

Clark, Damon; Del Bono, Emilia

Working Paper

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

IZA Discussion Papers, No. 8617

Provided in Cooperation with:

IZA – Institute of Labor Economics

Suggested Citation: Clark, Damon; Del Bono, Emilia (2014) : The Long-Run Effects of Attending an Elite School: Evidence from the UK, IZA Discussion Papers, No. 8617, Institute for the Study of Labor (IZA), Bonn

This Version is available at:

<https://hdl.handle.net/10419/106553>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

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

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

IZA DP No. 8617

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

Damon Clark
Emilia Del Bono

November 2014

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*

Discussion Paper No. 8617
November 2014

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0

Fax: +49-228-3894-180

E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

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

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 positive effects on 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

Corresponding author:

Damon Clark
Department of Economics
UC Irvine
3283 SSPB
Irvine, CA 92617
USA
E-mail: clarkd1@uci.edu

^{*} 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 Bhalotra, 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-

¹In 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.

²To 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.

³Dustmann 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 0.8 years 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 35 percent increase in completed years of post-compulsory education and a 56 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 (almost 60 percent of the control group mean) and doubled the probability of degree receipt. These effects likely reflect the

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.

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 had positive, although imprecisely estimated, effects on income and wages (of 16 percent and 7 percent respectively) and significantly decreased completed fertility (by almost 0.4 children). It is interesting that elite school attendance had these large impacts on the women in our sample. To the extent that they can be attributed to the additional education received by women that attended elite school, these effects are even more interesting. That is because we have little credible evidence regarding the causal effects of education on women’s labor market and fertility outcomes, especially for women born in the 1950s and after.⁴ As Goldin (2006) and Goldin et al. (2006) have argued, changing social norms and contraceptive technology (i.e. the introduction of oral contraception) meant that these women made labor market, marriage and fertility choices in a different environment to that faced by their mothers and grandmothers. It follows that the effects of education on the outcomes of this generation of women may

⁴On labor market outcomes, an obvious point of comparison is the effects of the British compulsory schooling reforms. Yet estimates of the 1947 and 1972 reforms provide conflicting evidence as to whether or not women’s wage increased: Devereux and Hart (2010) estimate zero returns for the 1947 reform; Grenet (2013) estimates much larger effects of the 1972 reform (about 6 percent). On marriage, Lefgren and McIntyre (2006) use the quarter-of-birth instrument and find “suggestive” evidence of small effects. On fertility, Geruso and Royer (2014) find no effects of the second UK compulsory schooling reform on completed fertility, a finding in line with Black et al. (2008) and Monstad et al. (2008) who study compulsory schooling reforms in Norway and the US. Currie and Morretti (2003) examine fertility responses to a different education margin, college in the US, but their focus is on infant health and they have only a crude measure of fertility (parity of a birth conditional on being observed to give birth at least once).

differ from those on the outcomes of earlier generations.

Third, for men, our point estimates suggest that elite school attendance had no effect on income or wages and no effect on fertility or marriage. The former estimates are somewhat imprecise, which may be because our income measure is banded, so we cannot rule out positive effects almost as large as conventional estimates of the return to a year of completed full-time education (i.e., 5-10 percent). Nevertheless, we speculate that effects might be small 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.⁵ 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

⁵The 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 eleven 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).

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 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).⁶ In Aberdeen in the 1960s, there were three elite schools and 15 non-elite schools, three of which

⁶In 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”.

were private^{7,8}

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”).

The assignment procedure was as follows:

- Students with assignment scores below 540 were assigned to a non-elite school.
- Students with assignment scores of 580 or more were assigned to an elite school unless assessed by their Head as “unsuitable” or “doubtful”.
- Students with scores between 560 and 579 were assigned to an elite school provided one of their intelligence scores was at least 112 and they were assessed by their Head as “suitable”.
- Students with assignment score between 540 and 559 (and any remaining students from the 560-579 group) were allocated to the remaining elite school places by the Appeals Subcommittee, which used their test scores as well as additional reports written by the school Head.

⁷In addition to these, there were two special needs secondary schools and a convent.

⁸Appendix 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.

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 less than 540 would be assigned to an elite school; (ii) most students with assignment scores of 560 or above would be assigned to an elite school; (iii) the fraction of students with assignment scores in the range 540-559 assigned to an elite school would be increasing in the assignment score (because a higher score likely meant a higher ranking among the borderline students considered by the Appeals Subcommittee). Our data, discussed in more detail below, are broadly consistent with this hypothesis. First, the distribution of assignment scores is as expected, with mean close to 500 (see Figure 1). Second, as seen in Figure 2, the relationship between school assignment and assignment scores has the expected pattern.⁹ The circles in this graph show the fraction of students attending an elite school for a 10-point interval of the score; the solid line is the probability of attending an elite school as predicted by a regression of elite school attendance on a third-order polynomial in the score and variables expected to predict elite school assignment, namely a dummy variable for scores in the borderline area (540-559), a dummy variable for scores to the right of the borderline area (560-), and interactions of these dummies and the 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 assignment scores in between.¹⁰ The few students with scores below 540 that report attending an elite school may have won an appeal against an initial non-selective assignment.¹¹ The few students

⁹What 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 particular, 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).

¹⁰Appendix 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.

¹¹Grounds for appeal would likely have included the child being unwell on a test day or missing time at school through illness or family circumstances.

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.¹²

Post-secondary options

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

¹²For 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.

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).¹³ The target population consisted of all students 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

¹³Aberdeen 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).

(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 9.¹⁴ 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 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

¹⁴This 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.

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.¹⁵ 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, since the assignment procedure changed in 1966/67, we exclude the one cohort that was subject to this new procedure. Fifth, 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.

Since we lose a large fraction of the sample to survey non-response, an obvious concern is that students assigned to elite schools were more or less likely to respond to the postal survey, and that this biases our estimates. To assess this possibility, Appendix A Figure 1 graphs the relationship between assignment score and survey response. The graph 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. For a more formal as-

¹⁵Grade 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.

assessment, we regressed survey response on a smooth (third-order polynomial) function of the assignment score and the set of score variables predictive of elite school assignment: dummy variables for borderline scores (540-559) and higher-than-borderline scores (560-) and interactions of these dummies and the score.¹⁶ We cannot reject that this second set of variables have no influence on the probability of survey response. As a further test, we checked that for students in grade 7 in 1962 (for whom we know secondary school assignment in March 1964 whether or not they responded to the postal survey), there is no impact of elite school attendance on the probability that individuals reply to the survey.

Some of the students in our sample attended private secondary schools. As seen in the right panel of Appendix A Figure 1, there is a generally positive relationship between assignment scores and private school attendance, but a negative relationship among students with borderline scores. This suggests that while some individuals would have attended private school irrespective of the elite school outcome, others attended private school as a result of failing to gain entry to an elite school. This is especially true for boys.¹⁷ We estimate elite school impacts in the full sample and in samples that exclude respondents that attended private school. In the first case, the elite school estimate implicitly compares the effects of attending an elite school relative to the mix of non-elite and private school counterfactuals that individuals would have attended otherwise. In the second case, it implicitly compares the effects of attending an elite school relative to a non-elite counterfactual. The two sets of estimates turn out to be very similar, hence little hinges on whether or not privately-educated respondents are excluded.¹⁸

We chose to exclude privately-educated respondents from our main analysis for two reasons. First, since private schools and elite schools are likely to share

¹⁶The F statistic (associated p-value) is 0.61 (0.66) for men and 1.03 (0.39) for women.

¹⁷Only 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.

¹⁸Appendix A Table 3b contains a full set of estimates based on a sample that includes private school students. These should be compared to the estimates that appear in Tables 1-3.

similar characteristic (e.g., high-ability peers and more-qualified teachers) we think it is more interesting to consider the effect of attending an elite school relative to a non-elite (public) school. Second, , and as discussed in the next section, the inclusion of privately-educated students changes the relationship between assignment score and elite school assignment in a way which makes it much more difficult to implement one of our estimation strategies (note that private school students are excluded from the sample underlying Figure 2).

Appendix A Table 1b presents descriptive statistics for the “base sample” that includes survey non-respondents and privately-educated students and the “final sample” that excludes them. 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 $A \in \{a_0, a_1, \dots, a_S\}$ 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.

Assuming constant treatment effects, we can write:¹⁹

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

Even if we knew the form of $g_0(\cdot)$, 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., $Cov(D_i, u_{0i}) > 0$).

3.2 IV strategy

To identify the causal effect of attending an elite school, we need variation in D this is uncorrelated with the error term u_0 conditional on the function $g_0(A)$. To fix ideas, suppose we knew that $g_0(A)$ can be represented by a third-order polynomial in A . In that case, we can rewrite the last equation as:

$$Y_i = \beta_0 + D_i\tau + \beta_1 A_i + \beta_2 A_i^2 + \beta_3 A_i^3 + u_{0i}$$

Recall that Figure 2 depicts the probability of elite school assignment predicted using a third-order polynomial in the assignment score and four variables that we expected would be good predictors of elite school assignment: dummy variables for borderline scores (540-559) and higher-than-borderline scores (560-) and interactions of these dummies and the score. The predicted probability fits

¹⁹More 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”).

the actual probability well. Indeed, we can confidently reject the hypothesis that the effects of these four variables are jointly equal to zero in this model.²⁰ The upshot is that these four variables can generate the required variation in D and thereby allow us to identify the causal effects of D . In other words, these four variables are valid instruments that can be used in a two-stage least squares (2SLS) procedure. Jacob and Lefgren (2004) use similar IV methods to estimate the effects of summer school programs in a similar setting (i.e., given an assignment rule that generates a treatment probability that changes sharply through a small range of assignment scores).

Looking at Figure 2, we would expect these four instruments to have considerable explanatory power for D irrespective of the function chosen for $g_0(A)$. However, for the instruments to be valid (i.e., to generate consistent estimates of the treatment effect), $g_0(A)$ must be specified correctly.²¹ To see why, suppose that $g_0(A)$ was assumed constant. If the true function was increasing, then the 2SLS strategy would load this positive underlying outcome-score relationship onto the estimated treatment effect. In other words, the 2SLS estimates might suggest a positive elite school effect when the true effect might be zero. This is exactly the same issue that confronts researchers implementing regression discontinuity (RD) designs, at least those that capture the underlying relationship between the outcome and the assignment variable using “global polynomial” methods (i.e., low-order polynomials - see Lee and Lemieux, 2010).

3.3 Robustness to $g_0(A)$ specification

Like these global polynomial approaches to RD analysis, we also use low-order polynomials to proxy for $g_0(A)$. To ensure that our estimates are not biased by this choice, we implement four robustness tests that are commonly found

²⁰The associated F-statistics are well in excess of the thresholds for instrument relevance suggested by the literature (e.g., Stock and Yogo, 2005).

²¹We do not report overidentification test statistics because, 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 a check that the various instruments identify the same parameter.

in the RD literature. 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, we conduct falsification tests of the “effect” of elite school attendance on pre-determined 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.

3.4 Heterogeneous treatment effects

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 three steps in the argument, the second and third of which closely follow the argument developed by Angrist et al. (1996). First, we show that our 2SLS estimates are likely similar to those that would be obtained if we used a single instrument equal to the predicted probability of attending an elite school (i.e., $P(D_i = 1|A_i)$). Second, we show that under some additional assumptions, this estimator would identify a weighted average of score-specific local average treatment effects (LATEs): $E[Y_i(1) - Y_i(0)|D_i(a_s) - D_i(a_{s-1}) = 1, A_i = a_s]$.²² These LATEs capture the average effect among a particular

²²This 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

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. Third, we argue that this weighted average of score-specific LATEs would likely approximate the average effect among borderline students.

3.5 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 continuous. 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”. 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 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

for the typical borderline student.

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.²³ One disadvantage is that extrapolation may generate non-robust estimates (as discussed by Angrist and Rokkanen, 2012), particularly if the outcome-score relationship changes sharply through the borderline interval.²⁴²⁵ A second disadvantage is that even outside of the borderline range, the treatment probability is not exactly zero or one. This means that our estimates of $g_1(A_i)$ and $g_0(A_i)$ will conflate the true relationship between the outcome and the assignment scores (assuming this is specified correctly) with the changing probability of treatment outside the borderline range (assuming the treatment effect is non-zero). The resulting biases are likely small, since the treatment probability is always close to zero or one outside of the borderline range. Nevertheless, this is the main reason why we prefer our IV approach. It

²³We 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 argue identifies 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. We also generated an RD estimate using the following procedure. First, we estimated treatment effects at every score between 540 and 559 (using the same bandwidth used to estimate the effect at 550). Second, we generated a weighted average of these using as weights the standard errors of these estimates (we also weighted using the number of observations at the extrapolation point). Both sets of estimates are very close to those generated for scores of 550.

²⁴The IV strategy is not guaranteed to perform well in this case, since sharp changes in the outcome-score relationship in this range will be correlated with an increased probability of treatment making our estimates imprecise and sensitive to the specification of the treatment-score relationship (this is something we will check). However, no data is discarded, hence there is no extrapolation involved.

²⁵Angrist and Rokkanen (2012) also propose a conditional independence approach to estimation in an RD context. In our context, conditional independence is unlikely to be satisfied, hence we do not pursue this. For example, if we restrict the sample to students in the borderline area (for whom we have treatment variation) and regress social class on elite school attendance and the predicted probability of elite school attendance, where the probability is predicted using all of the other covariates (primary school, two prior test scores, assignment scores and so on), we find that elite school attendance has a strong “effect” on social class.

is also one reason why we prefer to estimate models on samples that exclude private school students: when they are included, the treatment probability is much lower than one among men with high assignment scores (Appendix A Figure 1 shows that the probability of attending a private school is high and increasing to the right of the borderline range). 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). We choose the bandwidth using a version of the cross-validation procedure proposed by Imbens and Lemieux (2008).²⁶ We also present estimates and confidence intervals for a wide range of bandwidths.²⁷ The bottom line is that the IV and RD approaches generate similar estimates, at least for the main outcomes.²⁸

Kinked Regression Discontinuity Design (KRD)

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. Although our strategy can be viewed as a heavily parameterized KRD strategy, we do not use the KRD procedures advanced in these papers.²⁹ This is because our

²⁶The only difference between the procedure we use and the procedure they propose is that we assess how well different bandwidths predict points that are 10 units away from the data range used. The intuition is that when using a bandwidth of 30 to predict outcomes at score 550 from above, we use data on the range [560,590]. To compare the performance of a bandwidth of 30 with a bandwidth of 40 (for example), we always need to assess how well they predict outcomes at points 10 units away from the data range used.

²⁷We 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 (results available upon request).

²⁸To push the RD approach further, we also generated estimates for every point in the borderline range (i.e., 540, 541,...559) and then took a weighted average of these. These estimates were very similar to those reported for the score of 550 (whether we used as weights the standard errors on the estimates or the number of observations associated with the score for which the effect was being calculated).

²⁹We thank an anonymous referee for noting that our strategy can be viewed as a heavily parameterized KRD design.

samples are relatively small and because 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). Although this means that the KRD approach is unlikely to shed much light on elite school effects, we produce KRD estimates for completeness. These can be found in Appendix A Table 7.

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 1 reports estimates of elite school effects on three measures of educational attainment. Our main measure (top panel) 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 A-level equivalent qualifications obtained at age 17 (middle panel) and degree or higher level qualifications (bottom panel).³⁰ For each of these outcomes we report least

³⁰In a previous version of this paper we analyzed a larger number of qualifications. To streamline the analysis we focused on A levels and degrees. Apart from O levels (for which

squares estimates (columns 1-2 and 5-6 for men and women, respectively) and 2SLS estimates (columns 3-4 and 7-8). In columns 1 and 5 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 2, 4, 6 and 8 we also include a set of covariates that take into account individual differences in demographic characteristics, socio-economic characteristics, and previous attainment.

Table 1 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). After controlling for covariates and a flexible function of the assignment score, this is 1.4 years. The 2SLS estimates are smaller than this, although statistically not different, and in the model including covariates (column 4) the estimated effect is 1.21 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 1.99), our preferred estimate (2SLS without covariates) implies that elite school attendance increased years of post-compulsory education by almost 60 percent.³¹

The results for women are qualitatively similar. As reported in Table 1, on average, women that attended elite schools completed around 2.9 additional years of full-time education (column 5). After controlling for assignment scores and other covariates, the estimated effect is 0.7 years. The 2SLS estimates are consistent with these numbers, ranging from 0.81 (column 7) to 0.67 (column 8) years. Since the control mean for women is 2.3 years, the implied effect size is about 35 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 3 and 4 and columns 7 and 8 of

we found no effects), these are arguably the better-known qualifications.

³¹This 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.

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).³² Third, our analysis passes a “falsification test”. Specifically, our estimates suggest that elite school attendance has no “effect” on pre-determined outcomes. This can be seen in Figure 4, which depicts the elite school “effect” on a pre-assignment test score (taken at age 9). It can also be seen in the first panel of Appendix A Table 2, which reports the estimated effect on years of post-compulsory schooling predicted using the extensive set of background characteristics available in our data.³³ Fourth, our estimates are similar to those derived from models that use different polynomial specifications (see Appendix A Table 3a).

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. The vertical line corresponds to the optimal bandwidth. These estimates are reported in Appendix Table 3a (columns 4 and 8 for men and women, respectively), and as one can see they are generally very close to and never statistically different from the 2SLS estimates.

The remaining panels of Table 1 report education impacts on A levels and degrees. For men, note that the control mean for university degree (bottom panel) is low (14 percent). This suggests that few borderline students 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 11 percentage points) and degree receipt (by 16 percentage points). The latter effect represents an increase of more than 100 percent of the control group mean. Since the academic track entails at least six years of

³²As 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.

³³By contrast, OLS estimates suggest an impact of elite school attendance on the predicted outcome. This highlights the importance of dealing with omitted variables bias in the OLS estimates.

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 almost one year of completed full-time education. For women, elite school attendance is shown to increase the probability of obtaining A-level qualifications or equivalent by around 23 percentage points - or 56 percent of the control group mean. There is no apparent impact on degree receipt.

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 2 reports estimates of elite school effects on gross annual income, employment, and imputed gross hourly wages. 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, and is expressed in banded intervals. We impute gross hourly wages using occupation-specific means of hourly gross wages from the *New Earnings Survey* (NES).³⁴

Given the large effects on education documented above, we might expect

³⁴In 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.

elite school attendance to generate significant impacts on income and wages. In fact, both our adjusted OLS and 2SLS point estimates suggest that for men, elite school attendance had no significant effects on annual income, hourly wages or the probability of employment, although neither can rule out positive effects commensurate with typical estimates of the returns to a year of education (e.g, 6% - Card, 2001).³⁵ In the next section we discuss what might account for this puzzling finding. For now, we note that the 2SLS estimates pass all of our robustness checks. In particular, models with covariates generate similar estimates (columns 3 and 4), the model appears to fit the data well (Figure 5), falsification checks do not reveal any impact of elite school attendance on measures of income predicted using control variables (Appendix A Table 2) and the RD estimates tell a remarkably similar story across a wide range of bandwidths (Appendix A Figure 5 and Appendix A Table 3a). Although the estimates appear to be somewhat sensitive to the order of the polynomial in assignment scores (Appendix A Table 3a), they remain statistically indistinguishable from zero.

By contrast, as reported in Table 2, we find that elite school attendance increased women’s annual income by around 16 percent. Although the estimate is not very precise, and is somewhat sensitive to the order of the polynomial (Appendix A Table 3a), it is robust to the inclusion or exclusion of the covariates and Figure 5 shows a good data fit. Only the RD estimates are suggestive of potentially smaller effects (at larger bandwidths), but even in this case the effects are still on the order of 8 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 7 percent (Table 2, bottom panel). On the other hand, this alone

³⁵Income 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.

cannot account for the large increase in annual income, so labor supply impacts seem likely. Although we find no elite school effects on whether women worked at all (middle panel in Table 2), this does not preclude 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 3 suggest no obvious effects of elite school attendance on fertility and marriage outcomes. Since we found that elite school attendance had no impact on labor market outcomes, this is not surprising. One possible caveat to this summary is the suggestion of a positive elite school effect on the number of children (top panel), although the 2SLS estimates are sensitive to the inclusion of covariates and the estimate is much lower when these are excluded. The magnitude of this effect is somewhat sensitive to the polynomial specification, while the RD strategy generates smaller effects (Appendix A Table 3a). It is also interesting that the unadjusted OLS estimates for these outcomes are small and statistically indistinguishable from zero.

For women, the estimates reported in Table 3 suggest that elite school attendance decreased the probability of having any children and decreased total fertility, although they do not reveal any obvious effect on the probability of being married or on the timing of fertility (not reported). 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 fertility estimates are somewhat imprecise and this is not surprising when we look at Figure 6. This reveals no obvious relationship between fertility and the assignment score. As a result, the estimated polynomial tracks this outcomes less well than it tracks other outcomes. An obvious question then

is whether our estimates are robust to change the order of the polynomial or using the RD strategy. As seen in Appendix A Table 3a, the estimates are somewhat sensitive to these choices, although all the point estimates suggest fertility reductions of at least 0.2 children. In our view, a reasonable conclusion is that elite school attendance decreased fertility for women and that effects were potentially large.

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.³⁶

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.

³⁶Individuals 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 direction of the women’s results: elite school attendance makes further education more attractive (by reducing its costs or increasing its 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. But even if the direction of the effects is as expected, the magnitude of the estimates is striking. In particular, although the estimates are not very precise, our results suggest income effects of at least 8 percent and effects on completed fertility greater than 0.2 children (taking the RD estimates as a lower bound).

Since elite school attendance increased completed education, these estimates may reflect the causal effect of women’s education on these outcomes. If so, they imply that these causal effects are larger than existing estimates suggest. To be more specific, as the estimated impact of elite school attendance on education is of the order of 0.8 years of schooling (Table 1, and Appendix Table A 3a), our analysis would imply an income return to one extra year of schooling of at least 10 percent, and a fertility response of more than 0.25 children per woman. As noted in the Introduction, studies of the income and fertility returns to education for women are typically based on compulsory schooling reforms. This suggests at least two explanations for the difference between those estimates and ours. First, education may have different effects on the types of students affected by compulsory schooling reforms (i.e., those that would otherwise drop out of school) than on the types of students affected by elite school attendance (i.e., those around the 80th percentile of the age-11 ability distribution).³⁷ Second, they might be explained by other dimensions of the elite schools quasi-experiment. Most obviously, placing girls in

³⁷Another explanation centers on differences in the content of the education provided (e.g., whether schools provided information on birth control), but there is little we can do to assess this. The time period studied is unlikely to explain the difference, since the second UK compulsory schooling reform studied by Grenet (2013) and Geruso and Royer (2014) affected individuals born only a few years after those that we study.

an elite school environment with other high-achieving girls may change their perceptions of women’s role in society and their decisions regarding career, marriage and family. Although it is difficult to make progress here, our analysis underlines the importance of additional research into the effects of women’s education on these outcomes.

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. Note first that our estimates are somewhat imprecise, such that we cannot quite rule out effects comparable to conventional estimates of the return to a year of completed education (5-10 percent). This imprecision may be due in part to the “banded” nature of income, which likely affects the estimates for men more than estimates for women.³⁸ If our point estimates are correct, however, one possible 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

³⁸ Only 3 percent of the women in our sample have income in the top band versus 20 percent for men.

(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. This is consistent with other (OLS) estimates

found in the UK literature (e.g., Grenet 2013; Devereux and Hart 2010), which reveal a gender gap of similar magnitude.

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. Although the estimates are imprecise, the point estimates in Appendix A Table 5 suggest 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. 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, it also shows that, on average men working in professional occupations earn considerably less than men working in managerial occupations.

The comparison with women is interesting. Again, although the estimates are imprecise, Appendix A Table 5 suggests 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.

Although we think that this vocational training hypothesis is both plausible and consistent with the facts, we recognize that it rests on important assumptions and leaves some questions unanswered. One important assumption is that elite school education increases the costs of vocational training. Were this not the case, we would expect elite school attendance to increase men’s income and wages. One additional question is why the gender differ-

ences associated with the extra education induced by elite school assignment are not found in analyses of the extra education induced by the British compulsory schooling laws (e.g., Grenet, 2013). An explanation consistent with the vocational training explanation for our findings here is that for the men affected by the compulsory schooling reforms, there was no displacement of vocational training. In analyses not reported here, we used the UK 2011 Census to examine whether the men affected by 1947 and 1972 compulsory schooling reforms were less likely to pursue apprenticeship training (the most common form of vocational training at the time). We find no evidence that they were. A second and more obvious explanation is that the effects of education might depend on the type of students and the types of courses involved. For example, education might have larger effects on the lower-ability men affected by the compulsory schooling reforms than on the higher-ability men affected by elite school. Alternatively, there might be larger returns to the lower-level courses (e.g., CSEs) pursued by individuals affected by the compulsory schooling laws than to the higher-level courses (e.g., A levels) pursued by individuals affected by elite schooling.

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 0.069. 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 our results.³⁹

39

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 it is consistent with some key facts. Since these facts undermine the most obvious alternative explanations, we view the vocational training hypothesis as the leading explanation for the results we find here.

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 cannot reject the hypothesis 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.⁴⁰ 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.⁴¹ 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 of 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 the late mid-1970s and beyond.

Our final point is that policy-makers might be advised to keep these in-

⁴⁰For 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.

⁴¹See Galindo-Rueda and Vignoles (2004) for an account of the debate surrounding elite schools in the UK during the 1960s.

stitutions 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 Assessments 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.
- Currie, J. and E. Moretti (2003), Mother's Education and the Intergenerational Transmission of Human Capital: Evidence from College Openings, *Quarterly Journal of Economics*, 118(4): 1495-1532.
- Deming, D. (2010), Better Schools, Less Crime?, *Quarterly Journal of Economics*, 126(4): 2063-2115.
- Devereux, P. and D. Hart (2010), Forced to be Rich? Returns to Compulsory Schooling in Britain, *Economic Journal*, 120: 1345-1364.
- 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.

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.

Geruso, M. and H. Royer (2014), The Impact of Education on Family Formation: Quasi-Experimental Evidence from the UK, mimeo, University of California, Santa Barbara.

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

Goldin, C., Katz L.F. and I. Kuziemko (2006), The Homecoming of American College Women: The Reversal of the Gender Gap in College, Journal of Economic Perspectives, 20: 133-156.

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

Grenet, J. (2013), Is Extending Compulsory Schooling Alone Enough to Raise Earnings? Evidence from French and British Compulsory Schooling Laws, Scandinavian Journal of Economics, 115(1): 176-210.

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 University 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.

Jacob, B. and L. Lefgren (2004), Remedial Education and Student Achievement: A Regression-Discontinuity Analysis, *Review of Economics and Statistics*, 81(1): 226-244.

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.

Lefgren, L. and F. McIntyre, 2006, The Relationship between Women's Education and Marriage Outcomes, *Journal of Labor Economics*, 24(4): 787-830.

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.

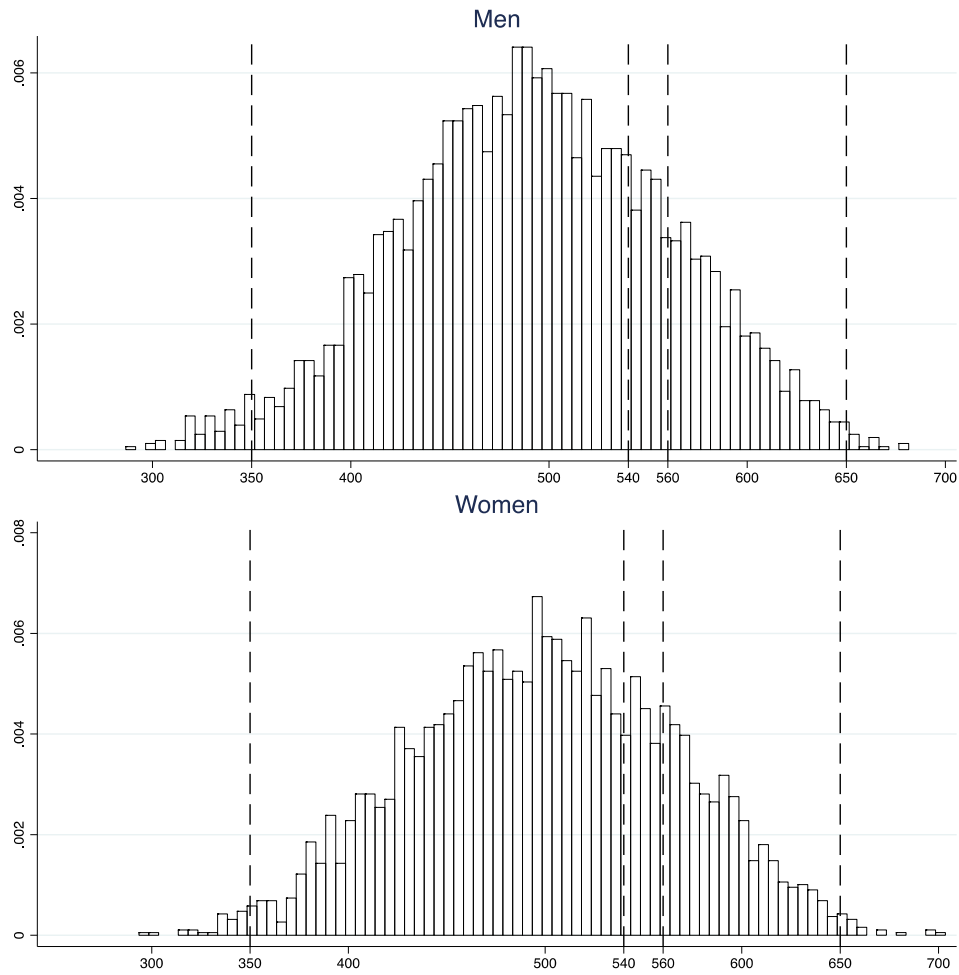
Steedman, H., Gospel, H. and P. Ryan (1998), Apprenticeship: A Strategy for Growth, A Special Report published by the Centre for Economic Performance, London School of Economics and Political Science.

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.

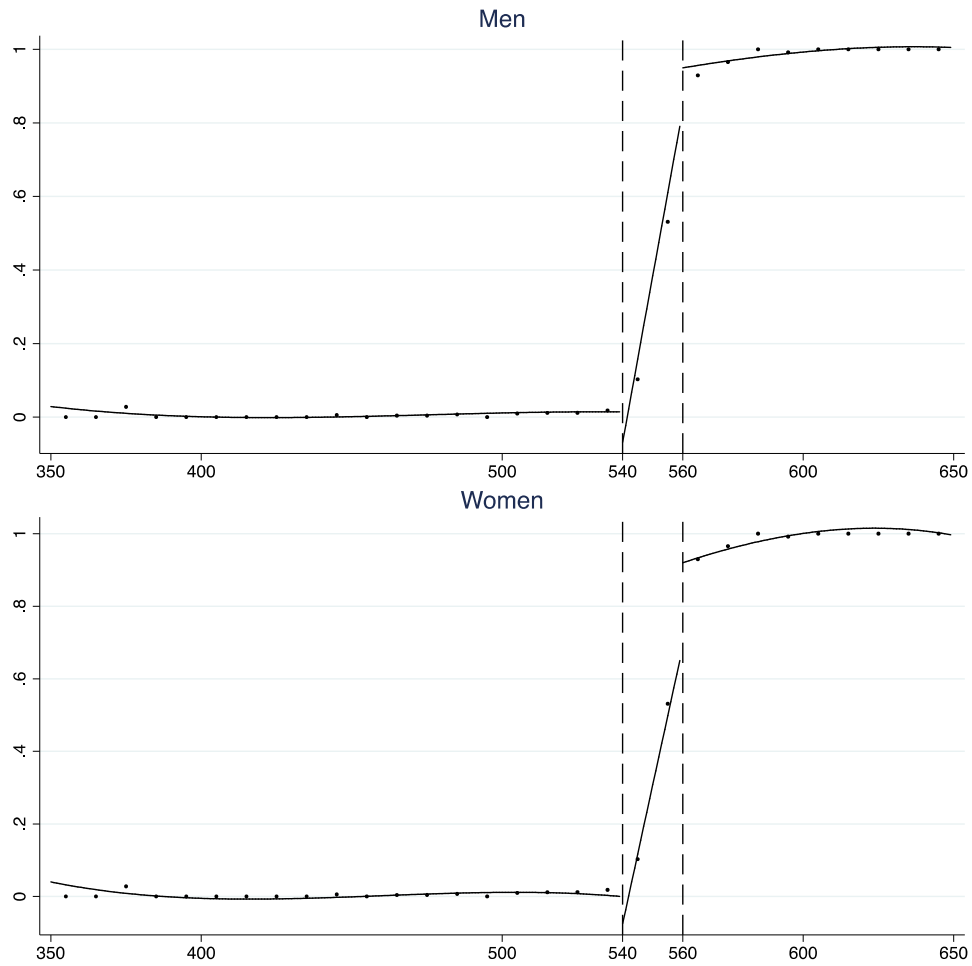
Willis, R. (1973), A New Approach to the Economic Theory of Fertility Behavior, *Journal of Political Economy*, 81: 514-64.

Figure 1: Distribution of assignment scores



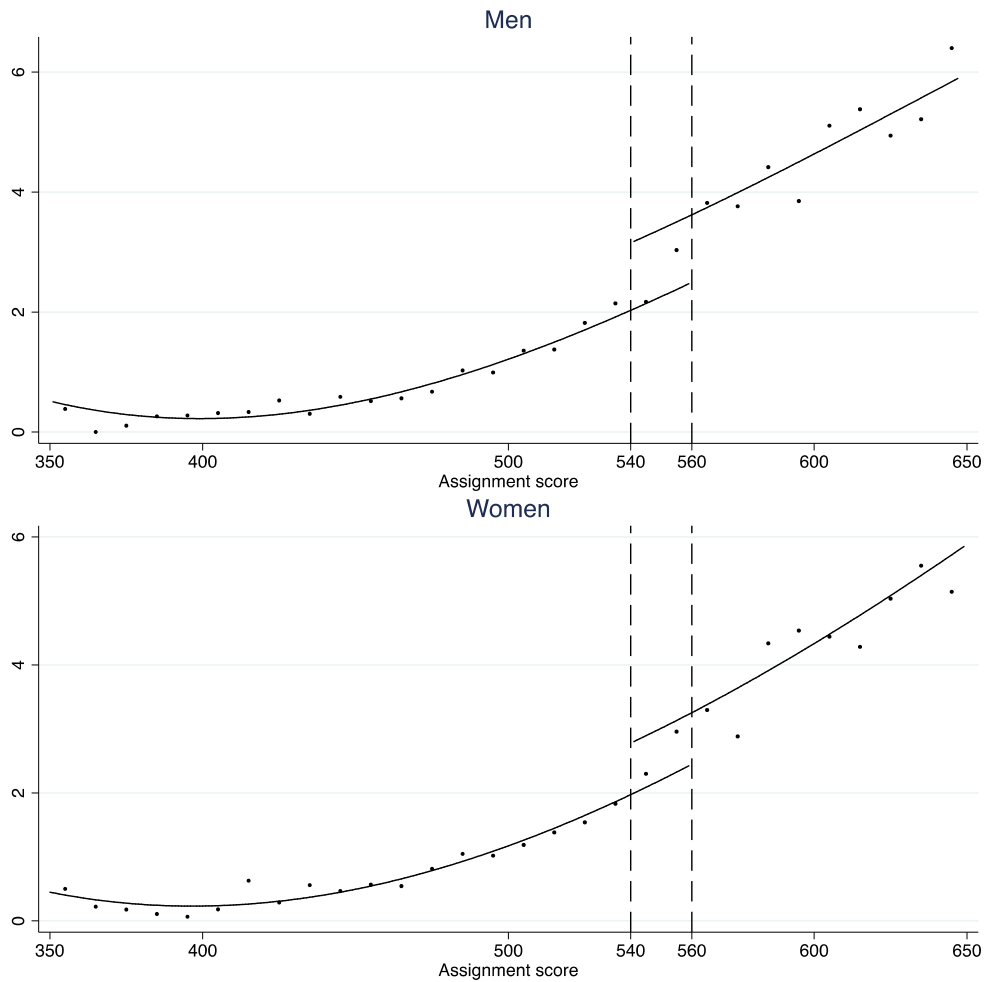
Notes: figures based on "base sample" (see text and Appendix Table 1b). Each bar is drawn over an interval defined by 5 values of the assignment score (350-354, etc). Vertical lines plotted at the following score values: 350-354, 540-544, 560-564, 650-654.

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



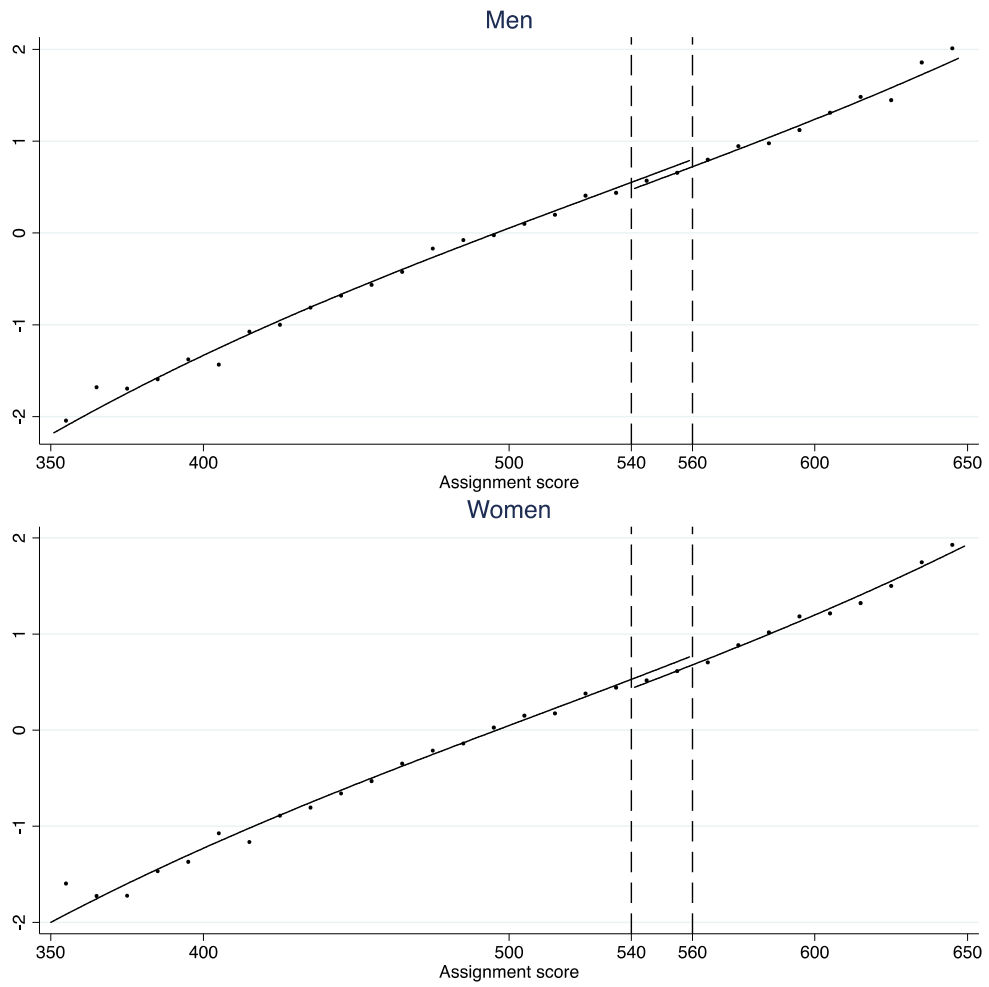
Notes: figure based on "final sample" (see text and Appendix Table 1b). Each circle represents the fraction of students in each cell that attended an elite school. Cells defined over 10 values of the assignment score (350-354, etc). The solid line represents the probability of elite school assignment using the "first stage" model described in the text.

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



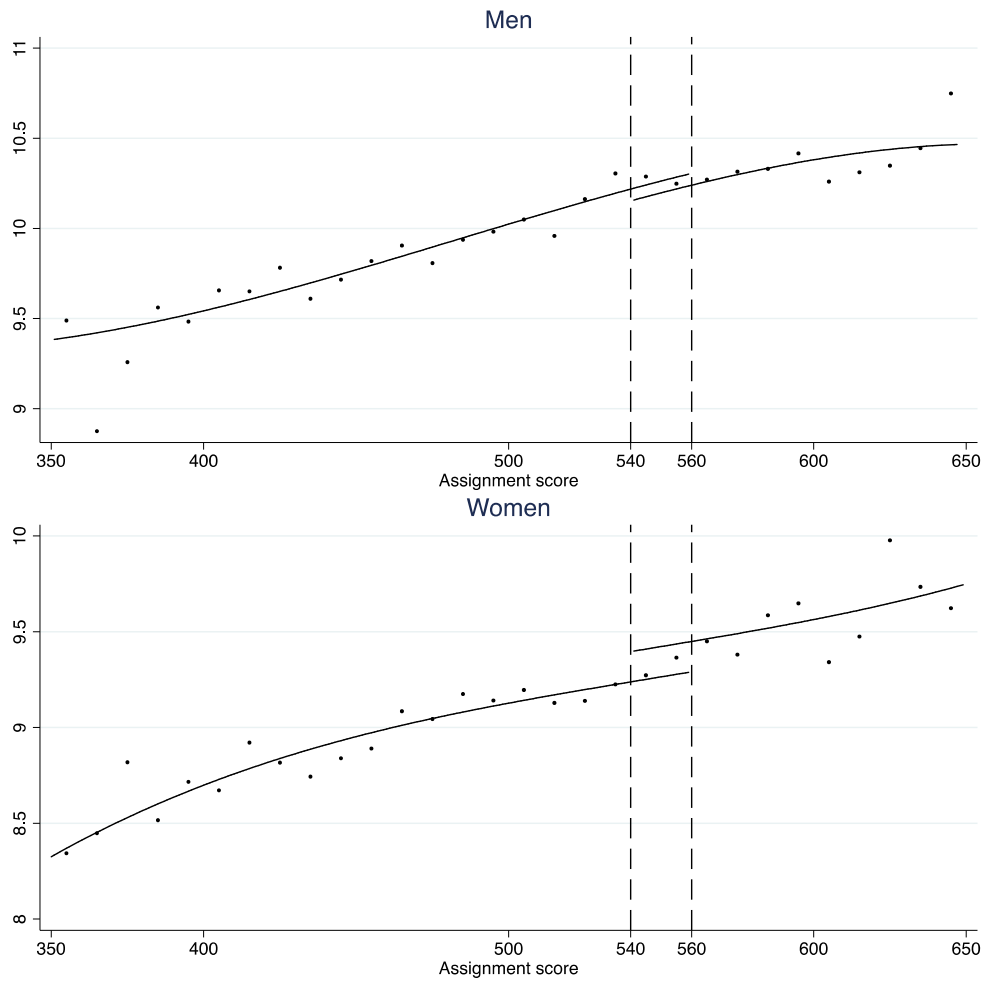
Notes: figure based on "final sample" (see text and Appendix Table 1b). The circles are outcome means corresponding to 10-score intervals (350-359, etc). The solid lines are outcomes predicted by the 2SLS model described in the text. The line on the left-hand side is the predicted outcome given non-elite school assignment, the line on the right-hand side is the predicted outcome given elite school assignment. The gap between these lines corresponds to the 2SLS estimate of the impact of attending an elite school reported in the "2SLS" column of Table 1.

Figure 4: Assignment score and pre-assignment test score at age 9



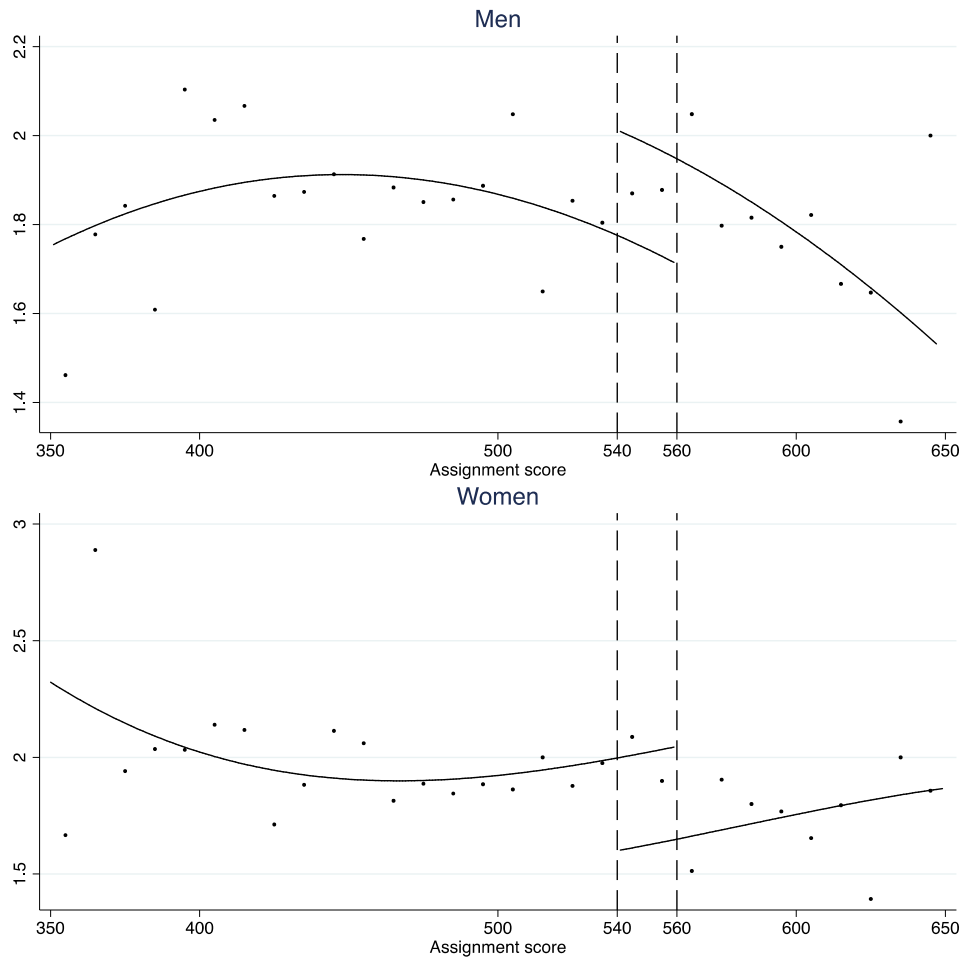
Notes: see notes to Figure 3.

Figure 5: Assignment score and log annual income



Notes: see notes to Figure 3.

Figure 6: Assignment score and number of children



Notes: see notes to Figure 3.

Table 1: Impact of elite school attendance on educational attainment

	Men				Women			
	OLS	OLS+X	2SLS	2SLS+X	OLS	OLS+X	2SLS	2SLS+X
Years of education	3.229 (0.151)	1.433 (0.229)	1.126 (0.279)	1.210 (0.279)	2.914 (0.135)	0.697 (0.165)	0.807 (0.248)	0.672 (0.241)
	Control mean=1.99, N=2072				Control mean=2.32, N=2397			
A levels	0.523 (0.0265)	0.201 (0.0427)	0.107 (0.0592)	0.128 (0.0554)	0.601 (0.0250)	0.243 (0.0381)	0.230 (0.0519)	0.198 (0.0508)
	Control mean=0.46, N=2072				Control mean=0.41, N=2397			
Degree	0.425 (0.0283)	0.179 (0.0411)	0.162 (0.0508)	0.157 (0.0475)	0.325 (0.0245)	0.0416 (0.0317)	0.0332 (0.0409)	0.0251 (0.0411)
	Control mean=0.14, N=2072				Control mean=0.18, N=2397			

Notes: the cells contain estimates (robust standard errors clustered by assignment score in parentheses) of the impact of attending elite school on the relevant outcomes. Estimates are produced for four sets of models. "OLS" (i.e., least squares models without additional covariates); "OLS+X" (i.e., least squares models with a full set of covariates including: dummies for school and grade attended in 1962, father's occupation, and mother's socio-economic status; a linear term of age within grade and fourth-order polynomials of test scores at ages 7 and 9.); "2SLS" (i.e., the 2SLS models described in text, with no covariates except a third-order polynomial in the assignment score); "2SLS+X" (i.e., the 2SLS models with the additional covariates listed above). The models are estimated on the subset of the "final sample" (defined in text and Appendix Table 1b) with assignment scores in the interval [350,650). "Control mean" refers to outcome mean among individuals with borderline assignment scores (540,559) that attended non-elite schools.

Table 2: Impact of elite school attendance on labor market outcomes

	Men				Women			
	OLS	OLS+X	2SLS	2SLS+X	OLS	OLS+X	2SLS	2SLS+X
Log Annual Income	0.422 (0.0322)	0.0296 (0.0560)	-0.0655 (0.0652)	-0.0521 (0.0660)	0.490 (0.0415)	0.222 (0.0812)	0.158 (0.0974)	0.172 (0.101)
	Control mean=10.23, N=2035				Control mean=9.27, N=2344			
Employed	0.0575 (0.0142)	-0.0134 (0.0245)	-0.0352 (0.0331)	-0.0257 (0.0328)	0.0528 (0.0168)	-0.0173 (0.0311)	-0.00911 (0.0370)	-0.0123 (0.0396)
	Control mean=0.94, N=2072				Control mean=0.87, N=2397			
Imputed log hourly wage	0.356 (0.0230)	0.0668 (0.0396)	-0.0279 (0.0511)	-0.00798 (0.0544)	0.356 (0.0250)	0.0625 (0.0369)	0.0741 (0.0490)	0.0679 (0.0458)
	Control mean=2.37, N=1862				Control mean=1.90, N=2110			

Notes: see notes to Table 1.

Table 3: Impact of elite school attendance on fertility and marriage

	Men				Women			
	OLS	OLS+X	2SLS	2SLS+X	OLS	OLS+X	2SLS	2SLS+X
Number of children	-0.0286 (0.0603)	0.223 (0.104)	0.236 (0.157)	0.315 (0.156)	-0.214 (0.0497)	-0.296 (0.104)	-0.397 (0.127)	-0.379 (0.126)
	Control mean=1.86, N=2070				Control mean=2.07, N=2395			
Any children	-0.0185 (0.0213)	0.0195 (0.0385)	0.0113 (0.0529)	0.0277 (0.0536)	-0.0689 (0.0169)	-0.0541 (0.0319)	-0.0834 (0.0415)	-0.0774 (0.0413)
	Control mean=0.86, N=2067				Control mean=0.90, N=2395			
Currently married	0.00979 (0.0266)	-0.0232 (0.0472)	-0.0359 (0.0625)	-0.0296 (0.0640)	0.00769 (0.0193)	0.00100 (0.0374)	0.0517 (0.0471)	0.0460 (0.0486)
	Control mean=0.86, N=2067				Control mean=0.76, N=2392			

Notes: see notes to Table 1.

Appendixes

Appendix A: Additional Figures and Tables

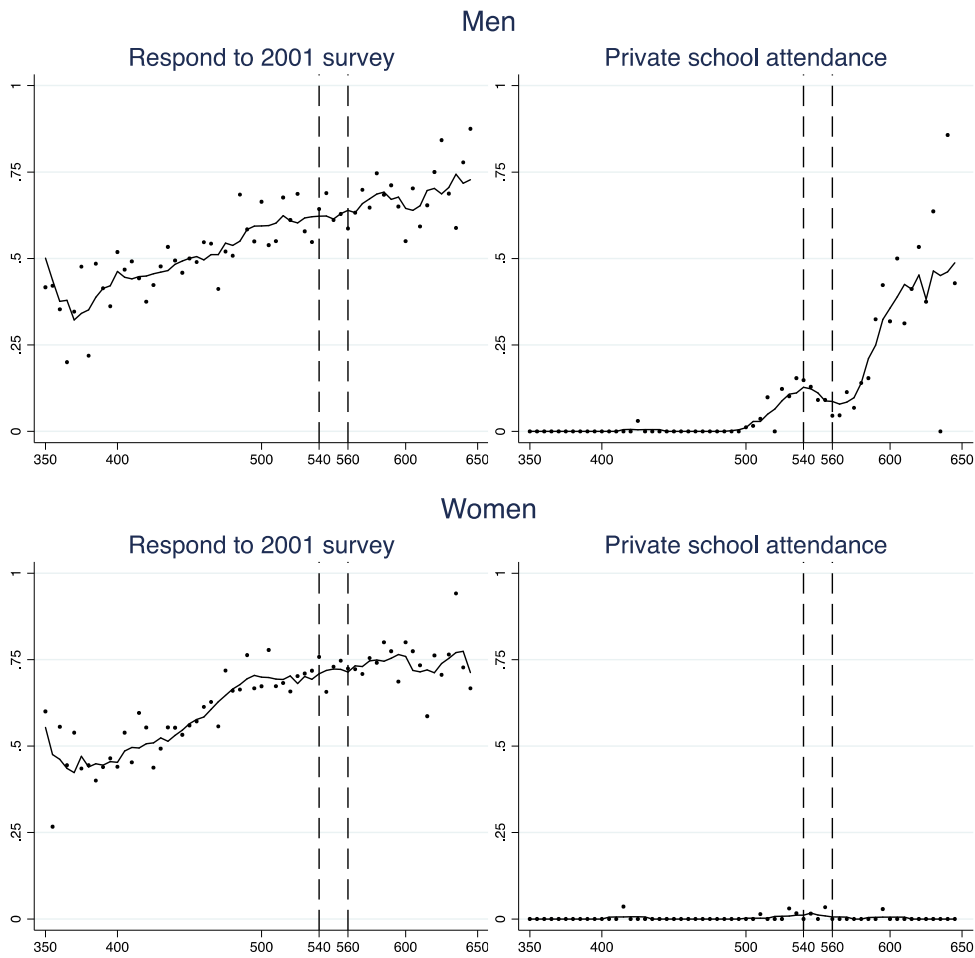
Appendix B: Aberdeen Cohort and Labour Force Survey Comparison

Appendix C: The Case of Heterogeneous Treatment Effects

Appendix D: A Model of School Quality with Vocational Training

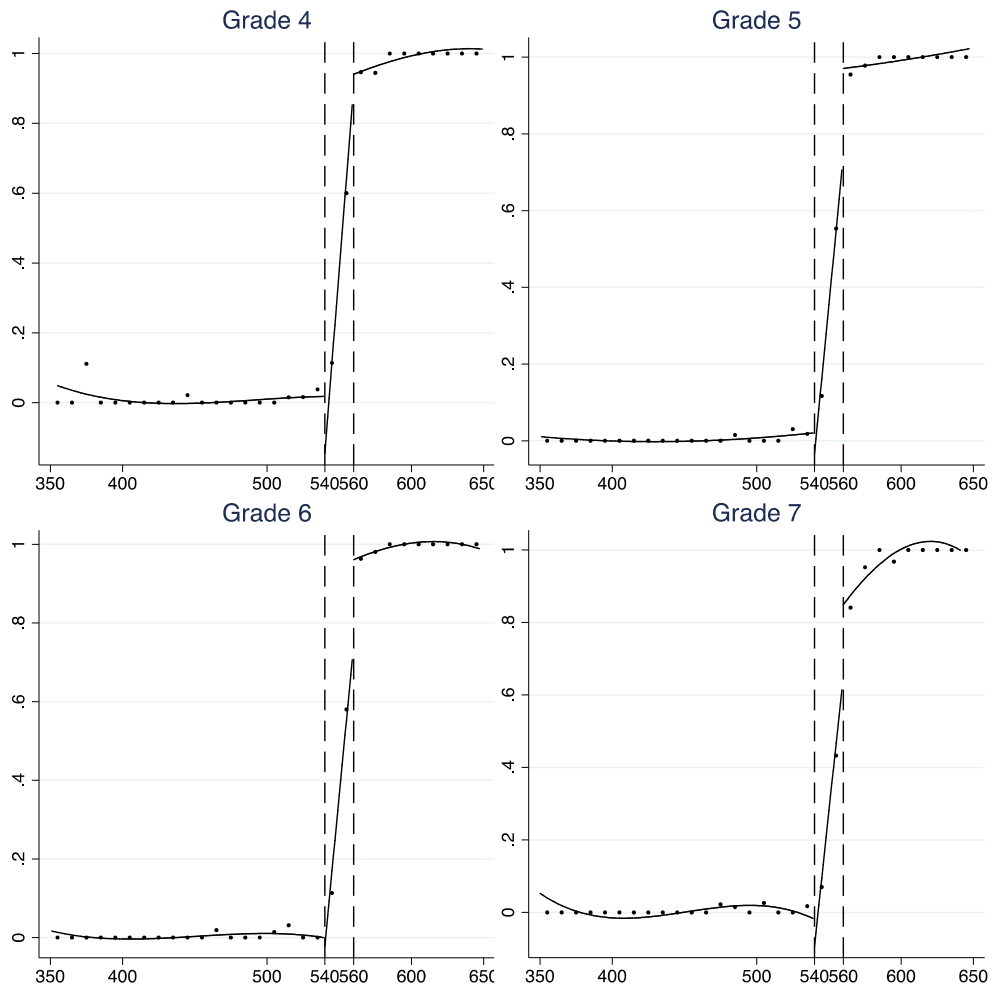
Appendix A: Additional Figures and Tables

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



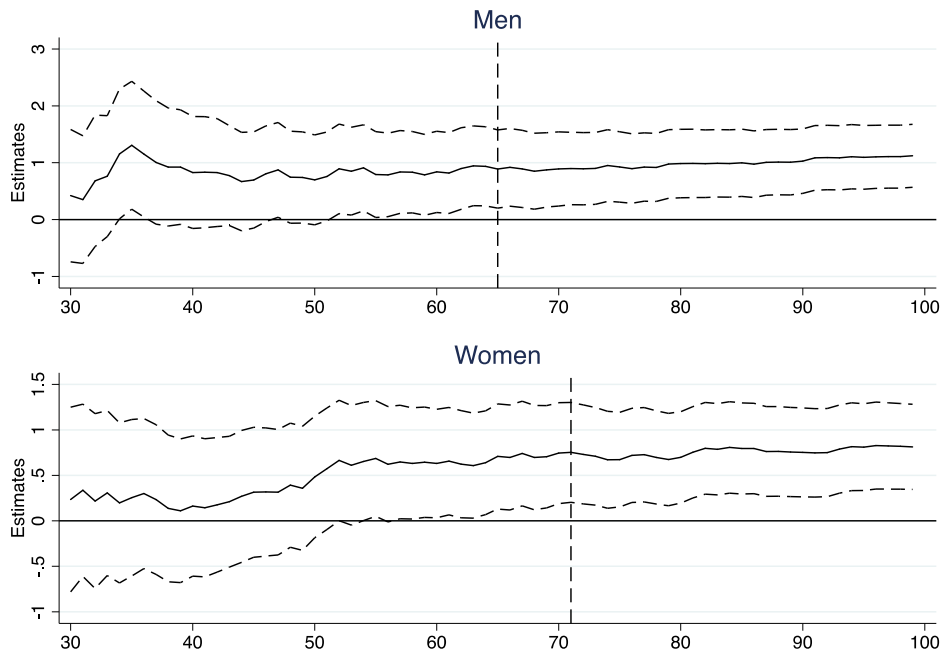
Notes: figures based on "base sample" (see text and Appendix Table 1b). Circles represent cell-specific outcome means. Cells defined over 5-score intervals (350-354, etc). Vertical lines plotted for cells 540-544 and 560-564.

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



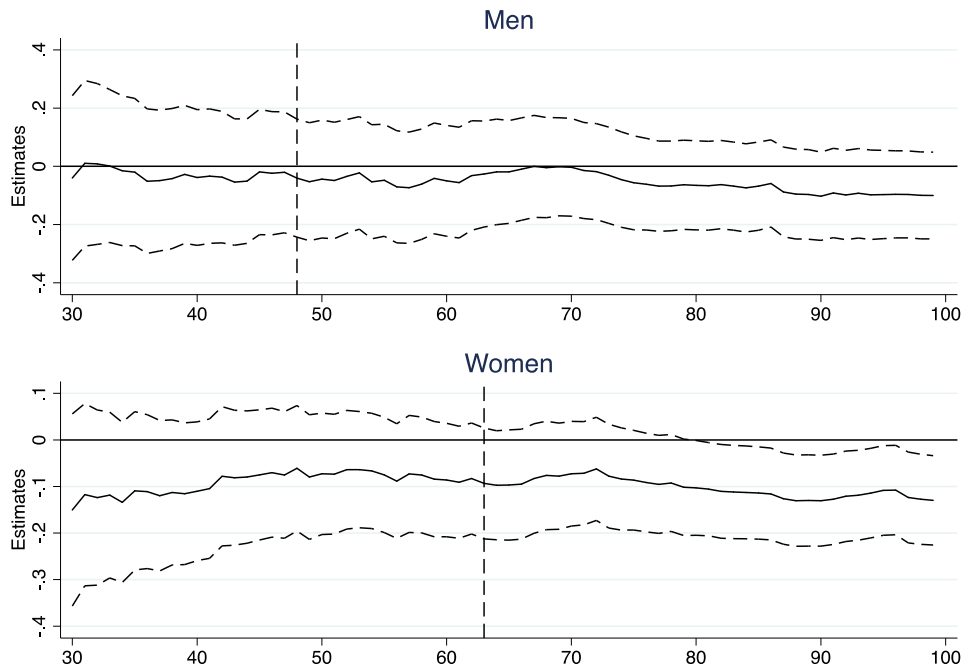
Notes: see notes to Figure 2.

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



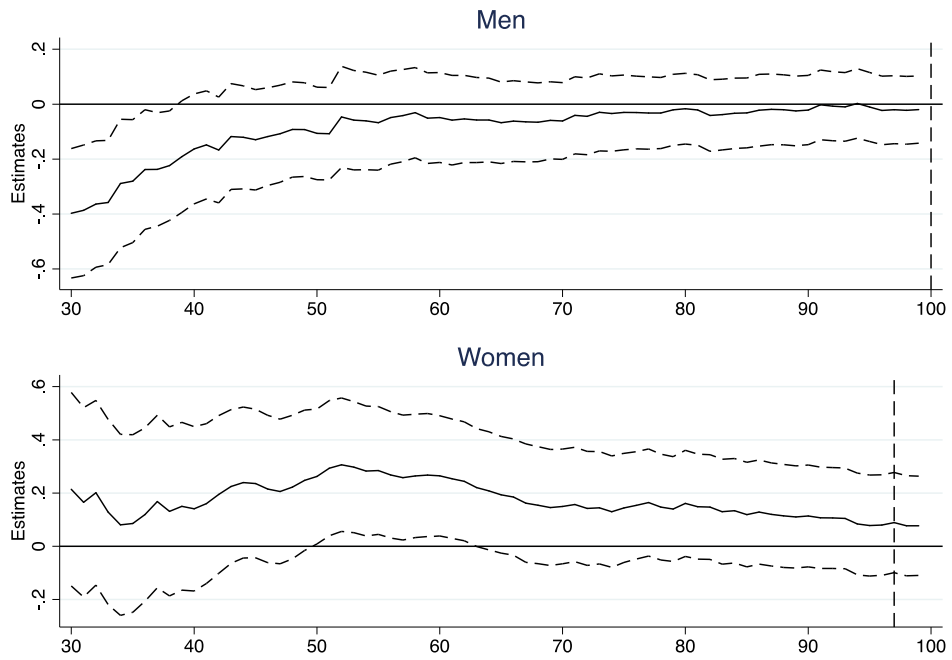
Notes: figures show RD estimates (and associated 95% confidence interval) of the impact of attending an elite school on years of post-compulsory schooling for bandwidths 30 to 100. The vertical line is the optimal bandwidth, as suggested by a cross-validation procedure described in the text.

Appendix A Figure 4: RD estimates - pre-assignment test score at age 9



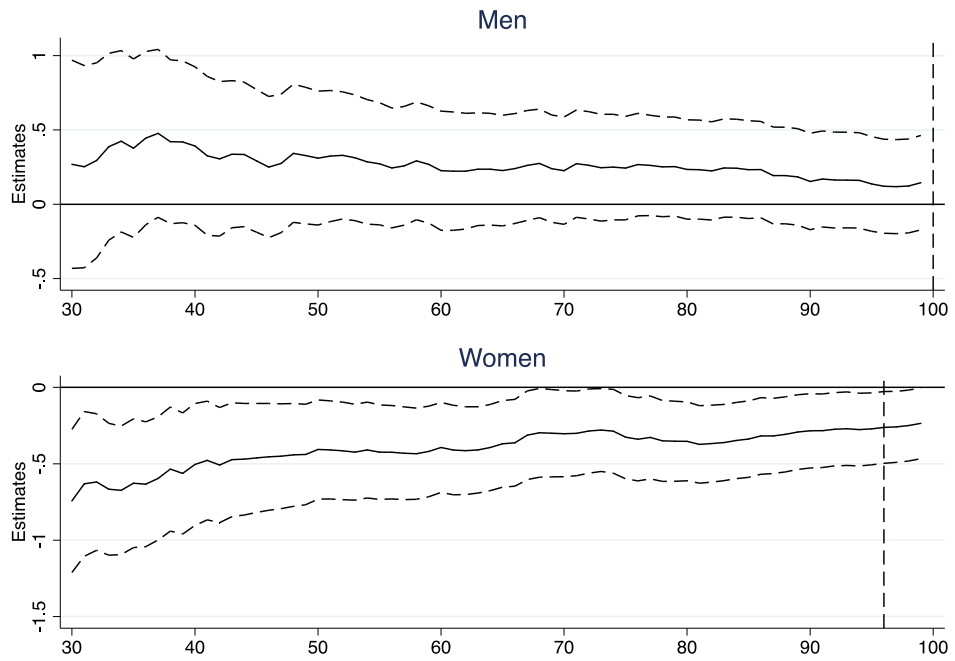
Notes: see notes to Appendix A Figure 3.

Appendix A Figure 5: RD estimates - log annual income



Notes: see notes to Appendix A Figure 3.

Appendix A Figure 6: RD estimates - number of children



Notes: see notes to Appendix A Figure 3.

Appendix A Table 1a: sample selection and subsample definitions

	N	
Original sample	12150	
Did not move during 1962-64 (358)		May have left Aberdeen
Not private primary school (255)		Some of these didn't take test
Not RC primary school (299)		Most took test, but stayed in RC school
Not RC private primary school (81)		Same arguments as previous two
Not primary elite school already (504)		Always stay in elite school
Not primary special school (167)		Some of these didn't take test
Not primary school outside Aberdeen (12)		Attend secondary outside Aberdeen
Not missing grade info (47)		Because we want to split by grade
Not missing assignment info (172)		Need to construct IV
Not missing test scores at ages 7, 9 (309)		Need for covariate controls
Not secondary outside Aberdeen (123)		No secondary info
Base sample	9893	
Not in grade 3 in 1962 (1992)		Different assignment procedure
Not missing 2001 survey responses (3151)		Needed for outcomes
Not private secondary school (250)		Want to focus on elite vs non-elite
Final sample	4528	

Notes: Description of sample selection procedure applied to original data.

Appendix A Table 1b: Descriptive Statistics

	Base sample		Final sample					
	mean	N	mean	N	Non-elite		Elite	
					mean	N	mean	N
Male	0.519	9893	0.464	4528	0.480	3489	0.412	1039
Age (Dec 1962)	9.716 (1.441)	9893	10.203 (1.158)	4528	10.216 (1.160)	3489	10.162 (1.151)	1039
<i>Father's social class</i>		9893		4528		3489		1039
Other (unemployed, disabled, etc.)	0.052		0.048		0.052		0.035	
Unskilled manual	0.17		0.161		0.184		0.085	
Semi-skilled manual	0.148		0.146		0.164		0.085	
Skilled manual, other	0.2078		0.211		0.226		0.163	
Skilled manual, requiring apprent.	0.2521		0.260		0.243		0.316	
Other non manual	0.1083		0.115		0.093		0.189	
Intermediate/Technical	0.0512		0.049		0.033		0.103	
Professional	0.0107		0.010		0.005		0.026	
<i>Grade in 1962</i>		9893		4528		3489		1039
Grade 3	0.205		0.000		0.000		0.000	
Grade 4	0.2039		0.260		0.261		0.253	
Grade 5	0.202		0.251		0.252		0.246	
Grade 6	0.1939		0.247		0.242		0.266	
Grade 7	0.1948		0.242		0.245		0.235	
Test7	-0.020 (0.924)	9893	0.059 (0.887)	4528	-0.182 (0.797)	3489	0.867 (0.671)	1039
Test9	-0.040 (0.955)	9893	0.057 (0.905)	4528	-0.234 (0.758)	3489	1.037 (0.627)	1039
Assignment score (grades 4-7)	495.4 (66.5)	8509	503.3 (62.8)	4528	479.1 (48.1)	3489	584.6 (29.1)	1039
Replied to 2001 survey	0.595	9893	1	4528	1	3489	1	1039
Went to private secondary school	0.040	5885	0	4528	0	3489	0	1039
Elite School	0.216	5885	0.216	4528	0	3489	1	1039

Notes: see Appendix Table 1a for derivation of each sample. Missing assignment scores among some "base sample" students because assignment scores are not available for students in grade 3 in 1962.

Appendix A Table 2: Falsification tests of elite school impacts

	Men			Women		
	OLS	OLS+X	2SLS	OLS	OLS+X	2SLS
Years of education	2.015 (0.0948)	0.225 (0.102)	-0.00445 (0.131)	2.128 (0.0951)	0.179 (0.0871)	0.0626 (0.117)
A levels	0.380 (0.0168)	0.0415 (0.0177)	-0.00124 (0.0228)	0.396 (0.0168)	0.0286 (0.0148)	0.00865 (0.0197)
Degree	0.246 (0.0127)	0.0345 (0.0134)	0.00938 (0.0174)	0.259 (0.0125)	0.0227 (0.0112)	0.0167 (0.0149)
Log Annual Income	0.369 (0.0171)	0.0242 (0.0181)	-0.0100 (0.0221)	0.359 (0.0174)	-0.00905 (0.0145)	-0.0330 (0.0212)
Employed	0.0582 (0.00391)	-0.00274 (0.00486)	-0.00491 (0.00623)	0.0521 (0.00359)	-0.00810 (0.00366)	-0.0149 (0.00501)
Imputed hourly wage	0.271 (0.0124)	0.0163 (0.0122)	-0.0183 (0.0153)	0.279 (0.0123)	0.0110 (0.0106)	-0.00472 (0.0141)
Number of children	-0.0801 (0.00968)	-0.0463 (0.0148)	-0.0443 (0.0200)	-0.105 (0.00851)	-0.0522 (0.0130)	-0.0363 (0.0195)
Any children	-0.0325 (0.00311)	-0.00847 (0.00457)	-0.00496 (0.00601)	-0.0395 (0.00301)	-0.0136 (0.00430)	-0.00777 (0.00631)
Currently married	0.0159 (0.00339)	-0.00408 (0.00500)	-0.00483 (0.00655)	0.0130 (0.00320)	-0.00195 (0.00456)	-0.00966 (0.00588)

Notes: see notes to Table 1. The dependent variables in these models are those predicted by a regression of the relevant outcomes on dummies for father's occupation, mother's socio-economic status, school and grade attended in 1962, relative age within the school-grade and third-order polynomials in the age-7 and age-9 test scores.

Appendix A Table 3a: Impact of elite school attendance: robustness checks I

	Men				Women			
	2SLS			RD	2SLS			RD
	Order 2	Order 3	Order 4		Order 2	Order 3	Order 4	
Years of education	1.136 (0.279)	1.126 (0.279)	0.947 (0.333)	0.867 (0.279) 1159 (65)	0.820 (0.256)	0.807 (0.248)	0.631 (0.298)	0.756 (0.248) 1494 (71)
A levels	0.111 (0.059)	0.107 (0.059)	0.070 (0.076)	0.109 (0.059) 1194 (69)	0.245 (0.053)	0.230 (0.052)	0.126 (0.066)	0.247 (0.052) 1832 (100)
Degree	0.161 (0.051)	0.162 (0.051)	0.155 (0.059)	0.129 (0.051) 1318 (78)	0.021 (0.043)	0.033 (0.041)	0.021 (0.050)	0.0241 (0.041) 1832 (100)
Log Annual Income	-0.061 (0.065)	-0.065 (0.065)	-0.145 (0.085)	-0.022 (0.065) 1522 (100)	0.144 (0.099)	0.158 (0.097)	0.315 (0.128)	0.078 (0.097) 1769 (97)
Employed	-0.036 (0.033)	-0.035 (0.033)	0.023 (0.045)	-0.001 (0.033) 754 (40)	-0.016 (0.036)	-0.009 (0.037)	0.033 (0.052)	0.011 (0.037) 1618 (79)
Imputed hourly wage	-0.022 (0.052)	-0.028 (0.051)	-0.112 (0.062)	0.037 (0.051) 1358 (95)	0.072 (0.050)	0.074 (0.049)	0.079 (0.059)	0.073 (0.049) 1619 (100)
Number of children	0.238 (0.158)	0.236 (0.157)	0.132 (0.198)	0.161 (0.157) 1542 (100)	-0.381 (0.123)	-0.397 (0.127)	-0.581 (0.166)	-0.238 (0.127) 1793 (96)
Any children	0.011 (0.053)	0.011 (0.053)	-0.034 (0.066)	0.035 (0.053) 1314 (78)	-0.088 (0.042)	-0.083 (0.041)	-0.139 (0.052)	-0.058 (0.041) 1816 (98)
Currently married	-0.036 (0.063)	-0.036 (0.063)	-0.080 (0.078)	0.023 (0.066) 1337 (80)	0.057 (0.046)	0.052 (0.047)	0.031 (0.063)	0.028 (0.0434) 1784 (95)

Notes: see notes to Table 1. The 2SLS estimates correspond to the "2SLS" estimates in Tables 1-3. Those estimates used a third-order polynomial hence are reproduced in the "Order 3" column. The "Order 2" and "Order 4" columns report estimates obtained with second- and fourth-order polynomials respectively. The RD estimates are based on the regression discontinuity procedures described in the text. We report sample sizes and bandwidths used (in parentheses) underneath the RD estimates. Bandwidths were chosen using a cross-validation procedure described in the text.

Appendix A Table 3b: Impact of elite school attendance: robustness checks II

	Men				Women			
	OLS	OLS+X	2SLS	2SLS+X	OLS	OLS+X	2SLS	2SLS+X
Years of education	2.862 (0.157)	0.587 (0.190)	0.817 (0.272)	0.942 (0.276)	2.896 (0.135)	0.645 (0.167)	0.775 (0.253)	0.657 (0.244)
	Control mean=2.40, N=2238				Control mean=2.38, N=2406			
A levels	0.469 (0.0269)	0.0908 (0.0314)	0.0927 (0.0550)	0.106 (0.0528)	0.598 (0.0252)	0.233 (0.0381)	0.223 (0.0523)	0.192 (0.0509)
	Control mean=0.52, N=2238				Control mean=0.41, N=2406			
Degree	0.379 (0.0293)	0.0965 (0.0343)	0.149 (0.0485)	0.159 (0.0470)	0.323 (0.0246)	0.0327 (0.0319)	0.0340 (0.0413)	0.0268 (0.0413)
	Control mean=0.19, N=2238				Control mean=0.19, N=2406			
Log Annual Income	0.390 (0.0322)	0.0431 (0.0479)	-0.0934 (0.0687)	-0.105 (0.0692)	0.488 (0.0415)	0.219 (0.0807)	0.151 (0.0975)	0.168 (0.101)
	Control mean=10.24, N=2200				Control mean=9.26, N=2353			
Employed	0.0564 (0.0140)	0.00837 (0.0197)	-0.0363 (0.0336)	-0.0254 (0.0331)	0.0517 (0.0168)	-0.0222 (0.0309)	-0.0104 (0.0369)	-0.0135 (0.0395)
	Control mean=0.94, N=2238				Control mean=0.87, N=2406			
Imputed hourly wage	0.320 (0.0236)	0.0170 (0.0296)	-0.0265 (0.0482)	-0.00911 (0.0500)	0.353 (0.0252)	0.0527 (0.0376)	0.0718 (0.0495)	0.0666 (0.0458)
	Control mean=2.37, N=2003				Control mean=1.90, N=2119			
Number of children	-0.0130 (0.0595)	0.180 (0.0806)	0.243 (0.151)	0.323 (0.150)	-0.213 (0.0498)	-0.293 (0.104)	-0.385 (0.127)	-0.366 (0.127)
	Control mean=1.81, N=2335				Control mean=2.06, N=2404			
Any children	-0.0123 (0.0209)	0.0338 (0.0296)	0.0135 (0.0517)	0.0256 (0.0518)	-0.0678 (0.0170)	-0.0497 (0.0326)	-0.0792 (0.0420)	-0.0737 (0.0417)
	Control mean=0.85, N=2232				Control mean=0.89, N=2404			
Currently married	-0.0742 (0.0468)	0.0351 (0.0647)	0.0601 (0.106)	0.0909 (0.108)	-0.0771 (0.0471)	-0.00684 (0.0926)	-0.0744 (0.115)	-0.0445 (0.117)
	Control mean=0.84, N=2233				Control mean=0.77, N=2401			

Notes: see notes to Table 1. Estimates reported here are based on the same models as in Tables 1-3, but here the sample includes students that attended private schools. Hence attending a non-elite school means attending a non-elite public school or a private school.

Appendix A Table 4: Education impacts by SES

	Low-SES				High-SES			
	OLS	OLS+X	2SLS	2SLS+X	OLS	OLS+X	2SLS	2SLS+X
Years of education	2.826 (0.156)	0.908 (0.215)	0.769 (0.262)	0.791 (0.261)	2.983 (0.129)	1.079 (0.203)	0.918 (0.257)	0.829 (0.256)
	Control mean=2.05, N=2524				Control mean=2.36, N=1945			
A levels	0.538 (0.0291)	0.180 (0.0438)	0.144 (0.0533)	0.128 (0.0547)	0.534 (0.0235)	0.245 (0.0392)	0.182 (0.0512)	0.165 (0.0522)
	Control mean=0.41, N=2524				Control mean=0.46, N=1945			
Degree	0.331 (0.0258)	0.0844 (0.0353)	0.0695 (0.0424)	0.0686 (0.0422)	0.372 (0.0240)	0.108 (0.0312)	0.0997 (0.0442)	0.0672 (0.0442)
	Control mean=0.15, N=2524				Control mean=0.16, N=1945			
Log Annual Income	0.450 (0.0460)	0.144 (0.0691)	0.205 (0.103)	0.136 (0.0951)	0.339 (0.0423)	0.129 (0.0739)	-0.00356 (0.0990)	0.00740 (0.0939)
	Control mean=9.67, N=2469				Control mean=9.71, N=1910			
Employed	0.0759 (0.0181)	0.0110 (0.0302)	0.0407 (0.0396)	0.0570 (0.0428)	0.0281 (0.0153)	-0.0313 (0.0289)	-0.0540 (0.0353)	-0.0719 (0.0357)
	Control mean=0.91, N=2524				Control mean=0.88, N=1945			
Imputed hourly wage	0.310 (0.0299)	0.0115 (0.0429)	0.00502 (0.0573)	-0.0342 (0.0529)	0.314 (0.0222)	0.0860 (0.0378)	0.0484 (0.0480)	0.0550 (0.0477)
	Control mean=2.07, N=2239				Control mean=2.13, N=1733			
Number of children	-0.122 (0.0595)	-0.126 (0.108)	-0.248 (0.149)	-0.225 (0.153)	-0.0929 (0.0468)	-0.0614 (0.0933)	-0.0545 (0.117)	-0.0240 (0.119)
	Control mean=2.05, N=2523				Control mean=1.88, N=1942			
Any children	-0.0469 (0.0206)	-0.0198 (0.0357)	-0.0509 (0.0497)	-0.0391 (0.0516)	-0.0369 (0.0173)	-0.0449 (0.0332)	-0.0590 (0.0441)	-0.0600 (0.0452)
	Control mean=0.89, N=2522				Control mean=0.88, N=1940			
Currently married	-0.0876 (0.0554)	0.0862 (0.102)	0.0204 (0.132)	0.0521 (0.134)	-0.0363 (0.0419)	-0.0270 (0.0650)	0.00398 (0.109)	0.00510 (0.105)
	Control mean=0.79, N=2518				Control mean=0.84, N=1941			

Notes: see notes to Table 1. 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	OLS+X	2SLS	2SLS+X	OLS	OLS+X	2SLS	2SLS+X
Managers	0.129 (0.0255)	0.0275 (0.0469)	-0.0571 (0.0582)	-0.0316 (0.0597)	0.0576 (0.0169)	0.0148 (0.0328)	-0.0274 (0.0371)	-0.0314 (0.0369)
	Control mean=0.37, N=2072				Control mean=0.11, N=2397			
Professionals & asoc. prof.	0.276 (0.0289)	0.101 (0.0501)	0.0379 (0.0616)	0.0504 (0.0645)	0.297 (0.0233)	0.0612 (0.0338)	0.106 (0.0515)	0.111 (0.0507)
	Control mean=0.24, N=2072				Control mean=0.26, N=2397			
Clerical & secretarial	-0.0209 (0.0114)	-0.0229 (0.0228)	0.000653 (0.0264)	0.00127 (0.0282)	-0.0385 (0.0238)	-0.125 (0.0409)	-0.164 (0.0547)	-0.166 (0.0561)
	Control mean=0.08, N=2072				Control mean=0.42, N=2397			
Craft and related	-0.185 (0.0143)	-0.0841 (0.0253)	0.00983 (0.0316)	-0.0208 (0.0342)	-0.0105 (0.00530)	0.00767 (0.00991)	0.0296 (0.0106)	0.0234 (0.0115)
	Control mean=0.13, N=2072				Control mean=0.01, N=2397			
Personal and sales	-0.0256 (0.0160)	-0.00494 (0.0315)	0.00499 (0.0376)	0.00155 (0.0382)	-0.161 (0.0184)	0.0365 (0.0348)	0.0666 (0.0437)	0.0783 (0.0454)
	Control mean=0.08, N=2072				Control mean=0.16, N=2397			
Other	-0.170 (0.0175)	-0.0106 (0.0277)	0.0119 (0.0389)	0.0142 (0.0404)	-0.130 (0.0113)	0.00818 (0.0187)	0.00774 (0.0271)	0.000657 (0.0293)
	Control mean=0.09, N=2072				Control mean=0.04, N=2397			

Notes: see notes to Table 1.

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

	% degree	% A-levels	Mean annual income	Mean log annual income
Men (N=1828)				
Managers	0.23	0.51	38445.41	10.50
Prof. and associate prof.	0.46	0.67	33310.70	10.34
Clerical and secretarial	0.03	0.25	20690.68	9.83
Craft and related	0.01	0.26	23724.19	9.99
Personal and sales	0.02	0.28	21943.75	9.90
Other	0.03	0.14	20606.82	9.81
Women (N=1927)				
Managers	0.23	0.43	24743.02	9.95
Prof. and associate prof.	0.43	0.63	21576.21	9.84
Clerical and secretarial	0.04	0.23	13680.78	9.40
Craft and related	0.04	0.08	9846.15	9.12
Personal and sales	0.03	0.13	9209.89	8.93
Other	0.01	0.04	7112.30	8.73

Notes: Means for those in "final sample" (see text and Appendix Table 1b) that were 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 A Table 7: Kinked Regression Discontinuity Estimates

Years of education		Log annual income		Number of children	
Lower kink	Upper kink	Lower kink	Upper kink	Lower kink	Upper kink
Panel A: Men					
1.412	1.706	-0.321	-0.176	0.287	-0.00934
(0.703)	(0.850)	(0.197)	(0.215)	(0.394)	(0.544)
N=403	N=353	N=399	N=352	N=403	N=352
Panel B: Women					
1.52	3.619	0.44	0.871	-0.918	-1.798
(0.675)	(1.307)	(0.322)	(0.511)	(0.376)	(0.495)
N=496	N=442	N=486	N=435	N=495	N=442

Notes: each estimate corresponds to a kinked regression discontinuity estimate. "Lower kink" refers to estimate based on the kink at assignment score 540 and uses data in the score interval [520,560); "upper kink" refers to estimate based on the kink at assignment score 560 and uses data in the score interval [540,580). Otherwise, the samples are the same as those underlying the estimates in Tables 1-3. The estimate is obtained using 2SLS where the excluded instrument is the interaction of score and a dummy for score greater than or equal to the kink point. Robust standard errors clustered by score in parentheses.

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)	LFS 01/02 (2)	Aberdeen	LFS 01/02 (1)	LFS 01/02 (2)
		All UK	Born in Scotland		All UK	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	2078	3,093	20,850	2149

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)	LFS 01/02 (2)	Aberdeen	LFS 01/02 (1)	LFS 01/02 (2)
		All UK	Born in Scotland		All UK	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)	LFS 01/02 (2)	LFS 01/02 (1)	LFS 01/02 (2)
	All UK	Born in Scotland	All UK	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	2,131

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)		LFS 01/02 (2)	
			All UK		Born in Scotland	
	MEN					
Years of post-compulsory education	0.072		0.069		0.085	
	(0.004)		(0.003)		(0.010)	
Left ft education at 15 or earlier		-				
Left ft education at 16		0.200		0.144		0.053
		(0.027)		(0.020)		(0.064)
Left ft education at 17		0.266		0.308		0.325
		(0.039)		(0.028)		(0.081)
Left ft education at 18		0.413		0.313		0.332
		(0.041)		(0.032)		(0.099)
Left ft education at 19+		0.458		0.539		0.579
		(0.028)		(0.023)		(0.073)
Observations	2152	2152	4,142	4,142	407	407
	WOMEN					
Years of post-compulsory education	0.128		0.110		0.129	
	(0.006)		(0.005)		(0.013)	
Left ft education at 15 or earlier		-				
Left ft education at 16		0.186		0.165		0.086
		(0.035)		(0.027)		(0.086)
Left ft education at 17		0.326		0.293		0.198
		(0.041)		(0.035)		(0.094)
Left ft education at 18		0.368		0.419		0.271
		(0.056)		(0.037)		(0.123)
Left ft education at 19+		0.801		0.755		0.816
		(0.036)		(0.031)		(0.084)
Observations	2316	2316	4,595	4,595	459	459

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:

$$g_1(A_i = a_s) - g_0(A_i = a_s) = \tau \quad \text{if} \quad M + 1 \leq s \leq R$$

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 $A \in \{a_0, a_1, \dots, a_S\}$ where $a_s - a_{s-1} > 0$. Assume that $P(D_i = 1|A_i = a)$ is such that:

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

where $0 < M < R < S$.

Instrument

Consider the instrument $Z_i = P(\widehat{D}_i = 1|A_i) \in \{0, P(\widehat{a}_{M+1}), P(\widehat{a}_{M+2}), \dots, P(\widehat{a}_{R-1}), 1\}$ which has $K + 1 = 1 + R - M$ points of support.¹ 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

¹In practice we use as instruments the four variables discussed in the text. But we could have used as an instrument the predicted probability of elite school attendance. If these four instruments are valid (conditional on the $g_0(A)$ function) then this predicted probability would also be a valid instrument. That is, it would have predictive power for the treatment variable D and it would be uncorrelated with the outcome error conditional on $g_0(A)$ and assuming that function was specified correctly.

borderline range:

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

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 $\lim_{A_i \rightarrow a_R^-} E[Y_i|A_i] = E[Y_i|A_i = a_R] = g_0(a_R)$) 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:

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

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:

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

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:

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

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):

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

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 $v_s - v_{s-1} = -p\theta[P(\widehat{a}_l) - E(\widehat{P})]$. Then, defining a^* such that $P(\widehat{a}^*) = E(\widehat{P})$, 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 $\tau^* \sim E[Y_i(1) - Y_i(0)|D_i(a^*) - D_i(a^* - 1) = 1, A_i = a^*] \sim E[Y_i(1) - Y_i(0)|A_i = a^*] \sim \tau$.²

²There 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

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 $s \in \{None\text{elite}, \text{Elite}\}$, this choice is made to maximize the following utility function:³

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

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.

³Card and Krueger (1996) assume that:

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

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

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 $Cov(\theta_i, c_i) < 0$. 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:

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

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.⁴

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

⁴It 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 $c_i \in [0, c_M]$.⁵ 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 $\Delta_i \equiv \ln y_{Ei} - \ln y_{Ni}$:

⁵Otherwise, 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 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.⁶ 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:

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

⁶To construct an example in which the net effect is negative, suppose $c_i \sim U[0, \bar{c}]$ where $\bar{c} > c_{M(N)}$, such that:

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

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

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

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:

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

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.