DISCUSSION PAPER SERIES

I Z A

IZA DP No. 3182

Selective Schools and Academic Achievement

Damon Clark

November 2007

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Selective Schools and Academic Achievement

Damon Clark
*University of Florida
and IZA*

Discussion Paper No. 3182
November 2007

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 3182
November 2007

ABSTRACT

Selective Schools and Academic Achievement*

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 university enrollment, evidence suggesting they may have important longer run impacts.

JEL Classification: C21, I21

Keywords: selective schools, education, instrumental variables

Corresponding author:

Damon Clark
Department of Economics
University of Florida
PO Box 117140
Gainesville, FL 32611-7140
USA
E-mail: damon.clark@cba.ufl.edu

* I thank Liangliang Jiang and Matt Masten for excellent research assistance and David Card, Ken Chay, Julie Cullen, David Figlio, Jason Fletcher, Caroline Hoxby, Larry Kenny, David Lee, Tom Lemieux, Jens Ludwig and seminar participants at the University of Florida, University of California, Berkeley, University of British Columbia, New York University, NBER Spring Education Meetings 2007, Society of Labor Economics Meetings 2007, Lancaster University, Bristol University and the Institute of Education 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. I also thank Ofer Malamud for help with the university enrolment data. Partial funding for this work was provided by a National Academy of Education/Spencer Post-Doctoral Research Fellowship.

1 Introduction

In several European countries including France, Germany, Holland, Italy and parts of the UK, different types of students are tracked into different types of high school based on academic ability as perceived at the end of primary school. In the US, 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 either a range of courses (such as the International Baccalaureate) or a more specialized field (such as Science or I.T.), can only be followed by students that have satisfied 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, since they add to the wider literature on the effects of attending different types of high school. One strand of this literature, begun by Coleman, Hoffer, and Kilgore (1982), considers the impacts of attending a private high school in the US. Coleman, Hoffer, and Kilgore (1982) found that private schools improved test scores; more recent evidence points to small test score effects (Altonji, Elder, and Taber (2005)) but larger effects on longer-run outcomes (Evans and Schwab (1995), Neal (1997) and Altonji, Elder, and Taber (2005)). The

[1] Hoffman (2007) reports that in 2005-06, of the 97,382 schools in the US, 2,796 were magnet schools. Although the number of academically selective magnet high schools is not reported, Planty, Provasnik, and Daniel (2007) report that in 2002-3, of the 16,500 high schools in the US, 390 offered the International Baccalaureate. According to figures from the International Baccalaureate Organization, the number of high schools offering the IB program increased by more than 50% between 2002 and 2007, from 390 to 555 (personal communication).

[2] Although there are no formal national requirements for teachers teaching in magnet schools or teaching IB programs in the US, because these advanced courses cover college-level material, teachers will usually be expected to have at least a Bachelors degree in a related field. The College Board offer similar guidelines for Advanced Placement (AP) teachers. Milewski and Gillie (2002) report that 69.7% of AP teachers hold a masters degree, compared to 38% of teachers overall.

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

other strand of this literature considers the impact of attending a high-achieving public school, defined as one with more able students. Coleman, Hoffer, and Kilgore (1982) 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, Jacob, and Levitt (2005), Cullen, Jacob, and Levitt (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 the difference between peer quality in selective and non-selective schools is extreme, a linear-in-means peer effects model - a model of student achievement that was an additive linear function of own ability and school-mean ability - would predict large selective school test score impacts. As such, analysis of these impacts provides a powerful test of this linear-in-means specification, the workhorse model in the education literature (Hoxby and Weingarth (2005)). Second, 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 could affect these outcomes without affecting test scores. Some, such as peer effects on college enrollment, could operate in a world in which the high school curriculum was 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]

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, were established fairly recently. This is especially true in the US, where the number of students in magnet programs tripled between 1982 and 1991 (Steel and Levine (1996)) and where programs differ by curricula

[4] Coleman, Hoffer, and Kilgore (1982) showed that both offer a disproportionate number of Advanced Placement courses. More recently, Klopfenstein (2004) documented large differences in the number of Advanced Placement courses offered by schools.

[5] **?** show that choice of high school course can have large impacts on earnings. **?** argue that many colleges upweight the high school GPAs of students with Advanced Placement credits.

and admissions procedures.[6] Second, 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 US districts provide the ideal tool for dealing with this problem, but lotteries are rarely used outside of the US.[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 US, 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% 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. Since private school enrollment is low (around 7%) and non-selective 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, 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 generated a sharp change in the probability of attending selective school over a narrow range of assignment test scores and I use instrumental variables and regression discontinuity methods to exploit this sharp change and 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 had small effects on standardized test scores. Ordinary least squares estimates suggest effects of around one third of a standard deviation. While these

[6]Based on an analysis of the National Educational Longitudinal Study (NELS), Gamoran (1996) finds that students in city magnet schools (stand-alone and school-within-school) enjoyed higher science, reading and social studies test scores then students in urban comprehensive public high schools.

[7]Ballou, Goldring, and Liu (2006) study magnet schools using administrative data from a school district that uses lotteries to allocate magnet school places, but they conclude that attrition biases are too large to make strong statements about their effects.

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

are close to those found in the previous UK literature (Sullivan and Heath (2004)), instrumental variables estimates suggest effects of around one sixth of a standard deviation, with confidence intervals failing to rule out zero effects and consistently ruling out effects larger than the least squares estimates. This implies that for these outcomes, attending a selective school was less beneficial than having high-SES or highly-educated parents, since both social class and parental education are more strongly correlated with test outcomes even conditional on assignment test scores. The test score effects are especially small in relation to the difference in peer quality between selective and non-selective school students, which is twice as large as the difference in peer quality between students entering high- and low-achievement high schools in Chicago. Despite the absence of any test score effects, I present evidence consistent with the hypothesis that selective schools improved longer run outcomes. In particular, I find that selective schools had large effects on course-taking, enabling students to pursue advanced courses only offered in selective schools and increasing the probability that they studied academic courses available in all types of schools. Given UK rules governing progression into post-compulsory education, these course effects would be expected to generate effects on university enrollment and I provide evidence that suggests this is the case.

My finding of small test score effects despite huge differences in peer quality is in line with evidence on the test score effects of attending private and high-achieving public schools in the US. Since my findings suggest these results may generalize to other countries and other settings in which across-school peer differences are even larger, they add weight to the view that high school peers and teachers have more subtle effects than those implied by linear-in-means specifications of the education production function. Although I argue that the selective schools that I study likely improved longer run outcomes it would not be sensible to generalize from these results to the likely long-run effects of other types of selective schools 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 could also improve longer run outcomes without improving test scores. This is consistent with the difference between test score and longer run outcomes found for US private schools. Moreover, 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 Conceptual Framework

Before describing UK selective schools and estimating their effects, I sketch a simple conceptual framework that helps clarify the mechanisms underpinning selective school effects and the caveats that must accompany their interpretation. Consistent with the definition introduced above, I assume selective schools offer an advanced academic curriculum and admit only able students. To be more precise, assume that student ability, a, is distributed according to a standard uniform distribution and that students wishing to attend selective schools must demonstrate they have ability greater than a^*. By assuming that a fraction p of eligible students are admitted to selective school, I cover the case in which access is rationed by lottery (as in some US districts) and the case in which all eligible students are admitted (as in the UK, described in more detail below).

Consider first the selective school impact on standardized test scores. This will depend on the education production function which, for expositional purposes, we can assume depends on some function of own ability and the ability of school peers: $T_{ij} = f(a_i, a_{-ij})$, where T_{ij} is the test score of student i in school j, a_i is student i's own ability and a_{-ij} is the vector of abilities of all other students in the school. As noted above, in the special case in which this function is linear in means ($T_{ij} = b_0 + b_1 a_i + b_2 \overline{a_{-ij}}$) and peer effects are important ($b_2 > 0$), selective schools will have large test score impacts. These can be written as: $E(selective|a_i \in [a^*, 1]) = b_2 E[\overline{a_{-ij}}|sel] - E[\overline{a_{-ij}}|non] = \frac{b_2}{2}[\frac{a^*}{a^*+(1-p)(1-a^*)}] > 0$, where the last step involves substituting expected mean ability in the two schools calculated using the properties of the standard uniform distribution. Notice that for a given cutoff, the difference in mean ability is increasing in the fraction that attend the selective school and, for a given fraction, increasing with the cutoff. In the case in which $p = 1$ (all eligible students attend selective schools), the difference is $\frac{1}{2}$.

With this algebra in mind, it is obvious that the finding of small selective school effects should be

[9] One explanation put forward to explain the combination of strong parental demand and weak academic achievement effects is that these schools confer non-academic benefits (Cullen, Jacob, and Levitt (2006) show that students that opt-out have lower self-reported disciplinary incidences and arrest rates). Another is that academic achievement might improve for students whose parents choose with academic achievement in mind (Hastings, Kane, and Staiger (2006) find some evidence for this explanation among elementary school students in Charlotte-Mecklenberg).

interpreted as a rejection of the hypothesis that peer effects are both important ($\beta_2 > 0$) and linear in means ($f(a_i, a_{-ij}) = \beta_0 + \beta_1 a_i + \beta_2 \overline{a_{-ij}}$). Two simple counter-examples demonstrate that non-positive selective school test score effects could be consistent with differently structured yet important peer effects. Suppose first that test scores depend on own ability and the ability of the most able student in the school ($T_{ij} = c_0 + c_1 a_i + c_2 \max(a_{-ij})$). These effects, labelled "shining light" by Hoxby and Weingarth (2005), could be driven by more able students inspiring less able students to increase effort and improve performance. In that case, provided $p > 0$ and there were similar numbers of high-ability students in selective and non-selective schools, selective school effects would be zero despite peer effects being important ($c_2 > 0$). Suppose next that peer effects depend on own ability relative to school-mean ability ($T_{ij} = d_0 + d_1 a_i + d_2(a_i - \overline{a_{-ij}})$). These effects, labelled "invidious comparison" by Hoxby and Weingarth (2005), would be consistent with individuals drawing confidence from being of high relative ability. In that case, as specified in this example, selective school effects would be negative ($-\frac{d_2}{2}[\frac{a^*}{a^*+(1-p)(1-a^*)}] < 0$) despite peer effects being important ($d_2 > 0$). Of course these different effects need not operate in isolation: selective schools could, for example, exert positive linear-in-means effects on all selective school students but negative invidious comparison effects on selective school students of lower-than-average ability. The general point is that small selective school test score effects reject a particular model of peer effects, not their existence. Although the same point could be made in the context of private and high-achieving public school comparisons - changes in the first moment of the peer ability distribution will be associated with changes in other characteristics of the distribution including maximum peer ability and relative peer ability - those concerns are especially relevant in the selective schools context, in which selective school student ability is bounded from below. They are especially relevant in the UK selective schools context, in which all eligible students are admitted to selective schools ($p = 1$).

Consider next the selective school impact on college enrollment. To simplify, consider college enrollment decisions assuming that all students graduate high school. This is reasonable, since we can only identify selective school effects for students eligible for selective schools, all of whom will presumably graduate high schools of any type with high probability (this is true in the UK data). Assume students enroll in college when some (normalized) utility index is positive, and write this as: $U^*_{ij} = V[Q(a_i, (1 + A_j)T_{ij})] - C(a_i, a_{-ij})$.

The function $C(.)$ represents the consumption costs of attending college, abstracting from tuition costs. The function $Q(.)$ measures college quality (or selectivity) and is allowed to depend on student ability (as measured for example by SAT score) and high school GPA (assumed correlated with standardized test scores), mediated by A_j, an indicator for whether school j offers an advanced academic curriculum. The function $V(.)$ maps college quality to utility. Assuming this function is increasing, students will choose the highest quality college to which they are admitted; provided the highest-quality college generates positive utility, they will attend college.

This framework points to two mechanisms by which selective schools can influence college enrollment prospects independent of test score effects. First, because the cost function $C(.)$ is allowed to depend on the abilities of other students in the school, the consumption costs of college enrollment can be influenced by peers. Second, the high school curriculum can, depending on $Q(.)$, influence college quality. This would not be the case if college admission decisions are based solely on ability (e.g., SAT scores) but would be the case if colleges up-weight the GPAs of students that have studied advanced curricula. For the US, Klopfenstein and Thomas (2006) argue that many colleges do upweight high school GPA to take account of Advanced Placement courses. As discussed below, in the period that I study, UK college admission rules were even more tightly linked to the school curriculum, since there was no equivalent to the SAT and high school courses set students on a narrow path to university degrees. An important implication of the dependence of longer run selective school effects on the function $Q(.)$ is that these effects will depend on the interaction between the particular curriculum offered and college admission rules. As such, longer-run selective schools effects will be setting-specific. For example, it may be that UK selective schools have different effects on longer run outcomes in the 1990s than they did in the 1970s (the period from which my data are drawn); it may be that US magnet schools offering science programs have different effects than US magnet schools offering the IB program.

Finally, consider the selective school impact on labor market earnings. For students that attend college, there will be selective school effects on earnings independent of selective school effects on test scores provided earnings are associated with college selectivity or quality (as found by Hoekstra (2007) although not Dale and Krueger (2002)). In addition, there will be selective school earnings effects for students that do not attend college provided advanced academic curricula attract a labour market premium (as found by Betts and Rose (1996)). The overall impact of selective schools on earnings includes both of these effects plus the selective school effect on college enrollment. Again, the selective schools context is not the only one in which longer run effects can operate independent of effects on high school test scores, and the literature on US private schools is consistent with such a divergence. Nevertheless, since the high school curriculum is an important channel through which these effects may flow, and since selective schools are defined in terms of the advanced academic curriculum that they provide, the divergence between test score and longer run outcomes may be especially marked in this context. Moreover, since curricula and college admission rules may differ across selective school settings, it is clear that these longer run effects may be setting-specific.

3 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 some scholarships in exchange for some financial support from the local school district; after the Act, they received all of their funding from the district, were not allowed to charge fees and were required to admit students on the basis of academic potential assessed in grade five (at age ten or eleven).[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 non-selective schools, known as modern schools, remained broadly unchanged, the important caveat being that where they previously educated all but those students that won scholarships to grammar schools, they now educated all students that "failed" the test.[11,12]

Most of the attention in this paper is focused on the East Ridings, a large school district in the north-east of England and the only one for which assignment test scores are available.[13] Until 1974, students were assigned to the district's grammar schools on the basis of two pieces of information: a district-wide

[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 contained a mixture of elite Grammar and non-elite Modern schools.

[12] 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 US.

[13] The schools were gradually integrated after 1974, but all of the students in these data had left school by that time.

assignment test taken in grade five and a primary school order of merit supplied by the primary school Head Teacher.[14] On the basis of this information, a district-wide 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 places and passed the names of borderline students to a second "Assessors Panel". This gave detailed consideration to borderline students, those that had been referred for special investigation and those that 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 where 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. Since primary rank and assignment test score is likely to be highly but not perfectly correlated, and since some parents may have turned down a selective school place, it is not surprising that the relationship is as presented in Figure 1a: 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, since there were in effect separate competitions by sex and cohort.[15] The next section discusses the econometric implications of this assignment procedure in more detail.

The four grammar schools in East Ridings 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, students with assignment test scores to the right of the marginal 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 single-sex, gave principals (Head Teachers) a large measure of autonomy and employed teachers who were, on average, better qualified. 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

[14] 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 Education 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.

[15] The data used to generate Figure 1 are described in more detail below.

attended Universities rather than teacher training colleges.[16] Along other dimensions including teacher age, teacher experience and teacher turnover, across-school differences were much smaller.[17] The third panel of Figure 1b suggests there was no consistent difference in the amount of homework assigned to students: according to the estimates of Miles and Skipworth (1974), boys with higher assignment test scores attended schools that assigned slightly more homework, girls with higher assignment test scores attended schools that assigned slightly less.

Although this paper focuses on these twenty selective and non-selective schools, the East Ridings had two "bilateral" high schools and three private schools. 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 (which contained selective and non-selective tracks) were geographically isolated and probably not a realistic option for most of the children that eventually attended schools in the sample.[18] Private school may have been a realistic option for wealthier parents, and since I do not have data on private school students, one might worry that a disproportionate number of these will be children from high-SES families that narrowly failed to secure a selective school place. There is some anecdotal evidence to support this story and the relationship between assignment test score and SES seen in Figure 1a is also consistent with it.[19] There are however two reasons for thinking that while data on students in private schools would be useful, they would almost certainly have no effect on the main results. First, private schools enrolled only around 8% of the district's students. Among the students we are most concerned about - those that narrowly failed to secure a selective place - this fraction may be even smaller.[20] Second, while my estimated selective school effects are small (at least for test scores), this type of phenomenon would bias estimated selective school impacts upwards: if it is true that students otherwise bound for non-selective

[16] According to a former master at one of the boys elite 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.

[17] Results available upon request.

[18] 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 elite school). The School District explicitly stated that these schools would not compete for students with existing schools

[19] Anecdotal evidence from private correspondence with Michael Mortimore. To the extent that SES is higher among marginal students that attend elite rather than non-elite schools (confirmed in regression-based models of the relationship between assignment test score, SES and the probability of attending elite school), it suggests that some high-SES students that failed to secure an elite school place may have enrolled in private school.

[20] This number is based on data for individual schools from 1964 edition of the Education Authorities and Schools Directory. In particular, I calculate the total number of students in private schools in the East Ridings in this year (1345) and assume a drop out rate of 20% in order to arrive at a number per grade of around 200 (around 8% of the per-grade sample of around 2500). This could be an under-estimate if some students attend private schools in another district and an over-estimate if a greater number of non-district residents attend these private schools.

schools choose private school, then the true selective school effects will be even smaller.

Like other UK grammar schools, the East Ridings grammar schools provided an advanced academic curriculum. Students were expected to stay in school until the end of tenth grade (one year later than the compulsory leaving age) and, in the final two grades, pursue a set of externally evaluated 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 non-selective schools, practice varied by ability: lower-ability students would be expected to leave at the compulsory leaving age and would, in their final grade, take non-examined lower-level courses;[21] high-ability students would be encouraged to stay the extra year and pursue the same type of academic courses as selective school students. All non-selective schools offered Mathematics and English Language courses, but few offered more advanced courses such as English Language, Physics and Chemistry, and none offered advanced courses such as Additional Mathematics and Classics. Appendix Figure 1, based on data discussed below, supports this taxonomy of courses into those routinely offered by non-selective schools (English), those occasionally offered by selective schools (English Literature) and those never offered by selective schools (Advanced Mathematics and Classics). Since 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. Below, I use the full sample of students to estimate the causal effects of selective schools on various course-taking outcomes.

Students wishing to continue in education beyond grade ten could pursue two-year "A level" courses and, if they applied and were accepted, attend university or Higher Education colleges. Students' choice of university degree would depend on A level choices. In turn, students' A level choices would, to a large extent, be determined by their choice of high school course. Rarely would a student apply to study physics at university without studying physics at A level; rarely would a student choose an A level physics course had they not taken and passed the high school physics course. As a result, it is likely that selective school effects on high school course-taking influenced the type of university course that students applied for and

[21] These were called Certificates of Secondary Education (CSE).

the probability that they applied for a university course. The structure of the UK selective school system is illustrated in the Web Appendix.

4 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 Ridings, it was based on these test scores and a primary school order of merit. To motivate the strategy used to identify effects in this context suppose, for expositional purposes, that the true relationship between selective school attendance and some outcome y_i is:

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

where d_i is a dummy variable indicating whether or not individual i attended a selective school, a_i is ability and ε_i is a random disturbance term. Since ability is unobserved, least squares estimation of equation (1) will not identify the causal effect of attending selective school. Instead, estimates will probably be upward-biased because of the positive correlation between selective school attendance and ability. This will be the case even after controlling for observable family characteristics thought to proxy ability, such as socioeconomic status, since there is likely to 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. In the setting considered here, the same problem might also arise if we also control for these family characteristics and also the assignment test score, since 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. This is the familiar ability bias associated with least squares estimation of the return to years of schooling, except that in this case we are considering a particular type of schooling.

4.1 Identification

The key to identifying the causal effect of selective school attendance is the use of across-t variation in the probability of attending selective school, where t refers to the assignment test score. To see this more clearly, we can decompose ability into the part that can be predicted by the assignment test score (t_i) and the primary school rank (p_i) and the part that cannot:

$$a_i = E(a_i|t_i, p_i) + u_i$$

where u is by definition independent of t and p. Substituting into (1) gives:

$$y_i = \beta_0 + \beta_1 d_i + E(a_i|t_i, p_i) + u_i + \varepsilon_i$$

and taking expectations with respect to t_i gives:

$$E(y_i|t_i) = \beta_0 + \beta_1 P(d_i = 1|t_i) + E(a_i|t_i)$$

where we use the law of iterative expectations and exploit the properties of ε and u. Since $E(a_i|t_i)$ is a function of t_i we can write:

$$\overline{y_t} = \beta_0 + \beta_1 P(d = 1|t) + f(t) + \nu_t \tag{2}$$

where ν_t is sampling error hence independent of $P(d = 1|t)$. Equation (2) then relates cell-mean outcomes to the cell probability of selective school attendance and the control function $f(.)$.

If the function $P(d = 1|t)$ was known, least squares estimation of equation (2) would identify the causal effects of interest provided three conditions held. Although the function $P(d = 1|t)$ is not known, it is still useful to consider these three conditions. First, $f(.)$ would have to be specified correctly, or else there could be a correlation between the misspecification error and $P(.)$, leading to biased estimates of β_1. In practice, since $f(.)$ would probably be specified as a low-order polynomial such as a quadratic, underlying ability would have to be a smooth function of t over the range in which $P(.)$ varied. To return to the issue of private schools, the concern is that since a fraction of the high-ability students that were not admitted to selective schools might have enrolled in private schools (hence will not be in my data), there may in the sample be a sharply increasing relationship between ability and t that is not picked up by $f(.)$.

Second, the selective school treatment effect would have to be constant over the range in which $P(.)$

varied. In equation (1), the implicit assumption is that this effect is constant across the whole range of assignment test scores (since β_1 is a constant), but the assumption could be relaxed so that the effect was a function of t provided $\beta_1(t)$ was constant over the assignment test score interval in which $P(.)$ varied. Both the simple case and this more plausible case ensure there are no omitted variables biases introduced by a correlation between selective school returns and $P(.)$. This more plausible case also implies however that the selective school effects being identified are local to students on the selective school margin, and cannot therefore be generalized to students at other points of the assignment test score distribution. This interpretation issue is not unique to the setting considered here: all instrumental variables estimates must be interpreted as local average treatment effects when effects are heterogenous (Angrist, Imbens, and Rubin (1996)). In the discussion below, I consider the implications of interpreting selective school effects as local to marginal students. Finally, to ensure that estimates are reasonably precise, $P(.)$ and $f(.)$ would have to be only weakly correlated. Since $f(.)$ will vary across outcomes, this implies that this might be an effective means of identifying selective school effects on some outcomes but not others.

4.2 Instrumental Variables strategy

As noted above, we do not know the precise form of $P(.)$ hence could not estimate equation (3) even if all three conditions held. There is, however, a simple alternative strategy that also uses across-t variation in $P(.)$ and also identifies the selective school effects under the same three conditions. This strategy involves working with data at the individual level (i.e., equation (1)), proxying a_i with $f(t_i)$ and instrumenting d_i with the predicted probability of attending selective school $P(\widehat{d_i = 1}|t_i)$ estimated for example by a logit model. Wooldridge (2002) discusses this procedure in more technical detail (Procedure 18.1, Chapter 18) and notes its advantages and disadvantages. The main advantage is that when $f(t_i)$ is specified correctly and treatment effects are constant (the first two conditions discussed above), $\widehat{P(.)}$ is a valid instrument for d_i. Note that this is true even if the probability model is misspecified (e.g., a logit is used when the true model suggests a probit would be more appropriate); note also that the standard errors in the second stage do not need to be corrected for the sampling variability inherent in estimating the predicted probability (Wooldridge (2002)). This predicted probability is also likely to be a powerful instrument, since the correlation between attending selective school d_i and the predicted probability of attending selective school $\widehat{P(.)}$ will be high. The disadvantage is that since identification is via functional form (the difference in the shape of $\widehat{P(.)}$ relative to $f(.)$), IV estimates will only be precise when $\widehat{P(.)}$ and $f(.)$ are not collinear (the third condition discussed above).

I use this instrumental variables strategy to identify the effects of selective school attendance on test scores. I predict the probability of attending selective school using a logit model with a fourth-order polynomial in assignment test scores. This turns out to be a powerful instrument, since the predicted probability tracks the empirical probability of selective school attendance very closely. Moreover, this model can be expanded to include sex and cohort interactions to take account of the separate selective school competitions taking place each year. This instrument will be valid provided $f(.)$ captures the underlying relationship between assignment test score and the outcome. My choice of $f(.)$ is guided by the data, and for the most part is specified as a quadratic in assignment test scores. I also present the results of experiments with alternative functional forms and, for the most part, the estimates remain broadly unchanged.

Regarding the correlation between $f(.)$ and $P(.)$ and the precision of these estimates, this will, as noted above, depend on the outcome being considered. In the case of test scores, the instrumental variables estimates are consistently small and precise. The precision derives from the contrast between the predicted probability of selective school attendance $\widehat{P(.)}$, which increases sharply over a narrow interval of assignment test scores (as seen in Figure 1), and the shape of $f(.)$, which appears smooth across the entire range of assignment test scores. The absence of a sharp change in outcomes despite a sharp change in the probability of selective school attendance suggests the absence of large selective school effects on these outcomes.

4.3 Regression Discontinuity strategy

There is a close connection between equation (2) and the "pure" and "fuzzy" regression discontinuity (RD) approaches. In particular, the standard RD framework with discrete running variable (in this case assignment test score) is a special case of (2) in which $P(.)$ jumps from zero to one at a particular threshold t^R. In that case, equation (2) can be rewritten:

$$\overline{y_t} = \beta_0 + \beta_1 I(t \geq t^R) + f(t) + \nu_t \tag{3}$$

where $I(.)$ is an indicator for whether the assignment test score falls to the right or left of the threshold t^R. Assuming $f(.)$ is specified correctly (the first condition noted above), least squares estimates of β_1 can be interpreted as the corresponding jump in outcomes at t^R. This holds without additional assumptions on the nature of treatment effect heterogeneity (Lee and Card (2006)). The "fuzzy" RD is a special case of equation (2) in which there is a discontinuity in the probability of treatment at the threshold t^R, but where this discontinuity can be smaller than unity. In that case, least squares estimates of β_1 continue to identify the corresponding jump in outcomes, but estimates of this jump must be scaled up by the inverse of the estimated first-stage discontinuity and the standard errors corrected to reflect the sampling variation inherent in this first-stage estimate (Lee and Card (2006)).

I estimate selective school effects on course-taking using a variation on these RD approaches. Specifically, I exploit the sharp change in the probability of selective school attendance over a relatively narrow interval of test scores $[t^L, t^R]$ by discarding data in this marginal interval and specifying an equation similar to (3) in which $I(.)$ indicates whether the assignment test score falls to the right or the left of this interval. Again, provided $f(.)$ is specified correctly, β_1 can be interpreted as the jump in outcomes at t^R. Again, since the probability of selective school attendance increases by less than unity over this interval, the interpretation of the estimated jump in outcomes must take this "first stage discontinuity" into account.

The key drawback of this approach relative to the standard RD approaches is that this marginal interval is to some extent arbitrary. I deal with this by using the data to define it. Specifically, I find the sex- and cohort-specific assignment test score that comes closest to predicting a probability of selective school attendance of one half and define the marginal area to be the eight-point symmetric interval around this (the probability of attending selective school is estimated to increase by roughly 0.85 when moving from one side of this eight-point interval to the other). There is a clear trade-off between choosing a narrow interval to minimize the number of observations dropped from the sample and a wider interval to maximize the size of the first-stage discontinuity in the probability of selective school attendance. Although the eight-point interval seemed to do the best job of balancing this trade off, this choice is to some extent arbitrary and so I present robustness checks based on intervals of varying width.

4.4 IV versus RD approaches

Both the IV and RD approaches identify selective school effects by exploiting across-t variation in $P(.)$. As a result, they will work well under similar conditions and I estimate all of the outcomes using both approaches. However, I present the main test score results using the IV approach and the main course results using the RD approach. The test score results are almost identical across the two approaches but I present the IV results because these exploit all of the data hence are expected to be more precise. The course outcomes are typically binary (such as the probability of taking any courses), and since these take us out of the linear IV framework, I present the RD results. I check their robustness using IV methods, although these are now based on maximum likelihood models that make stronger distributional assumptions (Wooldridge (2002), Chapter 15.7.3). Ultimately, the qualitatively important results in this paper - that selective schools have small test score and large course-taking effects - are clear from visual inspection of the graphs and apparent in the least squares estimates.

Finally, since the probability of university enrollment is correlated with the probability of attending selective school, neither approach works well and I instead fall back on simple probit models of university enrollment. One must be careful not to give these a causal interpretation, but they can provide useful suggestive evidence on the longer run effects of attending an selective school.

5 Data

The East Ridings database was assembled by Miles and Skipworth (1974) as part of their project "The Correlates of Academic Performance in Secondary Schools" and was subsequently deposited in the U.K. Data Archive (study numbers 49, 120 and 197). The database consists of "test" files and "course" files.[22]

5.1 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

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

assignment test taken in grade five and a measure of socioeconomic status (SES).[23] 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 (NFER)) and were intended to provide information on performance across the full range of abilities. In a Web Appendix, I provide examples of the types of questions asked in these tests. 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 ten tests.[24] The Web Appendix also plots the distribution of these scores.

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 will be 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 will bias the estimates upwards. Second, there will be students in my sample that did not take the assignment test, either because they were in the district in grade five but did not take the test, or because they moved into the district after grade five.[25] Since 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 heterogenous 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 will be 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

[23] Since 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 elite schools incorrectly categorize some low-SES students as mid-SES students (because most student are at least mid-SES), a comparison of low-SES students in elite and non-elite schools will be biased in favor of the latter. I present results with and without SES controls.

[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 US accountability movement.

[25] According to Secondary and Further Education Sub-Committee (1964, p.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 score suggests no obvious differences between students with and without missing scores.

previously suggests that around 10% of these 13 year old cohorts do not appear in the test data two years later.[27] This may because these students are not in school when the tests are administered (e.g., because they were ill) or because they had 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% of the data in any of these fashions has almost no impact on the results. Only attrition among the worst-performing marginal non-selective school students or the best-performing marginal selective school students could lead to selective school effects being under-estimated. Second, although I have no data on students with missing scores on all tests, an analysis of students with missing scores on some tests but not others (around 3%-5%) 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 3.

5.2 The Course Files

The "course" file contains information on advanced courses pursued in tenth grade. These data are available for six cohorts of students pursuing these courses – those in grade ten 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 3. These two files overlap for only one cohort of students, that taking the tests in grade nine in 1969 and pursuing courses in grade ten in 1970 (Appendix Table 3 Panel C).[30]

[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-elite school that did not take part in the study. The remaining difference - roughly 200-300 students per cohort - accounts for about 10% of the expected cohort size.

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

[29] While there are no missing values in the test files, I presume 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 may have moved into the area after grade ten or switched schools after grade ten. I drop these students from the matched

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 selective school university enrollment effects.[32] Appendix Table 3 Panel C shows that overall university enrollment is around 4%. This appears low, but is not inconsistent with other university enrollment data.[33]

6 The Impacts of Selective Schools

In this section I report the estimated impacts of attending a selective school. I discuss the impacts on test scores before turning to other outcomes related to the longer run effects of selective schools.

6.1 Test Scores

Table 1a considers the impact of selective school attendance on math test scores. Column (1), which reports least squares estimates without any regression adjustment, confirms the expected large difference in mean scores between selective and non-selective school students. This difference amounts to around 1.5 standard deviations.[34] The second column presents least squares estimates that regression-adjust for a quadratic in the assignment test scores. This results in a substantial reduction in the selective school effect. The estimate is further reduced when cohort, birth month and sex are added (column (3)) and reduced again when socioeconomic status terms are added (column (4)).

The least squares estimate in column (4) can be compared with the mean outcome among non-selective

data.

[31] university records more accurately report which students are observed in university three or four years after they began their course, hence I define attendance as being enrolled for at least three years. Since drop-out 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 which does not always uniquely identify students. The restriction reduces the number of students that might have attended college, hence 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 to twelfth grade.

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

[34] Since the standard deviation in these scores is 5.563 (see panel A of Table 1b).

school students scoring just below the range over which the probability of selective school attendance is increasing (roughly the assignment test score range [112,116]) and with the overall standard deviation in math scores. These statistics are presented in the heading to panel A of Table 1b and suggest an effect of around 13% of this mean and around a third of the math score standard deviation. Since this is the cumulated effect of four years of selective school, one could argue that it is a small one (I interpret these results below). It is also likely to be an upward-biased estimate of the true causal effect of attending an selective school, since the other piece of information determining selective school attendance (primary school order of merit) is likely positively correlated with omitted variables that positively impact math test scores.

As discussed above, estimates based on an instrumental variables strategy that uses as an instrument the predicted probability of attending selective school will be robust to this concern. This estimate is presented in column (5). The predicted probability is derived from a logit model that predicts selective school attendance using a fourth-order polynomial in assignment test scores. Not surprisingly, the predicted probability is highly correlated with actual selective school attendance (the coefficient is close to one with a t-statistic in excess of 30). Also as expected, the resulting instrumental variables estimate (-0.258) is smaller than the least squares estimate reported in column (4) and is in fact negative. Although less precise than the least-squares estimate, the estimate suggests that while a zero effect of selective school attendance cannot be ruled out, effects larger than one third of a standard deviation can be. The same basic pattern of estimates – large mean differences, small least-squares estimates and even smaller instrumental variable estimates – is repeated for Reading, Science and IQ test scores in columns (1)-(5) of Panels B, C and D of Table 1b. In each case, zero effects cannot be ruled out, while effects larger than one third of a standard deviation can be.[35]

Since the estimates presented in column (5) of Tables 1a and 1b did not make use of any covariates, they can be given a graphical interpretation. In particular, the estimated selective school effects can be viewed as the difference between the predicted outcomes in the treatment and non-treatment regime (since the probability of selective school attendance increases from almost zero to almost one over this range, it is only in this range that the two lines can be distinguished on the graphs). As seen in Figure 2, the two predicted

[35] Math estimates are relatively less precise in part because these are based on only two of the three cohorts used in the English test score analyses.

outcomes are often indistinguishable (i.e., the two lines lie almost on top of one another) and only for the IQ test does there seem to be a positive gap. While this IQ effect is statistically significant (see column (5) of panel D of Table 1b), it is small relative to the standard deviation in IQ test scores and, as I discuss below, at the top end of the range of estimates produced by my robustness checks. As well as enabling one to visualize the effect sizes, the graphical presentation is a useful means of assessing the adequacy of the assignment test control function, in this case a quadratic. Visual inspection of these graphs suggests that the control function used to capture the relationship between the assignment test score and the outcome fits the data well, although I present robustness checks based on alternative functional forms below.

The estimates in column (5) do not take account of the sex- and cohort-specific nature of the assignment process (i.e., separate competitions by sex and cohort), hence the full power of the IV approach was not exploited. Moreover, since outcomes were not regression adjusted for sex and cohort, these estimates may have been biased by correlations between variables that determine selective school attendance conditional on assignment test score (such as being in a particular cohort) and outcomes. The estimates reported in column (6) deal with this by predicting the probability of selective school attendance using assignment test scores interacted with sex and cohort and including sex and cohort in the outcome equation. As expected, the estimates in column (6) of Table 1b are more precise. They are also slightly larger, although none increase by more than one point and SES controls (column (7)) reduce them slightly. As a result, the overall conclusion remains: effects larger than one third of a standard deviation can be ruled out, while zero effects cannot be.

To probe the robustness of these results I assess whether they are sensitive to the assumptions underlying Tables 1a and 1b. For comparison, column (1) of Appendix Table 1 reports the estimates obtained from the baseline model (column (6) of Table 1b). This model was based on the subsample of students with non-missing outcomes and assignment test scores in the [90,140] interval. This sample restriction was imposed so that the estimated control function was not influenced by outliers in the left tail of the assignment test score distribution and the model used a quadratic function to control for assignment test scores. Columns (2)-(4) of Appendix Table 1 report the results obtained when each of these restrictions is relaxed: column (2) reports the estimates obtained when missing outcomes are coded as zero, column (3) reports the estimates obtained using the full range of assignment test scores and column (4) reports estimates based on a more flexible assignment test score control function (cubic). Coding non-missing outcomes as zero has no obvious effect on the results and results are similar when a more flexible control function is used. Estimates are smaller when the full sample of assignment test scores are used and in some cases are negative and significant.[36] This suggests that the decision to eliminate the left tail of the assignment test score distribution is a sensible one. Not surprisingly, estimates based on a regression discontinuity approach - with a quadratic (column (5)) or a cubic (column (6)) - generate similar point estimates and further reinforce the conclusion that these selective schools had small effects on standardized test scores. These regression discontinuity estimates are graphed in Web Appendix Figure 1.

Although I interpret these results in the next section, a potential concern is that they 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, Figure 2 shows that marginal students obtained neither very low or very high scores (the horizontal lines in those graphs represent the maximum score). Second, the almost linear relationship between these scores and the assignment test scores suggests that students of higher ability at age eleven were still of higher ability in grade ten and were able to demonstrate this on these tests. Third, both of these findings are consistent with the intentions underlying the design of these tests, namely to test the full range of grade nine abilities. Fourth, estimated effects on particular quantiles of the test score distribution do not point to significant selective school effects. In particular, there is no consistent evidence for effects at the 25th or the 75th percentile, as one might expect if these tests were too hard or too easy to detect differences in the middle of the ability (conditional on the assignment test score) distribution.[37] Interestingly, when the dependent variable is changed to questions attempted rather than questions correctly answered, selective school effects are again small (Appendix Table 1, column (3)).

[36] Heuristically, this is because the control function must be "curved" to fit the left tail of the assignment test score distribution and "linear" to fit the right part of the distribution. These requirements can only be reconciled via negative elite school effects.

[37] Regression discontinuity estimates (standard errors) at the 25th percentile are -0.784 (1.345) for math, -0.781 (0.792) for reading, -1.141 (0.732) for science and -0.0145 (1.488) for IQ. Similar estimates at the 75th percentile are 0.727 (1.052), -0.483 (0.560), -1.330 (0.601) and 1.254 (1.444).

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 (reported in Web Appendix Tables 1c and 1d). While point estimates are sometimes larger for boys than for girls on the Math test, the point estimates are not large (around one third of a standard deviation) and the same pattern is not repeated across the other tests. Similarly, looking at estimated effects by SES category, there are no consistent patterns either within particular tests (i.e., both low- and high-SES students outperform mid-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-students in Reading and Science). The overall conclusion then is that when assessed in terms of test score performance, selective school effects are small.

6.2 Other Outcomes

Unfortunately, I do not have information on earnings. I do have information on university enrollment, but as discussed above, I cannot estimate these effects using RD-IV methods and must instead use probit models. Before presenting these results, I first discuss selective school effects on course-taking since, as discussed above, we would expect course-taking outcomes given the university admission rules facing the students in this sample. I present the results for four course outcomes: any course, any academic course, Mathematics course and number of courses. The first is an indication of high school dropout, since students that 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.[38] The second and third are outcomes that we would not expect selective schools to affect mechanically. In other words, while no non-selective schools offered Advanced Mathematics or Classics, all offered some academic courses and all offered a Mathematics course. The fourth is related to the requirements these students must meet in order to continue on to post-compulsory education.[39]

[38] Personal communication with non-elite school teachers suggests few students would stay beyond the compulsory school age and puruse only lower-level courses. Since I cannot be sure that this did not apply to some students, and since the elite school effect on the probability of taking at least one course is small (around 8 percentage points), I choose not to interpret these as high school graduation effects.

[39] The formal requirement was that students pass five of these courses. I do not focus on this outcome because the requirement was not a hard one and because this outcome mixes course-taking with course performance. Since selection into course-taking

Figure 3 illustrates the relationship between these normalized assignment test scores and mean outcomes. Since these course-taking outcomes are binary, the linear IV approach used in the previous section is no longer appropriate, and I choose instead to estimate selective school effects using an RD-probit approach. As a first step to implementing this approach, I define a sex- and cohort-specific eight-point marginal interval and normalize the assignment test scores on it. To the left of the marginal interval I graph mean outcomes among non-selective school students; to the right I graph mean outcomes among selective school students. In the marginal interval, over which the probability of selective school attendance increases from around 0.1 to around 0.9, I present mean outcomes for both sets of students. Although students with assignment test scores in this marginal interval are excluded from the sample used to generate the RD estimates, it is interesting to compare outcomes among selective and non-selective school students with the same assignment test scores. Focusing on these mean outcomes (the selective school means are represented by open circles), it seems that a higher fraction of selective than non-selective school students pursue at least one course and at least one academic course. Differences between selective and non-selective school students are even more marked with respect to the fraction pursuing a math course and the mean number of courses pursued. Above the first three graphs in Figure 3 are the associated RD-probit estimates. In each case, the outcome is modeled as a function of a dummy for being on the right-hand side of the marginal interval and a linear spline in the (normalized) assignment test score. I use a spline because the data in Figure 3 suggest that the relationship between assignment test scores and outcomes may be different among selective and non-selective school students (I check the robustness of the results to this assumption). Predicted outcomes from these models are then superimposed on these graphs, and the estimated effect is the difference between the predicted outcomes for students in selective and non-selective schools with (normalized) assignment test scores of zero.

These regression discontinuity estimates suggest effects of around 8 percentage points for the probability of taking at least one course (academic or any), around 15 percentage points for the probability of taking a math course and around two academic courses on the total number of courses pursued. These estimates can be compared with mean outcomes among non-selective school students with scores to the left of this interval (75%, 57% and 3.5 courses respectively). These numbers are reported in the panel headings of Table 2a, which

means that I cannot estimate course performance, I choose to examine the total number of courses taken.

also contrasts these RD estimates (presented in column (4)) with those based on probit and linear regression models that do not exclude students in the marginal area (columns (1)-(3)) and other RD estimates that include sex (column (5)) and SES (column (6)) controls. As expected, the point estimates reported in columns (1)-(3) are typically larger than those report in column (4), which are in turn of comparable magnitude to those reported in columns (5) and (6). Again, dealing with selection biases reduces the estimated selective school effects, although the confidence intervals around these estimates overlap. This suggests that these effects do not rely on the precise method used to estimate them and the robustness analysis in Appendix Table 2a shows that the positive estimates are robust, even though the precise magnitudes vary across the particular robustness checks. As seen in the bottom-right graph in Figure 3, in panel D of Table 2a and in Appendix Table 2b, similarly firm conclusions can be drawn regarding the selective school impact on the total number of courses taken, which is to almost double them.[40]

Tables 2b and 2c allow selective school impacts on course-taking to differ across sex and SES group (to save space I present only the outcomes showing the largest mean differences; estimates for all outcomes are reported in Web Appendix Tables 2b and 2c). Unlike the impacts on standardized test scores, these Tables reveal some interesting patterns. In particular, it seems that selective school effects are largest for students with the lowest base course-taking rates (i.e., the lowest course-taking rates in non-selective schools). For example, among students in non-selective schools, the probability of taking a math course is significantly smaller for girls than boys. In turn, selective school effects on the probability of taking a math course are larger for girls. Among students in non-selective schools, both course-taking outcomes are consistently lower among low-SES students. In turn, the estimated selective school effects are consistently larger. Indeed, for low-SES students, the impact of attending a selective school is to more than double the probability of taking a math course and more than double the total number of courses pursued. While the point estimates are not always precise enough to reject the hypothesis that the various effects are the same across subgroup, this pattern is a consistent one.

[40]Although this is the combined effect on the probability of taking any courses and on the number of courses taken given that at least one is taken, Web Appendix Figure 2 provides evidence as to the importance of the second effect. This graph uses all cohorts of students appearing in the course files, including those that cannot be matched to the test files. Since this is a selected sample of students that pursue at least one course, the first graph is not presented and I do not base any estimates on these graphs. Nevertheless, they provide suggestive evidence that the differences seen in **Figure 3** are at least partly driven by differences in course-taking behavior among students that pursue at least one course.

To summarize, there appear to be large selective school effects on course-taking. This is true for outcomes not thought to be mechanically affected by selective school attendance (any academic course, Mathematics course) and for outcomes that may in part be mechanically affected by selective school attendance but which are likely to affect progression to post-compulsory education (total number of courses pursued). To assess whether or not these effects are associated with post-compulsory enrollment, Figure 4 graphs university enrollment against assignment test scores for the subsample of students for whom university enrollment data is available. As seen in Figure 4, university enrollment is low among this sample of students, and positive only for students with high assignment test scores.[41] For this reason, instrumental variables and regression discontinuity estimates of selective school effects are imprecise and sensitive to the function used to control for trends in outcomes by assignment test scores. This is especially true for women, among whom enrollment is low. For men, the estimates presented in Table 3 are consistent with strong selective school effects that are large relative to the sample mean enrollment probabilities.

7 Interpretation and Discussion

The selective school impact on test scores measures the cumulated effects of the selective school treatment over four years of schooling - from grade six to grade nine. Considered in this way, the effects seem surprisingly small. The least squares estimates point to total effects of only one third of a standard deviation, equivalent to moving students less than ten percentile points up the overall test score distribution; the IV estimates suggest that these are at the top of the range of possible effects, with the IV point estimates around one sixth of a standard deviation and not statistically distinguishable from zero.

It is useful to compare these effect sizes with those found in the previous literature and with the estimated effects of other student characteristics. The study most closely related is Sullivan and Heath (2004), which uses the nationally representative National Child Development Study (NCDS) to analyze the impacts of various student and school characteristics on mathematics test scores and other outcomes. For the purposes of assessing the external validity of these results, the study is attractive, since the NCDS cohort is only two

[41] There are several reasons for low non-elite 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 colleges. This may be why measured university enrollment is twice as high among boys than girls in this sample.

years younger than the cohorts observed in these data and the mathematics test administered was the same as the one administered in the East Ridings. With this in mind, it is interesting to note that their least squares estimates are similar to the least squares estimates obtained here.[42] While there are some important differences between the two studies, this suggests the East Ridings findings may be similar to those that would be found in other parts of the country.[43] It is also interesting to compare these least squares selective school estimates with least squares estimates of the effects of other family characteristics. For example, Sullivan and Heath (2004) estimate that conditional on assignment test score, the effect of having professional parents that left school at age 19 or older (relative to unskilled parents without post-compulsory schooling) is 50% larger than the effect of attending a selective school.[44] Although these least squares estimates cannot be given a causal interpretation, they suggest the effects of attending selective school may be smaller than the effects of having high-SES and highly-educated parents, even conditional on the assignment test score. This conclusion is consistent with the evidence presented above since, conditional on assignment test score, my preferred test score models typically produced larger coefficients on the high-SES parents dummy variable than on the selective school variable.

There are several explanations for the discrepancy between these effects and those that would be predicted by a simple education production function in which student achievement is a linear function of peer quality. One class of explanation centers on the particular schools studied and the tests used. First, because the selective schools are single sex, peer effects may differ from those that would be observed in mixed-sex schools. Yet while there is evidence to suppose that peer effects might differ across boys and girls (**?**), it is hard to reconcile the hypothesis of strong linear-in-means peer effects in mixed-sex schools with small selective school effects for both boys and girls. Second, because the tests were "low stakes," one could argue that students and schools would not care about them. Yet there is a strong correlation between ability (as demonstrated on the assignment test) and these test scores, and even if students exerted less effort than they would on a high-stakes test, we would still expect peer effects to generate selective school test score

[42] They find that elite schools are associated with a 6.03 point increase in percentage of questions answered correctly. Using this dependent variable and the same specification as presented in column (4) of Table 1a, I find an effect of 6.42 (standard error 1.36).

[43] The main differences are that the NCDS models are estimated using the full set of districts in England and Wales and include a richer set of family background controls.

[44] Sullivan and Heath (2004), Table 4, model 2.

impacts. Indeed, these low-stakes tests may be attractive, since they ensure the incentives to perform well do not differ across schools and students. Third, because I do not find strong effects on these scores, one might worry that they do not measure learning, but rather some permanent characteristic impervious to school influences. This seems unlikely. While one of these tests was designed to measure IQ, the Math, Reading and Science tests were designed to test material common to the curricula studied by students at both non-selective and selective schools.

A second explanation is that peer effects may be weaker in high schools. This could be because older students are less susceptible to peer influences or because there is more within-school grouping at high school, so that students spend most of their time with peers of similar ability. Miles and Skipworth (1974) provide evidence of these practices in this setting, although only in the later high school grades. Certainly, the most convincing evidence in support of academic peer effects is at the elementary level (Hoxby (2000)).[45] Of course this research design only identifies effects among students close to the selective school cutoff. As such, if peer effects do not correspond to the linear-in-means specification commonly assumed in the literature, and if they are weaker here than at other points of the ability distribution, a selective school analysis may not capture the most important peer effects. As discussed above however, since the linear-in-means specification is the workhorse model, it is nevertheless interesting that I find limited support for it.

The discrepancy between the effects estimated here and those predicted by an education production function in which student achievement is a linear function of teacher quality may be easier to explain, since evidence on the relationship between teacher effectiveness and characteristics such as aptitude, college selectivity and experience is mixed.[46] Although it would be surprising if aptitude and subject-area expertise did not improve students' performance on advanced courses, it would be less surprising if teachers with these characteristics were no more effective at teaching the basic material captured by standardized tests. More generally, one can imagine that selective schools are more focussed on advanced outcomes and devote fewer resources to more basic material. This would be consistent with East Ridings evidence suggesting that in the early grades in which this material would have been taught, selective school students were set the same

[45] Although I know of no studies focussed exclusively on high schools, Sacerdote (2001) does not find peer effects on academic outcomes among college students.

[46] Summers and Wolfe (1977) find that teacher quality is positively related to undergraduate college selectivitiy; Hanushek (1971) finds that teacher quality is positively related to verbal ability; Rivkin, Hanushek, and Kain (2005) find no relationship between teacher quality and whether or not teachers hold masters degrees.

amount of homework as non-selective school students.

In contrast to my test score findings, I find much larger effects on outcomes likely to influence the longer-run effects of attending an selective school, including course-taking and university enrollment. As discussed above, some of these course-taking findings are not surprising, since it is well known that non-selective schools did not offer advanced academic courses. Other course-taking effects, such as the selective school effect on the probability of taking a Mathematics course are less easily explained in this way, and could instead be due to other factors. For example, it may be that while non-selective schools offer these courses, the costs of taking them were nevertheless higher in non-selective schools, perhaps because non-selective schools rationed access to these courses or because of non-selective school peer effects. It may also be that non-selective school students anticipate (relatively) small gains to pursuing these courses, perhaps because non-selective school teachers were (relatively) less qualified to teach them. To obtain evidence on this last point, one could use data on the (externally evaluated) grade obtained on each course to estimate selective school impacts on course performance. Given the extent of selection into course-taking however, these are likely to 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, Gregg, and McConnell (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 mechanism underpinning these curricula effects, there is reason to expect them to impact longer-run educational outcomes. This expectation is based on the rules governing progression into post-compulsory education and the evidence for university enrollment effects found here. There is also reason to expect these effects to generate longer run selective school impacts on earnings, based on both the evidence relating additional years of education to earnings (see for example Oreopoulos (2006)) and on the estimated earnings returns to advanced qualifications. 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%. Since these effects are found for cohorts exposed to selective schools and cohorts not exposed to selective schools, they are not simply capturing the impacts of attending an selective school.[47] Although they may in part be driven by ability biases, Dearden, Ferri, and Meghir (2002) show that the returns to these types of qualifications remain significant when test scores at ages seven, eleven and sixteen are held constant.

8 Conclusions

The paper estimated the effect of attending a selective school in a particular UK school district. Selective school places were allocated to the top 20% of the district's students, and this assignment rule generated dramatic differences in peer quality across schools. These differences, in addition to differences in teacher qualifications, might lead one to expect these schools to improve the test scores of the students that 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 US (Altonji, Elder, and Taber (2005) and Cullen, Jacob, and Levitt (2006)) and suggests these results may generalize to other countries and other settings in which across-school differences in peer and teacher quality are even larger. In turn, this suggests that peer and teacher quality have more subtle 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 is not surprising that 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 US private schools literature and would explain why UK selective schools appear popular with parents.[48] More generally, because selective schools typically offer

[47] Since this variation in elite school exposure operates at the cohort level (there are large returns for cohorts of all ages), the implicit assumption is that there are no offsetting cohort effects.

[48] 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.

an advanced curriculum, they are likely to exhibit differences between their effects on test scores and their effects on longer run outcomes; because this 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.

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., G. W. IMBENS, AND D. B. RUBIN (1996): "Identification of Causal Effects Using Instrumental Variables," *Journal of the American Statistical Society*, 91(434), 448–458.

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 Organisatio 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, 2006.

BETTS, J., AND H. ROSE (1996): "The Effect of High School Courses on Earnings," *Review of Economics and Statistics*, 86(2), 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*. Basic Books, Inc., New York, NY.

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.

Dale, S. B., and A. B. Krueger (2002): "Estimating the Payoff to Attending A More Selective College: An Application of Selection on Observables and Unobservables," *The Quarterly Journal of Economics*, 117(4), 1491–1527.

Dearden, L., J. Ferri, and C. Meghir (2002): "The Effect of School Quality on Educational Attainment and Wages," *The Review of Economics and Statistics*, 84(1), 1–20.

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. (1971): "Teacher Characteristics and Gains in Student Achievement," *American Economic Review*, 61, 280–288.

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

Hoekstra, M. (2007): "The Effect of Attending the Flagship State University on Earnings: A Discontinuity-Based Approach," mimeo, University of Pittsburgh.

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.Washington, DC: National Center for Education Statistics. NCES 2007-354.

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

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

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

Klopfenstein, K., and K. Thomas (2006): "The Link Between Advanced Placement Experience and Early College Success," mimeo, Texas Christian University, 2006.

Lee, D. S., and D. Card (2006): "Regression Discontintuity Inference with Specification Error," NBER Technical Working Paper 322, March 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 2006.

Maurin, E., and S. McNally (2006): "Selective Schooling," London School of Economics, mimeo.

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," College Board Research Report No. 2002-10.

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

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.Washington, DC: National Center for Education Statistics. NCES 2007-065.

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

Sacerdote, B. (2001): "Peer Effects with Random Assignment: Results for Dartmouth Roommates," *Quarterly Journal of Economics*, 116(2), 681–704.

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

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.

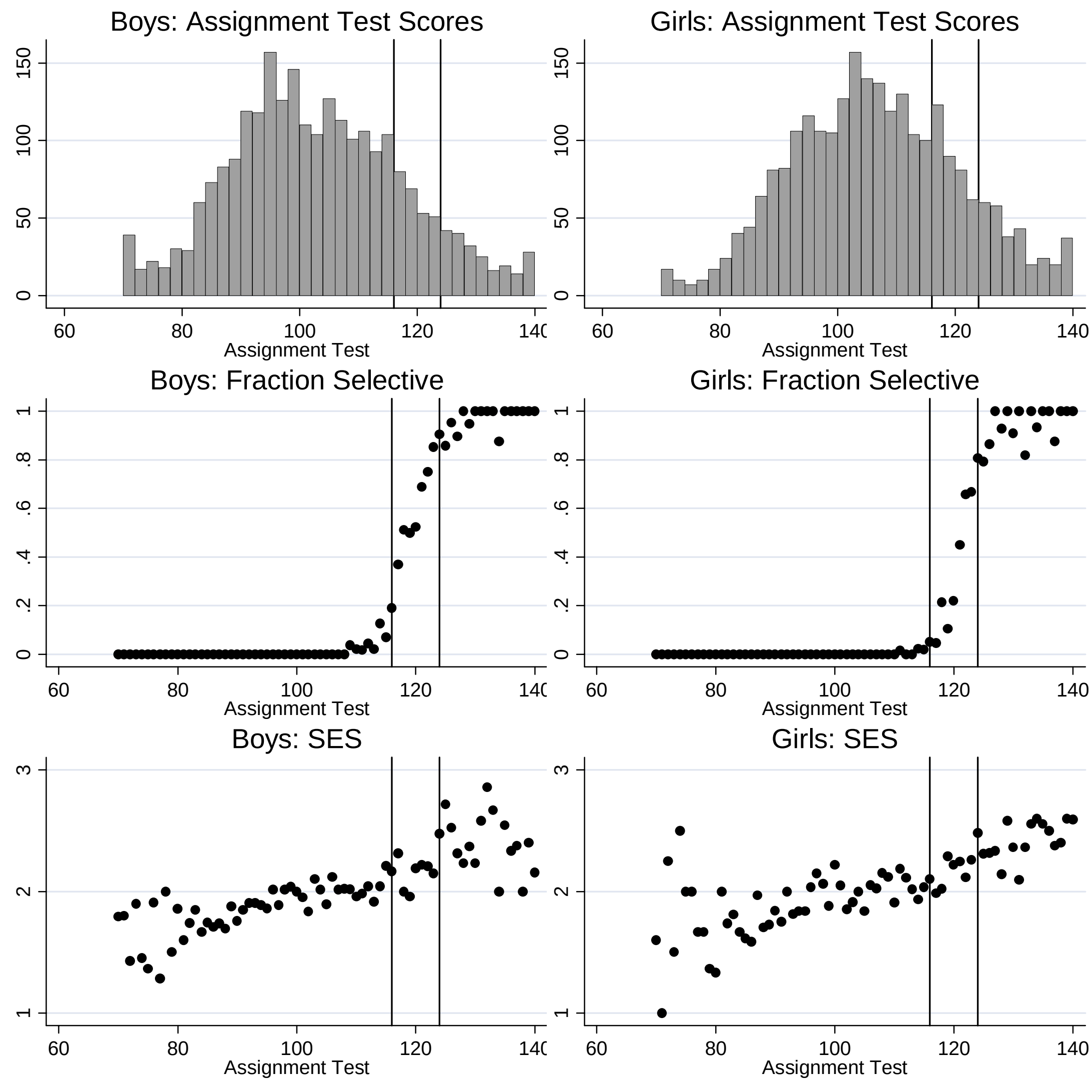
Sullivan, A., and A. Heath (2004): "Students' academic attainments and British state and private schools," mimeo, Department of Sociology, University of Oxford, Oxford, England.

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

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

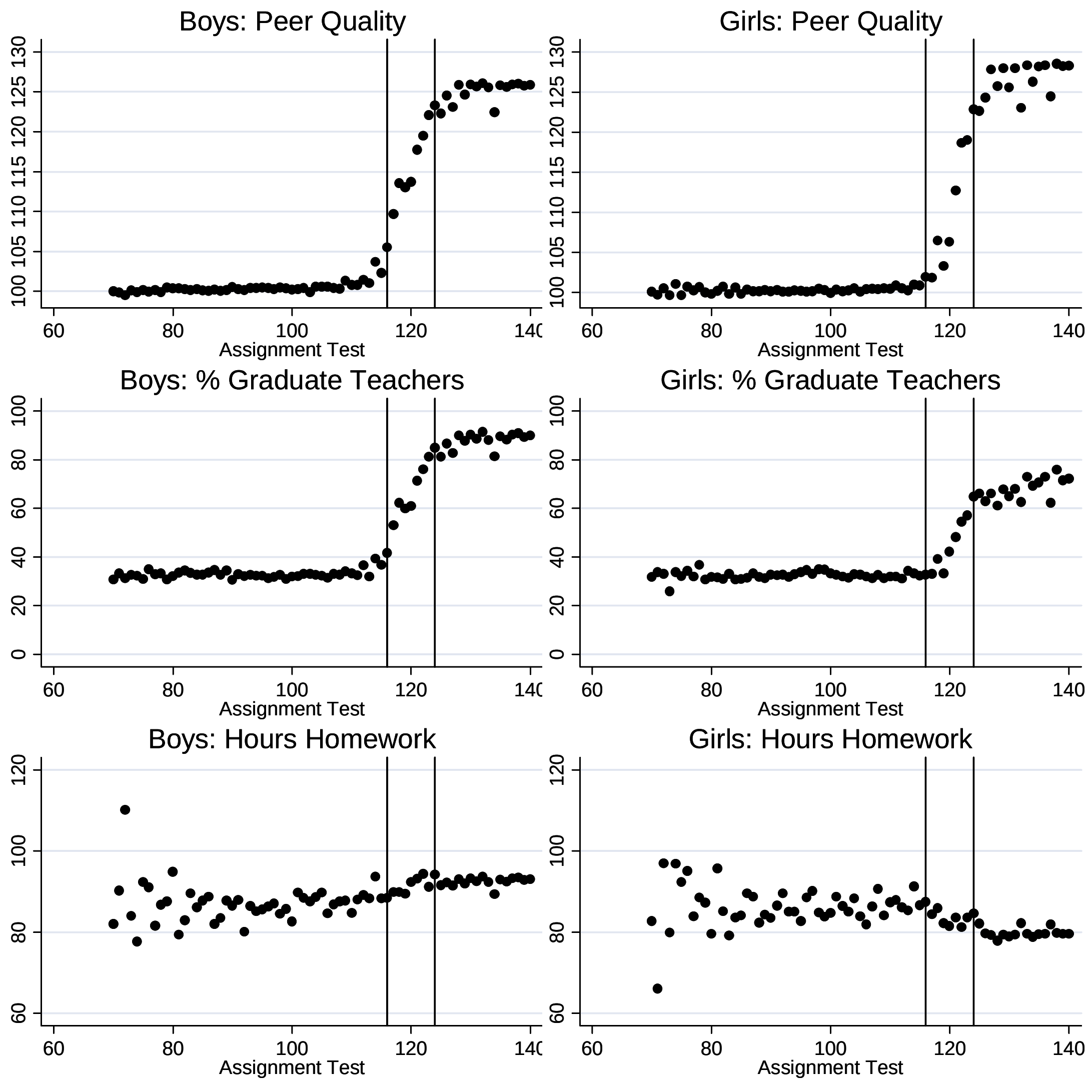
Wooldridge, J. (2002): *Econometric analysis of cross section and panel data*. The MIT Press, Cambridge, MA.

Figure 1a: Assignment test scores, fraction in selective schools, socio-economic status (SES)



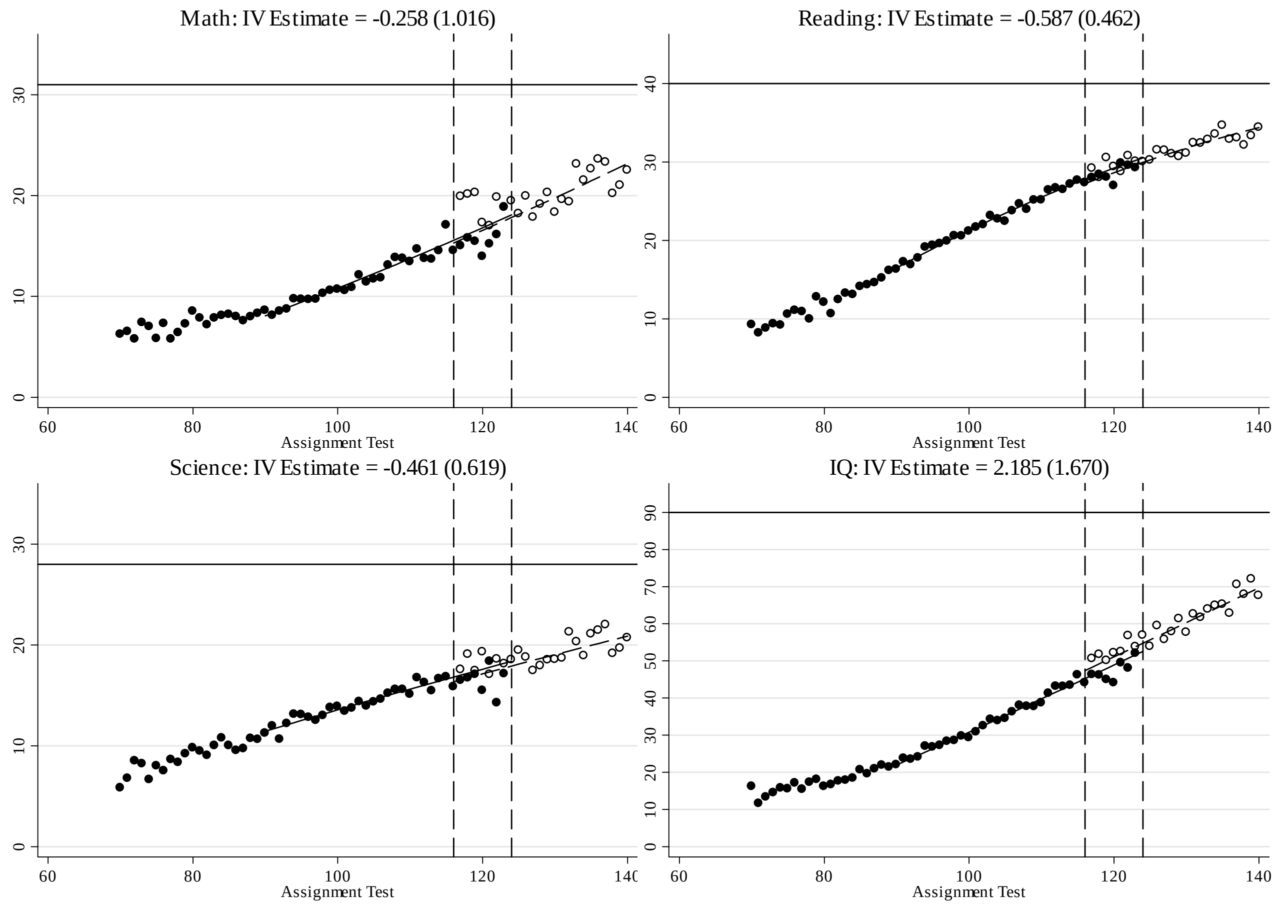
Notes: In top two panels, sample includes all students in test data with non-missing assignment test scores. In bottom panel, sample includes all students with non-missing assignment test score and non-missing SES.

Figure 1b: School characteristics by assignment test score



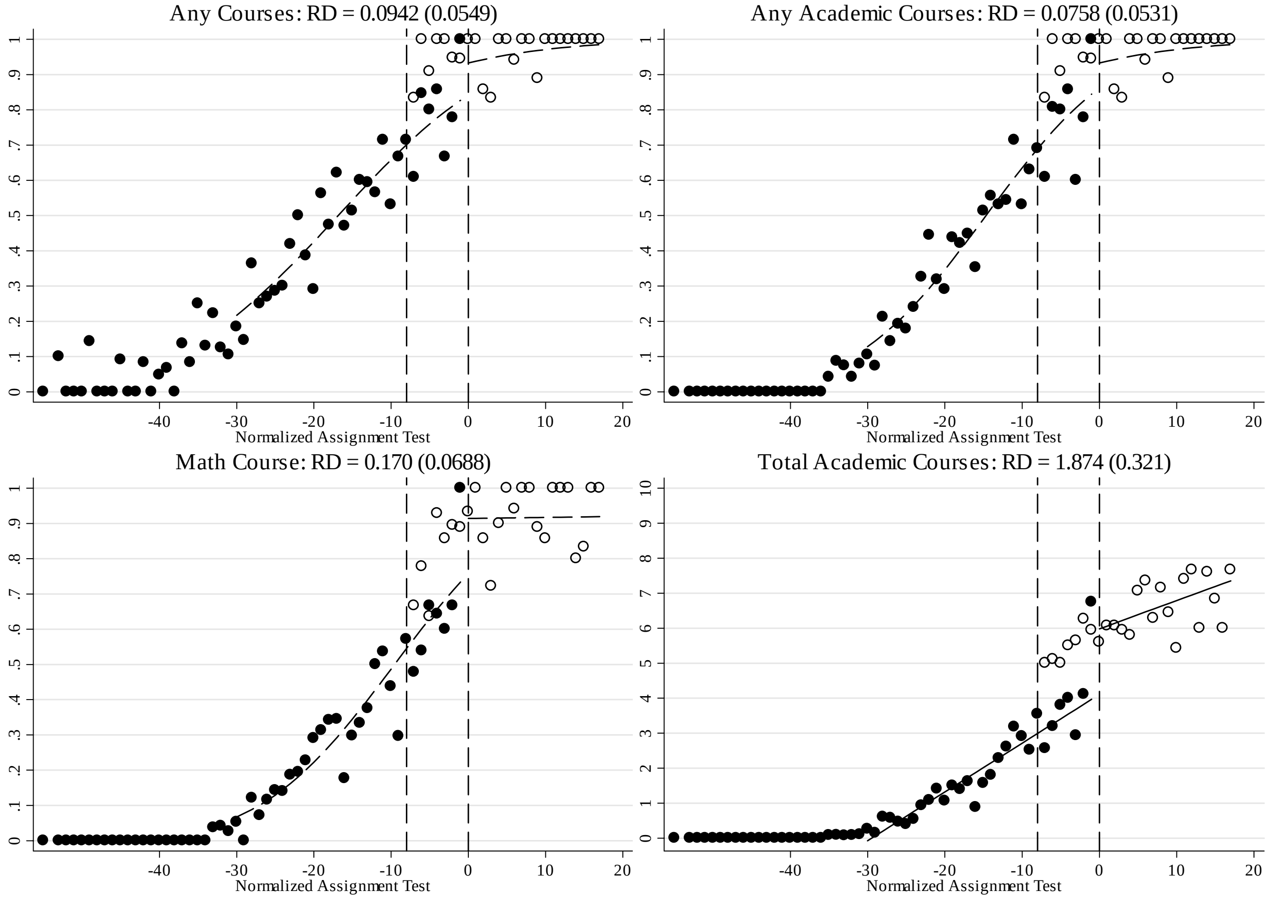
Notes: In top panel, peer quality measured by average assignment test score of students in test data; in second and third panels % graduate teachers and hours homework calculated by Skipworth and Miles (1974).

Figure 2: Impact of selective schools on grade nine test scores



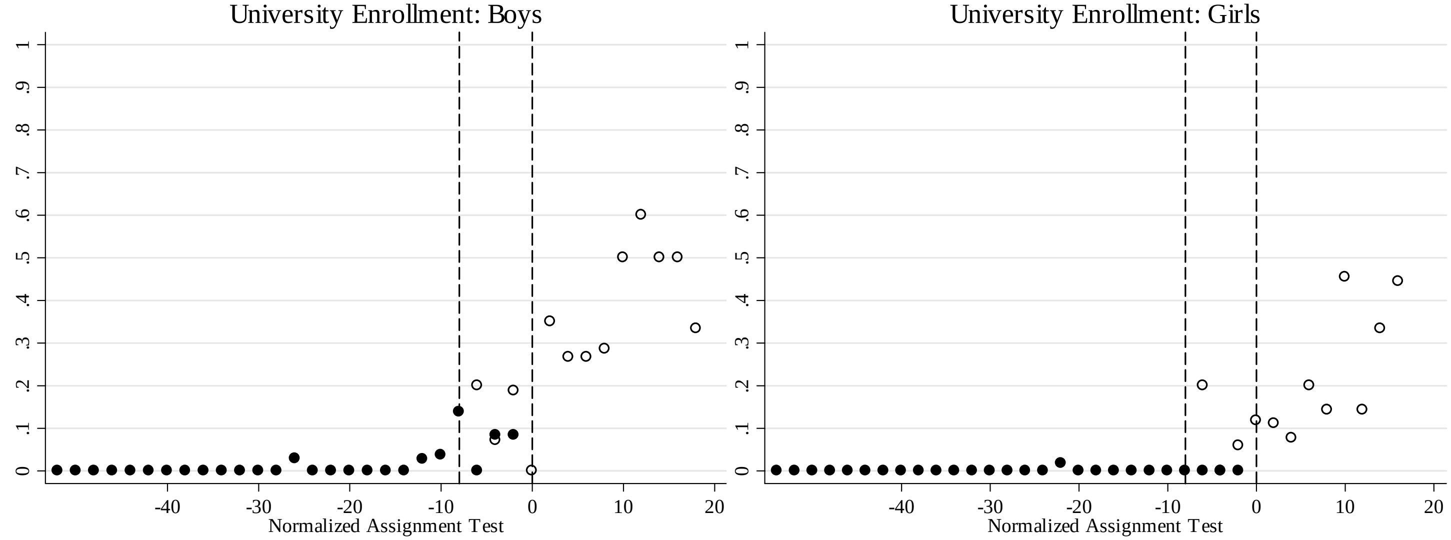
Notes: Solid (open) circles are mean scores among non-selective (selective) school students with the corresponding assignment test score. The horizontal line represents the maximum score on this test. The broken vertical lines indicate the “marginal interval” over which the probability of attending an selective school given the assignment test score rises from around 0.2 to 0.9. The superimposed solid (and broken) lines represent the predicted test score from the instrumental variables model described in the text (which includes no covariates). The vertical distance between the solid and the broken lines in the marginal area is the IV estimate reported above each panel.

Figure 3: Impact of selective school on advanced course-taking in grade ten



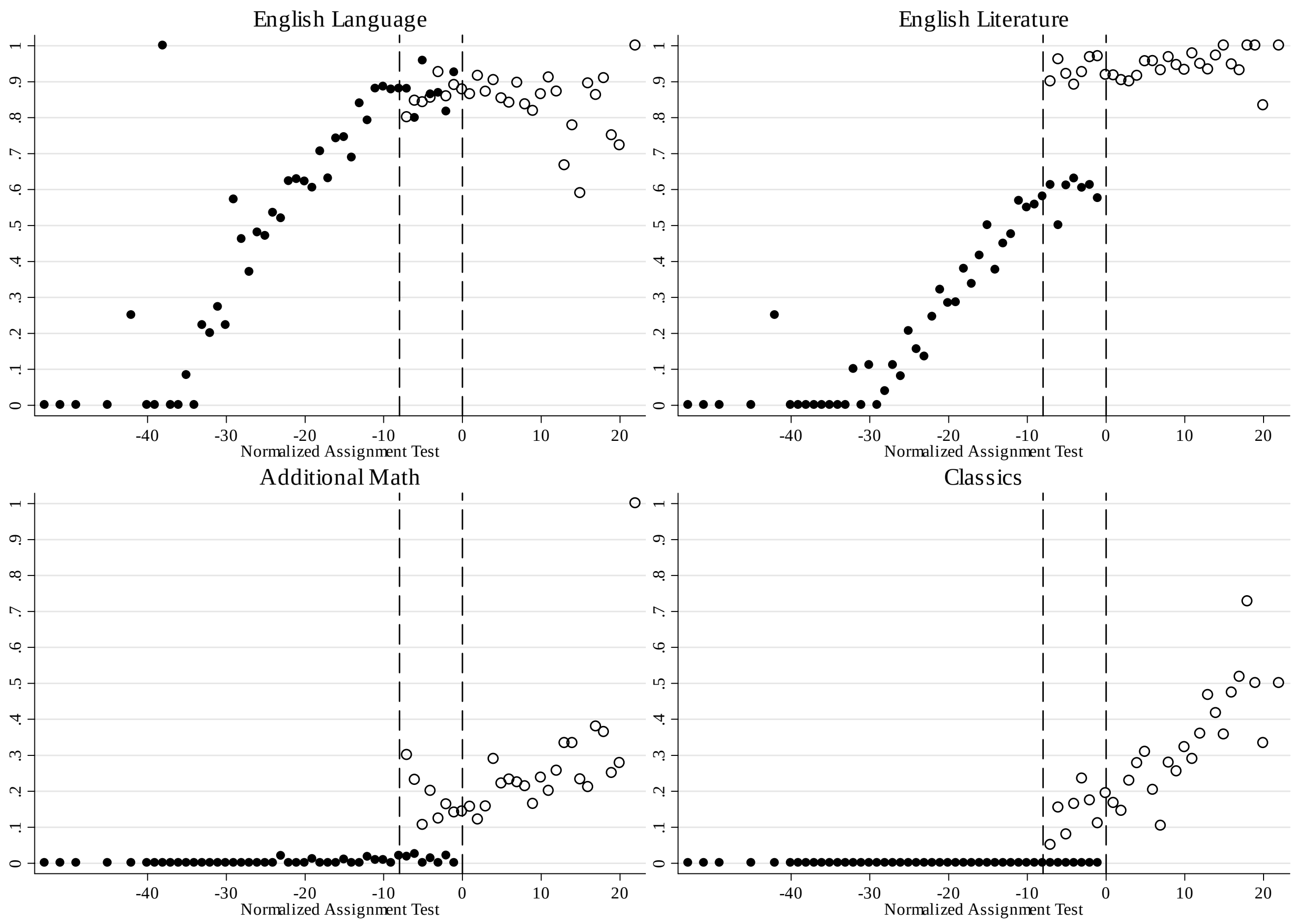
Notes: Solid (open) circles are mean scores among non-selective (selective) school students with the corresponding assignment test score. The assignment test score is normalized to be four at the point at which the predicted probability of selective school attendance is closest to 0.5. Over the eight-point interval around this (indicated by the broken vertical lines), the probability of attending an selective school given the assignment test score rises from around 0.2 to 0.9. The superimposed line represents the predicted test score from the regression discontinuity probit model described in the tezt (which does not include covariates). The vertical distance between these lines corresponding to a standardized test score of zero is the RD estimate reported above each panel.

Figure 4: Fraction enrolling in University by assignment test score and school type



Notes: Solid (open) circles are mean scores among non-selective (selective) school students with the corresponding assignment test scores (each circle is the average over two test scores). The assignment test score is normalized to be four at the point at which the predicted probability of selective school attendance is closest to 0.5. Over the eight-point interval around this (indicated by the broken vertical lines), the probability of attending an selective school given the assignment test score rises from around 0.2 to 0.9.

Appendix Figure 1: Fraction taking specific advanced courses by assignment test score and school type



Notes: Solid (open) circles are mean scores among non-selective (selective) school students with the corresponding assignment test score. The assignment test score is normalized to be four at the point at which the predicted probability of selective school attendance is closest to 0.5. Over the eight-point interval around this (indicated by the broken vertical lines), the probability of attending an selective school given the assignment test score rises from around 0.2 to 0.9.

Table 1a: The impact of selective schools on math test scores

	OLS				IV		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Selective	7.890	2.559	1.996	1.927	-0.258	0.997	0.897
	(0.575)	(0.463)	(0.413)	(0.408)	(1.027)	(0.860)	(0.858)
Assign_Test		0.588	0.565	0.563	0.122	0.402	0.396
		(0.142)	(0.128)	(0.129)	(0.182)	(0.147)	(0.149)
Assign_Test^2		-0.156	-0.136	-0.137	0.080	-0.054	-0.052
		(0.065)	(0.058)	(0.059)	(0.089)	(0.071)	(0.072)
Sex*Cohort			YES	YES		YES	YES
SES-Mid				0.277			0.275
				(0.243)			(0.246)
SES-High				0.768			0.795
				(0.237)			(0.238)
R^2	0.27	0.42	0.47	0.47	0.40	0.46	0.46
N	2690	2690	2690	2690	2690	2690	2690

Notes: Sample includes all students with non-missing math score and assignment test scores in the [90,140] interval.

Table 1b: The impact of selective schools on test scores

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Math (mean=14.977, sd=5.563)							
Selective	7.890	2.559	1.996	1.927	-0.258	0.997	0.897
	(0.575)	(0.463)	(0.413)	(0.408)	(1.027)	(0.860)	(0.858)
Panel B: Reading (mean=27.236, sd=7.572)							
Selective	8.277	0.763	0.548	0.483	-0.588	-0.543	-0.662
	(0.756)	(0.316)	(0.324)	(0.317)	(0.467)	(0.410)	(0.421)
Panel C: Science (mean=16.361, sd=4.560)							
Selective	4.662	1.044	0.605	0.536	-0.461	-0.177	-0.263
	(0.408)	(0.391)	(0.380)	(0.376)	(0.626)	(0.529)	(0.513)
Panel D: IQ (mean=44.340, sd=15.340)							
Selective	23.470	4.082	3.207	3.101	2.185	2.933	2.749
	(1.860)	(0.683)	(0.645)	(0.645)	(1.676)	(1.340)	(1.321)

Notes: Columns refer to specifications in Table 1a. Samples include all students with non-missing test scores and assignment test scores in the interval [90,140]. Mean refers to mean among non-selective school students with assignment test scores in the interval [112,116].

Table 2a: Impact of selective schools on advanced course-taking (matched data)

	"Marginal Area" included (N=1244)			"Marginal Area" excluded (N=1039)		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: P(any course); Non-selective mean=0.783						
Selective/Dummy	0.120	0.130	0.130	0.094	0.100	0.093
	(0.053)	(0.055)	(0.053)	(0.055)	(0.056)	(0.055)
Boy		-0.019	-0.022		-0.030	-0.033
		(0.012)	(0.013)		(0.017)	(0.018)
Mid-SES			0.047			0.036
			(0.018)			(0.020)
High-SES			0.022			0.026
			(0.015)			(0.013)
Panel B: P(academic course); Non-selective mean=0.764						
Selective/Dummy	0.139	0.146	0.143	0.076	0.078	0.066
	(0.055)	(0.057)	(0.054)	(0.053)	(0.054)	(0.056)
Boy		-0.013	-0.017		-0.025	-0.030
		(0.012)	(0.013)		(0.016)	(0.018)
Mid-SES			0.061			0.047
			(0.021)			(0.025)
High-SES			0.040			0.044
			(0.018)			(0.020)
Panel C: P(math course); Non-selective mean=0.575						
Selective/Dummy	0.209	0.199	0.185	0.170	0.166	0.146
	(0.047)	(0.049)	(0.048)	(0.069)	(0.069)	(0.073)
Boy		0.045	0.043		0.015	0.014
		(0.020)	(0.021)		(0.014)	(0.014)
Mid-SES			0.137			0.073
			(0.035)			(0.030)
High-SES			0.138			0.081
			(0.039)			(0.030)
Panel D: Number of academic courses; Non-selective mean=3.578						
Selective/Dummy	1.752	1.768	1.684	1.874	1.897	1.756
	(0.180)	(0.181)	(0.186)	(0.321)	(0.337)	(0.340)
Boy		-0.065	-0.092		-0.195	-0.210
		(0.116)	(0.114)		(0.135)	(0.129)
Mid-SES			0.770			0.652
			(0.150)			(0.143)
High-SES			0.793			0.813
			(0.201)			(0.176)
Panel E: "First Stage" P(Attend Selective School)						
Dummy				0.860	0.857	0.851
				(0.0317)	(0.0332)	(0.0339)

Notes: Panels A-C use probit models, panel D uses a linear probability model. In each panel, specifications (4)-(6) mirror specifications (1)-(3) with the "marginal area" excluded hence are regression discontinuity estimators. Specifications (1)-(3) same as specifications (1)-(3) in Table 1a. Mean refers to mean among students with normalized assignment test scores in marginal area attending non-selective schools. Estimation sample includes all students with normalized assignment test scores greater than -30. Robust standard errors clustered by Adj_Test_11 in parentheses.

Table 2b: Impact of selective schools on advanced course-taking (matched data): by sex

	Full sample (N=1244)			Marginal area excluded (N=1039)		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: P(math course) by sex; mean(boys)=0.705; mean(girls)=0.514						
Boys	0.044	0.049	0.020	0.117	0.147	0.124
	(0.057)	(0.063)	(0.056)	(0.082)	(0.079)	(0.086)
Girls	0.395	0.419	0.419	0.215	0.199	0.181
	(0.072)	(0.069)	(0.071)	(0.076)	(0.078)	(0.081)
Panel B: Total academic courses by sex; mean(boys)=3.852; mean(girls)=3.388						
Boys	0.839	0.794	0.663	1.712	1.818	1.662
	(0.343)	(0.363)	(0.362)	(0.601)	(0.605)	(0.587)
Girls	2.705	2.744	2.691	2.091	2.045	1.911
	(0.256)	(0.256)	(0.254)	(0.383)	(0.430)	(0.000)

Notes: See notes to Table 2a. Each row reports results of models estimated separately for boys and girls. In full sample, N=814(girls), N=772(boys).

Table 2c: Impact of selective schools on advanced course-taking (matched data): by SES

	Full sample (N=1244)			Marginal area excluded (N=1039)		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: P(Math course) by SES: mean(low)=0.3; mean(mid)=0.512; mean(high)=0.511						
Low-SES	0.406	0.482		0.439	0.459	
	(0.103)	(0.089)		(0.123)	(0.122)	
Mid-SES	0.101	0.073		0.088	0.099	
	(0.060)	(0.064)		(0.063)	(0.058)	
High-SES	0.170	0.165		0.130	0.119	
	(0.074)	(0.076)		(0.148)	(0.149)	
Panel B: Total Aca courses by SES: mean(low)=1.900; mean(mid)=3.010; mean(high)=3.021						
Low-SES	1.943	2.057		3.073	3.150	
	(0.539)	(0.477)		(1.169)	(1.136)	
Mid-SES	1.709	1.760		1.544	1.725	
	(0.380)	(0.380)		(0.492)	(0.489)	
High-SES	1.383	1.312		1.677	1.514	
	(0.382)	(0.366)		(0.718)	(0.733)	

Notes: See notes to Table 2a. Each row reports results of models estimated separately for low-, medium- or high-SES students. Specification (3) excluded since each row reports estimates for single SES category hence no SES controls. In full sample, N=422 (low-SES), N=732 (med-SES), N=422 (high-SES).

Table 3: The Impact of Selective Schools on University Enrollment: Boys (Mean=6%)

	Probit Estimates				IV Estimates		
	(1)	(2)	(4)	(5)	(6)	(7)	(8)
Selective	0.199	0.057	0.057	0.031	0.052	0.062	0.044
	(0.050)	(0.088)	(0.090)	(0.082)	(0.154)	(0.156)	(0.148)
Test_11		0.020	0.020	0.018	0.016	0.015	0.015
		(0.004)	(0.004)	(0.004)	(0.007)	(0.007)	(0.007)
Month			YES	YES		YES	YES
SES-Mid				0.215			0.058
				(0.092)			(0.029)
SES-High				0.330			0.124
				(0.127)			(0.051)
N	359	359	359	359	359	359	359

Notes: Sample includes all students with Test_11 scores in the [105,140] interval.

Appendix Table 1: Test scores: Robustness checks

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	IV estimates					RD estimates	
	Baseline model	Missing Scores=0	Dep Var: Questions Attempted	Full Sample	Cubic control function	Baseline Model	Cubic control function
Panel A: Math (mean=12.229, sd=5.563)							
Selective	0.919	1.297	0.241	0.058	0.651	-0.062	-0.304
	(0.860)	(0.992)	(0.937)	(0.876)	(1.003)	(0.712)	(1.043)
Panel B: Reading (mean=22.562, sd=7.572)							
Selective	-0.642	-0.359	0.102	-3.473	-0.792	-0.377	-0.760
	(0.409)	(0.817)	(0.995)	(1.222)	(0.428)	(0.486)	(0.607)
Panel C: Science (mean=14.167, sd=4.560)							
Selective	-0.274	1.322	1.936	-0.324	-0.175	-0.698	-0.684
	(0.509)	(0.862)	(1.245)	(0.507)	(0.557)	(0.509)	(0.640)
Panel D: IQ (mean=35.636, sd=15.340)							
Selective	2.697	1.657	-0.675	-3.013	2.055	0.645	-0.372
	(1.354)	(1.551)	(1.970)	(2.515)	(1.357)	(1.521)	(1.676)

Notes: Column (1) refer to specification (5) in Table 1a, which is represented by the graphs in Figure 2. In this baseline specification, sample includes all those with non-missing test scores and assignment test scores in [90,140] interval and model controls for a quadratic in the assignment test score. Specifications (2)-(7) consider various departures from these assumptions. Column (3) corresponds to the specification in column (1), but the dependent variable is the number of questions answered (i.e., correct plus incorrect).

Appendix Table 2a: Advanced course-taking: robustness checks

	Probit Models					RD-IV	ML-IV
	(1) Base	(2) Narrow interval	(3) Wide interval	(4) Full sample	(5) Quadratic spline	(6)	(7)
	Panel A: P(Any Courses)						
Selective	0.093	0.110	0.065	0.0871	0.248	0.0824	0.180
	(0.055)	(0.043)	(0.064)	(0.0517)	(0.228)	(0.0702)	(0.068)
	Panel B: P(Academic Course)						
Selective	0.066	0.084	0.033	0.0386	0.240	0.0708	0.152
	(0.056)	(0.044)	(0.063)	(0.0501)	(0.192)	(0.0670)	(0.071)
	Panel C: P(Math Course)						
Selective	0.146	0.162	0.140	0.112	0.336	0.268	0.327
	(0.073)	(0.058)	(0.088)	(0.066)	(0.211)	(0.082)	(0.075)
N	1039	1096	981	1383	1039	1039	1244

Notes: Each column refers to a different specification, starting with the base specification (column (6) of Table 2a). Column (2) uses a marginal interval of width six points; column (3) an interval of width ten points. Column (4) uses the full sample of data and column (5) uses a quadratic spline in the (normalized) assignment test score. Column (6) uses an RD-IV strategy (see text) with first stage coefficient (standard error) 0.857 (0.0230)); column (7) estimates uses an ML-IV strategy (see text).

Appendix Table 2b: Advanced Course-Taking: Robustness Checks

	Linear Regression Models					LPM-IV	IV
	(1) Base	(2) Narrow interval	(3) Wide interval	(4) Full sample	(5) Quadratic spline	(6)	(7)
Selective	1.756	1.749	2.065	2.804	2.406	2.047	2.731
	(0.340)	(0.282)	(0.362)	(0.332)	(0.563)	(0.607)	(0.449)
N	1039	1096	981	1383	1039	1039	1244

Notes: Each column refers to a different specification, starting with the base specification (column (6) of Table 2a). Column (2) uses a marginal interval of width six points; column (3) an interval of width ten points. Column (4) uses the full sample of data and column (5) uses a quadratic spline in the (normalized) assignment test score. Column (6) uses an RD-IV strategy (see text) with first stage coefficient (standard error) 0.857 (0.0230)); column (7) uses an IV strategy (see text).

Appendix Table 3: Descriptive statistics

	All schools			Selective schools			Non-selective 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	0.340	0.474	6219	0.340	0.474	926	0.340	0.474	5293
Cohort 8	0.346	0.476	6219	0.308	0.462	926	0.352	0.478	5293
Boy	0.497	0.500	6219	0.473	0.500	926	0.502	0.500	5293
SES	2.982	0.940	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	6.760	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	57.810	11.365	889	31.642	12.210	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	18.750	3.298	597	13.436	4.209	3345
Panel B: Course Files									
Cohort 1	0.153	0.360	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	0.500	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	0.450	5572	0.203	0.402	1909	0.322	0.467	3663
Assignment	116.203	11.080	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	1.790	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
University|Pass>=5	0.303	0.460	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 University|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 non-unique) hence cannot be matched to the University data. The same applies to a small number of students in the merged test and course data.