
Control Mean	0.344	0.344	0.344	0.242	0.242	0.242
Observations	1,532	2,070	2,070	1,879	2,396	2,396
Overid. test	0.716	0.642	0.522	0.0224	0.0192	0.0235
Panel C: Degree or higher						
Elite school	0.143	0.149	0.162	0.0107	0.041	0.046
	(0.0501)	(0.047)	(0.054)	(0.0433)	(0.040)	(0.047)
F(Score)	2nd degree	3rd degree	4th degree	2nd degree	3rd degree	4th degree
Control Mean	0.136	0.136	0.136	0.176	0.176	0.176
Observations	1,532	2,070	2,070	1,879	2,396	2,396
Overid. test	0.39	0.399	0.425	0.410	0.336	0.312
Panel D: Log annual income						
Elite school	-0.0712	-0.062	-0.137	0.191	0.194	0.325
	(0.0618)	(0.065)	(0.081)	(0.0999)	(0.098)	(0.126)
F(Score)	2nd degree	3rd degree	4th degree	2nd degree	3rd degree	4th degree
Control Mean	10.227	10.227	10.227	9.271	9.271	9.271
Observations	1,511	2,029	2,029	1,840	2,335	2,335
Overid. test	0.109	0.328	0.564	0.791	0.791	0.728
Panel E: Any child						
Elite school	0.0217	0.025	-0.017	-0.117	-0.091	-0.148
	(0.0549)	(0.053)	(0.065)	(0.0416)	(0.041)	(0.049)
F(Score)	2nd degree	3rd degree	4th degree	2nd degree	3rd degree	4th degree
Control Mean	0.864	0.864	0.864	0.897	0.897	0.897
Observations	1,530	2,068	2,068	1,878	2,394	2,394
Overid. test	0.481	0.549	0.703	0.593	0.404	0.717
Panel F: Number of children						
Elite school	0.307	0.300	0.148	-0.489	-0.473	-0.623
	(0.151)	(0.148)	(0.181)	(0.107)	(0.108)	(0.132)
F(Score)	2nd degree	3rd degree	4th degree	2nd degree	3rd degree	4th degree
Control Mean	1.864	1.864	1.864	2.067	2.067	2.067
Observations	1,530	1,750	1,750	1878	2,394	2,394
Overid. test	0.855	0.577	0.851	0.649	0.360	0.747

Notes: See notes to Table 2. The estimates corresponding to "2nd degree" polynomials use a narrower range of data: assignment scores 460-650.

Appendix A Table 4: Education impacts by socio-economic status

	LOW SES					HIGH SES				
	OLS estimates			GMM estimates		OLS estimates			GMM estimates	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Years of post-compulsory education										
Elite school	2.825	0.944	0.908	0.755	0.802	2.983	1.196	1.079	1.009	0.886
	(0.156)	(0.229)	(0.215)	(0.261)	(0.261)	(0.129)	(0.195)	(0.203)	(0.236)	(0.235)
F(Score)		Y	Y	Y	Y		Y	Y	Y	Y
Control variables			Y		Y			Y		Y
Control Mean	2.054	2.054	2.054	2.054	2.054	2.361	2.361	2.361	2.361	2.361
Observations	2,521	2,521	2,521	2,521	2,521	1945	1945	1945	1945	1945
Overid. test				0.241	0.316				0.624	0.761
Panel B: No qualifications										
Elite school	-0.305	0.0217	0.0424	0.0742	0.0887	-0.207	0.00895	0.0154	0.0626	0.0621
	(0.0206)	(0.0290)	(0.0298)	(0.0363)	(0.0372)	(0.0168)	(0.0233)	(0.0242)	(0.0336)	(0.0347)
F(Score)		Y	Y	Y	Y		Y	Y	Y	Y
Control variables			Y		Y			Y		Y
Control Mean	0.0595	0.0595	0.0595	0.0595	0.0595	0.082	0.082	0.082	0.082	0.082
Observations	2,521	2,521	2,521	2,521	2,521	1,945	1,945	1,945	1,945	1,945
Overid. test				0.200	0.164				0.284	0.309
Panel C: Ungraded CSE, school leaving certificate, other qualifications										
Elite school	0.00688	-0.0835	-0.0723	-0.0779	-0.0621	0.0147	-0.0725	-0.0788	-0.0901	-0.104
	(0.0260)	(0.0409)	(0.0422)	(0.0579)	(0.0581)	(0.0211)	(0.0388)	(0.0396)	(0.0534)	(0.0552)
F(Score)		Y	Y	Y	Y		Y	Y	Y	Y
Control variables			Y		Y			Y		Y
Control Mean	0.321	0.321	0.321	0.321	0.321	0.23	0.23	0.23	0.23	0.23
Observations	2,521	2,521	2,521	2,521	2,521	1,945	1,945	1,945	1,945	1,945
Overid. test				0.546	0.641				0.00569	0.00784
Panel D: O-level or equivalent										
Elite school	0.360	0.0432	0.0275	-0.0468	-0.0566	0.265	-0.0180	-0.0179	-0.0616	-0.0433
	(0.0259)	(0.0448)	(0.0467)	(0.0518)	(0.0504)	(0.0235)	(0.0333)	(0.0334)	(0.0486)	(0.0495)
F(Score)		Y	Y	Y	Y		Y	Y	Y	Y
Control variables			Y		Y			Y		Y
Control Mean	0.774	0.774	0.774	0.774	0.774	0.795	0.795	0.795	0.795	0.795
Observations	2,521	2,521	2,521	2,521	2,521	1,945	1,945	1,945	1,945	1,945
Overid. test				0.756	0.757				0.124	0.101
Panel E: A-level or equivalent										
Elite school	0.538	0.203	0.180	0.144	0.132	0.534	0.268	0.245	0.201	0.178
	(0.0291)	(0.0442)	(0.0438)	(0.0527)	(0.0546)	(0.0235)	(0.0390)	(0.0392)	(0.0494)	(0.0506)
F(Score)		Y	Y	Y	Y		Y	Y	Y	Y
Control variables			Y		Y			Y		Y
Control Mean	0.262	0.262	0.262	0.262	0.262	0.32	0.32	0.32	0.32	0.32
Observations	2,521	2,521	2,521	2,521	2,521	1,945	1,945	1,945	1,945	1,945
Overid. test				0.236	0.201				0.128	0.184
Panel F: HNC, teaching, etc.										
Elite school	0.167	0.000617	-0.0176	-0.00176	-0.0191	0.172	0.0884	0.0738	0.0804	0.0683
	(0.0280)	(0.0455)	(0.0464)	(0.0579)	(0.0571)	(0.0234)	(0.0387)	(0.0403)	(0.0469)	(0.0476)
F(Score)		Y	Y	Y	Y		Y	Y	Y	Y
Control variables			Y		Y			Y		Y
Control Mean	0.327	0.327	0.327	0.327	0.327	0.303	0.303	0.303	0.303	0.303
Observations	2,521	2,521	2,521	2,521	2,521	1,945	1,945	1,945	1,945	1,945
Overid. test				0.154	0.133				0.901	0.842
Panel G: Degree or higher										
Elite school	0.331	0.0931	0.0843	0.0718	0.0701	0.372	0.138	0.108	0.0996	0.0673
	(0.0258)	(0.0368)	(0.0353)	(0.0417)	(0.0414)	(0.0240)	(0.0315)	(0.0312)	(0.0424)	(0.0414)
F(Score)		Y	Y	Y	Y		Y	Y	Y	Y
Control variables			Y		Y			Y		Y
Control Mean	0.155	0.155	0.155	0.155	0.155	0.164	0.164	0.164	0.164	0.164
Observations	2,521	2,521	2,521	2,521	2,521	1,945	1,945	1,945	1,945	1,945
Overid. test				0.303	0.389				0.61	0.792

Notes: See notes to Table 2. Low-SES is defined as father's social class 1 through 4 (lower skilled, semi-skilled, unskilled manual or not working); high-SES is defined as father's social class 5 through 8 (skilled manual occupation that required an apprenticeship, non-manual, intermediate or professional occupation).

Appendix A Table 5: Impact of elite school attendance on occupation

	MEN					WOMEN				
	OLS estimates			GMM estimates		OLS estimates			GMM estimates	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Managers										
Elite school	0.118	-0.007	0.001	-0.083	-0.063	0.056	0.027	0.020	-0.021	-0.029
	(0.027)	(0.050)	(0.050)	(0.063)	(0.064)	(0.018)	(0.035)	(0.035)	(0.039)	(0.039)
F(Score)		Y	Y	Y	Y		Y	Y	Y	Y
Control variables			Y		Y			Y		Y
Control Mean	0.400	0.400	0.400	0.400	0.400	0.112	0.112	0.112	0.112	0.112
Observations	1,826	1,826	1,826	1,826	1,826	1,926	1,926	1,926	1,926	1,926
Overid. test				0.0418	0.0654				0.566	0.44
Panel B: Professionals and associate professionals										
Elite school	0.281	0.119	0.116	0.059	0.065	0.306	0.059	0.050	0.115	0.115
	(0.031)	(0.054)	(0.055)	(0.065)	(0.069)	(0.025)	(0.037)	(0.038)	(0.052)	(0.050)
F(Score)		Y	Y	Y	Y		Y	Y	Y	Y
Control variables			Y		Y			Y		Y
Control Mean	0.243	0.243	0.243	0.243	0.243	0.28	0.28	0.28	0.28	0.28
Observations	1,826	1,826	1,826	1,826	1,826	1,926	1,926	1,926	1,926	1,926
Overid. test				0.837	0.825				0.946	0.892
Panel C: Clerical and secretarial										
Elite school	-0.026	-0.027	-0.027	-0.004	-0.012	-0.056	-0.093	-0.091	-0.141	-0.133
	(0.012)	(0.023)	(0.024)	(0.028)	(0.030)	(0.026)	(0.045)	(0.046)	(0.060)	(0.061)
F(Score)		Y	Y	Y	Y		Y	Y	Y	Y
Control variables			Y		Y			Y		Y
Control Mean	0.078	0.078	0.078	0.078	0.078	0.406	0.406	0.406	0.406	0.406
Observations	1,826	1,826	1,826	1,826	1,826	1,926	1,926	1,926	1,926	1,926
Overid. test				0.324	0.194				0.589	0.455
Panel D: Craft and related										
Elite school	-0.190	-0.065	-0.078	0.020	-0.006	-0.010	0.009	0.007	0.012	0.012
	(0.016)	(0.025)	(0.026)	(0.033)	(0.035)	(0.005)	(0.008)	(0.008)	(0.006)	(0.007)
F(Score)		Y	Y	Y	Y		Y	Y	Y	Y
Control variables			Y		Y			Y		Y
Control Mean	0.113	0.113	0.113	0.113	0.113	0.007	0.007	0.007	0.007	0.007
Observations	1,826	1,826	1,826	1,826	1,826	1,926	1,926	1,926	1,926	1,926
Overid. test				0.0247	0.0285				0.373	0.328
Panel E: Personal and sales										
Elite school	-0.025	0.004	0.009	0.022	0.019	-0.177	0.002	0.013	0.036	0.044
	(0.017)	(0.031)	(0.031)	(0.037)	(0.037)	(0.018)	(0.034)	(0.034)	(0.041)	(0.043)
F(Score)		Y	Y	Y	Y		Y	Y	Y	Y
Control variables			Y		Y			Y		Y
Control Mean	0.078	0.078	0.078	0.078	0.078	0.147	0.147	0.147	0.147	0.147
Observations	1,826	1,826	1,826	1,826	1,826	1,926	1,926	1,926	1,926	1,926
Overid. test				0.966	0.979				0.799	0.668
Panel F: Other										
Elite school	-0.158	-0.024	-0.020	0.015	0.020	-0.117	-0.003	0.001	0.013	0.009
	(0.018)	(0.028)	(0.028)	(0.041)	(0.042)	(0.012)	(0.020)	(0.022)	(0.027)	(0.030)
F(Score)		Y	Y	Y	Y		Y	Y	Y	Y
Control variables			Y		Y			Y		Y
Control Mean	0.087	0.087	0.087	0.087	0.087	0.049	0.049	0.049	0.049	0.049
Observations	1,826	1,826	1,826	1,826	1,826	1,926	1,926	1,926	1,926	1,926
Overid. test				0.508	0.553				0.326	0.368

Notes: See notes to Table 2.

Appendix A Table 6: Educational attainment and mean income by occupation

	% degree	% A-levels	Mean annual income	Mean log annual income
			MEN	
Managers	0.230	0.510	38,445	10.498
Prof. and associate prof.	0.456	0.671	33,311	10.341
Clerical and secretarial	0.033	0.248	20,691	9.835
Craft and related	0.014	0.256	23,724	9.991
Personal and sales	0.025	0.278	22,003	9.904
Other	0.029	0.142	20,519	9.809
Observations	1,826	1,826	1,826	1,826
			WOMEN	
Managers	0.227	0.427	24,743	9.954
Prof. and associate prof.	0.434	0.632	21,576	9.842
Clerical and secretarial	0.042	0.234	13,681	9.396
Craft and related	0.038	0.077	9,846	9.118
Personal and sales	0.028	0.130	9,210	8.933
Other	0.005	0.036	7,112	8.728
Observations	1,926	1,926	1,926	1,926

Notes: Sample includes all men and women in grades 4-7 in 1962 who reply to the 2001 survey, went to a non-private secondary school, are employed in 2001 and have non-missing occupational code. Percentage of individuals holding a degree-level qualification or an A-level qualification or equivalent by occupation shown. Also shown, mean values of annual income and log annual income (for all individuals and for those working as employees only) in each occupational category.

Appendix B: Aberdeen Cohort and Labour Force Survey Comparison

Appendix B Table 1: Aberdeen Study versus Labour Force Survey: age left full-time education

	MEN			**WOMEN**		
	Aberdeen	LFS 01/02 (1) All UK	LFS 01/02 (2) Born in Scotland	Aberdeen	LFS 01/02 (1) All UK	LFS 01/02 (2) Born in Scotland
<=14	1.40	2.75	2.03	1.03	2.80	1.25
15	40.98	30.02	39.56	37.67	31.02	40.15
16	21.01	29.42	25.61	23.54	28.13	24.35
17	8.31	7.89	9.17	12.54	9.61	10.59
18	7.50	8.36	7.50	5.92	9.49	5.32
19+	20.80	21.56	16.13	19.40	18.94	18.35
Observations	2,865	19,967	2069	3,093	20,850	2134

Note: The Aberdeen sample consists of the entire sample of individuals replying to the 2001 survey. The LFS 01/02 (1) sample consists of all individuals born between 1950 and 1955 in the last 2 quarters of 2001 and the first 2 quarters of 2002. The LFS 2001/02 (2) sample restricts the LFS 2001/02 (1) sample to all individuals born in Scotland. LFS data is weighted using sampling weights.

Appendix B Table 2: Aberdeen Study versus Labour Force Survey: highest educational qualification

	MEN			WOMEN		
	Aberdeen	LFS 01/02 (1) All UK	LFS 01/02 (2) Born in Scotland	Aberdeen	LFS 01/02 (1) All UK	LFS 01/02 (2) Born in Scotland
None	20.40	21.08	22.71	23.98	26.99	31.50
Other qual.	2.72	9.92	7.14	3.51	9.38	5.41
Low CSEs	0.88	1.56	1.09	4.26	5.74	2.50
O level or equivalent	22.48	23.17	20.42	29.26	23.15	18.85
A level or equivalent	14.75	8.50	11.84	11.38	8.31	13.83
HNC, teaching, etc.	20.05	16.80	20.84	13.58	13.15	15.03
Degree	18.71	18.96	15.97	14.04	13.27	12.87
Observations	2,833	19,888	2069	3,049	20,792	2134

Note: see notes to Appendix B Table 1.

Appendix B Table 3: Labour Force Survey - percent with trade apprenticeship

	MEN		WOMEN	
	LFS 01/02 (1) All UK	LFS 01/02 (2) Born in Scotland	LFS 01/02 (1) All UK	LFS 01/02 (2) Born in Scotland
None	26.17	35.06	5.02	6.86
Other qual.	0.00	0.00	0.00	0.00
Low CSEs	15.00	32.22	2.81	9.14
O level or equivalent	46.06	57.73	6.04	6.85
A level or equivalent	20.18	12.51	3.21	5.13
HNC, teaching, etc.	62.89	71.31	9.16	7.00
Degree	10.21	7.24	3.34	3.16
Total	30.66	37.68	4.84	5.85
Observations	19,790	2,061	20,737	2,130

Note: In the LFS trade apprenticeships are recorded in a separate question and are not included among the qualifications listed above. Numbers show the percentage of individuals with the corresponding level of qualification who also hold a trade apprenticeship. LFS samples as described in Appendix B Table 1.

Appendix B Table 4: Aberdeen Study versus Labour Force Survey: returns to education - gross annual income (Aberdeen), gross weekly pay (LFS)

	Aberdeen		LFS 01/02 (1) All UK		LFS 01/02 (2) Born in Scotland	
MEN						
Years of post-compulsory education	0.072		0.075		0.090	
	(0.004)		(0.001)		(0.007)	
Left ft education at 15 or earlier		-		-		-
Left ft education at 16		0.200		0.176		0.155
		(0.027)		(0.009)		(0.048)
Left ft education at 17		0.266		0.308		0.361
		(0.039)		(0.013)		(0.064)
Left ft education at 18		0.413		0.411		0.356
		(0.041)		(0.014)		(0.076)
Left ft education at 19+		0.458		0.595		0.648
		(0.028)		(0.010)		(0.050)
Observations	2152	2152	27,665	27,665	727	727
WOMEN						
Years of post-compulsory education	0.128		0.123		0.123	
	(0.006)		(0.002)		(0.011)	
Left ft education at 15 or earlier		-		-		-
Left ft education at 16		0.186		0.208		0.096
		(0.035)		(0.012)		(0.062)
Left ft education at 17		0.326		0.367		0.321
		(0.041)		(0.016)		(0.071)
Left ft education at 18		0.368		0.504		0.368
		(0.056)		(0.016)		(0.098)
Left ft education at 19+		0.801		0.829		0.832
		(0.036)		(0.013)		(0.063)
Observations	2316	2316	29,032	29,032	794	794

Note: Cells show least squares estimates of the returns to years of post-compulsory education. Dependent variable is log gross annual income for the Aberdeen sample and the log of the gross weekly wage for the LFS samples. Samples restricted to employees in work at the time of the survey. The Aberdeen sample consists of the entire sample of individuals replying to the 2001 survey. LFS samples as described in Appendix B Table 1.

Appendix C: The Case of Heterogeneous Treatment Effects

In this Appendix we argue that if treatment effects are heterogeneous, then under additional assumptions, the IV estimator described in section 4 will likely approximate the average treatment effect among borderline students.

C.1 Treatment Effects

In principle, treatment effects could vary among students with the same assignment score. In addition, average treatment effects could vary across scores. Other things equal, we might expect treatment effects to be increasing in assignment scores (e.g., if higher-ability students gain more from elite schools), although that need not be the case. For a given score, we might expect treatment effects to be higher for students selected into elite schools than for students not selected into elite schools (e.g., if administrators assign students based on their perceived suitability for elite school). Since the probability of treatment is seen to be increasing in scores in the borderline range (e.g., see Figure 2), then among the treated students, we might expect average treatment effects to be decreasing with scores. For example, nearly all students with scores of 559 are selected, but very few students with scores of 541 are selected. It is possible that, on average, treatment effects are larger among the latter group.

We will assume that for students with scores in the borderline range, the expected treatment effect conditional on the assignment score is uncorrelated with the assignment score:

$$g_1(A_i = a_s) - g_0(A_i = a_s) = \tau \qquad if \qquad M + 1 \leq s \leq R$$

This does not imply that treatment effects are the same for all students with a given score, or that average treatment effects among the treated students with score a equal average treatment effects among the treated students with score

a'. Instead, it is consistent with the idea that students have imprecise control of their scores, and that in the borderline range, the score is uncorrelated with the treatment gain. The extreme case in which it holds is if scores are randomly assigned within the borderline range.

Under this and some other assumptions, we will argue that the IV estimator will likely approximate τ, the average treatment effect among borderline students. Again, note that this is not the average treatment effect among borderline students selected into treatment (the treatment on the treated).

Treatment Assignment

Assume that the assignment score A is a scalar with S points of support, such that $A \in \{a_0, a_1..a_S\}$ where $a_s - a_{s-1} > 0$. Assume that $P(D_i = 1|A_i = a)$ is such that:

$$\begin{aligned} P(D_i = 1|A_i = a_s) &= 0 \; if \; s \leq M \\ P(D_i = 1|A_i = a_s) - P(D_i = 1|A_i = a_{s-1}) &> 0 \; if \; M+1 \leq s \leq R \\ P(D_i = 1|A_i = a_s) &= 1 \; if \; s \geq R \end{aligned}$$

where $0 < M < R < S$.

Instrument

Consider the instrument $Z_i = P(D_i \widehat{=} 1|A_i) \in \{0, P(\widehat{a_{M+1}}), P(\widehat{a_{M+2}}), ..P(\widehat{a_{R-1}}), 1\}$ which has $K + 1 = 1 + R - M$ points of support. Assume that the predicted probability of treatment is zero for scores to the left of the borderline range, one for scores to the right of the borderline range and increasing within the borderline range:

$$\begin{aligned} P(D_i = \widehat{1|}A_i = a_s) &= 0 \; if \; s \leq M \\ P(D_i = \widehat{1|}A_i = a_s) - P(D_i = \widehat{1|}A_i = a_{s-1}) &> 0 \; if \; M+1 \leq s \leq R \end{aligned}$$

$$P(D_i = \widehat{1|}A_i = a_s) \quad = \quad 1 \; if \; s \geq R$$

C.2 IV estimator

Given these assumptions, we can write the outcome equation as:

$$\begin{aligned}
Y_i &= g_0(A_i) + [g_1(A_i) - g_0(A_i) - \tau_i]1(A_i \geq a_R) + D_i\tau_i + \{D_i(u_{1i} - u_{0i}) + u_{0i}\} \\
&= f(A_i) + D_i\tau_i + \{D_i(u_{1i} - u_{0i}) + u_{0i}\} \\
&= f(A_i) + D_i\tau^* + \{D_i(\tau_i - \tau^*) + D_i(u_{1i} - u_{0i}) + u_{0i}\}
\end{aligned}$$

where $f(A_i)$ is a continuous function (since $lim_{A_i \to a_R^-} E[Y_i|A_i] = E[Y_i|A_i = a_R] = g_0(a_R))$ and $\tau_i = E[Y_i(1) - Y_i(0)|x_{M+1} \leq X_i \leq x_R]$. Provided we can proxy for $f(A_i)$, IV estimation using Z_i as an instrument will identify τ^*, where τ^* solves $Cov[Z_i, D_i(\tau_i - \tau^*)] = 0$ such that:

$$\tau^* \quad = \quad \frac{Cov(Z_i, D_i\tau_i)}{Cov(Z_i, D_i)}$$

To derive the numerator, we use arguments similar to those used to prove Theorem 2 in Imbens and Angrist (1994). Specifically:

$$\begin{aligned}
E[D_i\tau_i|Z_i = z_m] - E[D_i\tau_i|Z_i = z_k] &= E[D_i(z_m)\tau_i|Z_i = z_m] - E[D_i(z_k)\tau_i|Z_i = z_k] \\
&= E[(D_i(z_m) - D_i(z_k))\tau_i] \\
&= E[\tau_i|D_i(z_m) - D_i(z_k) = 1]P[D_i(z_m) - D_i(z_k) = 1] \\
\frac{E[D_i\tau_i|Z_i = z_m] - E[D_i\tau_i|Z_i = z_k]}{P(z_m) - P(z_k)} &= E[Y_i(1) - Y_i(0)|D_i(z_m) - D_i(z_k) \\
&= = 1, x_{M+1} \leq X_i \leq x_R] \equiv \alpha_{z_m,z_k}
\end{aligned}$$

where $P(z_m) = P(D_i = 1|z_m)$ and similarly for other values of Z.

It can then be shown that if $z_m > z_l > z_k$, then:

$$\alpha_{z_m,z_k} \quad = \quad \frac{P(z_m) - P(z_l)}{P(z_m) - P(z_k)}\alpha_{z_m,z_l} + \frac{P(z_l) - P(z_k)}{P(z_m) - P(z_k)}\alpha_{z_l,z_k}$$

such that:

$$\begin{aligned} E[D_i\tau_i|Z_i = z_m] - E[D_i\tau_i|Z_i = z_k] &= \alpha_{z_m,z_k}[P(z_m) - P(z_k)] \\ E[D_i\tau_i|Z_i = z_m] &= E[D_i\tau_i|Z_i = z_k] + \\ &\quad [P(z_l) - P(z_k)]\alpha_{z_l,z_k} + [P(z_m) - P(z_l)]\alpha_{z_m,z_l} \end{aligned}$$

Generally:

$$E[D_i\tau_i|Z_i = z_k] = E[D_i\tau_i|Z_i = z_0] + \sum_{l=1}^{k}[P(z_l) - P(z_{l-1})]\alpha_{z_l,z_{l-1}}$$

We can now express the numerator as:

$$\begin{aligned} Cov[Z_i, D_i\tau_i] &= E[D_i\tau_i(Z_i - E(Z_i))] \\ &= \sum_{l=0}^{K}\pi_l E[D_i\tau_i|Z_i = z_l](Z(z_l) - E(Z)) \\ &= \pi_0 E[D_i\tau_i|Z_i = z_0](Z(z_0) - E(Z)) + \sum_{l=1}^{K}\pi_l E[D_i\tau_i|Z_i = z_l](Z(z_l) - E(Z)) \\ &= \sum_{l=1}^{K}\pi_l(Z(z_l) - E(Z))[\sum_{k=1}^{l}[P(z_k) - P(z_{k-1})]\alpha_{z_k,z_{k-1}}] \\ &= \sum_{k=1}^{K}\alpha_{z_k,z_{k-1}}[P(z_k) - P(z_{k-1})]\sum_{l=k}^{K}\pi_l(Z(z_l) - E[Z]) \\ &= \sum_{k=1}^{K}\alpha_{z_k,z_{k-1}}[P(z_k) - P(z_{k-1})]\sum_{l=k}^{K}\pi_l(z_l - E[Z]) \\ &= \sum_{k=1}^{K}\alpha_{z_k,z_{k-1}}w_k \end{aligned}$$

where $\alpha_{z_k,z_{k-1}}$ as defined above and:

$$\begin{aligned} w_k &\equiv [P(z_k) - P(z_{k-1})]\sum_{l=k}^{K}\pi_l[z_l - E(Z)] \\ \pi_l &= P(Z = z_l) \end{aligned}$$

A similar argument establishes that the denominator can be expressed:

$$\begin{aligned} Cov(Z_i, D_i) &= \sum_{k=1}^{K} [D_i | Z_i = z_k] \sum_{l=k}^{K} \pi_l [z_l - E(Z)] \\ &= \sum_{k=1}^{K} [P(D=1|z_k) - P(D=1|z_{k-1})] \sum_{l=k}^{K} \pi_l [z_l - E(Z)] \\ &= \sum_{k=1}^{K} w_k \end{aligned}$$

It follows that:

$$\begin{aligned} \tau^* &= \sum_{k=1}^{K} \alpha_{z_k, z_{k-1}} p_k \\ p_k &\equiv \frac{w_k}{\sum_{k=1}^{K} w_k} \end{aligned}$$

where $p_k \geq 0$ and $\sum_{k=1}^{K} p_k = 1$.

We can express this in terms of assignment scores rather than values of the instrumental variable (i.e., functions of the assignment scores):

$$\begin{aligned} \tau^* &= \sum_{s=M+1}^{R} \alpha_{A_s, A_{s-1}} q_s \\ \alpha_{A_k, A_{k-1}} &= E[Y_i(1) - Y_i(0) | D_i(a_s) - D_i(a_{s-1}) = 1, A_i = a_s] \\ q_s &= \frac{v_s}{\sum_{s=M+1}^{R} v_s} \\ v_s &= [P(D=1|a_s) - P(D=1|a_{s-1})] \sum_{l=s}^{R} \theta_l [\widehat{P(a_l)} - E(\widehat{P})] \\ \theta_l &= P(A_i = a_l) \quad if \quad l < R \\ &= P(A_i \geq a_l) \quad if \quad l = R \end{aligned}$$

where $q_s \geq 0$ and $\sum_{s=M+1}^{R} q_s = 1$.

C.3 Interpretation

This expression tells us that the IV estimator will be a weighted average of LATEs. The LATEs are the average treatment effects among students that would be selected at score a but not score $a-1$ (i.e., the marginal students). The weights are a not-easily-interpreted function of scores. To show that this expression will likely approximate τ, we make the following argument:

1. First, it seems reasonable to suppose that, conditional on the score, assignment is positively correlated with treatment effect (i.e., correlated with a student's person-specific treatment gain $u_i(1) - u_i(0)$). This implies that the marginal student assigned with a 541 score will have a large positive value of $u_i(1) - u_i(0)$, the marginal student assigned with a 559 score will have large negative value and the marginal student assigned with a mid-range score will have a value close to zero

2. Second, it seems reasonable to suppose that the weights v_s will be inverse-U shaped, taking a maximum for mid-range scores. For example, if $P(D=1|A)$ is linear in the borderline range, such that $P(D=1|a_s) - P(D=1|a_{s-1}) = p$, and if A is distributed uniformly over this range, such that $\theta_l = \theta$ for $A_i < a_l$, then $v_s - v_{s-1} = -p\theta[\widehat{P(a_l)} - E(\widehat{P})]$. Then, defining a^* such that $\widehat{P(a^*)} = E(\widehat{P})$, this will be positive over $a \in [a_m, a^*]$, negative over $a \in (a^*, a_R]$ and decreasing everywhere.

If the LATEs and the weights are symmetric about a^*, it follows that $\tau^* \sim E[Y_i(1) - Y_i(0)|D_i(a^*) - D_i(a^* - 1) = 1, A_i = a^*] \sim E[Y_i(1) - Y_i(0)|A_i = a^*] \sim \tau$.[1]

[1] There are two reasons why the weights will not be symmetric, although these work in opposing directions. First, since the scores are approximately normally distributed, with mean to the left of the borderline range, there is a higher likelihood of observing scores in the left-hand part of the borderline range (i.e., higher θ_l). Second, we know that $\theta_R = P(A_i \geq a_R)$, which will be larger than all other score-specific densities.

Appendix D: A Model of School Quality with Vocational Training

In this Appendix we present a simple model to support the argument made in section X, that for men, the existence of vocational training likely explains the absence of elite school effects on income. We begin with a baseline model without vocational training. This is adapted from the model that Card and Krueger (1996) used to examine the labor market implications of attending different school systems (our focus is on different types of school within the same system). We then introduce vocational training into the model.

D1: Baseline Model without Vocational Training

Modifying the Card and Krueger (1996) model slightly, we assume that individuals that have reached the compulsory school leaving age choose between leaving school and continuing in academic education for a further A years. We assume that for individual i that attended school type $s \in \{Nonelite, Elite\}$, this choice is made to maximize the following utility function:[2]

$$\begin{aligned} U(y_{is}, A_{is}) &= lny_{is} - f(A_{is}) \\ lny_{is} &= \theta_i + \theta_s + b_s^A A_{is} + u_{is} \\ f(A_{is}) &= \gamma_s^A c_i A_{is} + \frac{k}{2} A_{is}^2 \end{aligned}$$

where y_{is} is annual earnings, θ_i is person-specific ability and c_i is the person-specific cost of academic education. We make the standard assumption that

[2]Card and Krueger (1996) assume that:

$$\begin{aligned} U(y_{is}, E_{is}) &= lny_{is} - f(E_{is}) \\ lny_{is} &= a_i + b_s E_{is} + u_{is} \\ f(E_{is}) &= c_i E_{is} + \frac{k}{2} E_{is}^2 \end{aligned}$$

where E is total years of education (including compulsory school years).

$Cov(\theta_i, c_i) < 0$. The remaining parameters capture the effect of school type on the productivity of the (compulsory) years spent in school (θ_s), the return to additional years of academic education (b_s^A) and the cost of additional years of academic education ($\gamma_s^A c_i$).

It seems reasonable to allow the return to additional schooling to depend on the type of school attended: as Card and Krueger (1996) noted, a high-quality education may improve a student's ability to benefit from additional education. There are two reasons why it seems reasonable to allow the costs of additional education to depend on school type. First, since some of the post-compulsory education that took place in our setting occurred within the elite schools (i.e., students from non-elite schools had to transfer in), this may have created additional costs for non-elite students. Second, more generally and more plausibly, while the majority of elite school students stayed in academic education, the majority of non-elite school students did not, such that it might have been less costly for elite students to comply with default behavior than for non-elite school students to defy it (e.g., because of the costs of being separated from friends).

Maximization reveals the optimal schooling choice to be $A_i^* = max\{\frac{b_s^A - \gamma_s^A c_i}{k}, 0\}$ and maximized utility to be:

$$U(y_{is}, A_{is}^*) = \theta_i + \theta_s + max\{\frac{(b_s^A - \gamma_s^A c_i)^2}{2k}, 0\} + u_{is}$$

Proposition 1 summarizes three implications of this model.

Proposition 1

Assume the following conditions hold:

C1: The returns to academic education are higher for students that attended an elite school ($b_E^A > b_N^A$).

C2: The cost of academic education is lower for students that attended an elite school ($\gamma_E^A < \gamma_N^A$).

In that case:

1. There is some cost cutoff below which all individuals will pursue some academic education and above which no individuals will pursue any academic education. Among the students that pursue academic education, the length of academic education is decreasing in cost.
2. Elite school students will pursue more post-compulsory education.
3. Elite school students will obtain higher wages.

Proof

The first claim follows from inspection of the expression for A^*. The second follows from this expression and the assumption that $E[c_i]$ is the same for elite and non-elite students (among the borderline students). The third follows from substituting this expression into the equation for wages.

D.2: Vocational Training

We introduce vocational training by allowing students to choose between two post-secondary tracks: academic and vocational. Conditional on choosing the vocational track, we assume students solve a maximization problem similar to the one presented above, but with parameters b_s^A and γ_s^A replaced with parameters b_s^V and γ_s^V. We make the following assumptions on these parameters:

A1: $b_s^A > b_s^V$

A2: $\frac{b_s^A}{\gamma_s^A} < \frac{b_s^V}{\gamma_s^V}$

The first will ensure that the lowest-cost individuals (in expectation the most-able individuals) will choose academic training. The second will ensure that students on the margin of choosing vocational training over leaving school without pursuing any education will prefer vocational training to academic education.[3]

Proposition 2

1. Given assumptions A1 and A2, schooling decisions can be characterized by two cutoffs c_L and c_M. Students with $c_i < c_L$ will pursue academic education, with the length of academic education decreasing in cost; students with $c_L < c_i < c_M$ will pursue vocational training, with the length of vocational training decreasing in cost; students with $c_i > c_M$ will leave school without pursuing any vocational training or academic education.

2. An increase in b_s^A or a decrease in γ_s^A will increase the fraction of students that pursue academic education and decrease the fraction that pursue vocational training, with the fraction that leave school without pursuing any vocational training or academic education unchanged.

3. An increase in b_s^V or a decrease in γ_s^V will increase the fraction of students that pursue vocational training and decrease the fraction that leave school without pursuing any vocational training or academic education.

Proof

The proposition can be proved with reference to Figure 1. In particular, we

[3] It seems plausible to suppose the return to vocational training is lower than the return to academic education, since vocational training can be thought of as a combination of education and unskilled work. It seems plausible to suppose that the cost of vocational training is lower than the cost of academic education since vocational training pays a training wage.

can show that $c_M = \frac{b_s^V}{\gamma_s^V}$ (i.e., the type indifferent between vocational training and leaving school) and we know that $U(V_i^*; c_M) = 0$ while $U(A_i^*; \frac{b_s^A}{\gamma_s^A}) = 0$, where $c_M > \frac{b_s^A}{\gamma_s^A}$. We know that $U(A_i^*; 0) = \frac{(b_s^A)^2}{2k} > U(V_i^*; 0)$ and we can show that $U(A_i^*; c_i)$ and $U(V_i^*; c_i)$ cross at most once over the range $c_i \in [0, c_M]$.[4] The second and third parts of the Proposition then follow from Figure 1.

Proposition 3

If, in addition to conditions C1 and C2 and assumptions A1 and A2, we have the following condition:

C3: The costs of vocational training are higher for elite-school than non-elite school students

then:

1. Students assigned to elite school will pursue more post-compulsory academic education
2. Students assigned to elite school will pursue less vocational training
3. Assignment to an elite school need not increase wages

Proof

The first two claims follow immediately from Figure 2. The expected wage return to attending an elite school can be expressed as follows, where $\triangle_i \equiv lny_{Ei} - lny_{Ni}$:

[4]Otherwise, since the difference between them is continuous, and since it is positive when $c_i = 0$ and negative when $c_i = c_M$, there would be have to be two turning points. The first-order condition for a turning point demonstrates that there can be at most one value of c_i in this range.

$$
\begin{aligned}
E(\Delta_i) &= \int_0^{c_L(N)} [(b_E^A)^2 - \gamma_E^A b_E^A c_i - (b_N^A)^2 + \gamma_N^A b_N^A c_i] f(c_i) dc_i \\
&+ \int_{c_{L(N)}}^{c_{L(E)}} [(b_E^A)^2 - \gamma_E^A b_E^A c_i - (b^V)^2 + \gamma_N^V b^V c_i] f(c_i) dc_i \\
&+ \int_{c_{L(E)}}^{c_{M(E)}} [(b^V)^2 - \gamma_E^V b^V c_i - (b^V)^2 + \gamma_N^V b^V c_i] f(c_i) dc_i \\
&+ \int_{c_{M(E)}}^{c_{M(N)}} [-(b^V)^2 + \gamma_N^V b^V c_i] f(c_i) dc_i
\end{aligned}
$$

It is straightforward to construct examples in which the net effect is negative.[5] The intuition is that assignment to an elite school has ambiguous effects on human capital, increasing it for some (lower-cost) students that would anyway be inclined to academic study and decreasing it for other (higher-cost) students that would have pursued vocational training had they been assigned to the non-elite school.

D.3: Measured returns to education

An obvious question is whether the model can account for any of the other facts presented. We show that it can account for the lower return to academic education measured for men. To see why, note that:

$$
\begin{aligned}
E[lny_i|A_{is}] &= E(\theta_i|A_{is}) + \theta_s + b_s^A A_{is} + b^V E(V_{is}|A_{is}) \\
&= E(\theta_i|A_{is}) + \theta_s + b_s^A A_{is} + b^V (cons + r_{AV} A_{is}) \\
&= E(\theta_i|A_{is}) + \theta_s + (b_s^A + r_{VA} b^V) A_{is}
\end{aligned}
$$

[5]To construct an example in which the net effect is negative, suppose $c_i \sim U[0, \overline{c}]$ where $\overline{c} > c_{M(N)}$, such that:

$$
\begin{aligned}
E(\triangle_i) &= \frac{1}{k\overline{c}} \{ [(b_E^A)^2 - (b_N^A)^2] c_{L(N)} + [(b_E^A)^2 - (b^V)^2](c_{L(E)} - c_{L(N)}) - (b^V)^2 (c_{M(N)} - c_{M(E)}) \\
&+ [\gamma_N^A b_N^A - \gamma_N^V b^V] \frac{c_{L(N)}^2}{2} + [\gamma_E^V b^V - \gamma_E^A b_E^A] \frac{c_{L(E)}^2}{2} - \gamma_E^V b^V \frac{c_{M(E)}^2}{2} + \gamma_N^V b^V \frac{c_{M(N)}^2}{2} \}
\end{aligned}
$$

If $b^V = 0.08$, $b_N^A = 0.1$, $b_E^A = 0.12$, $\gamma_N^V = 0.28$, $\gamma_E^V = 0.35$, $\gamma_N^A = 1.4$, $\gamma_E^A = 1.3$, $\overline{c} = 0.3$, $k = 0.01$ and $\theta_E = \theta_N$, then it is simple to show that $E(\triangle_i) \sim -0.01$.

$$= E(\theta_i|A_{is}) + \theta_s + [b_s^A - \frac{E(V_{is}|A_{is}=0)}{E(A_{is}|A_{is}>0)} b^V] A_{is}$$

The presence of vocational training has two effects on the estimated returns to education. First, it generates the bias represented by the second term in square brackets. It can be seen that this will be zero if $b^V = 0$ (since $E(V_{is}|A_{is} = 0) = 0$), but positive otherwise. If c is distributed uniformly, then:

$$\begin{aligned} E(V_{is}|A_{is}=0) &= [\frac{b_s^V - \gamma_s^V(\frac{c_L+c_M}{2})}{k}]\frac{c_M - c_L}{\overline{c} - c_L} \\ E(A_{is}|A_{is}>0) &= [\frac{b_s^A - \gamma_s^A(\frac{c_L}{2})}{k}] \\ Bias &= b^V[\frac{b_s^V - \gamma_s^V(\frac{c_L+c_M}{2})}{b_s^A - \gamma_s^A(\frac{c_L}{2})}]\frac{c_M - c_L}{\overline{c} - c_L} \end{aligned}$$

Using the same parameters described above, it can be shown that this bias is on the order of 25 percent of the true return to academic education.

Second, vocational training weakens the ability bias generated by the first term $E(\theta_i|A_{is})$. Intuitively, that is because vocational training weakens the correlation between costs (hence ability) and academic education. Both forces imply that the measured returns to academic education will be smaller in the presence of vocational training.