Ability-Grouping and Academic Inequality: Evidence From Rule-Based Student Assignments

Clement (Kirabo) Jackson
Cornell University, ckj5@cornell.edu

"ABILITY-GROUPING AND ACADEMIC INEQUALITY: EVIDENCE FROM RULE-BASED STUDENT ASSIGNMENTS"

C. Kirabo Jackson[¥]
Cornell University (Draft date: 20 August, 2008)

In Trinidad and Tobago students are assigned to secondary schools after fifth grade based on achievement tests, generating large differences in school and peer quality. Using rule-based instrumental variables to address self-selection bias, I find that being assigned to a school with high-achieving peers has large positive effects on examination performance, particularly for girls. This suggests that ability-grouping (or school tracking) reinforces achievement differences by assigning the weakest students to schools that provide the least value-added. While students benefit from attending schools with brighter peers on average, the *marginal* effect is non-linear such that there are small benefits to attending high-achievement schools over average schools, while there are sizable benefits to attending average schools over low-achievement schools. This suggests that school ability-grouping may harm those consigned to low-achievement schools at the lower end of the achievement distribution.

I Introduction and Background

Students in Trinidad and Tobago take an exam at the end of fifth grade that is used to assign them to secondary school. Each student lists their secondary school choices, and a student's likelihood of being assigned to their first-choice school increases with their score. As a result, high-achieving students typically attend high-performing secondary schools while low-achieving students typically attend the poorest-performing schools. This school ability-grouping generates large differences in peer and input quality across schools with potentially strong implications for economic inequality.[1] Researchers have linked differences in school quality to differences in labor market outcomes [Card and Krueger (1992a;1992b), Betts (1995) and Grogger (1996)] and test scores to subsequent earnings [Murnane, Tyler and Willet (2000)] suggesting that school ability-grouping could reinforce achievement differences by assigning low-achievers to schools that provide the least value-added.

While there is no strong consensus in the literature, researchers generally find that ability

[¥] I am grateful for feedback received from Ron Ehrenberg, Roland Fryer, Kevin Hallock, Caroline Hoxby, Bob Hutchens, Clement Jackson, Lawrence Katz, Jordan Matsudaira and Henry Schneider. I am also grateful for useful comments received from participants of the Labor Economics workshop at Cornell University. I am deeply grateful to Marcia Riley and I would like to thank Yvonne Lewis, Rosaline Mendez and Simone Rawlins of the Trinidad and Tobago Department of Education Research and Evaluation for allowing me to access their data, their assistance and generosity. All errors are my own.

[1] According to the 2000 Census, fifty four percent of the working age populations' highest educational level is secondary school. Secondary school outcomes are monitored by prospective employers so that performance at secondary school is an important determinant of the economic welfare for the majority of the working age population.

grouping across schools combined with curricular differences across ability groups (school tracking) is associated with increased educational and socio-economic inequality [Atkinson, Gregg and McConnell (2006); Ariga, Brunello, Iwahashi and Rocco (2005); Brunello and Checci (2006); Hanushek and Woessmann (2007); Maurin and McNally (2007)]. However, there is little evidence on the mechanisms. It is unclear whether the increased inequality is due to (a) differences in school value-added due to peer and input quality across schools, (b) differences in curricula, or (c) efficiency gains that accrue differentially to students at the top/bottom of the ability distribution. Ability grouping across schools is currently practiced in several nations worldwide[2] and elements of ability grouping exist in the United States where several districts have highly selective secondary schools.[3] As such, understanding the potential link between ability-grouping and inequality should be important to policymakers and educators in several nations since these different possible mechanisms have different policy implications.

I use data from Trinidad and Tobago to deepen our understanding of the relationship between ability grouping and educational inequality by answering the following empirical questions: (1) Do students, on average, benefit from attending schools with higher-achieving peers on a range of academic outcomes? (2) Does ability-grouping across schools increase educational inequality, *on the margin*, by assigning low-achievement students to low value-added schools while assigning high-achievement students to high value-added schools? (3) Do the marginal effects differ for males and females? (4) Are the marginal effects non-linear (i.e. does attending a school with marginally higher-achieving peers have a larger effect at low or high peer achievement levels)? Trinidad and Tobago data are well suited to identifying school ability-grouping effects on the margin because: (a) the school ability-grouping system generates large meaningful differences in school and peer quality, (b) the student assignment mechanism creates exogenous variation in school attendance, (c) there is a national curriculum so that school effects are not confounded with curricular differences, and (d) all schools have homogenous student populations so that school effects are not confounded with a "homogeneous student" effect. This paper is focused on the distributional effects,[4] rather than the efficiency effects of

[2] This includes Austria, Germany, Japan, Hungary, the United Kingdom, the Slovak Republic, the Czech Republic, Jamaica, Barbados and others.

[3] Notable examples are Boston Latin School and Stuyvesant High School in New York City. There also exist magnet schools that admit students based on prior achievement.

[4] It is worth noting that the where skilled and un-skilled labor are imperfect substitutes in the production process, the distribution of educational outcomes has direct implications for macroeconomic growth.

ability-grouping and is the first paper to test for a particular mechanism – i.e. that ability-grouping exacerbates pre-existing achievement differences, on the margin, by consigning low-achievers to schools that produce the least value-added and *vice versa*.

The current evidence on ability-grouping relies on comparisons between observationally similar students in ability-grouped and non-ability-grouped school systems. As documented by Dustmann (2004) and argued by Manning and Pischke (2006), such evidence may not reflect causal relationships since students may select into schools based on *unobserved* characteristics that also affect outcomes. Unlike previous studies, to address this concern, I use rule-based instrumental variables in the spirit of Campbell (1969) and Angrist and Lavy (1999) based on the student school assignment rules used by the Ministry of Education. The assignment rules (described in section IV) are largely deterministic, non-linear, non-monotonic functions of student preferences and incoming test scores that lead to test score cut-offs for each school below which admission is unlikely. As suggested in Fisher (1976), I use the deterministic portion of the assignment rules to obtain rule-based assignments, which are complicated non-linear functions of test scores and preferences, as exogenous instruments while directly controlling for smooth functions of these same underlying covariates. This results in two sources of plausibly exogenous variation: (a) comparing outcomes of students with similar scores who just made and just missed the rule-based cut-off (while controlling for their scores); (b) comparing outcomes of students with the same scores who ended up at different schools because of their preferences (while directly controlling for their preferences). A unique feature of the data is that I can observe a student's desired schools so that I do not have to rely *solely* on the local variation due to the cut-offs. To show that my identification strategy is likely valid, I present falsification tests showing that, conditional on test scores and preferences, the instruments are not correlated with incoming student characteristics such as religion, gender, and primary school district.

The results indicate that there is student self-selection such that OLS estimates overstate the benefits to attending schools with higher-achieving peers. However, instrumental variables estimates show that students who attend schools with higher achieving peers are more likely to stay in school to take the secondary leaving exams, have high test scores, pass more exams, and earn the prerequisites for admission to tertiary education. This suggests that the link between ability-grouping and inequality, documented by various researchers, is not driven solely by student self-selection. This findings show that ability-grouping may increase educational inequality on a broad range of outcomes by reinforcing pre-existing achievement differences. To my knowledge, this paper is the first to document that consigning low-achievers to low value-added schools is one mechanism through which ability-grouping (and therefore tracking) may increase educational inequality. When I test for non-linear effects, I find that very high-achieving schools provide little value-added over average schools – a finding consistent with recent research on school quality documenting that while there are some improved outcomes, students' test scores are not improved by attending a high-achievement school [Cullen, Jacob and Levitt (2005;2006); Clark (2008)]. On the other hand, I find that schools with average achievement peers provide much more value-added than schools with low-achievement peers. This is consistent with Hanushek and Woessman (2007) and Atkinson, Gregg, and McConnell (2006) finding that tracking disproportionately hurts students at the bottom of the test score distribution. I find that girls benefit more from attending schools with higher-achieving peers than boys. This is similar to findings from North Carolina [Hastings, Kane and Staiger (2006a; 2006b)] and relates to a growing literature documenting larger treatment effects for girls than for boys.

The remainder of the paper is as follows: Section II discusses how ability-grouping could create differences in school value-added that would lead to increased educational inequality. Section III describes the Trinidad and Tobago education system and the data used. Section IV describes the empirical strategy used to identify school effects and the corresponding results. Section V presents concluding remarks.

II The School Value-Added Mechanism.

The main theoretical justification for ability grouping (both at the school and classroom level) is that a homogeneous student body leads to improved student outcomes by allowing for more student cohesion, greater teacher focus, and a curriculum and pace more closely tailored to the particular ability level of the students. Researchers have studied the distributional and efficiency effects of *classroom* ability-grouping, and the results are mixed [studies include Betts and Shkolnik (1999); Rees, Brewer and Argys (1999), Figlio and Page (1998, 2002); Hoffer (1992)]. Using experimental data, Duflo, Dupas and Kremer (2008) find that the classroom homogeneity created by ability-grouping may benefit both high and low-achieving students. Since I make cross-school comparisons in Trinidad and Tobago, where all schools have homogenous student populations, the estimated differences in school value-added associated with peer achievement will not be confounded with this “homogeneous classroom effect”.

Ability-grouping is often coupled with a dual education system where certain schools have an academic focus while others have a vocational focus. Studying the removal of vocational programs in favor of general education in Romania, Malamud and Pop-Eleches (2007) find that students are less likely to work in manual or craft-related occupations when they receive a general education. Fortunately, Trinidad and Tobago has a national curriculum so that differences in outcomes across schools will not be confounded with curricular differences. In this paper, I refer to ability-grouping as tracking with no curricular differences across schools.

The differences in school value-added in Trinidad and Tobago will largely reflect (a) differences in peer quality and (b) differences in input quality endogenous to peer quality. I discuss both, in turn, below. Since school ability-grouping groups students by achievement, it has a profound effect on the peers to which students are exposed. Ability grouping lowers the average peer quality for low-achievement students and increases the average peer quality of high-achievement students. In general, researchers have found that both high and low achieving students benefit from higher achieving peers and there is no consensus on whether such peer effects are non-linear.[5] If all students benefit from higher-achieving peers, high-achievement schools may improve student outcomes solely due to having brighter peers. Ability grouping, by creating greater inequality in peer inputs, may engender greater educational inequality. Note that non-linear peer effects (i.e. high achievers benefit more from high achieving peers than do low achieving students) have been used to justify ability-grouping.[6] Such non-linear peer effects relate to the efficiency of ability-grouping rather than the distributional effects. As such, non-linear peer effects, while a very important consideration in evaluating the total effect of ability-grouping, are beyond the scope of this paper.

[5] Several studies find that students tend to have better outcomes on average when their peers are brighter on average [Hoxby (2000), Hoxby and Weingarth (2005), Sacerdote (2001), Zimmerman (2003)] while others provide mixed evidence [Katz, Kling Liebman (2007), Angrist and Lang (2004), Burke and Sass (2006)].

[6] If peer effects are non-linear in means, (i.e. students are affected by the mean characteristics of their peers, but exert different effects on students with different characteristics) the distribution of students matters for both the dispersion *and* the average level of outcomes. For example, if high-achievement students are unaffected by peer characteristics, while low-achievement students benefit from high-achieving classmates, achievement grouping across schools would increase test score dispersion and reduce average test scores by lowering test scores for those tracked into low-achievement schools solely due to peer effects *ceteris paribus*. In this scenario, high achieving peers exert a positive externality on lower-achievement classmates that is not being exploited under tracking. If, on the other hand, high-achievement students benefit more from having high-achievement peers than low-achievement students do, the opposite would be true. While this is an important aspect of tracking, the efficiency implications of school tracking are beyond the focus of this paper.

Another reason ability-grouping could affect student outcomes is that, over time, the quality of school inputs may be endogenous to the quality of the peers. Schools with bright, motivated students may, over time, have more affluent alumni networks, and, as such, higher alumni donations leading to better facilities and better funding through greater political influence. In addition to these input differences, schools with brighter students may have better teachers since teachers may prefer teaching in schools with well-behaved, bright, motivated students.[7] While there are direct effects of average peer achievement, school ability-grouping could lead to differences in school inputs that are endogenous to average peer quality in schools.

In sum, ability grouping can affect student outcomes in a variety of ways. Since all Trinidad and Tobago schools are in an ability grouped education system and all share the same curriculum, most of the differences in value-added across schools with different levels of incoming peer achievement will be due to the differences in input quality and differences in peer achievement caused by ability grouping. As such, by using Trinidad and Tobago data, I hope to isolate one potential mechanism behind the ability-grouping effect. Specifically, I aim to identify the effect of ability-grouping on the margin (i.e. the effect of being assigned to a high ability school versus a low ability school).

III The Trinidad and Tobago Education System (and Data).

Trinidad and Tobago is a nation of roughly 1.3 million inhabitants consisting of twin islands in the southern Caribbean. Trinidad is 1,841 square miles in area while Tobago has an area of about 115 square miles.[8] Primary and secondary public school education is free and compulsory between the ages of six and twelve. There are eight educational school districts. The Trinidad and Tobago education system evolved from the English education system. Secondary school begins in first form (after grade five) and ends at fifth form (the equivalent of grade 10) when students take the Caribbean Secondary Education Certification (CSEC) examinations. These are the Caribbean equivalent of the British Ordinary levels (O-levels) examinations.[9] Students seeking to continue their education typically take five or more subjects, including

[7] Supporting this notion, Jackson (2008) finds that a quasi-exogenous repatriation of low-income black students into schools at the end of school desegregation was associated decreases in observable and unobservable teacher quality.
[8] Library of Congress Country Studies (Trinidad and Tobago)
[9] There are 31 CSEC subjects covering the range of purely academic subjects such as Physics, Chemistry and Geography, to the more work and vocationally related such as Technical Drawing and Principles of Business and Office Procedures.

English language and mathematics. As such, these are the two most commonly taken subjects. The CSEC examinations are accepted as an entry qualification for higher education in Canada, the United Kingdom and the United States. After taking the CSEC, students may continue to take the Caribbean Advanced Proficiency Examinations (CAPE), at the end of sixth form (the equivalent of grade 12), which is considered tertiary level education but is a prerequisite for admission to the University of the West Indies.[10] The CAPE is the Caribbean equivalent of the English Advanced Levels (A-Levels) examinations.

In Trinidad and Tobago, there are three types of secondary schools: Government schools, Government Assisted schools and Comprehensive schools. Government schools are secondary schools that provide instruction from form one (grade 6) through fifth form (grade 10) and often continue to upper sixth form (grade 12). These schools teach the national curriculum and are fully funded and operated by the Government. Government Assisted schools, often the more elite schools, are like Government schools but differ along a few key dimensions. They are run by private bodies (usually a religious board) and, while capital expenses are publicly funded, their teacher costs are not paid for by the Government. Along all other dimensions, Government and Government Assisted schools are identical. The third type of school, Comprehensive schools, are Government schools that were historically vocational in focus. In the past, students with low test scores after grade 5 were assigned to such schools and after 3 years took an exam to gain admission to a senior secondary school (or possibly a regular Government school) which would prepare them for the CSEC examinations. This system of schools has been phased out so that in 2000, the relevant sample period, all schools taught the same academic curriculum and only a handful of Comprehensive schools did not provide instruction through to the CSEC exams.[11]

III.1. Data and Summary Statistics: There is substantial variation in school and peer quality in Trinidad and Tobago. Students take the Secondary Entrance Assessment (SEA) examinations after grade five before entering secondary school. To illustrate the extent of ability-grouping across secondary schools, Table 1 summarizes the full SEA data for the cohort of 2000, broken

[10] This is the largest University in the Caribbean and is the primary institution of higher learning for those seeking to continue academic studies. There are a few small technical schools and teachers colleges across the Caribbean.
[11] In those few junior Comprehensive schools that do not provide instruction through to the CSEC exams the vast majority of students would attend the senior secondary school associated with their junior secondary school. For example, a typical student who is assigned to Arima junior secondary school will take the CSEC examinations at Arima senior secondary school provided they do not drop out of the system.

up by the schools' rankings in incoming SEA scores (i.e. the school with the highest average incoming total SEA scores is ranked first and the school with the lowest average total SEA scores is ranked last). In 2000 there were 31,593 SEA test takers of whom 22,876 students were linked to CSEC exam data in 123 secondary schools five years later (or four years for early takers).[12] The first three variables are the total SEA, SEA math and SEA English scores. These continuous test scores have been standard normalized to be mean zero with a variance of one. As one would expect, the top 40 ranked schools have higher SEA scores than lower-ranked schools. The average total SEA scores are 1.78 standard deviations higher at the top 40 schools than the bottom 40 schools. The difference between the top and bottom ranked school is a full 4.93 standard deviations. Schools ranked in the top 40 had students with over one standard deviation higher incoming math and English SEA scores than schools ranked between 41 and 80, which in turn had students with average math and English incoming scores over half a standard deviation higher than schools ranked below 80. Females make up slightly more than half of students in each school group, reflecting the gender imbalance in schools in Trinidad and Tobago.

Figure 1

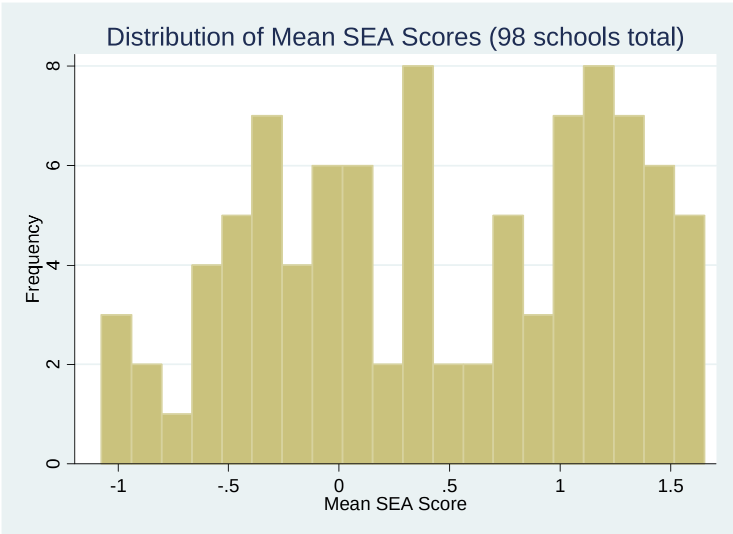


Figure 1 above shows the distribution of mean incoming total SEA scores in Trinidad and Tobago schools for 2000. The distribution of school peer achievement is somewhat bimodal, such that the plurality of schools have mean peer SEA scores between 1 and 1.6 standard deviations above the mean or between 0.2 standard deviation above and 0.6 standard deviations below the mean. While there are a few schools with mean SEA scores lower than 0.6 standard

[12] Students were matched based on name, gender and date of birth. The match rate was just over 70 percent which is consistent with the national high school dropout rate of one third.

deviations below the mean, and a few schools with mean scores between 0.2 and 1 standard deviation above the mean, most schools fall into the two ranges described above.

As one might expect, those schools that have the brightest peers also have the best outcomes. In 2000, Table 1 shows that 90 percent of students at schools ranked better than 40 took the CSEC exams compared to 75 percent for schools ranked 41 to 80, and 65 percent for schools ranked below 80. Students in the top 40 schools take several more exams and pass several more exams than students at lower ranked schools, such that the average student at a top 40 school takes 6.38 exams and passes 5.45 of them compared to taking 4.43 exams and passing 2.2 exams in schools ranked between 41 and 80 and taking 2.93 exams and passing only 1.03 at schools ranked below 80. Some of these differences reflect the fact that students who do not take the CSEC exams have no passes or exams attempted.[13] There are also large differences in math and English grades earned by these students on the CSEC exams. The CSEC is graded from 1 though 7 with 7 being the lowest score and 1 being the highest. A one point difference represents the difference between an A and a B. Students who have not taken the CSEC exams are given a grade of 7. Students at top 40 schools score on average 2.05 grade points better in math and 2.2 grade points better in the English CSEC exams than students in schools ranked 41 through 80. They also score 3.08 points better in math and 3 points better in English than students at schools ranked below 80. This three point difference is the distance from an A to a D, such that if the average student at a top 40 school earns a B, the average student at schools ranked between 41 and 80 earns a D and a student in a school ranked below 80 earns an F. The last outcome is obtaining a certificate. This variable denotes passing five CSEC subjects including math and English. This is a common prerequisite for continuing education. There are clear differences in this outcome across schools such that 70 percent of students at the top 40 schools earn a certificate, compared to only 18 percent at schools ranked between 41 and 80 and 5 percent at schools ranked below 80. Surprisingly, virtually no student who attends a school ranked below 80 satisfies the requirement to continue to sixth form (grades 11 and 12).

Table 1 documents that schools with the highest achieving students are on average smaller and disproportionately Government Assisted schools, while the schools with the weakest performing students are disproportionately Comprehensive schools. Roughly two thirds of the

[13] The regression estimates in Section IV explicitly take into account the fact that student with missing CSEC outcomes may not have received the lowest possible outcomes. The overall findings are robust to alternative assumptions about the possible performance of students who did not take the CSEC exams.

top 40 schools are Assisted while none are Comprehensive, and about one third of schools outside of the top 40 are Comprehensive schools. In Trinidad and Tobago, as in many nations, the schools that attract the brightest students typically have the best school resources. The one input for which there is aggregate data across school types is teachers. In 1999, 86 percent of teachers at Government Assisted schools had a bachelor's degree compared to 70 percent for Government schools and only 64 percent for Comprehensive schools. Similarly, 31 percent of Government Assisted school teachers had an education degree compared to 28 percent for Government schools and 12 percent for Comprehensive schools [National Institute of Higher Education and Science and Technology (1999)].

Table 1

Summary Statistics for 2000 SEA cohorts: By School Rank in Average Incoming Total SEA Scores

Rank Range (incoming peer scores)	1-40	41-80	81+
SEA Cohort Year		2000	
Normalized SEA Score Total (incoming)	1.26	0.12	-0.52
	(0.67)	(0.76)	(0.80)
Normalized SEA Score Math (incoming)	1.11	0.01	-0.59
	(0.68)	(0.83)	(0.77)
Normalized SEA Score English (incoming)	1.16	0.06	-0.51
	(0.67)	(0.81)	(0.81)
Female	0.53	0.52	0.55
	(0.50)	(0.50)	(0.50)
Take CSEC	0.90	0.75	0.65
	(0.30)	(0.43)	(0.48)
Exams Taken	6.38	4.43	2.96
	(2.37)	(2.82)	(2.69)
Exams Passed	5.45	2.26	1.03
	(2.61)	(2.43)	(1.73)
English Grade (7=lowest , 1=highest)	2.27	4.32	5.35
	(1.94)	(2.08)	(1.88)
Math Grade (7=lowest , 1=highest)	2.64	4.87	5.64
	(1.98)	(1.88)	(1.59)
Certificate*	0.70	0.18	0.05
	(0.46)	(0.38)	(0.22)
Admitted Cohort Size	179.24	389.18	544.75
	(150.87)	(232.58)	(203.32)
Government Assisted School	0.65	0.00	0.00
	(0.48)	(0.00)	(0.00)
Comprehensive School	0.00	0.34	0.34
	(0.00)	(0.47)	(0.47)
Observations	5337	10016	16240

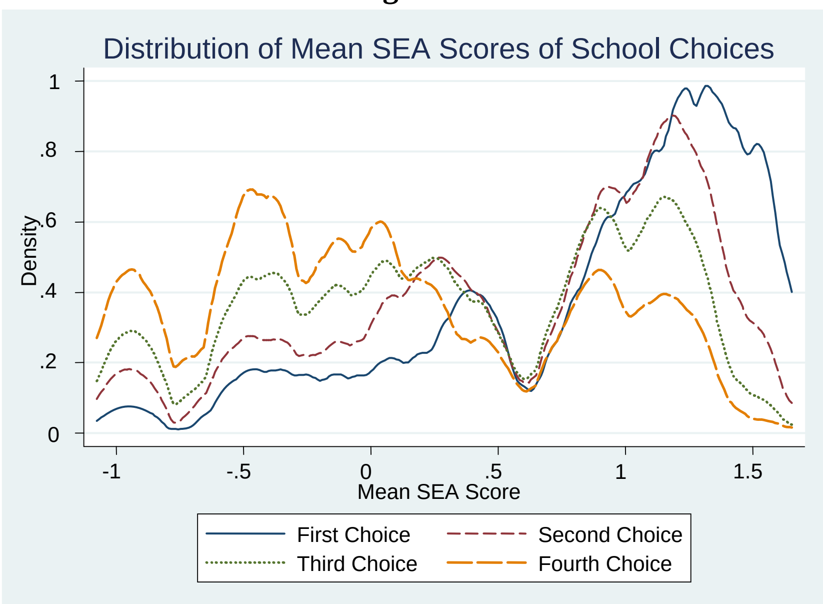
Standard deviations are reported below the sample means.
Note: Certificate denotes passing five CSEC exams including English and math. This is a prerequisite to most tertiary education institutions.

III.2. Student Assignment Rules (Algorithm): Due to a disparity between the number of secondary-school places and the number of school-age children, students compete for a limited number of premium slots. Unlike in many countries where private schools are often of higher perceived quality, private schools in Trinidad and Tobago account for a very small share of student enrollment and tend to serve those who "fall through the cracks" in the public system. After grade five, students take the SEA examinations. Each student lists four ordered secondary school choices. These choices and their score on the SEA are used by the Ministry of Education to assign them to schools using the following algorithm. Each secondary school has a predetermined number of open slots each year and these slots are filled sequentially such that the most highly subscribed/ranked school fills its spots first, then the next highly ranked school fills is slots and so on and so forth until all school slots are filled. This is done as follows: (1) Each student is tentatively assigned to their first choice school. The school that is oversubscribed with the highest "cut off" score fills its slots first. For example, suppose both school A and school B have 100 slots and 150 students list each of them as their top choice. If the 100th student at school A has a score of 93% (its "cut-off" score) while the 100th student at school B has a score of 89%, school A is ranked first and fills all its spots first. (2) Those filled school slots and the students who are assigned to the highest ranked school are removed from the applicant pool and the process is repeated, where a student's second choice now becomes their first choice if their first choice school has been filled.

This process is used to assign over 95 percent of all students, however, there is a group of students for whom this mechanism may not be used. Government Assisted schools (which account for about 16 percent of school slots) are allowed to admit 20 percent of their incoming class at the principal's discretion. As such, the rule is used to assign 80 percent of the students at these schools, while the remaining 20 percent are hand picked by the school principal before the next highest ranked school fills any of its slots. For example, suppose the highest ranked school has 100 slots and is a Government Assisted school. The top 80 students will be assigned to that school while the principal will be able to hand pick 20 other students who listed the school as their top choice. The remaining 20 students would be chosen based on family alumni connections, being relatives of teachers or religious affiliation (Since Government Assisted schools are often run by religious bodies). Only after all the spots (the assigned 80 percent and the hand-picked 20 percent) at the highest ranked school have been filled will the process be repeated for the remaining schools. As such, the school assignments are based partly on a deterministic function of student test scores and student preferences (which is beyond students' control after taking the SEA exams), and partly on the hand-picking of students by school principals (which can potentially be manipulated by students).

Since student preferences are an important part of the assignment process, it is important to better understand them. Students' school choices are based largely on their own perceived ability, geography, and religion. Specifically, higher ability students tend to have higher achievement schools in their list, students often request schools with the same religious affiliation as their own, and students typically list schools that are geographically close to their homes. Since Trinidad and Tobago is small, attending school far from home is uncommon but feasible. Figure 2 shows the distribution of the mean peer incoming SEA scores of students' school choices. As one would expect, the distribution of mean SEA scores of first choice schools is to the right of the second choice school which is to the right of the third choice schools which, in turn, is to the right of the fourth choice schools. In other words, students tend to put schools with higher achieving peers higher up on their preference ranking. In fact, on average the difference between the mean incoming SEA scores at a student's top choice school and second choice school is 0.277 standard deviations, between the top choice school and the third choice school is 0.531 standard deviations and between the top choice school and the fourth choice school is 0.82 standard deviations. Roughly 15 percent of students make their top choice school and for those students who did not, the difference in mean total SEA scores between their actual school and their top choice school is 0.87 standard deviations.

Figure 2



IV The Effect of Mean SEA Score of School on Student Outcomes

IV.1 Identifying School Effects (Naïve Baseline model): To estimate the effect of attending a school with higher achieving peers the basic empirical strategy is to compare the outcomes of students with similar incoming test scores at different schools using cross-sectional variation from the 2000 SEA cohort. For the naïve baseline specification, I model the outcome of student *i* at a school *s* with the following equation.

$$Y_{i,s} = SEA_i \cdot \beta + \overline{SEA}_s \pi + X_i \delta + \varepsilon_{i,s} \qquad [1]$$

In [1], $\overline{SEA}_s$ is the mean total SEA scores for incoming students at school *s*, SEA_i is a matrix of incoming test scores (the student's total SEA score, total SEA score squared, math SEA score, math SEA score squared, English SEA score and English SEA score squared), X_i includes student gender, religion, and primary school district, and $\varepsilon_{i,s}$ is the idiosyncratic error term. Naïve OLS estimates of π from [1] may be biased since (1) students who are unhappy with their initial school assignment can appeal and have their assignment changed, (2) students may be able to transfer across schools, and (3) Government Assisted schools can admit 20 percent of their incoming class at the discretion of the school principal. Since readers may worry that there is ample opportunity for students to self-select into secondary schools, I propose a rule-based instrumental variables strategy to deal with this potential endogeneity concern.

IV.2 Rule-Based Instrument: To remove self-selection bias from the actual school attendance, one would want the school assignment that would prevail if students could not self-select into schools. Such an assignment can be constructed by "tweaking" the school assignment mechanism to impose the deterministic portion of the assignment mechanism (described in section III) on *all* students. Since the deterministic portion of the assignment mechanism is used to assign most students to schools, the school assignments based on the "tweaked" assignment mechanism should be correlated with the schools students actually attend. However, since the deterministic portion of the assignment mechanism cannot be manipulated by students or school principals, the "tweaked" assignments should be uncorrelated with *unobserved* student characteristics such as motivation and ability conditional on student test scores and school choices. As such, I propose an instrumental variables strategy based on these "tweaked" assignments.

The rule-based instrumental variables strategy is in the spirit of Campbell (1969), Angrist and Lavy (1999) and Andrabi, Das and Khwaja (2007). I exploit the fact that the school attended, and therefore the mean SEA score of students at the school attended, is partly based on a deterministic function of the student's total SEA score and the student's school preferences. Since this deterministic function is non-linear and non-monotonic, it can be used as an instrument while directly controlling for smooth functions of the underlying covariates themselves [Fisher (1976)]. For each school student pair, I use the following rule for whether student *i* is assigned to school *s*.

$$Rule_{is} = \begin{cases} 0 \text{ if } SEA_i \leq c_s \\ 0 \text{ if } s \notin \{pref_i\} \\ 1 \text{ if } SEA_i > c_s \text{ and } j \in \{pref_i\} \text{ and } Rule_{is'} \neq 1\ \forall s' \succ_i s \end{cases} \quad [2]$$

Where c_s is the cut-off score for school *s*, $pref_i$ is the set of school choices for student *i* and $s' \succ_i s$ if student *i* prefers school *s'* over school *s*. The first condition captures that fact that schools use cut-off scores to assign students. The second condition captures that students are not assigned to schools that are not in their choice set. The third condition shows that a student is assigned to a school if they have a score above the schools cut-off, the school is in their choice set, and the student has not already been assigned to a school they prefer. $Rule_{is}$ is, in essence, the deterministic portion of the student assignment algorithm.

As discussed in section III, the cut-offs are unknown to educators, principals and students while the SEA exams are being taken – precluding any gaming of the cut-offs. The cut-offs are set to fill all of a school's available slots each year based on the assignment algorithm. Since I want the cut-off that would prevail in the absence of any self-selection or hand-picking, I model the cut-off score for school *s* as the cut off score that would prevail if all students were assigned to schools according to $Rule_{is}$. Specifically the cut-off for school *s* is

$$c_s = [SEA_{rank_{is}=T_s} \mid s \in \{pref_i\} \text{ and } Rule_{is'} \neq 1\ \forall s' \succ_i s] \quad [3]$$

Where $Rank_{is}$ is the rank of student *i* among those who are in the admission pool for school *s*, so that $SEA_{rank_{is}=T_s}$ is the SEA score of the T^{th} ranked students in school *s*'s applicant pool. T_s is the fixed capacity of school *s*.

The rule-based instrument is constructed sequentially as follows: (1) All secondary school sizes are taken as given,[14] (2) all students are tentatively assigned to their top choice school, (3) the school for which the first rejected student has the highest test score fills all its slots (with the highest scoring students who listed that school as their first choice), (4) the students who were rejected from the top choice school are sent back into the applicant pool and their second choice school becomes their first choice school, (5) Step 2 is repeated, excluding assigned students and assigned school slots until the lowest ranked school is filled. The *only* difference between how students are actually assigned and the "tweaked" rule-based assignment is that at step (3) the "tweaked" rule does not allow any students to be hand-picked while, in fact, some students are hand-picked by principals only at Government Assisted schools. The resulting $Rule_{is}$ variables correctly identify the school assignment for 15,200 students. Since students who list schools way above their score range will not be assigned based on their preferences there are 5,460 students with no simulated assignment. Among students assigned to schools within their choice set, the rule is correct about two thirds of the time. I do not identify school value-added for schools that no student is assigned to based on $Rule_{is}$. Of the 123 secondary schools in Trinidad and Tobago, 98 of them have students who are simulated to be assigned to them. The remaining 26 schools are schools that nobody lists in their preferences, either because they are new schools, or because they are undesirable.[15]

IV.3 Sources of Variation: Conditional on incoming test scores and preferences, $Rule_{is}$ captures two plausibly exogenous sources of variation. The first source comes from comparing the outcomes of students at different schools who score just above and just below a school's cut-off. The logic behind this source of variation is the familiar regression discontinuity logic. Specifically, the likelihood of attending one's preferred school increases in a sudden and discontinuous manner as one's score goes from below the cut-off to above the cut-off for that school. If the location of the cut-off is exogenous to student characteristics, one can reasonably attribute any discontinuous jumps in the outcomes as one's score goes from below to above the

[14] School sizes are not endogenous to the application process and are based on strict capacity rules. School sizes are determined before students are assigned to schools and based on their predetermined school sizes the algorithm is applied. As such, the number of students assigned to a particular school (even if they do not attend) is the actual number of predetermined slots at the school.

[15] Since students with low scores will be assigned to the local high school that has available space if they "fail" out of their choice schools, students have no incentive to list these schools if they believe they have a change of gaining entry to a higher ranked school.

cut-offs to the increased likelihood of attending one's preferred school. The second source of variation comes from comparing the outcomes of students with the same test scores at different schools because they have different school preference orderings. Since preferences are directly observed, and the cut-offs generate exogenous variation in school assignment among students with the same preferences, one can directly control for a student's preferences (a unique feature of the Trinidad and Tobago data). One benefit of this second source of variation is that I am not restricted to identifying the local causal effect of schools based only on students around the cut-offs, but can identify school effects based on variation from *all* students.

IV.4 Econometric Model: The rule-based school assignments are used as instruments to isolate exogenous variation in mean SEA scores of the actual schools attended in a two-stage least squares (2SLS) regression. Specifically, I estimate the following system of equations.

$$\overline{SEA}_s = SEA_{i,t-1} \cdot \beta_1 + \sum_{j=1}^{95} Rule_{ij} + X_i\delta_1 + \sum_{p=1} I_{i,p} \cdot \theta_p + \varepsilon_{i,s,t,1} \qquad [4]$$

$$Y_{i,s,t} = SEA_{i,t-1} \cdot \beta_2 + \overline{SEA}_s \pi_{s,2} + X_i\delta_2 + \sum_{p=1} I_{i,p} \cdot \theta_p + \varepsilon_{i,s,t,2} \qquad [5]$$

All variables are defined as in [1], $\overline{SEA}_s$ is the mean total SEA scores for incoming students at school *s*, $I_{i,p}$ is an indicator variable equal to 1 if a student's rank ordering are preference group *p* and equal to zero otherwise[16] and the rule based school assignment $Rule_{ij}$ is excluded in the second stage equation. All regressions include student demographics, the student's total SEA score, total SEA score squared, math SEA score, math SEA score squared, English SEA score and the English SEA score squared. Standard errors are clustered at the school level. The first stage yields an F-statistic on the excluded instruments of 12.47.

IV.5 Specification Tests (Falsification Exercise): I argue that since the rule-based instruments are deterministic non-smooth functions of student test scores and preferences, conditional on test scores and student preferences, the rule-based school assignments are exogenous to unobserved student attributes. While this should be true by construction, I test the

[16] Each preference group is defined by a distinct preference ordering of schools. All students who list schools A,B,C,D in that order form a group, while students who list schools B,A,C,D are in a different group because even though they have the same list of schools, the ordering is different. There are 4561 preference groups with more than one student.

validity of this assumption by seeing if the instruments are correlated with observable student characteristics *before* entering secondary school. If the rule-based instruments were correlated with observable student characteristics such as gender, religion, and primary school district, it would cast serious doubt on the assertion that they isolate exogenous variation that is free from self-selection bias. On the other hand, if the instruments are not correlated with these pre-treatment student characteristics it would lend credibility to the identifying assumption. In table 2, I present the F-test associated with the null hypothesis that the rule-based instruments are equal to zero in a regression that controls only for total SEA scores, total SEA scores squared and the preference indicator variables. I also show the coefficients and t-statistics associated with mean total SEA scores using the rule-based instruments as exogenous predictors in the first stage. Table 2 presents these results for both the outcomes of interest and the observable pre-treatment characteristics (student gender, student religion, and student primary school district).

Table 2

Falsification Tests

P-value of the test that all rule based school dummies are equal to zero in an OLS regression of the outcome on the rule-based indicator variables while controlling for each student's total SEA score, total SEA score squared and the preference indicator variables.

Specification	1	2	3	4	5	6	
outcome	Take CSEC	Number taken	Math grade	English grade	Number passed	Certificate	
P-value	**0.0005**	**0.0005**	**0.0225**	0.0855	**>0.0001**	**>0.0001**	
Specification	7	8	9	10	11		12
outcome	Religion 1	Religion 2	Religion 3	Religion 4	Religion 4		Female
P-value	0.9909	0.9975	0.9923	1	0.9817		0.0995
Specification	13	14	15	16	17	18	19
outcome	District 1	District 2	District 3	District 4	District 5	District 6	District 8
P-value	1	1	1	1	1	1	1

Coefficient on Mean total SEA scores at school from 2SLS regressions. The excluded instruments are the rule-based school dummies.

Specification	1B	2B	3B	4B	5B	6B	
outcome	Take CSEC	Number taken	Math grade	English grade	Number passed	Certificate	
Coef on Mean SEA	**0.117**	**1.167**	**-0.665**	**-0.566**	**0.882**	**0.203**	
se	**[0.025]***	**[0.165]***	**[0.135]***	**[0.146]***	**[0.146]***	**[0.025]***	
Specification	7B	8B	9B	10B	11B		12B
outcome	Religion 1	Religion 2	Religion 3	Religion 4	Religion 4		Female
Coef on Mean SEA	-0.003	0.003	-0.022	0.011	0.02		0.021
se	[0.015]	[0.014]	[0.025]	[0.025]	[0.026]		[0.022]
Specification	13B	14B	15B	16B	17B	18B	19B
outcome	District 1	District 2	District 3	District 4	District 5	District 6	District 7
Coef on Mean SEA	0.005	0	-0.001	0.002	0.005	0.002	-0.01
se	[0.008]	[0.005]	[0.006]	[0.009]	[0.006]	[0.006]	[0.009]

Robust standard errors in brackets.

+ significant at 10%; * significant at 5%; ** significant at 1%

Each column represents a separate regression for a different outcome. All regressions include indicator variables for each possible preference ordering and control for the SEA score and the square of the total SEA score.

Specifications 1 through 6 show that one can reject the null hypothesis that the rule-based instruments are uncorrelated with the secondary school outcomes at traditional levels. The 2SLS regression coefficients in specifications 1B through 6B show that mean total SEA scores, as predicted by the rule-based instruments, are strongly associated with improved secondary school outcomes. All of these estimates are significant at the 1 percent level. In stark contrast, specifications 7 through 19 show that one cannot reject the null hypothesis that the excluded instruments are uncorrelated with any of the pre-treatment student characteristics at the 5 percent level. The 2SLS regression coefficients in specifications 7B through 19B show that mean peer total SEA scores, as predicted by the rule-based instruments, are not associated with pre-treatment student characteristics. In fact none of the t-statistics have p-values smaller than 0.3. Table 2 shows that the instruments are strongly correlated with secondary school outcomes, but have no statistical relation to observed pre-treatment student characteristics. The fact that student religion, which is used by principals when hand-picking students, is not correlated with the instruments (note the p-values are all above 0.98) lends credibility to the exogeneity of the rule-based instruments.

IV.6 Main Results: Each cell in Table 3 presents the coefficient on mean incoming total SEA scores at the school for a separate regression defined by outcome, sample, and method. All regressions include incoming test scores, their quadratics, gender, primary school district, student religion, and student school preferences. As pointed out in Horowitz and Manski (1998) the true conditional expectation of the outcome for a given group is

$$E[Y \mid X] = E[Y \mid X, g=1] \cdot P(g=1 \mid X) + E[Y \mid X, g=0] \cdot P(g=0 \mid X) \qquad [6]$$

Where Y is the outcome, X is a matrix of observable characteristics and g is equal to 1 if the outcome is observed and 0 otherwise. While I observe g and X for all students, $E[Y \mid X, g=0]$ is not observed for school dropouts. While a naive approach would be to assign the lowest possible value to those with missing outcomes, equation [6] illustrates that doing so will overstate the conditional expectation for schools that have many non-takers. I deal with missing exam outcomes using three methods. First, I use the naive approach of assigning the lowest possible outcome to those who do not take the CSEC exams. Second, I use imputation – assign each missing outcome with the conditional expectation of the outcomes based on the instruments.[17] The third method is to use a control function – i.e. include the probability of selection conditional on the instruments as a conditioning variable [Angrist (1995)]. All three approaches yield qualitatively similar results. However, as one would expect, the control function and imputation methods yield smaller estimates.

Table 3

Coefficient on Mean Standardized Normalized Incoming SEA Scores at School on various outcomes

		-1	-2	-3	-4	-5	-6		
		Dependent Variable							
		Take CSEC	Exams Taken	English Grade	Math Grade	Exams Passed	Certificate	Method	Control for missing outcomes
1	All Students	0.07	1.43	-1.117	-0.96	1.47	0.198	OLS	None
		[0.037]+	[0.119]**	[0.097]**	[0.069]**	[0.128]**	[0.027]**		
2	All Students	0.119	1.181	-0.599	-0.642	0.863	0.188	2SLS	None
		[0.033]**	[0.156]**	[0.122]**	[0.113]**	[0.185]**	[0.027]**		
3	All Students	-	1.016	-0.405	-0.378	0.676	0.167	2SLS	Impute
		-	[0.100]**	[0.074]**	[0.117]**	[0.221]**	[0.032]**		
4	All Students	-	0.991	-0.465	-0.512	0.833	0.179	2SLS	Control function
		-	[0.152]**	[0.133]**	[0.106]**	[0.191]**	[0.028]**		
5	Female	0.127	1.458	-0.843	-0.828	1.182	0.211	2SLS	None
		[0.041]**	[0.202]**	[0.157]**	[0.157]**	[0.188]**	[0.033]**		
6	Female	-	1.333	-0.738	-0.758	1.176	0.209	2SLS	Control function
		-	[0.199]**	[0.160]**	[0.155]**	[0.189]**	[0.033]**		
7	Male	0.127	1.006	-0.526	-0.426	0.678	0.13	2SLS	None
		[0.030]**	[0.198]**	[0.147]**	[0.133]**	[0.180]**	[0.030]**		
8	Male	-	0.838	-0.452	-0.284	0.648	0.11	2SLS	Control function
		-	[0.229]**	[0.165]**	[0.137]*	[0.198]**	[0.029]**		

Robust standard errors in brackets. Standard errors are adjusted for clustering at the school level.
+ significant at 10%; * significant at 5%; ** significant at 1%
All regressions control for student religion, gender, primary school district and include the student's total SEA score, total SEA score squared, math SEA score, math SEA score squared, English SEA score, and English SEA score squared.
The 2SLS regressions also include indicator variables for the student preference. The excluded instruments are the rule-based school assignments. The full sample has 22031 observations. The sample of females has 11407 observations and the sample of males has 9126 observations. (note the male and female samples do not add up to the full sample since preference groups with only one student are dropped from the regression models).

Rows 1 through 4 report the regression results using both male and female students. The baseline OLS results in row 1 show that attending a school with higher mean incoming test scores is associated with an increased likelihood of taking the CSEC exams, and with greater success on these exams. On average, attending one's top choice school as opposed to ones fourth choice school is associated with a 0.82 standard deviation difference in mean peer achievement. On average, students' top choice schools have mean SEA scores 0.277 standard deviations

[17] Underlying this approach is the assumption that students with missing test scores would have performed *at best* as well as observationally identical students with non-missing test scores. The assumption underlies many selection models [i.e. Heckman (1976); Amemiya (1985)]. Imputation should result in lower bound estimates if this assumption is valid.

higher than their second choice school. As such, the OLS results indicate that just missing one's top choice school would be associated with peer test scores that are 0.277 standard deviations lower which would lead to a 1.9 point decrease in CSEC exam taking, taking 0.4 fewer exams, earning 0.309 grade points lower in English and 0.265 grade points lower in math, passing 0.41 fewer exams and being 5.5 points less likely to earn a certificate.

Row 2 shows the 2SLS regression results that use the rule-based assignments and condition on a student's preferences. The results indicate that attending one's second choice school as opposed one's first choice school (a 0.277 standard deviation decrease in mean peer total SEA scores) would, on average, be associated with a 3.3 point decrease in the likelihood of taking the CSEC exams, taking 0.33 fewer exams, earning 0.17 grade points lower in English and 0.18 grade points lower in math, passing 0.24 fewer exams and being 5.2 points less likely to earn a certificate. All coefficients are significant at the one percent level. Students attending their third and fourth choice schools would experience even larger differences in outcomes. Since only 15 percent of students attend their top choice school, these differences are sizable.

While the CSEC taking and certificate results are similar between the OLS and 2SLS models, the coefficients for the number of exams passed and the math and English grades are roughly two thirds smaller in the 2SLS than the OLS models – suggesting that self-selection bias led to overstated OLS results. Rows 3 and 4 show the 2SLS results while accounting for the fact that arbitrarily assigning the lowest value of the outcome could lead one to overstate the effects of schools. For both the imputation and control function methods this reduces the magnitude of the coefficients by roughly 15 percent. The imputation method, which should provide a lower bound on the estimates, shows that attending one's second choice school as opposed one's first choice school would, on average, be associated with taking 0.28 fewer exams, earning 0.11 grade points lower in English and 0.1 grade points lower in math, passing 0.19 fewer exams and being 4.6 points less likely to earn a certificate. These estimates are all statistically significant at the one percent level. Even these lower bound estimates are non-trivial.

IV.7 Differences by gender: The analysis by gender is motivated by a growing literature documenting that females often benefit from interventions while males are unaffected and in some cases perform worse.[18] Rows 5 through 8 show the 2SLS models for males and females separately. I present both the naive model that assigns the lowest possible outcome to those with missing test data and the results using the control function method. Columns 5 and 7 show clear differences by gender. Females are more responsive to differences in peer achievement across schools than males such that attending one's second choice as opposed to first choice school (a 0.277 standard deviation difference in mean SEA scores) for females would be associated with a 3.5 point decrease in the likelihood of taking the CSEC exams, taking 0.4 fewer exams, earning 0.23 grade points lower in English and 0.23 grade points lower in math, passing 0.33 fewer exams and being 5.8 points less likely to earn a certificate. For males the same change is only associated with a 3.5 point decrease in the likelihood of taking the CSEC exams, taking 0.28 fewer exams, earning 0.15 grade points lower in English and 0.12 grade points lower in math, passing 0.19 fewer exams and being 3.6 points less likely to earn a certificate. Rows 6 and 8, which include the predicted likelihood of taking the CSEC as a control, show even larger differences by gender such that the coefficients for females are about twice as large as those for males for all outcomes.

In sum, Table 3 shows that, on average, students have better educational outcomes when they attend schools with brighter peers. Even after taking student self-selection into account and dealing with missing outcomes for school dropouts, one can reject the hypothesis that attending schools with higher achieving peers has no effect for all outcomes at the one percent level. This is compelling evidence that school ability-grouping will tend to exacerbate pre-existing achievement differences between students (and possibly increase test score dispersion) by consigning low-achieving students to those schools that create the least value-added and high-achieving students to schools that produce the most value-added. Consistent with a growing literature, I find much stronger marginal effects for females than for males.

IV.8 Elite schools or bad Schools? (Non-linear effects): Proponents of school ability-grouping support ability-grouping based on the belief that it creates excellent schools at the top of the achievement distribution, while opponents of school ability-grouping are concerned that it creates an underclass of schools with high concentrations of low-achieving students that produce

[18] For example Kling, Ludwig, and Katz (2005); Anderson (2007); Angrist, Lang, and Oreopoulos (2007); Angrist and Lavy (2007).

very low value-added. Much research on school quality has focused on the effect of attending high-achieving or “elite” schools. Since the rule-based instruments provide exogenous variation in school attendance for *all* schools, I can test whether the benefits to attending a school with higher achieving peers, on average, is driven by benefits to elite schools at the top of the school achievement distribution or ill-effects to attending low-achievement schools at the bottom of the school achievement distribution.

To test for such non-linearly, I estimate the value-added associated with each individual school using the 2SLS model. Specifically I estimate the following.

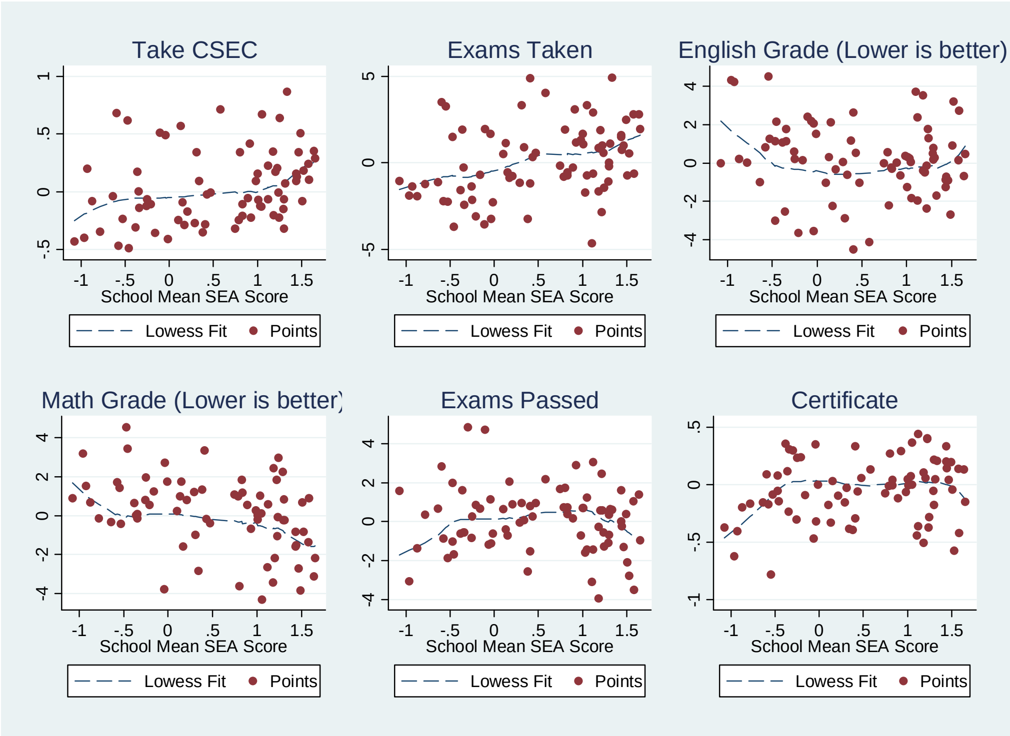
$$I_{i,s} = SEA_i \cdot \beta_1 + \sum_{j=1} Rule_{ij} + X_i\delta_1 + \sum_{p=1} I_{i,p} \cdot \theta_p + \varepsilon_{i,s,1} \qquad [7]$$

$$Y_{i,s} = SEA_i \cdot \beta_2 + \sum_{s=1} I_{i,s} \cdot \pi_{s,2} + X_i\delta_2 + \sum_{p=1} I_{i,p} \cdot \theta_p + \varepsilon_{i,s,2} \qquad [8]$$

All variables are defined as before, $I_{i,s}$ is an indicator variable equal to 1 if student *i* attends school *s*, $I_{i,p}$ is an indicator variable equal to 1 if a students rank ordering put them in preference group *p* and equal to 0 otherwise and the rule-based school assignments $Rule_{ij}$ are excluded in the second stage equation. All regressions include student demographics, the student’s total SEA score, total SEA score squared, math SEA score, math SEA score squared, English SEA score and the English SEA score squared. Six schools have weak first stages and create outlier estimates (outside the possible range of the data) and are as such omitted. The first stages for the remaining 92 school dummies are all strong with first stage F-statistics above 9. Once these school effects are estimated, by plotting school value-added against mean SEA scores, one can see graphically if schools with higher value-added have higher achieving peers and one can also see if this relationship is non-linear. If the relationship documented in Table 3 is driven by large benefits to attending average schools over low peer achievement schools, the relationship between school value-added and mean SEA score should be steepest at low mean SEA levels. If on the other hand, the large benefits on average are driven by “elite” schools at the top, then the relationship should be steepest at high mean SEA levels.

Figure 3 presents the scatter plot of the estimated school effects against the school mean total SEA scores. A few outlier points are not shown since they lay far outside the typical range of the estimates. Figure 3 shows that the relationship between value-added and incoming peer quality is not the same over all ranges of peer quality. For CSEC taking, the marginal effect is largest at low levels and high levels of peer quality, but relatively flat in the middle. There is a similar pattern for the number of exams taken. For both math and English grades, brighter peers are associated with better school effects at low peer achievement levels, but is flat in the middle of the distribution (note that lower values are better for math and English exam grades). Oddly, for math, at the very top of the peer quality distribution there is a positive relationship between mean SEA scores and math grades, while for reading, at the very top of the peer quality distribution, better peer appears to be associated with *worse* English performance. For the number of exams passed, as with the other variables, there is a positive relationship between school value-added and mean total SEA scores at low mean SEA score levels that becomes flatter in the middle of the distribution. Like the English grades, at high levels of peer SEA scores, higher achieving peers appear to be associated with worse outcomes on the margin as the slope becomes negative. For the likelihood of earning a certificate, as with all other outcomes, there is evidence of a steep positive association between higher achievement peers and school value-added at low mean SEA score levels, that levels off in the middle of the peer achievement distribution. As with English grades and the number of exams passed, the line slopes downward at the very top of the peer achievement distribution -evidence that at very high peer quality levels, attending schools with higher achieving peers may lead to worse outcomes.

Figure 3

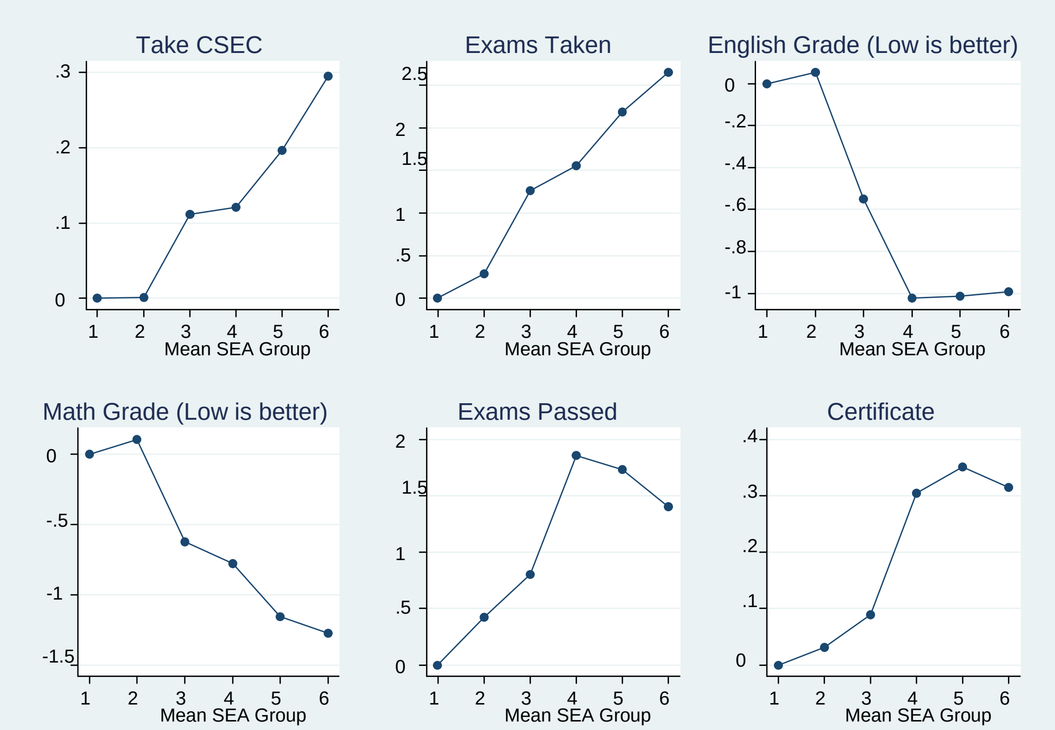


Two consistent patterns are evident in Figure 3. Specifically (1) there is a strong association between better outcomes and attending a school with higher-achieving peers at low levels of peer achievement, and (2) in the middle of the peer achievement distribution there is no strong relationship between higher school value-added and higher achieving peers. At the top of the peer quality distribution, it appears that there may be benefits to attending schools with brighter peers in math, but worse outcomes in English, the number of exams passed and obtaining a certificate. The findings would suggest that while thee are substantial gains to attending a school in the middle of the distribution vs. a school at the bottom of the SEA score distribution for all outcomes, there is little evidence that students have better outcomes at very high-achievement schools over attending schools in the middle of the distribution. In fact the evidence suggests that, on the margin, students may fare *worse* at the most selective schools than at less selective schools with average students.

Parametric approach: The plots in figure 3 are illuminating and suggestive. However, they are based on a noisy 2SLS procedure and do not present a statistical test for non-linearity. By imposing more parametric structure on the data, I can test for non-linearity and increase the statistical precision of the two-stage estimation procedure. Specifically I put schools into groups as follows: group 1 has schools with mean total SEA scores below -0.5, group 2 has schools with total mean SEA scores between -0.5 and 0, group 3 has schools with total mean SEA scores between 0 and 0.5, group 4 has schools with total mean SEA scores between 0.5 and 1, group 5 has schools with total mean SEA scores between 1 and 1.5, and group 6 has schools with total mean SEA scores above 1.5. I estimate the preferred 2SLS model shown in equations [7] and [8] where I use the SEA score ranges *en lieu* of the individual school indicator variables. Imposing parametric structure deals with weakly identified individual schools by putting them into groups and increases statistical precision. Figure 4 presents the coefficients on the school group dummy variables. The estimates are in table 4.

The results in Figure 4 are roughly consistent with the non-parametric results in Figure 3. For taking the CSEC exams and the number of CSEC exams taken, the coefficients on the group dummies are increasing as the mean SEA scores increase. For both these outcomes the sharpest increase appears to be between groups 2 and 3 (which covers mean SEA scores between -0.5 and 0.5). Unlike in figure 3, there is little evidence of a steep slope among schools with mean SEA scores below -0.5. The results for English and Math grades are illuminating. For both these outcomes, the marginal changes are small at the lower and upper end of the mean SEA distribution but are steepest in the middle between groups 2 and 5 (mean SEA scores between -0.5 and 1.5). For both outcomes, as with CSEC taking and exams taken, the largest marginal effects appear to be between groups 2 and 3 (mean SEA score between -0.5 and 0). Consistent with the findings in figure 3, for English grades, the slope between groups 4, 5 and 6 is slightly positive suggesting that marginal increases in mean total SEA scores for scores greater than 1 standard deviation above the mean may actually hurt English performance. For math however, there are apparent benefits to attending schools with brighter peers over all SEA score ranges.

Figure 4



Group 1: mean total SEA scores below -0.5, Group 2: mean total SEA scores between -0.5 and 0, Group 3: mean total SEA scores between 0 and 0.5, Group 4: mean total SEA scores between 0.5 and 1, Group 5: mean total SEA scores between 1 and 1.5, Group 6: mean total SEA scores above 1.5.

The results for the number of exams passed and earning a certificate are arguably the most important outcomes, since they cover the totality of what a student has achieved in secondary school. For both these outcomes, marginal increases in mean total SEA scores are associated with better outcomes between groups 1 and 4 (mean total SEA scores below 1 standard deviation above the mean). For both outcomes there appears to be a flattening out of the relationship between groups 4 5 and 6, such that marginal effects of improved peers is zero or possibly negative at mean SEA score ranges greater than 1 standard deviation above the mean. In other words, there may be no benefits (and possibly ill-effects) of attending a school with higher achieving peers if one is already at a school with mean SEA scores 1 standard deviation above the mean. While negative effects for elite school attendance may seem counterintuitive, it is possible that marginal students experience some school miss-match due to being the weakest of their school or class. Such a dynamic is documented by Zhang (2008) who finds that comparatively weak females have worse outcomes at "elite" schools while strong females benefit. It is also possible that the competitive environment at the most selective schools is deleterious to student outcomes.

Table 4

Tests for nonlinear effects. 2SLS regression results on school mean SEA scores range on various outcomes. Excluded instruments are the rule-based school assignments.

	Take CSEC	Exams Taken	English Grade	Math Grade	Exams Passed	Certificate
Mean SEA score Range						
-0.5 and 0	0.001	0.288	0.053	0.106	0.421	0.031
	[0.038]	[0.213]	[0.207]	[0.187]	[0.221]+	[0.027]
0 and 0.5	0.112	1.267	-0.549	-0.626	0.805	0.089
	[0.048]*	[0.229]**	[0.215]*	[0.186]**	[0.252]**	[0.032]**
0.5 to 1	0.121	1.558	-1.024	-0.779	1.858	0.304
	[0.051]*	[0.257]**	[0.282]**	[0.229]**	[0.299]**	[0.041]**
1 to 1.5	0.197	2.184	-1.012	-1.157	1.734	0.351
	[0.062]**	[0.280]**	[0.242]**	[0.222]**	[0.329]**	[0.049]**
above 1.5	0.295	2.648	-0.993	-1.273	1.4	0.314
	[0.094]**	[0.484]**	[0.334]**	[0.333]**	[0.442]**	[0.073]**

Robust standard errors in brackets. Standard errors are clustered at the school level.
+ significant at 10%; * significant at 5%; ** significant at 1%

All regressions control for student religion, gender, primary school district and include the student's total SEA score, total SEA score squared, math SEA score, math SEA score squared, English SEA score, and English SEA score squared, and indicator variables for the student preference. The excluded instruments are the rule-based school assignments. The sample has 22031 observations.

In sum, the evidence *does not* support the hypothesis that students benefit from attending the most selective schools on the margin. Figures 3 and 4 show that most of the benefits to attending a school with higher achieving peers occurs among schools with average or below average students and not to schools with mean total SEA scores greater than 1 standard deviation above the mean. This is a notable finding since this roughly represents the top third of schools (many of which are not considered particularly elite). This would imply that the most elite schools, the top 15 or 20 percent of schools, provide little value-added over the next highest 15 or 20 percent of schools. On the other hand, across a variety of outcomes, for the lowest two thirds of schools, student fare better when they attend schools with higher achieving peers. Insofar as much of the large differences in value-added occur at the bottom rather than the top of the peer achievement distribution, ability-grouping (and tracking) are likely to have larger negative effects for low-achieving students who get assigned to low-achievement schools than on higher achieving students who are assigned out of an elite school, into an average school.

V Conclusions

Studies have found that school ability-grouping is associated with increased inequality in educational attainment and wages. Since most evidence on school ability-grouping is based on the European experience where low-achievement tracks had vocational curricula and the removal of ability-grouping was coupled with several other institutional changes, it is not obvious that these changes were the result of ability-grouping or other mechanisms. Hanushek and Woessmann (2007), in a cross-country analysis, credibly isolate a tracking effect (ability grouping coupled with curricular differences) and document a link between early school tracking and increased test-score dispersion - driven largely by lower test scores at the bottom of the test score distribution. While this is an important finding there is little known about what aspect of tracking, or what underlying mechanisms drive this effect. Furthermore, since most of the existing evidence on ability-grouping is based on comparing observationally similar students across systems with and without ability-grouping, there is the lingering concern that some of these effects reflect endogenous student self-selection on *unobservable* characteristics.

To deepen out understanding of ability grouping (and tracking), I used Trinidad and Tobago data, where there are no curricular differences across schools, to identify an ability-grouping effect on the margin. While there are several ways in which ability grouping could lead to increased inequality, I test one particular mechanism. Specifically, I test whether ability-grouping tends to increase educational inequality by assigning students with low achievement to schools that provide the least value-added while assigning high-achieving students to schools that produce the most value-added. I exploit the explicit rules used by the Ministry of Education to assign students to secondary schools to implement a rule-based instrumentation strategy to remove self-selection bias that could affect my findings. I present falsification tests indicating that the identification strategy is likely valid. After taking self-selection bias into account, I show that those schools with the highest-achieving peers generally produce more value-added on several outcomes than schools with lower-achieving peers. This suggests that ability-grouping may lead to increased educational inequality on a broad range of academic outcomes specifically: test scores, staying in school, the number of examinations passed, and obtaining the requisite skills to apply for tertiary education. Importantly, these findings are probably not driven by endogenous student self-selection to schools on unobservable characteristics.

I also show that the marginal effect of attending a school with higher achieving peers is non-linear so that the benefits to attending schools with marginally brighter peers are larger at the lower end of the peer achievement distribution. In fact, there is some evidence that students may have worse outcomes, on the margin, from attending schools with brighter peers at high levels of peer achievement. This suggests that one mechanism through which ability-grouping may increase educational inequality is by the creation of low-achievement schools that produce low value-added rather than the creation of "elite" schools that produce very high value-added. This is consistent with the other researchers finding that tracking reduces test scores more at the lower end of the test score distribution. Adding to a growing literature documenting stronger benefits to interventions for females than for males, I find that females benefit more from attending schools with high-achievement peers than do boys on all outcomes. One implication of this result is that ability-grouping could increase the male-female test score gap. Given the growing concern that boys may be falling behind, particularly in the Caribbean, further research is needed to better understand these gender differences.

Since I estimate the effect of ability-grouping on the margin (i.e. the effect of being assigned to a school with marginally brighter peers), these findings do not speak to the efficiency implications of ability-grouping. However, the results suggest that school ability-grouping could have sizable distributional effects. If low-income students are more likely to have low incoming test scores, they will be systematically placed in schools that provide less value-added than those attended by more affluent higher-achieving students. As such, school ability-grouping would reinforce socioeconomic differences between students making it more difficult for children from low-income families to avoid having low-income families themselves. In a developing nation such as Trinidad and Tobago, where large proportions of the population do not pursue higher education, ability-grouping (and school tracking) may be undesirable as it may relegate productive students to secondary schools where they are almost assured to not pursue tertiary education.

References

Angrist, J. "Conditioning on the Probability of Selection to Control Selection Bias", (1995) , NBER Technical Working Paper No. 181. Cambridge, Massachusetts.

Angrist, J., and K. Lang, "How Important are Classroom Peer Effects? Evidence from Boston's Metco Program." (2004), Cambridge: Mass: National Bureau of Economic Research Working Paper 9263.

Angrist. J. and V. Lavy, "Using Maimonides' Rule To Estimate The Effect Of Class Size On Scholastic Achievement," (1999). The Quarterly Journal of Economics, MIT Press, vol. 114(2), pp 533-575.

Ariga, K., G. Brunello, R. Iwahashi, and L. Rocco, "Why Is the Timing of School Tracking So Heterogeneous?" (November 2005). Institute for the Study of Labor Discussion Paper No. 1854.

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

Betts, Julian. "Does School Quality Matter? Evidence from the National Longitudinal Survey of Youth" (1995) *The Review of Economics and Statistics*, Vol. 77, No. 2, pp. 231-250.

Betts, J. R., and J. L. Shkolnik, "The effects of ability-grouping on student math achievement and resource allocation in secondary schools." (1999), *Economics of Education Review, 19*, 1–15.

Betts, J. R., and J. L. Shkolnik, "Key difficulties in identifying the effects of ability-grouping on student achievement" (1999). *Economics of Education Review*, *Volume 19, Issue 1*, *Pages 21-26.*

Brunello, G. and Chechi, D. "Does School Tracking Affect Equality of Opportunity?" New International Evidence." (2006), Institute for the Study of Labor Discussion Paper 2346.

Burke M. and Sass, T. "Classroom Peer Effects and Student Achievement," (2006) Working Papers wp2006_02_02, Department of Economics, Florida State University.

Campbell, Donald T. "Reforms and experiments." (1969), American Psychologist 24:409-29.

Card, D. and Krueger, A. "School Quality and Black-White Relative Earnings: A Direct Assessment." (February 1992a), *Quarterly Journal of Economics*.

Card, D. and Alan B. Krueger "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States" *The Journal of Political Economy*, Vol. 100, No. 1 (Feb., 1992b), pp.1-40.

Clark, D. "Elite Schools and Academic Performance" (2007) Institute for the Study of Labor (IZA) Discussion Paper 3182.

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

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

Das, Jishnu, Tahir Andrabi, and Asim I. Khwaja (2006) "Students Today, Teachers Tomorrow: Identifying Constraints for the Provision of Education." Processed.

Duflo, Esther, Pascaline Dupas, and Michael Kremer "Peer Effects and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya" Working Paper

Dustmann, C. "Parental Background, Secondary School track Choice and Wages." (2004), *Oxford Economic Papers*. 56. 209-230.

Figlio, D. N. and Page, M. E. "School Choice and the Distributional Effects of Achievement Tracking: Does Separation Increase Inequality?," (May 2002) *Journal of Urban Economics*, Elsevier, vol. 51(3), pages 497-514.

Fisher, Gary. 1976. "The Identification Problem in Econometrics." Robert E. Kreiger Publishing Company, Huntington, New York.

Galindo-Rueda, F and Vignoles, A, "The Heterogeneous Effect of Selection in Secondary Schools: Understanding the Changing Role of Achievement," (2004) Institute for the Study of Labor Discussion Paper 1245

Gould E., Lavy V. and Passerman D. "Immigrating to Opportunity: Estimating The Effect of School Quality Using A Natural Experiment on Ethiopians in Israel" (May 2004), *Quarterly Journal of Economics*.

Grogger, Jeff: "School Expenditures and Post-Schooling Earnings: Evidence from High School and Beyond" (Nov., 1996), *The Review of Economics and Statistics*, Vol. 78, No. 4, pp. 628-637.

Hastings, Justine S., Thomas Kane and Douglas Staiger (2006b), "Gender, Performance and Preferences: Do Girls and Boys Respond Differently to School Environment? Evidence from School Assignment by Randomized Lottery," *American Economic Review Papers and Proceedings,* 96 (2): 232-236.

Hastings, Justine S., Thomas J. Kane and Douglas O. Staiger"Gender and Performance: Evidence from School Assignment by Randomized Lottery", *American Economic Review* 96(2), pp. 232-236.

Hastings Justine S. and Weinstein, Jeffrey M. "No Child Left Behind: Estimating the Impact on Choices and Student Outcomes" (April 2007), National Bureau of Economic Research Working paper 13009.

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

Heckman, J., Layne-Farrar, A. and Todd, P. "Does Measured School Quality Really Matter? An Examination of the Earnings-Quality Relationship." (1996), In Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success. G. Burtless, editor. Washington D.C. Brookings Institution.

Horowitz, Joel L. & Manski, Charles F., 1998. "Censoring of outcomes and regressors due to survey nonresponse: Identification and estimation using weights and imputations," *Journal of Econometrics*, Elsevier, vol. 84(1), pages 37-58, May.

Hoxby, C. "The Effects of Class Size and Composition on Student Achievement: New Evidence from Natural Variation." (1998), National Bureau of Economic Research WP 6869.

Hoxby, C."Peer Effects in the Classroom: Learning from Gender and Race Variation." (2000), National Bureau of Economic Research Working Paper 7867.

Hoxby, Caroline M., and Weingarth, G., (2006). Taking Race Out of the Equation: School Reassignment and the Structure of Peer Effects. Working Paper.

Malamud, Ofer and Pop-Eleches, Cristian: "General Education versus Vocational Training: Evidence from an Economy in Transition," (revise and resubmit at the Review of Economics and Statistics)

Jackson. C. Kirabo (2008). "Student Demographics, Teacher Sorting and Teacher Quality: Evidence from the End of School Desegregation" under submission

Katz, L. and J. Kling and R. Liebman. (2007) . "Experimental Analysis of Neighborhood Effects," *Econometrica*, Econometric Society, vol. 75(1), pages 83-119, 01.

Katz, L. and J. Kling and R. Liebman. "Experimental Analysis of Neighborhood Effects" (August 2005). NBER Working Paper No. W11577.

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

OECD, 2004, the PISA , Paris.

Maurin, E., and S. McNally. "Educational Effects of Widening Access to the Academic Track: A Natural Experiment" (2007). CEE Discussion Papers with number 0085.

Rangvid, B. S. "Educational Peer Effects: Regression Evidence from Denmark with PISA2000 data." (2003), Copenhagen: AKF Institute for Local Government Studies.

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

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

Tyler, John H., Richard J. Murnane and John B. Willett. "Do The Cognitive Skills Of School Dropouts Matter In The Labor Market?," Journal of Human Resources, 2000, v35(4,Fall),

Zimmerman, D. "Peer Effects in Academic Outcomes: Evidence from a Natural Experiment." (2003), *Review of Economics and Statistics*, Volume 85, Issue 1 (November).

Zhang, H. (2008). "Elite Schools and Student Academic Achievement: Evidence from A Randomized Natural Experiment in China"