An Article Submitted to

*The B.E. Journal of Economic Analysis & Policy*

Manuscript 1917

# Selective Schools and Academic Achievement

Damon Clark*

*University of Florida, National Bureau of Economic Research, and Institute for the Study of Labor, damon.clark@cba.ufl.edu

# Selective Schools and Academic Achievement*

Damon Clark

## Abstract

In this paper I consider the impact of attending a selective high school in the UK. Students are assigned to these schools on the basis of a test taken in primary school and, using data on these assignment test scores for a particular district, I exploit this rule to estimate the causal effects of selective schools on test scores, high school course taking and university enrollment. Despite the huge peer advantage enjoyed by selective school students, I show that four years of selective school attendance generates at best small effects on test scores. Selective schools do however have positive effects on course-taking and, more suggestively, university enrollment, evidence suggesting they may have important longer run impacts.

**KEYWORDS:** selective schools, test scores, instrumental variables

*I thank Liangliang Jiang and Matt Masten for excellent research assistance and Ofer Malamud for help with the university enrollment data. I thank David Card, Ken Chay, Julie Cullen, David Figlio, Jason Fletcher, Caroline Hoxby, Larry Kenny, David Lee, Tom Lemieux, Jens Ludwig and various seminar participants for helpful comments and suggestions. I also thank various former teachers and administrators that talked to me about the schools being studied in this paper, especially Michael Mortimore. Partial funding for this work was provided by the National Academy of Education and the Spencer Foundation.

# 1 Introduction

In several European countries, including France, Germany, Holland, Italy, and parts of the UK, students are tracked into different types of high school based on academic ability as perceived at the end of primary school. In the U.S., this type of academic tracking was traditionally confined to academically selective high schools in large cities, such as the Boston Latin School and New York's Stuyvesant High School. More recently and partly in response to the pressures imposed by school desegregation and accountability requirements, the number of school districts offering academically selective magnet high school programs has grown rapidly. These programs, which provide advanced education in a range of courses such as the International Baccalaureate (IB) or a more specialized field such as science or information technology (IT), can only be followed by students that have satisfied grade point average (GPA) as well as other requirements.[1] These various types of selective high school share three features: academically able students, a specialized academic curriculum, and, usually, highly qualified teachers.[2]

The presence of these schools raises two questions. First, what is their impact on the overall distribution of student outcomes? Second, what is their impact on the students that attend them? An answer to the first question would paint a complete picture of their impacts, but across-district variation in the presence of selective schools is likely confounded by other differences.[3]

An answer to the second question would pin down the partial impacts of selective schools. These partial impacts are an important part of the overall impact; they are also of independent interest because they add to the wider literature on the effects of attending different types of high school. One strand of this literature (begun by Coleman et al. 1982) considers the impacts of attending a private high school in the U.S. Coleman et al. found that private schools improved

[1] Hoffman (2007) reports that, in 2005–2006, there were 2,796 magnet schools out of 97,382 schools in the U.S. Although the number of academically selective magnet high schools is not reported, Planty et al. (2007) report that, in 2002–2003, of the 16,500 high schools in the U.S., 390 offered the IB program. According to figures from the International Baccalaureate Organization, the number of high schools offering the IB program increased by more than 50 percent between 2002 and 2007, from 390 to 555 (personal communication).

[2] There are no formal national requirements for teachers in magnet schools or for teaching IB programs in the U.S. However, because these advanced courses cover college-level material, teachers usually are expected to have at least a bachelor's degree in a related field. The College Board offers similar guidelines for Advanced Placement (AP) teachers. Milewski and Gillie (2002) report that 69.7 percent of AP teachers hold a master's degree compared to 38 percent of teachers overall.

[3] For the UK, this is demonstrated convincingly by Manning and Pischke (2006). Hanushek and Wössman (2006) offer perhaps the best approach to this problem, comparing test score progression across countries.

test scores. More recent evidence points to small test-score effects (Altonji et al. 2005) but larger effects on longer-run outcomes (Evans and Schwab 1995; Neal 1997; Altonji et al.). Another strand of this literature considers the impact of attending a high-achieving public school, defined as one with more able students. Coleman et al showed that test-score growth between tenth and twelfth grade was similar across high- and low-achieving public high schools. Compelling evidence from the Chicago Public Schools suggests that the causal effects of attending high-achieving public schools are, at best, small (Cullen et al. 2005, 2006).

A comparison of the defining characteristics of selective schools, private schools, and high-achieving public schools suggests that a selective school analysis could make two important contributions to this literature. First, because selective schools are defined in terms of characteristics that parents may associate with private and high-achieving public schools—able students and an academically oriented curriculum—an analysis of longer-run selective school effects could help to explain why parents incur costs to enroll their children in private and high-achieving public schools despite evidence pointing to small test-score effects.[4] It seems reasonable to suppose that parents care at least as much about college and labor-market prospects as they do about high school test scores, and there are several channels through which selective schools can affect these outcomes without affecting test scores. Some, such as peer effects on college enrollment, can operate in a world in which the high school curriculum is standardized. Others could be the result of the specialized curriculum offered by selective schools to the extent that this attracts a premium in the labor market or is favored by college admission rules.[5] Second, although the selective school "treatment" goes beyond an able peer group, the difference between peer quality in selective and nonselective schools is extreme, and an analysis of selective school impacts may shed light on whether and how peer effects operate.

At present, we know little about the impacts of attending selective schools, a reflection of the challenges facing researchers attempting to estimate them. First, programs are often heterogeneous and, in many cases, have been established fairly recently. This is especially true in the U.S. where the number of students in magnet programs tripled between 1982 and 1991 (Steel and Levine 1996) and where programs differ by curricula and admissions procedures.[6] Second,

[4] Coleman et al. showed that both offer a disproportionate number of AP courses. More recently, Klopfenstein (2004) documented large differences in the number of AP courses offered by schools.

[5] Betts and Rose (1996) show that the choice of a high school course can have large impacts on earnings. Klopfenstein and Thomas (2006) argue that many colleges upweight the high school GPAs of students with AP credits.

[6] Based on an analysis of the National Educational Longitudinal Study (NELS), Gamoran (1996) found that students in city magnet schools (standalone and school-within-school) enjoyed

researchers often lack a convincing research design, without which it is difficult to disentangle the causal effects of these programs from the superior abilities of the students that attend them. The lotteries used by some U.S. districts provide the ideal tool for dealing with this problem, but lotteries are rarely used outside of the U.S.[7] Instead, high school assignments are typically based on discussions between parents, primary schools, and high schools.[8]

In this paper, I evaluate the impact of attending selective schools in the UK. There are at least three reasons to be interested in the UK setting. First, in contrast to the U.S., selective schools in the UK, known as grammar schools, are well-defined and well established. In particular, they have since at least the 1940s taught an advanced academic curriculum to the top 20 percent of students in districts where they operate. More generally, the UK high school setting is a relatively clean one in that there is no grade retention and no dropout before the end of the grade in which students reach the minimum school leaving age. Because private school enrollment is low (around 7 percent) and nonselective schools are relatively homogenous, the estimated effect of attending a selective high school for four years can be given a straightforward interpretation. Second, in contrast to other European countries, students must, by law, be selected objectively. In general, this means that assignment is almost always dominated by an assignment test taken in the final primary school grade. In the particular district that I study, which, to the best of my knowledge, is the only one for which these assignment test scores are available, assignment was based on an assignment test and a primary school order of merit. I show that this assignment mechanism generates a sharp change in the probability of attending selective school over a narrow range of assignment test scores, and I exploit this sharp change to produce credible estimates of the causal effects of attending selective schools on a range of outcomes. These outcomes include standardized test scores, detailed information on course-taking, and, for a subset of students, university enrollment.

I find that attending a selective school for four years has small effects on standardized test scores. Ordinary least squares estimates suggest effects of around one-third of a standard deviation, close to those found in the previous UK literature (Sullivan and Heath 2004). Estimates that exploit the sharp relationship between assignment test scores and the probability of attendance suggest smaller effects of around one-sixth of a standard deviation with confidence intervals

---

higher science, reading, and social studies test scores than students in urban comprehensive public high schools.

[7] Ballou et al. (2006) studied magnet schools using administrative data from a school district that used lotteries to allocate magnet school places, but they concluded that attrition biases were too large to make strong statements about the effects.

[8] See Brunello and Checchi (2007) for a review of the European evidence. Some papers (e.g., Dustmann 2004) consider the role parents play in this decision in mediating the relationship between family background and later outcomes.

failing to rule out zero effects and consistently ruling out effects larger than the least squares estimates. Despite the absence of any test-score effects, I present evidence consistent with the hypothesis that selective schools improve longer-run outcomes. In particular, I find that attending a selective school increases the probability that students pursue academic courses offered in all types of schools and increases the probability that students pursue academic courses (e.g., classics), which are only offered in selective schools. Given UK rules governing progression into post-compulsory education, these course effects are expected to generate effects on university enrollment, and I provide suggestive evidence that this is the case.

With regard to test scores, my finding of small effects is in line with evidence on the test-score effects of attending private and high-achieving public schools in the U.S. As such, this suggests that these results may generalize to other countries and other settings in which across-school peer differences are even larger. As regards longer-run outcomes, although I find that these were positively affected by selective school attendance, it would not be sensible to generalize from these results to the likely long-run effects of selective school attendance in other settings. Among other things, these will depend on the interaction between the curriculum offered and the downstream rules governing college admissions. A more appropriate and more general conclusion is that other varieties of selective school can also improve longer-run outcomes without improving test scores. This is consistent with the difference between test-score and longer-run outcomes found for U.S. private schools. To the extent that parents associate high-achieving public schools with selective school characteristics, it could partly explain why parents choose these schools despite evidence pointing to small test-score effects.[9]

## 2 Selective High Schools in the UK

Selective high schools in the UK, known as grammar schools, were established by the 1944 Education Act. Before the Act, grammar schools were a class of private schools that offered a number of scholarships in exchange for some financial support from the local school district. After the Act, the grammar schools received all their funding from the district, were not allowed to charge fees, and were required to admit students on the basis of academic potential assessed in grade

[9] One explanation put forward to explain the combination of strong parental demand and weak academic achievement effects is that these schools confer nonacademic benefits (Cullen et al., 2006, show that students who opt out have lower self-reported disciplinary incidences and arrest rates). Another is that academic achievement may improve for students whose parents choose with academic achievement in mind. Hastings et al. (2006) found some evidence for this explanation among elementary school students in the Charlotte-Mecklenburg Schools in North Carolina.

five (at age 10 or 11).[10] At its simplest, this involved all students in a district taking a test (the "Eleven plus") with the grammar places going to the top-scoring students. The nonselective schools, known as modern schools, remained broadly unchanged, the important caveat being that, where they previously educated all but those students who won scholarships to grammar schools, they now educated all students that "failed" the test.[11]

Most of the attention in this paper is focused on the East Riding of Yorkshire, which is the only district for which assignment test scores are available. These data cover several cohorts of students leaving high school in the early 1970s, a few years before the schools in this district were integrated. The East Riding is a school district in the northeast of England, which, at least on observables, seems fairly representative of school districts in the UK (see Web Appendix Table 1a). In particular, the district is close to average size, is only slightly more affluent than average (as measured by district education, socioeconomic status, and housing tenure), and had a fairly typical share of students in selective schools. It is also worth noting that least squares estimates of selective school effects in this district are close to those obtained using nationally representative data from a similar time period (see Sullivan and Heath).[12]

Students were assigned to the district's grammar schools on the basis of two pieces of information: a districtwide assignment test taken in grade five and a primary school order of merit supplied by the primary school head teacher.[13] On the basis of this information, a districtwide order of merit was compiled by district administrators and passed to an "Allocations Panel" composed of teachers and district officers. This panel decided which students could be immediately awarded

[10] Some grammar schools rejected this bargain and became completely independent.

[11] Roman Catholic and Church of England schools are part of the UK public education system and also contain a mixture of grammar and non-grammar schools. By 1965, opposition to this system had grown, and the Labour Government passed a law requiring school districts to move to a "comprehensive" system in which school assignment was based on factors other than academic potential. The transition from tracked to comprehensive systems took place only gradually. Districts that abolished the tracked system typically replaced it with a neighborhood assignment system similar to those commonly used in the U.S.

[12] The dimension on which the district looks least typical is population density. The district is more sparsely populated than most because it includes a large rural area. On balance, this is probably helpful because it means that students would have attended the local selective school. In turn, this means that there was no ranking of selective schools such that students with marginal test scores (used for identification) attended lower-ranked selective schools (confirmed in the data and in personal communications with former teachers).

[13] This primary school order of merit was adjusted so as to enable comparisons across primary schools, although it is not clear exactly how this adjustment was done (Secondary and Further Sub-Committee 1964). For the purposes of the empirical strategy discussed below, the important point is that eligibility was based on two pieces of information, only one of which is in my dataset.

places and passed the names of borderline students to a second "Assessors Panel." This process gave detailed consideration to borderline students, that is, those who had been referred for special investigation and those who showed a disparity between the assignment test score and the primary school order of merit. Once the remaining selective school places had been determined, students were allocated to the school (of relevant type) closest to their parents' home. Parents were able to request a different school, and the district promised to honor these preferences whenever possible.

If test scores and primary rank were perfectly correlated and selective school places were never turned down, one would expect the relationship between the assignment test score and the probability of selective school attendance to be discontinuous. As primary rank and assignment test score are likely to be highly but not perfectly correlated and because some parents may have turned down a selective school place, it is not surprising that the relationship is as presented in Figure 1a below—not discontinuous but with the probability of attendance increasing sharply over a narrow interval. The change is even sharper when plotted separately by sex and cohort because there were in effect separate competitions by sex and cohort. The data used to generate Figures 1a and 1b are described in more detail below. The next section discusses the econometric implications of this assignment procedure in more detail.

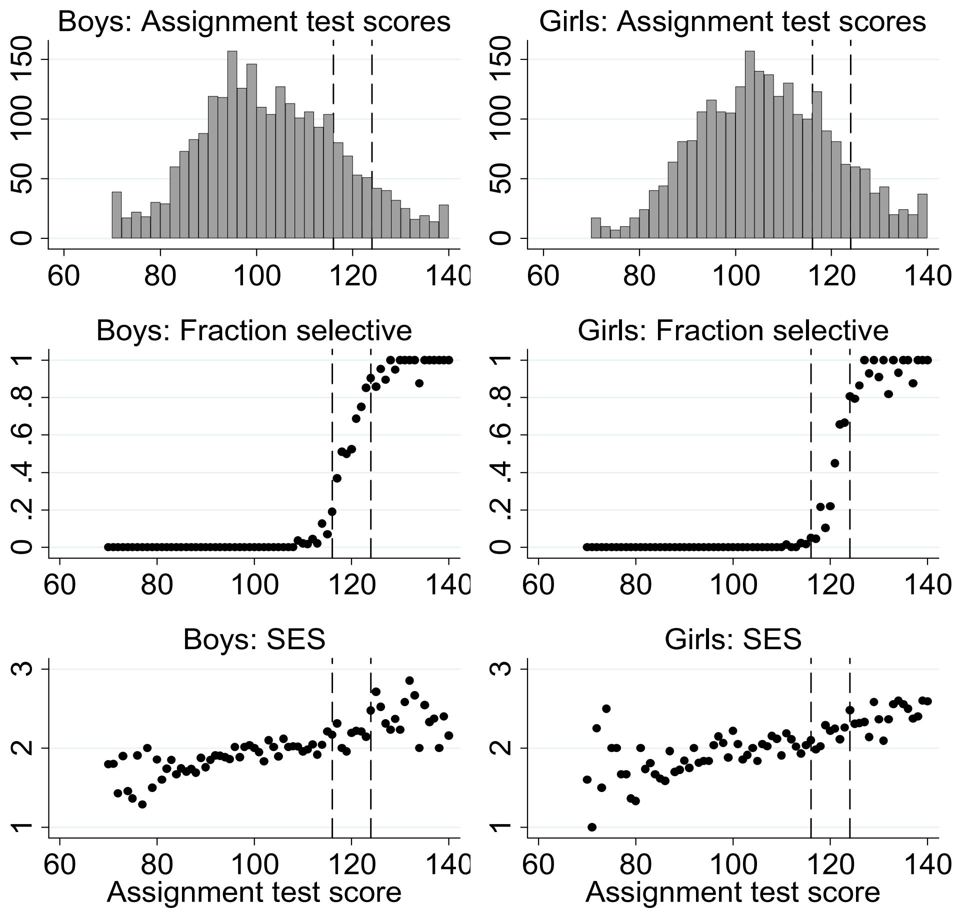
The four grammar schools in the East Riding were typical of grammar schools across the UK. First, the assignment mechanism ensured peer quality was high: as seen in the top panel of Figure 1b below, students with assignment test scores to the right of the marginal test-score interval attended schools in which peer quality was roughly two standard deviations higher than in the schools attended by students with scores to the left of this interval. Second, in keeping with their private school traditions, the schools were relatively small, were single-sex, gave head teachers a large measure of autonomy, and employed teachers who were, on average, better qualified (see Web Appendix Table 1b for national-level information on the first two of these characteristics). This can be seen in the middle panel of Figure 1b, which shows that students with assignment test scores to the right of the marginal interval attended schools in which teachers were more likely to have attended universities (as opposed to teacher training colleges).[14] Along other dimensions, including teacher age, teacher experience, and teacher turnover, across-school differences were much smaller.[15] The third panel of Figure 1b suggests that there was no consistent difference in the amount of homework assigned to students. According to the estimates of Miles and Skipworth (1974),

[14] According to a former master at one of the boys' grammar schools, nearly all of the teachers in this school were graduates of Oxford or Cambridge Universities (personal communication). See Mortimore (1999) for a fascinating history of this school.

[15] Results available upon request from the author.

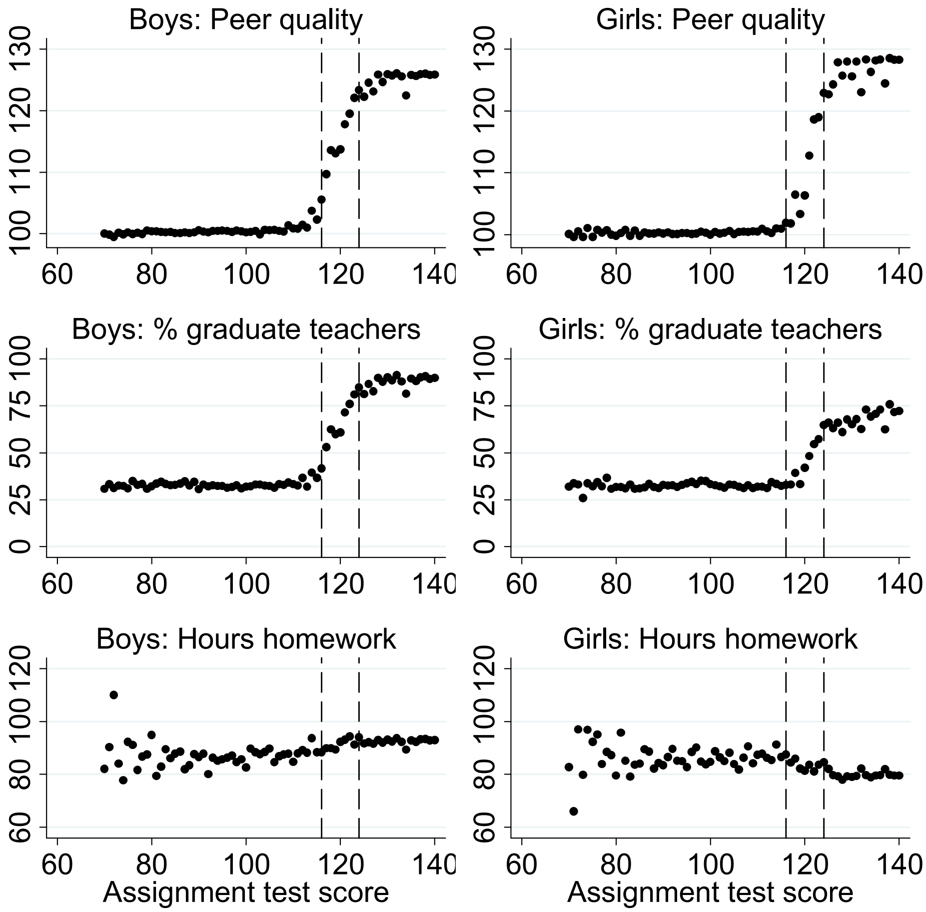
boys with higher assignment test scores attended schools that assigned slightly more homework, whereas girls with higher assignment test scores attended schools that assigned slightly less.

**Figure 1a: Assignment Test Scores**
**Selective School Attendance and Socioeconomic Status**




Notes: The graphs are based on the sample of students in the test files with non-missing assignment test scores: N=2452 (boys), N=2499 (girls). Socioeconomic status is coded as low (1), middle (2), or high (3). The solid circles represent mean outcomes for each assignment test score. The broken vertical lines indicate the "marginal" assignment test scores (over which the probability of attending a selective school conditional on the assignment test score increases from around 0.1 to 0.9).

**Figure 1b: School Characteristics by Assignment Test Score**




Notes: The graphs are based on the sample of students in the test files with non-missing assignment test scores: N=2452 (boys), N=2499 (girls). In the top panel, peer quality is measured by average assignment test score of students in test data. In the second and third panels, the percent graduate teachers and hours of homework are shown as calculated by Miles and Skipworth (1974). The solid circles represent mean outcomes for each assignment test score. The broken vertical lines indicate the "marginal" assignment test scores (over which the probability of attending a selective school conditional on the assignment test score increases from around 0.1 to 0.9).

Although this paper focuses on these 20 selective and nonselective schools, the East Riding had three private and two "bilateral" high schools (which had separate tracks within the same school). Ignoring these would cause problems if selective school eligibility influenced whether parents enrolled their children in them (hence whether they appeared in the sample), but the bilateral schools were geographically isolated and probably not a realistic option for most of the children that eventually attended schools in the sample.[16]

Private school may have been a realistic option for wealthier parents, and it may have been especially attractive to wealthy families whose children failed to secure a place in a selective school. Because I do not have data on private school students, this phenomenon could cause me to overestimate selective school impacts. The intuition is that private schools might attract students with below-marginal-assignment test scores who were, conditional on these scores, better on unobservables. Removing these students could make the below-marginal students look worse than the above-marginal students, biasing upwards the estimated selective school effects. However, there are three reasons for thinking that my conclusions are not distorted by this phenomenon. First, private school enrollment in the secondary grades is fairly low in England, around 7 percent nationwide (see Web Appendix Table 2), and I calculate around 8 percent in the East Riding.[17] Second, around one-half of secondary-grade students enrolled in private schools were likely enrolled in private schools in the primary grades (see Web Appendix Table 2). This implies that at most 4 percent of secondary-grade students could have moved to private schools in response to the selective school outcome. Third, students enrolled in private schools in the secondary grades are generally of high ability: nationally representative data show mean ability to be comparable to that among selective school students with a slightly thicker left tail (see Web Appendix Figure 1). All of this suggests that the fraction of students enrolling in private school in response to a bad selective school outcome was likely very small. Moreover, it is important to stress that my estimated test-score effects are very small and my estimated course-taking effects relatively large. These types of private school responses suggest that the true test-score effects might be even smaller and seem unlikely to account for a large part of the course-taking effects.

Like other UK grammar schools, the East Riding grammar schools provided an advanced academic curriculum. Students were expected to stay in school until the end of the tenth grade (one year later than the compulsory leaving age), and in grades nine and ten they pursued a set of externally evaluated

---

[16] These schools were built in the 1950s in response to strong population growth in a geographically isolated part of the district (from which it was difficult to commute even to the nearest grammar school). The School District explicitly stated that these schools would not compete for students with existing schools.

[17] This number is based on data for individual schools from the 1964 edition of the Education Authorities and Schools Directory. In particular, I calculate the total number of students in private schools in the East Riding in this year (1,345) and assume a dropout rate of 20 percent in order to arrive at a number per grade of around 200 (around 8 percent of the per-grade sample of around 2,500). This could be an underestimate if some students attended private schools in another district and an overestimate if a greater number of non-district residents attended these private schools.

academic courses (O Levels). All students were expected to study a core set of courses, including mathematics, English language, one foreign language, and two sciences. Higher-ability students could study more advanced courses such as additional mathematics, Latin, and Greek. In the nonselective schools, practice varied by ability: lower-ability students were expected to leave at the compulsory leaving age and would, in their final grade, take non-examined lower-level courses,[18] whereas high-ability students were encouraged to stay the extra year and pursue the same type of academic courses as selective school students. All nonselective schools offered basic courses, such as mathematics and English language courses, but fewer offered more advanced courses, such as English literature, physics, and chemistry, and none offered courses such as additional mathematics and the classics. Based on data discussed below, Figure 2 in the Appendix supports this taxonomy of courses into those routinely offered by nonselective schools (English), those occasionally offered by nonselective schools (English literature), and those never offered by nonselective schools (advanced mathematics and the classics). Because these graphs are based on the (selected) sample of students observed taking at least one course (discussed in more detail below), one cannot interpret these across-school differences as the causal effects of selective schools, and I estimate course-taking effects on a more representative set of course outcomes below. Nevertheless, this taxonomy is useful to keep in mind when interpreting estimates of selective school effects on course outcomes.

Students wishing to continue in education beyond grade 10 could pursue two-year "A Level" courses, and, if they applied and were accepted, they could attend university or higher education colleges. A student's choice of university degree would depend on that student's A Level choice. In turn, a student's A Level choice would be determined, to a large extent, by the student's choice of high school course. Rarely would a student apply to study physics at university without studying physics at the A Level; rarely would a student choose an A Level physics course had the student not taken and passed the high school physics course. As a result, it is likely that selective school effects on high school course-taking influence the type of university course that students apply for and the probability that they apply for a university course. The structure of the UK selective school system is illustrated in the Web Appendix.[19]

# 3 Empirical Strategy

As noted above, assignment to UK selective schools was based primarily on the scores obtained in primary school assignment tests. In the East Riding, it was based on these test scores and a primary school order of merit. To motivate the

[18] These were called Certificates of Secondary Education (CSE).
[19] For those interested, see the author's website.

strategy used to identify the causal effects of attending a selective school in this context, suppose that the true relationship between selective school attendance and some outcome $y_i$ is:

$$y_i = \beta_0 + \beta_1 d_i + \varepsilon_i\,, \tag{1}$$

where $d_i$ is a dummy variable indicating whether individual $i$ attended a selective school and $\varepsilon_i$ includes omitted variables and a random disturbance term. Least squares estimation of equation (1) will likely be upward biased because of the positive correlation between selective school attendance and this error term. This will be the case even after controlling for observable family characteristics, such as socioeconomic status, as there will likely be a positive correlation between the component of ability that cannot be predicted by family characteristics, the probability of selective school attendance, and the future outcome. It will also be the case after controlling for the assignment test score because there will be a component of ability (the primary school ranking) that is correlated with the probability of selective school attendance and perhaps with future outcomes.

The key to identifying the causal effect of selective school attendance is the use variation in the probability of attending selective school that is only driven by the assignment test score, denoted *t*. To see this more clearly, we can decompose the error term into the part that can be predicted by the assignment test score $(t_i)$ and the remainder that cannot:

$$\begin{aligned}\varepsilon_i &= E(\varepsilon_i \mid t_i) + u_i \\ &= f(t_i) + u_i\,,\end{aligned}$$

where the conditional expectation is written as a function of *t* and *u* is mean independent of *t*. Substituting into (1) gives:

$$y_i = \beta_0 + \beta_1 d_i + f(t_i) + u_i\,. \tag{2}$$

Assuming that $f(.)$ is correctly specified, consistent estimates of $\beta_1$ can be derived by exploiting variation in $d_i$ driven by variation in $t_i$.

In the strict regression discontinuity design, *d* is a step function in *t*: for *t* less than some threshold $t^R, d$ equals zero and for *t* greater than $t^R, d$ equals one. In that case, $d_i$ can be replaced with the relevant indicator function and equation (2) can be estimated by least squares. In the fuzzy regression discontinuity design, the treatment does not necessarily switch on at this threshold, but the probability of treatment does jump (i.e., is discontinuous) at this threshold. In that case, the same indicator function can be used as an instrument for the treatment, and an instrumental variables procedure can recover consistent estimates of $\beta_1$ (see Imbens and Lemieux 2008).

In the setting considered here, the probability of treatment changes sharply over a narrow interval in test scores. To exploit this sharp change, I use the predicted probability of treatment (as a function of assignment scores) as an instrument for treatment. The instrument is valid whether or not this first-stage model is specified correctly (e.g., a logit is used when the true model suggests a probit is more appropriate), and the second-stage standard errors do not need to be corrected for the sampling variability inherent in estimating the predicted probability (Wooldridge 2002, proc. 18.1). I must assume that the effect of treatment is constant over the interval in which the probability of treatment changes, but this interval is narrow; hence, this seems like a relatively mild assumption. The more important caveat is that, as in all regression discontinuity designs, the estimates must be interpreted as local to these marginal units.

Although this strategy can, in principle, recover consistent estimates of $\beta_1$ for students on the selective school margin, there are three practical challenges to the implementation of this approach. First, a functional form must be chosen for $f(.)$. I choose the most flexible function that fits the data, which in this case is a fourth-order polynomial in the assignment test scores (I present results obtained with other polynomials). Second, a bandwidth must be chosen within which the model is estimated. If all of the data are used, the model is fit using data a long way from the marginal interval, and the predicted probability through the interval would likely be misspecified; if too narrow a range is used, sample sizes will be small, and estimates will be imprecise. This is the usual bias-efficiency trade-off discussed by Imbens and Lemieux, and I choose the bandwidth based on a version of their cross-validation criterion.[20] Third, the chosen methods must deal with the binary nature of some of the outcome variables studied. I use bivariate probit models to estimate the impacts on binary outcomes. These make stronger assumptions on the underlying data-generating process than those required by the instrumental variables approach, but they take account of the inherent nonlinearities in these models. In any case, the large estimated impacts on binary outcomes such as advanced course enrollment are usually evident from visual inspection of the relevant graphs.

[20] Specifically, I choose the range of data that does the best job of fitting the data as test scores approach the marginal interval (from both sides).

# 4 Data

The East Riding database was assembled by Miles and Skipworth as part of their project "The Correlates of Academic Performance in Secondary Schools" and was subsequently deposited in the UK Data Archive (study numbers 49, 120, and 197). The database consists of "test" files and "course" files.[21]

## 4.1 The Test Files

For three cohorts of students (those in grade nine in 1969, 1970, and 1971), the main-student file (the "test" file) contains basic demographic information (date of birth and gender), the scores obtained on the assignment test taken in grade five, and a measure of socioeconomic status (SES).[22] It also contains the scores obtained on the tests given to students in their final year of compulsory schooling (grade nine). These tests were designed by an independent testing company (the National Foundation for Educational Research or NFER) and were intended to provide information on performance across the full range of abilities. In the Web Appendix,[23] I provide examples of the types of questions asked in these tests and plot the distribution of these scores. These tests are similar to those used in the National Child Development Study (the mathematics test is identical). As these items demonstrate, the tests are comparable to contemporary standardized tests such as the Florida FCAT grade 10 tests.[24]

Ideally, I would like the test file sample to include information on all students in the relevant cohorts that took an assignment test in grade five. In fact, the data cover all students in the sampled schools that took one or more of the grade nine tests. There are three reasons why these two populations may differ. First, there are students that took the assignment test but do not appear in my data because they attended a school that is not in my sample (e.g., a private school). As discussed above, only a small number of students attend private schools, and any sample selectivity biases the estimates upwards. Second, there are students in my sample that did not take the assignment test either because they were in the

[21] The database also includes a "schools" file used to produce the bottom two panels of Figure 1b.

[22] As SES was reported by the school, observed SES may not be completely predetermined, and one must be cautious about presenting results that condition on it. For example, if governing body schools incorrectly categorize some low-SES students as mid-SES students (because most students are at least mid-SES), a comparison of low-SES students in governing body and nongoverning body schools will be biased in favor of the latter. I present results with and without SES controls.

[23] For those interested, see the author's website.

[24] Florida is chosen because it is one of few states to make these items publicly available. It is also a large state that has been closely associated with the U.S. accountability movement.

district in grade five but did not take the test or they moved into the district after grade five.[25] Because all of my analyses are based on students for whom I observe assignment test scores, this is only an issue to the extent that selection into test-taking and heterogeneous-treatment effects mean these estimates should be interpreted as those relevant to the sample that took the test (and scored close to the selective school cutoff).[26]

Finally, there are students that took the assignment test and attended one of the schools in my sample but do not appear in my data because they did not take any of the grade nine tests. A comparison of the test-taking samples with published data on the number of 13-year olds in the sampled schools two years previously suggests that around 10 percent of these 13-year old cohorts do not appear in the test data two years later.[27] This may be because these students were not in school when the tests were administered (e.g., because they were ill) or because they left the district before reaching age 15. While these and other explanations for attrition could lead to biased estimates of the selective school effects, there are at least two reasons to suppose that any attrition problems are not large. First, because the estimated effects reflect the impact of selective schools on students at the selective school margin, these are almost completely unaffected by attrition among the lowest-performing students, the highest-performing students, or a random sample of students. Trimming a further 10 percent of the data in any of these fashions has almost no impact on the results. Only attrition among the worst-performing marginal nonselective school students or the best-performing marginal selective school students could lead to selective school effects being underestimated. Second, although I have no data on students with missing scores on all tests, an analysis of students with missing scores on

[25] According to Secondary and Further Education Sub-Committee (1964, 108), all students took the test: "Each pupil in the appropriate age group, together with a few really outstanding pupils in the age group below, take two verbal reasoning tests in their own schools." However, it is plausible to imagine that a small number of students would have been excluded from the test either at the request of their parents or for other reasons. In addition, a small number of secondary school students may have moved into the district after taking a test elsewhere.

[26] A regression of outcomes on demographics, school fixed effects, and a dummy variable for missing assignment test scores suggests no obvious differences among students with and without missing scores.

[27] The published data come from the "Statistics of Education: Schools" volumes 1967, 1968, and 1969. The number of 13-year olds in these years is 2,960, 3,033, and 3,157, and the corresponding test file samples are 2,353, 2,474, and 2,584. From the Secondary and Further Education Sub-Committee (1964), we know that around 200 students included in the first set of numbers attended private schools or schools outside of the county; around 100 students included in the first set of numbers attended the one non-governing body school that did not take part in the study. The remaining difference—roughly 200–300 students per cohort—accounts for about 10 percent of the expected cohort size.

some tests but not others (around 3–5 percent) does not support this.[28] Moreover, I show that coding missing scores as zero has ambiguous effects on the results, increasing the selective school estimates on some outcomes but not others.[29] Descriptive statistics associated with these test files are presented in Panel A of Appendix Table 1.

## 4.2 The Course Files

The "course" files contain information on advanced courses pursued in tenth grade. These data are available for six cohorts of students pursuing these courses—those in grade 10 in 1965 through 1970—and they contain detailed information on all of the advanced courses pursued by these students, including the grade obtained on the end-of-course exam (graded from one to nine). Descriptive statistics for these data are reported in Panel B of Appendix Table 1. These two files overlap for only one cohort of students, those taking the test in grade nine in 1969 and pursuing courses in grade 10 in 1970 (see Panel C Appendix Table 1).[30]

I can also match this cohort to administrative university data, to calculate the fraction of all students that subsequently enroll in university.[31] To improve the quality of this match, I restrict attention to the set of cohort members that pass five or more academic courses. If selective schools have positive causal effects on the probability that students reach this threshold, this could bias selective school university enrollment effects downwards. In fact, as shown below, I find large positive effects on university enrollment.[32] Appendix Table 1 Panel C shows that

[28] Linear models of the probability of having missing math or reading scores suggests that higher-ability students (measured by assignment test scores) are slightly less likely to have missing scores. The effect of a governing body school on these probabilities is positive but always small (less than half of a percentage point) and not statistically significant.

[29] Although there are no missing values in the test files, I presume that test scores are missing when the sum of correct and incorrect answers is zero.

[30] As the table shows, a small fraction of students appear in the course data but cannot be matched back to the test data. These students may have moved into the area after grade 10 or switched schools after grade 10. I drop these students from the matched data.

[31] University records more accurately report which students are observed in university three or four years after they begin their course; hence, I define attendance as being enrolled for at least three years. As the dropout rate from UK universities is low, this is not an important decision. University enrollment data come from the ESRC Data Archive (the Universities Statistical Record, Study Number 3456).

[32] The match is done via school, sex, and date of birth, a combination that does not always uniquely identify students. The restriction reduces the number of students that might have attended college; hence, it increases the probability that students are uniquely identified in these data. Passing five or more academic courses was typically the minimum standard needed to advance beyond tenth grade.

overall university enrollment is around 4 percent. This appears low, but it is consistent with other university enrollment data.[33]

# 5 The Impacts of Selective Schools

In this section, I report the estimated effects of attending a selective school. I discuss test-score effects before turning to effects on other outcomes related to the possible longer-run effects of attending a selective school.

## 5.1 Test Scores

Figure 2 illustrates the key ideas behind the test-score analysis. For boys in one cohort of the test files, the graphs plot mean test scores by assignment test score and school type.[34] For students with marginal assignment test scores, I plot means for both school types; for students with non-marginal scores, there are means for only one type.[35]

The least squares estimate above each graph is from a regression of the outcome on a selective school dummy and a fourth-order polynomial in the assignment test score. As this estimate exploits variation in school type conditional on assignment test scores, it is approximately equal to the difference between selective and nonselective school means for marginal students.[36] The concern is that conditional on assignment test scores there is positive selection into selective schools.

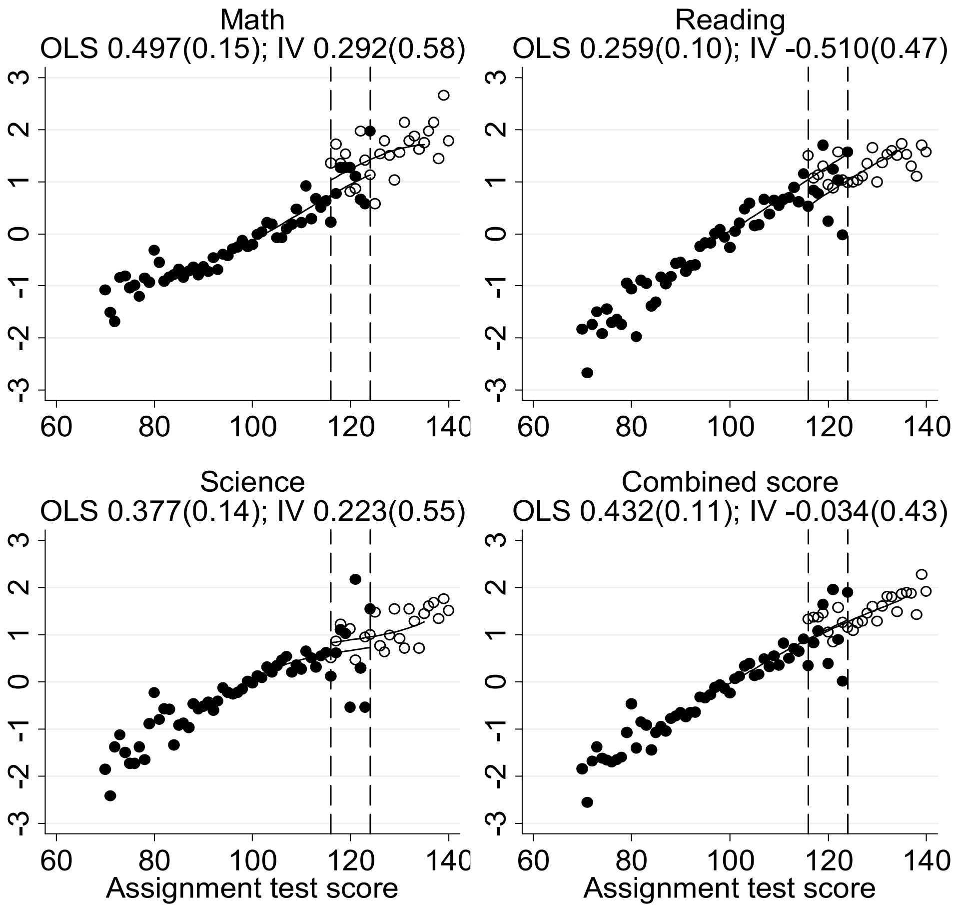
[33] The overall percentage of students aged 19 and attending university (defined to exclude further education institutions such as Teacher Training Colleges) in 1969 was 13 percent. This fraction includes students who attended private school, enrolled in college before/after their cohort, and dropped out of college before completing three years. These reasons and the fact that the East Riding is a relatively poor district can account for the difference.

[34] I use boys in the third cohort for illustration purposes. This cohort is only significant to the extent that it is the one for which I also have course-taking data.

[35] There are a small number of students with below-marginal scores that attend selective schools and a small number with above-marginal scores that attend nonselective schools. These are included in the samples on which the estimates reported in the graphs are based but are not presented on the graph. Excluding these observations makes the graphs easier to interpret.

[36] My least squares estimates are very similar to those generated by least squares regression of the outcome on a selective school dummy and a full set of assignment test score fixed effects. Suppose the effect of selective school conditional on assignment test score was constant in the marginal interval and suppose that the probability of attending selective school is either zero or one conditional on any test score outside of that interval (approximately correct). In that case, it can be shown that this regression will identify the "treatment effect" in the marginal interval (which is likely to be a biased estimate of the causal effect of selective school attendance).

**Figure 2: The Impact of Selective Schools on Grade Nine Test Scores: Evidence from One Competition**




Notes: The graphs are for boys in the third cohort of the test files with non-missing outcome data: N=520 (math), 522 (reading), 506 (science), 486 (combined score). The solid (open) circles are outcome means among nonselective (selective) students with the corresponding assignment test score. The broken vertical lines indicate the "marginal" assignment test scores (over which the probability of attending a selective school conditional on the assignment test score increases from around 0.1 to 0.9). The superimposed solid (and broken) lines represent the outcomes for nonselective (selective) school students predicted by IV estimates of the relationship between the outcome, selective school attendance, and a fourth-order polynomial in the assignment test score estimated on the subsample of students with assignment test scores in the [95,135] range (selective school attendance is instrumented; the excluded instrument is the probability of attending a selective school predicted by a logit model featuring a fourth-order polynomial in the assignment test score). The IV estimate of the impact of attending a selective school (reported above each graph, robust standard error in parentheses) is the estimated coefficient on a selective schools dummy in a regression of the outcome on this dummy and a fourth-order polynomial in the assignment test score (using the same sample of students).

The instrumental variables estimate above each graph is that obtained when the selective school dummy is instrumented with the estimated probability of attending a selective school. Because this estimated probability is based on a flexible function of the assignment test score (i.e., a logit model with a fourth-order polynomial in the assignment test score), these instrumental variables estimates exploit variation in selective school attendance across assignment test scores. Intuitively, the instrumental variables estimates will attribute positive causal effects to attending a selective school when the relationship between outcome and assignment test score tracks the relationship between selective school probability and assignment test score (i.e., is sharply increasing over the marginal test scores). To give a visual impression of the IV estimates, I use the IV estimates to superimpose onto the graphs the fitted relationship between assignment test score and outcomes. Over the marginal assignment test scores, where these fitted lines overlap, the selective school effect is the vertical distance between them; away from these marginal assignment test scores, where there is only one fitted line, a comparison of the fitted lines and the raw data tells us how well the IV estimates capture the underlying relationship between assignment test scores and outcomes.

The graphs suggest that instrumental variables estimates are smaller than least squares estimates, but the instrumental variables estimates are fairly imprecise. As seen in the graphs, this is because I am trying to capture the underlying relationship between assignment test scores and outcomes with a flexible function fit using relatively few data points. To add more observations and generate more precise estimates, I assume treatment effects are constant across cohort and sex (later relaxed) and pool data across the six cohort-by-sex groups in the dataset (i.e., boys and girls in the three cohorts in the test files).

Because selective schools are single sex, each cohort-by-sex group represents a separate selective schools competition. As such, two adjustments must be made to the pooled model. First, the first stage of this model must be estimated separately for each group (i.e., the assignment test score polynomial must be interacted with the cohort-by-sex dummies in the first stage). To see why, note that the IV estimates can only be given a causal interpretation when the monotonicity assumption is satisfied (i.e., when the estimated probability of treatment is weakly increasing in the predicted probability of treatment); otherwise, the estimates will be attenuated. Pooling groups without adding interactions to the first stage would likely violate this monotonicity condition.

A second concern is that because students in the same competition may have been subject to the same shocks the errors will be correlated within groups. To deal with this concern, I experimented with clustering standard errors at the group level. Although clustering should bias conventional standard downwards, the group-clustered standard errors were typically smaller than unclustered standard errors (see the Appendix Table 2). As clustered standard errors are known to be biased when group sizes are small (in this case I have six), I present the more conservative unclustered standard errors.[37]

These pooled estimates are reported in Table 1. The first column reports least squares estimates without any regression adjustment and confirms the expected large difference in mean scores between selective and nonselective school students—around one standard deviation. The second column reports least squares estimates that regression adjust for a fourth-order polynomial in the assignment test scores. These estimates are much smaller than those reported in column (1), and they are further reduced after regression adjusting for sex-by-cohort dummies, column (3), and an indicator of socioeconomic status, column (4).

The instrumental variables estimates are presented in column (5). Not surprisingly, the predicted probabilities used as instruments are highly correlated with actual selective school attendance (the coefficient is close to one with a t-statistic in excess of 30). Looking at the first three rows (for math, reading, and science), the instrumental variables are nearly always smaller than the least squares estimates. Typically, zero effects cannot be rule out while the confidence intervals are close to excluding the least squares estimates. Adjusting for socioeconomic status, column (6), has little impact on these estimates; hence, it does not change this conclusion. Since the estimates for all three tests are consistent with zero effects and consistent with each other, I report the estimated effects on the combined score in the fourth row. The instrumental variables estimate is close to zero and more precisely estimated. In the final row, I report estimates on IQ test scores. These are not expected to be affected by selective school attendance and the estimates are consistent with zero selective school effects.

[37] Fortunately, most of the variation in selective school attendance occurs within rather than between groups. This is reflected in the intragroup correlations in both the selective school variable (which is relevant to the standard errors on the least squares estimates) and the first-stage fitted values (relevant to the standard errors on the instrumental variables estimates), which are both less than 0.02. This suggests that the biases associated with robust standard errors are likely to be small, especially since the intragroup correlations in the outcomes are also small, typically less than 0.02. See Angrist and Pischke (2009) for an instructive discussion of these issues.

**Table 1: The Impact of Selective Schools on Grade Nine Test Scores**

| | OLS | | | | IV | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| **Panel A: Math (N=2251)** | | | | | | |
| Selective | 1.223 | 0.635 | 0.476 | 0.466 | 0.195 | 0.265 |
| | (0.048) | (0.089) | (0.083) | (0.083) | (0.153) | (0.159) |
| f(test, test2, test3, test4) | | Y | Y | Y | Y | Y |
| Sex-by-cohort indicators | | | Y | Y | Y | Y |
| Socioeconomic status | | | | Y | | Y |
| **Panel B: Reading (N=3362)** | | | | | | |
| Selective | 0.910 | 0.158 | 0.114 | 0.110 | -0.197 | -0.168 |
| | (0.027) | (0.048) | (0.049) | (0.049) | (0.103) | (0.127) |
| **Panel C: Science (N=2165)** | | | | | | |
| Selective | 0.847 | 0.368 | 0.219 | 0.206 | 0.111 | 0.083 |
| | (0.043) | (0.081) | (0.080) | (0.080) | (0.190) | (0.182) |
| **Panel D: Combined score (N=2063)** | | | | | | |
| Selective | 1.115 | 0.406 | 0.267 | 0.257 | -0.050 | -0.049 |
| | (0.036) | (0.068) | (0.064) | (0.064) | (0.114) | (0.113) |
| **Panel E: IQ (N=3377)** | | | | | | |
| Selective | 1.299 | 0.303 | 0.224 | 0.220 | 0.142 | 0.141 |
| | (0.032) | (0.050) | (0.049) | (0.048) | (0.102) | (0.102) |

Notes: The samples include all students in the test files with non-missing outcomes and assignment test scores in the [95,135] interval. Sample sizes were smaller for math and science because these were administered to only two of the three cohorts in the test files. The estimates reported in columns (1)–(4) of Panel A are the estimated coefficients from a regression of the outcome on a dummy for selective school attendance and the covariates described in the rows under these estimates. The estimates reported in columns (5)–(6) of Panel A are the estimates obtained when selective school attendance is instrumented using the probability of secondary school attendance predicted using a logit model featuring a fourth-order polynomial in the assignment test score. Robust standard errors are reported in parentheses. The estimates in panels B–E are based on the same specifications as those reported in Panel A.

To probe the robustness of these results, I assess whether they are sensitive to the modeling decisions made in generating them. The results of these checks are reported in Appendix Table 2. I begin by experimenting with the polynomial used to capture the underlying relationship between the assignment test score and the outcome. Note that the polynomial choice cannot be assessed visually because the pooled IV estimates cannot be represented in a graph (unlike the IV estimates for a single-sex-by-cohort group, as graphed in Figure 2). The baseline estimates were obtained from a model that captured this relationship via a fourth-order polynomial in the assignment test score. In columns (2) and (3), I report the results obtained when third- and fifth-order polynomials are used;

column (1) reproduces the baseline estimates, that is, column (6) of Table 2. The third-order polynomial, column (2), generates estimates in line with the baseline estimates, as does a second-order polynomial (not shown). The fifth-order polynomial, column (3), generates much less robust estimates, and it seems that the data cannot support such a flexible specification (a sixth-order specification generates similarly imprecise estimates, although these are closer to the preferred estimates). Second, I experiment with the bandwidth used to select the estimation sample. The baseline estimates are based on data in the [95,135] interval, which was at the conservative end of the range suggested by a cross-validation procedure along the lines suggested by Imbens and Lemieux.[38] A less conservative choice ([90,140]) does not change the estimates in any systematic way, and it is interesting that the estimated effect on the combined score is even closer to zero when this less conservative choice is made. A more conservative choice ([100,130]) tends to increase the estimates, but the cross-validation exercise suggests that this would be much too narrow a bandwidth. Even based on this choice, the selective school effect on the combined test outcome is less than 0.1 and not statistically distinguishable from zero. Third, I experiment with adding students with missing scores to the estimation sample (and assigning them a score of zero) and changing the dependent variable to the total number of questions attempted (including those with missing scores for whom I assign zero attempts). Neither change has any systematic effects on the results: the estimates suggest that selective schools have small impacts however the sample and dependent variable are defined.

A potential concern is that these estimates are influenced by ceiling effects. These could prevail if the tests were so easy that a large fraction of marginal students obtained maximum scores or so hard that a large fraction of marginal students obtained very low scores. In either case, selective school students would be unable to demonstrate their superior learning, and selective school effects would be biased towards zero. In fact, there are several reasons to think that this is not a concern here. First, histograms of the raw test scores show that few students achieved a maximum score (see Web Appendix Figure 1). Second, this is in part because these tests were designed with ceiling effects in mind and were intended to test the full range of grade-nine abilities. Third, estimated effects on particular quantiles of the test-score distribution do not point to significant selective school effects (not shown, details upon request). In particular, there is no consistent evidence for effects at the 25th or the 75th percentile of the outcome distribution, as one might expect if these tests were too

[38] Specifically, for each point to the left of the marginal interval, I use this point and the n observations to the left of it (to a left limit of 86) to predict the point to the right of it. I then calculate the average prediction error associated with each bandwidth (i.e., size n). I do the same using data to the right of the marginal interval (to a right limit of 139).

hard or too easy to detect differences in the middle of the ability (conditional on the assignment test score) distribution.

Although the evidence strongly suggests that average selective school effects are small, it may be that selective school effects differ for boys and girls and across SES groups. To examine this possibility, I estimated models that interact the selective school variable with gender and SES. There are no obvious trends to emerge from these estimates (see Web Appendix Tables 3a and 3b). The least squares estimates are larger for boys, but this could represent difference patterns of selection. This pattern is much weaker in the instrumental variables estimates, none of which are distinguishable from zero. Similarly, an analysis of selective school effects by SES reveals no consistent patterns within tests (i.e., both low- and high-SES students are associated with larger selective school effects than medium-SES students on the math test) or across tests (i.e., selective school effects are greatest for low-SES students in math and smallest for low-SES students in reading and science). The overall conclusion then is that when assessed in terms of test-score performance, selective school effects appear small. I discuss these results in more detail below.

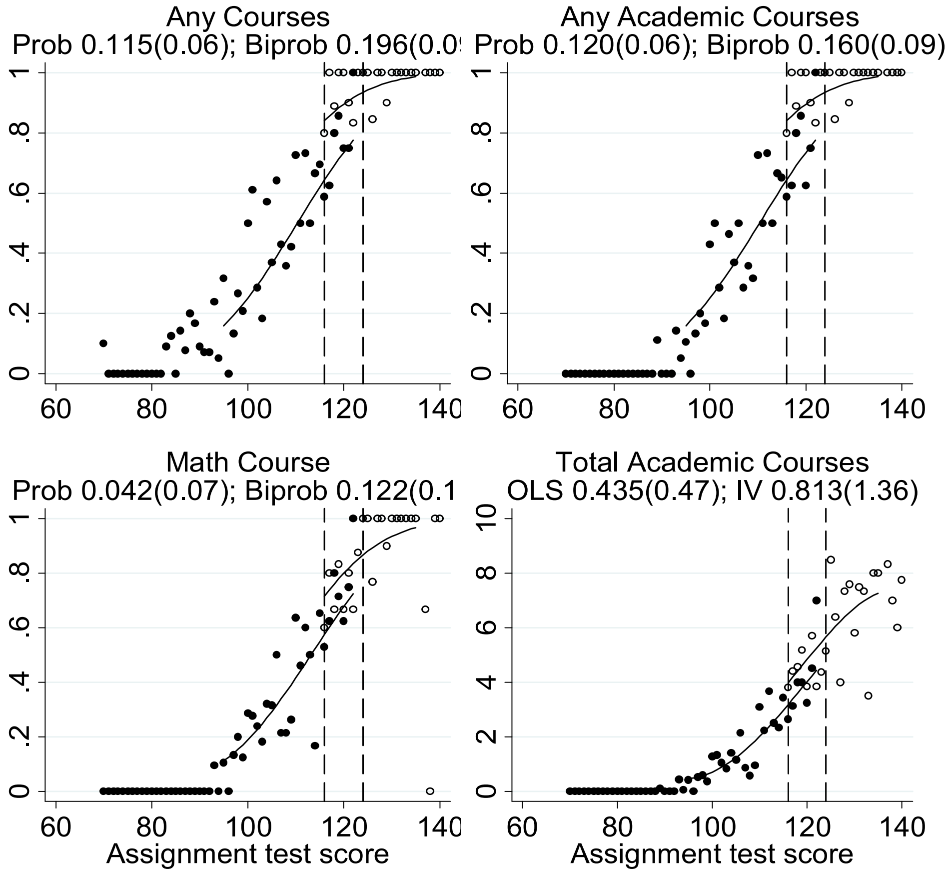
## 5.2 Other Outcomes

Unfortunately, I do not have information on earnings. I do however have extensive information on course-taking outcomes and information on university enrollment. As noted above, course-taking outcomes are likely to shape post-secondary education opportunities, which are likely to shape earnings. I present results for four course outcomes: whether any course was pursued, whether any academic course was pursued, whether a mathematics course was pursued, and the number of academic courses pursued. The first is an indication of high school dropout because students who took no courses very likely dropped out at the end of ninth grade (the compulsory school leaving age) not at the end of tenth grade when these courses finished.[39] The second is an indication of the extent to which selective schools oriented students towards an academic curriculum, and the third is a specific academic course thought to be especially highly valued on the labor market (Murnane et al. 1995). The fourth is a useful summary measure of course-taking activity and is important for gaining access to post-compulsory education.[40]

[39] Personal communication with nonselective school teachers suggests few students would stay beyond the compulsory school leaving age and pursue only lower-level courses (not available in these data). Because I cannot be sure that this did not apply to some students, I choose not to interpret these as high school dropout effects.

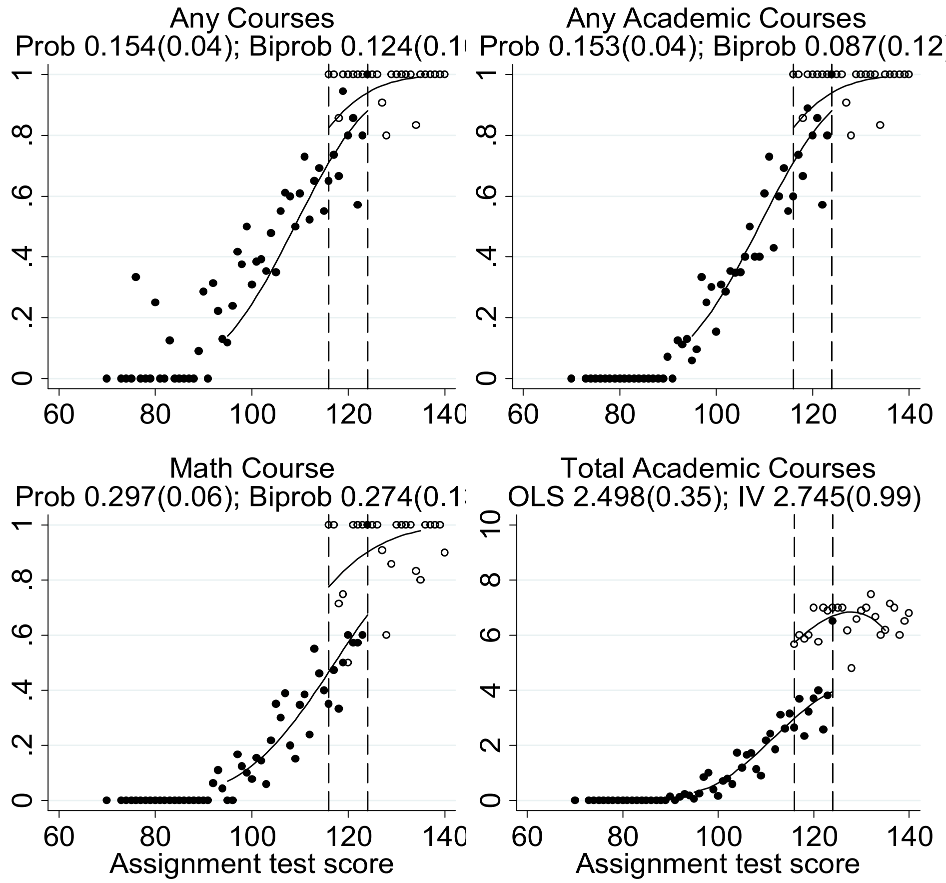
[40] Students are typically required to pass five or more academic courses if they wish to continue into post-compulsory education. I do not focus on this requirement as it does not appear to be a strict one and because it mixes course-taking with course performance. Selection into course-taking means that I cannot estimate course performance effects.

**Figure 3a: The Impact of Selective Schools on Advanced Course-taking in Grade 10: Boys**




Notes: The graphs are based on data on boys in the matched test-course file (N=530). The solid (open) circles are outcome means among nonselective (selective) students with the corresponding assignment test score. The broken vertical lines indicate the "marginal" assignment test scores (over which the probability of attending a selective school conditional on the assignment test score increases from around 0.1 to 0.9). For outcomes except "Total Academic Courses," the superimposed solid (and broken) lines represent the outcomes for nonselective (selective) school students predicted by a bivariate probit model of the relationship between the outcome, selective school attendance and the assignment test score estimated on the subsample of students with assignment tests scores in the [95,135] range (the variable in the selective schools equation excluded from the outcome equation is the probability of selective school attendance predicted by a logit model featuring a fourth-order polynomial in the assignment test score). The bivariate probit (Biprob) estimate reported above each graph (robust standard error in parentheses) is the vertical distance between these predicted outcomes at an assignment test score of 119 (where the conditional probability of selective school attendance is closest to 0.5). The probit (Prob) estimate reported above each graph (robust standard error in parentheses) is the estimated selective school impact derived from a probit model of the relationship between the outcome, selective school attendance and the assignment test score evaluated at an assignment test score of 119. For "Total Academic Courses," the relevant details are the same as those in the notes to Table 2.

**Figure 3b: The Impact of Selective Schools on Advanced Course-taking in Grade 10: Girls**




Notes: The graphs are based on data on girls in the matched test-course files (N=625). The probit and bivariate probit estimates of selective school impacts are evaluated at an assignment test score of 121 (where the conditional probability of selective school attendance is closest to 0.5). Otherwise, the relevant details are those described in the notes to Figure 3a.

It is also useful as there is no ceiling on the possible size of the selective school effects on this outcome. This is not true of selective school effects on binary outcomes such as any courses and any academic courses.

Since the analysis suggests that selective school effects on course-taking differ by sex, I report separate estimates for boys and girls. As I have course data for only one cohort of students, the sex-specific analyses refer to students in a single selective schools competition. This implies that the estimates can be presented graphically; Figures 3a and 3b plot the raw data and the estimated selective school effects.

The first three course-taking outcomes are binary. For these, I estimate selective school effects using probit and bivariate probit methods. The probit model deals with the binary outcomes and is valid provided the latent error term is normally distributed and uncorrelated with selective school attendance (conditional on the covariates). The bivariate probit model deals with binary outcomes and is valid provided the latent error term and latent selection error term are jointly normally distributed. Identification of the selective school effects is then achieved via the inclusion of a variable in the latent selection equation that is not also in the latent outcome equation (Wooldridge, sec. 15.7.3). In this case, the excluded variable is the predicted probability of selective school attendance.

Unlike the IV estimates of selective school effects on test scores, these bivariate probit estimates are sensitive to the specification of the underlying relationship between outcomes and assignment test scores. This is not a consequence of the estimation strategy: a linear IV strategy generates similar estimates and displays similar sensitivity. Instead, it is a consequence of the strong positive relationship between assignment test scores and these binary outcomes. Intuitively, this makes it hard to discern whether this relationship is best characterized as smooth and increasing or increasing with a sharp change over the marginal assignment test scores. To retain a degree of precision, I estimate the probit and bivariate probit models using a linear specification for the assignment test score.[41] The probit models are robust to more flexible specifications, but more flexible bivariate probit models generate much larger point estimates and much larger standard errors.

Probit estimates suggest that selective school attendance increases the probability of pursuing at least one course and at least one academic course by around 10–15 percent. These estimates are robust to functional form (up to a fourth-order polynomial in the assignment test score) and other covariates (i.e., socioeconomic status). They are also similar for boys and girls, as are baseline course-taking rates among marginal students in nonselective schools (around 70–75 percent in each case). The bivariate probit estimates are less precise than the probit estimates and the point estimates differ for boys and girls. Against this, the bivariate probit estimates cannot reject the probit estimates, and models pooled by sex generate bivariate probit estimates close to the probit estimates. Given the robustness of the probit estimates, the fact that the more flexible bivariate probit estimates generate larger estimates and the close correspondence of probit and bivariate probit estimates on other course-taking outcomes, a reasonable conclusion is that, for both boys and girls, selective school attendance likely has

[41] As these models are nonlinear, this linear specification will generate a nonlinear relationship between the assignment test score and the probability of observing these various outcomes.

large impacts on the probability that students pursue at least one course and at least one academic course.

One interpretation of these estimates is that, although nonselective schools allow students to choose the extent to which the curriculum is academic, selective schools enforce an academic curriculum. Estimates of selective school effects on the probability of pursuing a math course are consistent with this interpretation because these selective school effects are largest for the groups least likely to pursue a mathematics course in nonselective schools. For example, girls in nonselective schools are 25 percentage points less likely to pursue a math course than are boys in nonselective schools; selective school attendance increases the probability that girls pursue a math course by 30 percentage points and increase the probability that boys pursue one by only five percentage points. A similar pattern is revealed when the data are cut by socioeconomic status: course-taking is lower among the lower SES groups in nonselective schools, and selective school effects appear to be larger for these groups (not shown).

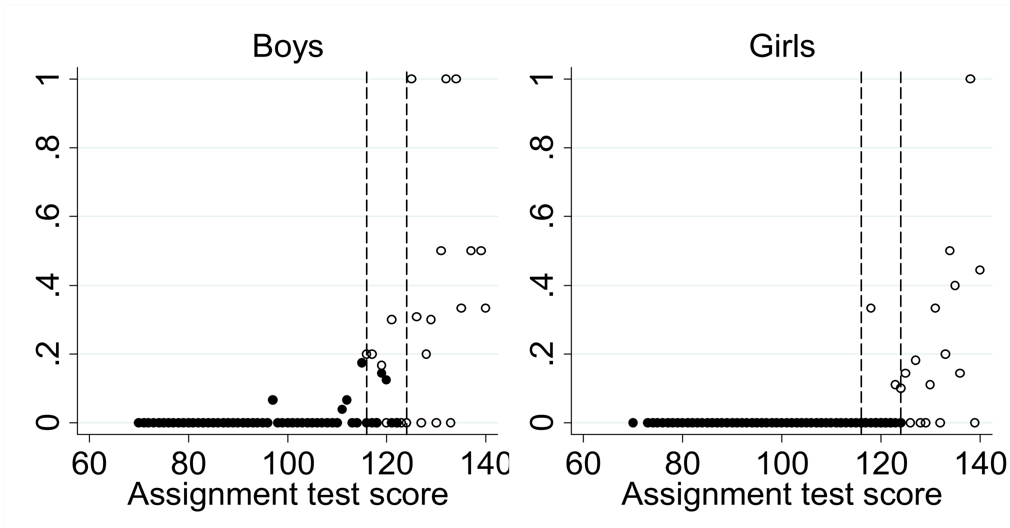
With regard to the overall number of academic courses pursued, selective school effects are again larger for girls. This time, however, these effects are not compensating for low levels of course-taking among girls in nonselective schools. Instead, these effects appear to be driven by the relatively larger number of academic courses taken by girls in selective schools. This could be because boys' selective schools allow students to pursue some nonacademic courses whereas girls' selective schools do not. Note that, since this outcome can be viewed as continuous, these estimates are obtained from linear IV models identical to those used to estimate selective school effects on test scores. As the graphs in Figures 3a and 3b suggest, these estimates are robust to changing the observation window and the assignment test score polynomial.

To summarize, there appear to be large selective school effects on course-taking. The obvious interpretation is that selective schools enforce a curriculum with a strong academic component. As such, selective school attendance has the largest effect on course-taking among students who would not pursue an academic curriculum in nonselective schools. To assess whether these course-taking effects have consequences for university enrollment, Figure 4 graphs university enrollment against assignment test scores for the subsample of students for whom university enrollment data are available. As expected, university enrollment is low among this sample of students and positive only for students with high assignment test scores.[42] Among girls, no nonselective school students

[42] There are several reasons for low nongoverning body school means (and low means in general). First, the particular district being studied may have had lower means than average (at a time when university enrollment was generally low). Second, the university enrollment data do not include the full range of higher education institutions, excluding for example teacher training

are observed to enroll in university; hence, probit models cannot be estimated. Among men, probit estimates are imprecise, but they suggest selective school effects on the probability of enrolling in university of between two and five percentage points, roughly equal to mean university enrollment among nonselective school students with marginal assignment test scores.[43]

**Figure 4: University Enrollment by School Type and Assignment Test Score**




Notes: The graphs are based on data on students in the matched test-course file (N=530 for boys, N=625 for girls). The solid (open) circles are mean university enrollment rates among students with the corresponding assignment test score. The broken vertical lines indicate the "marginal" assignment test scores (over which the probability of attending a selective school conditional on the assignment test score increases from around 0.1 to 0.9).

## 6 Interpretation and Discussion

In assessing the size of these effects, it is important to remember that selective schools expose students to more able peers, better qualified teachers, and a more advanced curriculum. On the one hand, this means that one must be careful not to compare these estimates with other estimates of the impact of only one of these characteristics. On the other hand, because selective schools expose students to more of each characteristic and as these characteristics are generally assumed to

---

colleges. This may be why measured university enrollment is twice as high among boys as girls in this sample.

[43] For example, a probit model for university enrollment that controls only for the assignment test score generates a selective school effect (standard error) of 0.0472 (0.057); adding a socioeconomic status control reduces this to 0.027 (0.051).

improve academic outcomes, the selective school estimates can be interpreted as upper bounds on the contribution of each of these characteristics; these upper bounds can be compared with estimates from the literature.

With regard to the test-score estimates, the most natural benchmark would seem to be the literature on peer effects. That is partly because, if is not obvious that these other selective school characteristics (better qualified teachers and a more advanced curriculum) should improve test scores, evidence on the relationship between teacher effectiveness and characteristics such as aptitude, college selectivity, and experience is mixed. Even if these types of teachers taught advanced courses more effectively, they may be no more effective at teaching the basic material captured by standardized tests.[44] Indeed, it is possible that these non-peer characteristics could generate negative selective school effects if resources such as teacher quality are devoted to the best students in the nonselective schools. It is also because the across-school difference in peer composition is extreme.

Although it is difficult to identify the causal effects of peers, recent studies have used arguably exogenous variation in peer composition to estimate linear-in-means peer effects (i.e., the impact of average peer ability on own ability). One strategy, devised by Hoxby (2000) uses variation arising from natural cohort-to-cohort variation in peer composition. The estimates reported by Hoxby imply that a standard deviation increase in average peer ability would improve individual student achievement by between 0.1 and 0.55 standard deviations; Sacerdote (n.d.) reviews a large number of subsequent studies and reports effects that fall roughly within this range. A second strategy, used by Angrist and Lang (2004) and Hoxby and Weingarth (2005) exploits variation in peer composition driven by school policies. Angrist and Lang exploit a Boston policy that bused predominantly black students from Boston into predominantly white schools in the surrounding districts. When interpreted as ability peer effects, their least squares estimates suggest effects of around 0.25; their instrumental variables estimates are smaller and consistent with no impacts. Hoxby and Weingarth exploit a district reassignment policy that generated large-scale changes in peer composition. Using these changes, they estimate linear-in-means peer effects of around 0.35, consistent with the previous literature. They then explore alternative models of peer effects, finding most support for “Boutique” and “Focus” models in which students benefit from the addition of students of similar ability. As they note, this could be because this peer structure allows teachers to tailor lessons to these students.

[44] Summers and Wolfe (1977) find that teacher quality is positively related to undergraduate college selectivity; Hanushek (1971) finds that teacher quality is positively related to verbal ability; Rivkin et al. (2005) find no relationship between teacher quality and whether or not a teacher holds a master’s degree.

Applied to the setting considered here, where the across-school difference in peer quality is around two standard deviations, a linear-in-means effect of 0.25 (towards the bottom of the range of previous estimates) would suggest selective school effects of around 0.5 standard deviations per year. In fact, my least squares estimates suggest that the impact of attending selective schools for four years is around one-half as large. My instrumental variables estimates are even smaller and not statistically distinguishable from zero. There are at least three reasons why these selective school estimates may be low in relation to these linear-in-means peer effect estimates. First, there may have been ability grouping within the schools in my sample. In that case, the across-school difference in peer quality could dramatically overstate the difference in classroom peer quality experienced by marginal students in nonselective and selective schools. While this type of ability grouping may have occurred in later grades, Miles and Skipworth note that in the early grades the sample schools did not group students this way. Second, it may be that linear-in-means peer effects are weaker in high schools, perhaps because older students are less susceptible to the influences of peers. Although the peer-effects literature has focused on elementary schools, where peer groups are easier to identify, Angrist and Lang present evidence consistent with peer effects weakening in more advanced grades. Finally, as implied by the Hoxby and Weingarth analysis, peer-effect estimates based on a linear-in-means specification may overstate selective school effects because this specification does not capture the structure of high school peer effects. Other peer-effects models (e.g., Boutique) have no obvious implications for the size and sign of the selective school effect. If these are better able to capture the true structure of high school peer effects, it is less surprising that my selective school estimates are small.

There are no obvious benchmarks for the estimated effects on course-taking. As noted above, the importance of this outcome is shaped by the rules governing progression to post-secondary education. To a large extent, these rules are setting specific. However, it is worth considering why selective schools could affect course-taking without affecting test scores. First, although peers appear not to affect test scores, they might reduce the social costs of pursuing various course (and college) options. Consistent with this possibility, Sacerdote notes that in the context of higher education peers appear to exert a stronger influence on social outcomes, such as drinking and joining a fraternity, than on academic outcomes, such as GPA. Alternatively, selective school students might pursue more courses because they anticipate (relatively) larger gains to pursuing them. These gains might be driven by selective school teachers, who are (relatively) more qualified to teach them.

To assess whether selective schools improve course outcomes, one could estimate selective school impacts on the (externally evaluated) grade obtained on each course. Given the extent of selection into course-taking, however, these estimates would likely be biased against selective schools. Alternatively, one can look at evidence from settings in which between-school variation in the number of courses taken is smaller. In particular, Maurin and McNally (2006) study an expansion of the selective school track in Northern Ireland and find a large increase in the probability of passing five or more courses at grade C or above, and Atkinson et al. (2006) use age-in-grade to instrument selective school attendance and find that selective school attendance has a significant impact on average course points. Assuming that selective schools do not affect the number or type of courses taken, these results would be consistent with positive selective school effects on course performance.

Whatever the mechanisms driving these curricula differences, they are the most obvious explanation for any selective school effects on university enrollment. This is because students cannot enroll in university without passing the relevant A Level exams and because it is difficult to pass the relevant A Level exams without first passing the relevant O Level exams. These curricula differences are also likely to generate longer-run selective school impacts on earnings, based on both the evidence relating additional years of education to earnings (see, e.g., Oreopoulos 2006) and on the estimated earnings returns to advanced qualifications.[45]

It is not clear whether these course-taking effects would be as large in a contemporary setting. In particular, there has been a large increase in the fraction of students staying in education beyond the compulsory school leaving age and a large increase in the fraction of students attending university in the UK. If lower-ability students have higher educational expectations, selective school attendance might have a less dramatic effect on course-taking and other longer-run outcomes. Against that, some selective school characteristics seem time invariant. First, schools that are assigned the highest-achieving students will always be able to offer courses that are not viable in schools assigned lower-achieving students. Second, these schools are always likely to attract the most qualified teachers, as they are needed to teach these courses and because these schools are likely to be more attractive to the most qualified teachers. Third, to the extent that there are social peer effects these will always generate selective school effects on outcomes that might include course-taking. As such, for as long as secondary school course-taking shapes post-compulsory educational opportunities (as it still does), we might not be surprised to see effects on longer-run outcomes.

[45] For example, in a careful study of the returns to various UK qualifications, McIntosh (2006) estimates a wage return to passing five or more courses (versus one to four courses) of over 20 percent. Because these effects are found for cohorts exposed to selective schools and cohorts not exposed to selective schools, these estimates are not picking up selective school effects.

# 7 Conclusions

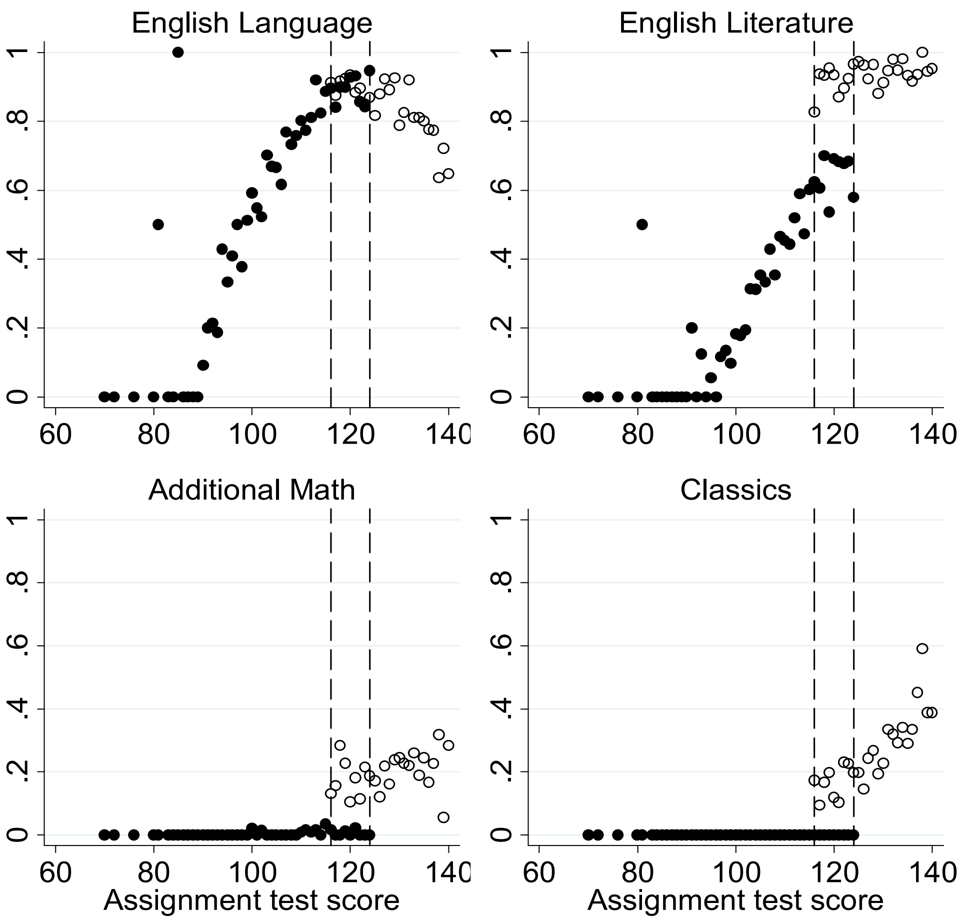
In this paper, the estimated effects of attending a selective school in a particular UK school district are discussed. Selective school places were allocated to the top 20 percent of the district's students, and this assignment rule generated dramatic differences in peer quality across schools. These differences, in addition to the variability in teacher qualifications, might lead one to expect these schools to improve the test scores of the students who attended them. In contrast, I find small effects on a range of standardized test scores. This is consistent with papers showing small test-score effects of attending private high schools and high-achieving public high schools in the U.S. (Altonji et al.; Cullen et al. 2006) and suggests that these results may generalize to other countries and other settings in which across-school differences in peer and teacher quality are even larger. Among other things, this implies that able high school peers have subtler effects than those implied by linear-in-means specifications of the education production function.

Despite these small test-score effects, I show that selective schools had large impacts on course-taking and, more suggestively, on university enrollment. Given the university admission rules facing students in my sample, it would not be surprising if course-taking effects were associated with university enrollment effects; given other estimates from the UK literature, it would not be surprising if these course and university effects had a longer-run impact on labor market earnings. The combination of large effects on longer-run outcomes despite small effects on test scores is consistent with the U.S. private schools literature and would explain why UK selective schools appear popular with parents.[46] More generally, because selective schools typically offer an advanced curriculum, they are likely to exhibit differences between their effects on test scores and their effects on longer-run outcomes. Because this advanced curriculum and its interaction with college admission rules may vary across settings, these longer-run effects are also likely to be setting specific. The obvious implication is that, where possible, these types of evaluations should consider effects on longer-run outcomes as well as those on outcomes more readily available in standard administrative datasets.

[46] Evidence supportive of the claim that these schools are popular comes from the vigorous campaigns to prevent their closure and anecdotal evidence of the efforts expended by parents to ensure their children meet the admission requirements.

# Appendix

**Figure 1: Advanced Course-taking by Assignment Test Score and School Type**




Notes: The graphs are based on data for students in the course files with non-missing assignment-test-score information (N=4004). Solid (open) circles are mean scores among nonselective (selective) school students with the corresponding assignment test score. The broken vertical lines indicate the "marginal" assignment test scores (over which the probability of attending a selective school conditional on the assignment test score increases from around 0.1 to 0.9).

**Table 1: Descriptive Statistics**

| | All Schools | | | Selective Schools | | | Nonselective Schools | | |
|---|---|---|---|---|---|---|---|---|---|
| | Mean | sd | N | mean | sd | N | mean | sd | N |
| **Panel A: Test Files** | | | | | | | | | |
| Cohort 6 | 0.314 | 0.464 | 6219 | 0.352 | 0.478 | 926 | 0.307 | 0.461 | 5293 |
| Cohort 7 | 0.34 | 0.474 | 6219 | 0.34 | 0.474 | 926 | 0.34 | 0.474 | 5293 |
| Cohort 8 | 0.346 | 0.476 | 6219 | 0.308 | 0.462 | 926 | 0.352 | 0.478 | 5293 |
| Boy | 0.497 | 0.5 | 6219 | 0.473 | 0.5 | 926 | 0.502 | 0.5 | 5293 |
| SES | 2.982 | 0.94 | 6219 | 3.481 | 0.911 | 926 | 2.894 | 0.917 | 5293 |
| Assign Missing | 0.204 | 0.403 | 6219 | 0.217 | 0.412 | 926 | 0.202 | 0.401 | 5293 |
| Assignment | 104.248 | 14.582 | 4951 | 126.942 | 6.76 | 725 | 100.355 | 11.737 | 4226 |
| Peer | 104.359 | 9.677 | 6219 | 127.077 | 1.636 | 926 | 100.384 | 1.856 | 5293 |
| IQ | 35.565 | 15.276 | 5930 | 57.81 | 11.365 | 889 | 31.642 | 12.21 | 5041 |
| English | 22.658 | 7.481 | 5892 | 31.115 | 4.296 | 872 | 21.189 | 6.921 | 5020 |
| Science | 10.873 | 6.184 | 5914 | 18.702 | 5.373 | 876 | 9.512 | 5.231 | 5038 |
| Math | 14.241 | 4.506 | 3942 | 18.75 | 3.298 | 597 | 13.436 | 4.209 | 3345 |
| **Panel B: Course Files** | | | | | | | | | |
| Cohort 1 | 0.153 | 0.36 | 5572 | 0.178 | 0.383 | 1909 | 0.139 | 0.347 | 3663 |
| Cohort 2 | 0.148 | 0.355 | 5572 | 0.154 | 0.361 | 1909 | 0.145 | 0.352 | 3663 |
| Cohort 3 | 0.152 | 0.359 | 5572 | 0.165 | 0.371 | 1909 | 0.145 | 0.352 | 3663 |
| Cohort 4 | 0.165 | 0.371 | 5572 | 0.158 | 0.365 | 1909 | 0.169 | 0.375 | 3663 |
| Cohort 5 | 0.185 | 0.388 | 5572 | 0.167 | 0.373 | 1909 | 0.195 | 0.396 | 3663 |
| Cohort 6 | 0.197 | 0.398 | 5572 | 0.179 | 0.383 | 1909 | 0.206 | 0.405 | 3663 |
| Boy | 0.476 | 0.499 | 5572 | 0.476 | 0.5 | 1909 | 0.475 | 0.499 | 3663 |
| SES | 3.214 | 1.445 | 5572 | 3.128 | 1.355 | 1909 | 3.259 | 1.488 | 3663 |
| Assign missing | 0.281 | 0.45 | 5572 | 0.203 | 0.402 | 1909 | 0.322 | 0.467 | 3663 |
| Assignment | 116.203 | 11.08 | 4004 | 126.714 | 6.815 | 1522 | 109.758 | 7.763 | 2482 |
| Peer | 115.833 | 8.262 | 5572 | 126.749 | 2.198 | 1909 | 110.145 | 2.614 | 3663 |
| Attempts | 4.659 | 2.728 | 5572 | 7.128 | 1.633 | 1909 | 3.373 | 2.259 | 3663 |
| Passes | 2.952 | 2.656 | 5572 | 5.138 | 2.528 | 1909 | 1.812 | 1.899 | 3663 |
| Avg Points | 4.243 | 1.79 | 5084 | 5.025 | 1.696 | 1909 | 3.773 | 1.677 | 3175 |
| Pass>=5 | 0.286 | 0.452 | 5572 | 0.608 | 0.488 | 1909 | 0.119 | 0.324 | 3663 |
| UnivPass>=5 | 0.303 | 0.46 | 783 | 0.366 | 0.482 | 558 | 0.147 | 0.355 | 225 |
| **Panel C: Merged Test and Course Files (cohort 6)** | | | | | | | | | |
| 1 - Drop Out | 0.475 | 0.499 | 1947 | 0.938 | 0.242 | 321 | 0.384 | 0.486 | 1626 |
| Pass >=5 | 0.138 | 0.345 | 1947 | 0.576 | 0.495 | 321 | 0.052 | 0.221 | 1626 |
| University | 0.035 | 0.184 | 1913 | 0.188 | 0.391 | 293 | 0.0074 | 0.086 | 1620 |

Notes: The Univ|Pass>=5 denominator is smaller than the total number of students with pass>=5 because some of those have birth dates missing (or in a small number of cases are not unique) and hence cannot be matched to the university data. The same applies to a small number of students in the merged test and course data.

**Table 2: The Impact of Selective Schools on Grade Nine Test Scores: Robustness Checks**

| | Base specification | Change polynomial: 3rd order | Change polynomial: 5th order | Widen test score range [90,140] | Narrow test score range [100,130] | Add obs w/ missing scores (=0) | Dep var: questions attempted |
|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| | | | | **Panel A: Math** | | | |
| Selective | 0.265 | 0.177 | 0.514 | 0.343 | 0.428 | -0.238 | 0.247 |
| | (0.159) | (0.150) | (0.204) | (0.176) | (0.203) | (0.164) | (0.183) |
| | | | | **Panel B: Reading** | | | |
| Selective | -0.168 | -0.211 | -0.029 | -0.225 | -0.148 | -0.153 | -0.276 |
| | (0.127) | (0.101) | (0.219) | (0.144) | (0.142) | (0.169) | (0.135) |
| | | | | **Panel C: Science** | | | |
| Selective | 0.083 | -0.059 | -0.017 | -0.055 | 0.054 | 0.002 | 0.087 |
| | (0.182) | (0.135) | (0.143) | (0.144) | (0.186) | (0.179) | (0.185) |
| | | | | **Panel D: Combined score** | | | |
| Selective | -0.049 | -0.073 | 0.253 | -0.005 | 0.060 | -0.153 | 0.058 |
| | (0.113) | (0.112) | (0.204) | (0.152) | (0.149) | (0.169) | (0.155) |
| | | | | **Panel E: IQ** | | | |
| Selective | 0.141 | 0.129 | 0.171 | 0.144 | 0.255 | -0.189 | 0.019 |
| | (0.102) | (0.101) | (0.106) | (0.176) | (0.137) | (0.147) | (0.130) |

Notes: The Univ|Pass>=5 denominator is smaller than the total number of students with pass>=5 because some of those have birth dates missing (or in a small number of cases are not unique) and hence cannot be matched to the university data. The same applies to a small number of students in the merged test and course data.

## References

Altonji, J., T. Elder, and C. Taber. 2005. Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools. *Journal of Political Economy* 113 (1): 151–184.

Angrist, J., and K. Lang. 2004. Does School Integration Generate Peer Effects? Evidence from Boston's Metco Program. *American Economic Review* 94 (5): 1613–1634.

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

Atkinson, A., P. Gregg, and B. McConnell. 2006. The Result of 11 Plus Selection: An Investigation Into Opportunities and Outcomes for Pupils in Selective Schools. Centre for Market and Public Organisation Working Paper 06/150, Bristol University, England.

Ballou, D., E. Goldring, and K. Liu. 2006. Magnet Schools and Student Achievement. Paper presented at the annual meeting of the American Educational Research Association.

Betts, J., and H. Rose. 1996. The Effect of High School Courses on Earnings. *Review of Economics and Statistics* 86:497–513.

Brunello, G., and D. Checchi. 2007. Does School Tracking Affect Equality of Opportunity? New International Evidence. *Economic Policy* 52:781–861.

Coleman, J., T. Hoffer, and S. Kilgore. 1982. *High School Achievement: Public, Catholic and Private Schools Compared*. New York: Basic Books, Inc.

Cullen, J., B. Jacob, and S. Levitt. 2005. The Impact of School Choice on Student Outcomes: An Analysis of the Chicago Public Schools. *Journal of Public Economics* 89 (5–6): 729–760.

———. 2006. The Effect of School Choice on Student Outcomes: Evidence from Randomized Lotteries. *Econometrica* 74 (5): 1191–1230.

Dustmann, C. 2004. Parental Background, Secondary School Track Choice and Wages. *Oxford Economic Papers* 56:209–230.

Evans, W., and R. Schwab. 1995. Finishing High School and Starting College: Do Catholic Schools Make a Difference? *Quarterly Journal of Economics* 110 (4): 941–974.

Gamoran, A. 1996. Student Achievement in Public Magnet, Public Comprehensive, and Private City High Schools. *Educational Evaluation and Policy Analysis* 18: 1–18.

Hanushek, E.A. 1971. Teacher Characteristics and Gains in Student Achievement. *American Economic Review* 61:280–288.

Hanushek, E.A., and L. Wössmann. 2006. Does Educational Tracking Affect Performance and Inequality? Differences-in-Differences Evidence across Countries. *The Economic Journal* 116 (510): C63–C76.

Hastings, J.S., T.J. Kane, and D.O. Staiger. 2006. Preferences and Heterogeneous Treatment Effects in a Public Choice School Lottery. NBER Working Paper 12145 (April).

Hoffman, L. 2007. Numbers and Types of Public Elementary and Secondary Schools From the Common Core of Data: School Year 2005-06. U.S. Department of Education NCES 2007-354. National Center for Education, Washington, D.C.

Hoxby, C.M. 2000. Peer Effects in the Classroom: Learning from Gender and Race Variation. NBER Working Paper No. 7867 (August).

Hoxby, C.M., and G. Weingarth. 2005. Taking Race out of the Equation: School Reassignment and the Structure of Peer Effects. Mimeo, Harvard University.

Imbens, G.W., and T. Lemieux. 2008. Regression discontinuity designs: A guide to practice. *Journal of Econometrics* 142 (2): 615–635.

Klopfenstein, K. 2004. The Advanced Placement Expansion of the 1990s: How Did Traditionally Underserved Students Fare? *Education Policy Analysis Archives*, vol. 12, no. 68.

Klopfenstein, K., and K. Thomas. 2006. The Link Between Advanced Placement Experience and Early College Success. Mimeo, Texas Christian University, 2006.

Manning, A., and J.-S. Pischke. 2006. Comprehensive versus Selective Schooling in England and Wales: What Do We Know. NBER Working Paper 12176 (April).

Maurin, E., and S. McNally. 2006. Selective Schooling. Mimeo, London School of Economics.

McIntosh, S. 2006. Further Analysis of the Returns to Academic and Vocational Qualifications. *Oxford Bulletin of Economics and Statistics* 68 (2): 225–251.

Miles, H.B., and G.E. Skipworth. 1974. Final Report to The Social Science Research Council on an Investigation of Some Correlates of Academic Performance of Pupils in Secondary Schools. University of Hull, Department of Educational Studies.

Milewski, G.B., and J.M. Gillie. 2002. What are the Characteristics of AP Teachers? An Examination of Survey Research. The College Board Research Report No. 2002-10, College Entrance Examination Board, New York.

Mortimore, M. 1999. *Bridlington School: A History*. Driffield, England: Alan Twiddle Publishing.

Murnane, R.J., J.B. Willett, and F. Levy. 1995. The Growing Importance of Cognitive Skills in Wage Determination. *Review of Economics and Statistics* 77:251–266.

Neal, D. 1997. The Effects of Catholic Secondary Schooling on Educational Achievement. *Journal of Labor Economics* 15 (1): 98–123.

Oreopoulos, P. 2006. Estimating Average and Local Average Treatment Effects of Education when Compulsory Schooling Laws Really Matter. *The American Economic Review* 96:152–174.

Planty, M., S. Provasnik, and B. Daniel. 2007. High School Coursetaking: Findings from The Condition of Education 2007. U.S. Department of Education NCES 2007-065, National Center for Education, Washington, D.C.

Rivkin, S., E. Hanushek, and J. Kain. 2005. Teachers, Schools and Academic Achievement. *Econometrica* 73 (2): 417–458.

Sacerdote, B. Forthcoming. Peer Effects in Education: How might they work, how big are they and how much do we know thus far? *Journal of Economic Perspectives*.

Secondary and Further Education Sub-Committee. 1964. Allocation to Secondary Schools, 1964: Memorandum by Chief Education Officer. Appendix to Minute No. 10, East Riding Secondary and Further Education Sub-Committee (September 4).

Steel, L., and R. Levine. 1996. Educational Innovation in Multicultural Contexts: The Growth of Magnet Schools in American Education. American Institutes for Research, Palo Alto, CA.

Sullivan, A., and A. Heath. 2004. Students' academic attainments and British state and private schools. Mimeo, Department of Sociology, University of Oxford, Oxford, England.

Summers, A., and B. Wolfe. 1977. Do Schools Make a Difference? *American Economic Review* 67:639–652.

Wooldridge, J. 2002. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: The MIT Press.