The Australian National University

Centre for Economic Policy Research

*DISCUSSION PAPER*

# Do Single-Sex Classes Affect Exam Scores? An Experiment in a Coeducational University

Alison L. Booth
The Australian National University
University of Essex

Lina Cardona-Sosa
University of Essex

Patrick Nolen
University of Essex

DISCUSSION PAPER NO. 679

February 2013

**ISSN: 1442-8636**
**ISBN: 978-1-921693-63-2**

# Do Single-Sex Classes Affect Exam Scores? An Experiment in a Coeducational University*

Alison L. Booth, Lina Cardona-Sosa, and Patrick Nolen
Australian National University and University of Essex

February 4, 2013

We examine the effect of single-sex classes on the pass rates, grades, and course choices of students in a coeducational university. We randomly assign students to all-female, all-male, and coed classes and, therefore, get around the selection issues present in other studies on single-sex education. We find that one hour a week of single-sex education benefits females: females are 7% more likely to pass their first year courses and score 10% higher in their required second year classes than their peers attending coeducational classes. We find no effect of single-sex education on the probability that a female will take technical classes and there is no effect of single-sex education for males. Furthermore we are able to examine potential mechanisms driving the single-sex effect for females. We find that the results are consistent with a reduction in stereotype threat for females and are not due to a potential tracking effect.

Keywords: single-sex, education, gender, experiment
JEL Classification: C91, C92, J16, J33

*Corresponding author: Patrick Nolen, Department of Economics, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, United Kingdom. Email: pjnolen@essex.ac.uk

# I. Introduction

Women are under-represented in the highest levels of mathematics, the physical sciences, and engineering (NAS 2006), despite the fact that more women now attend university than men.[1] Furthermore, there is a gender gap in standardized mathematics tests scores,[2] especially at the top of the distribution.[3] Single-sex education is sometimes viewed as a mechanism to decrease these gaps. This is because advocates of single-sex education claim that educating females in single-sex classes will increase their test scores, especially in mathematics and science, and that females taught in single-sex classes will then be more likely to take technical courses than their counterparts educated in mixed gender classes.

Policy makers, in seeking to decrease gender gaps in test scores and increase female representation in high-tech sectors of the workforce, have begun to listen to advocates of single-sex education and to allow expansion of publicly funded single-sex schools. According to the National Association for Single Sex Public Education (NASSPE), in 2002 there were only about a dozen US public schools offering single-sex classes; as of 2010 there were 540, of which 91 were all-girl or all-boy schools. In 2006, the US federal government eased regulatory restrictions on single-sex education that had been enforced under title IX, and now allows districts to create single-sex schools and have single-sex classes in publicly funded schools as long as enrollment is voluntary. In the UK, where single-sex education has been on the decline,[4] top education leaders have been calling for more government support for single-sex education (Lydall, 2012).

The expansion of single-sex education and calls for its support have occurred in spite of scientific evidence showing few, if any, benefits of single-sex education: a 2005 US Department of Education systematic review found 'minimal' evidence supporting single-sex education; Smithers and Robinson (2006) along with Tompson and Ungerleider (2004) argue that observed benefits of single-sex education are due to student selection into schools and Jackson (2012) shows that only females who have a strong preference to attend single-sex schools show any benefit from single-sex education; Halpmen et. al. (2011) go further stating

1. In the US in the 1960s there were 1.55 males for every female undergraduate but by 2003 there were 1.30 females for every male undergraduate (Goldin et. al. 2006).

2. A gender gap in standardized mathematics tests does exist but it varies by country (Else-Quest et. al. 2010, Guiso et. al 2008) and some argue it may not be large enough to be of any practical importance (Hyde et. al. 2008). Fryer and Levitt (2010), however, document a substantial gender gap in mathematics in the US; after six-years of education there is a 0.2 of a standard deviation gender gap in test scores. That gap is roughly half as large as the black-white test score gap.

3. What is agreed on is that, at the top of the distribution, there is a large and significant gap in mathematics test scores. For example, there is a 2:1 to 1 male-female ratio among students scoring 800 on the math SAT (Ellison and Swanson, 2010).

4. Despite the decline of single-sex education in the UK, over 60% of the top 50 secondary schools, based on GCSE results, were single-sex in 2011 according to the BBC league tables.

"there is no well-designed research showing that single-sex education improves students' academic performance" and argue that "there is evidence that sex segregation increases gender stereotyping and legitimizes institutional sexism." While there is a lack of evidence showing benefits from single-sex education, there is a small literature showing that the proportion of females in a classroom has an effect on educational and economic outcomes. Lavy and Schlosser (2011) show that, as the proportion of females in a classroom increases, the cognitive outcomes of both males and females improve. Schneeweis and Zweimüller (2012) show that, when a female studies in a class with a higher proportion of females, she is more likely to choose to study in a technical school later on. Individuals' attributes – such as willingness to behave competitively or to take risks – also affect outcomes; for example, math test scores are shown to be influenced by competitive behavior (Niederle and Vesterlund, 2010). Moreover, a growing recent literature explores how competitive and risk-taking attributes vary in response to cultural factors. Gneezy et. al. (2009) show that females in matrilineal societies are more likely to compete than males in that society, and compete as much as men in a patriarchal society. Booth and Nolen (2012a, 2012b) show that all-female environments make females more competitive and less risk averse.

When considering this literature alongside the research on single-sex education, the following questions arise: Is completely gender-segregated education taking things too far, perhaps increasing gender stereotyping and legitimizing institutional sexism? Might single-sex classes within a coeducational environment be a useful alternative way forward? Would single-sex classes in the technical subjects benefit women without adversely affecting men? We address the last two of these questions in the present paper. Specifically, we examine the effect of single-sex classes on test scores and course choices using a field experiment conducted at a high-ranked coeducational economics program in the UK. We randomly assigned first-year economics students into all-female, all-male, or coeducational classes. The students attend a coeducational university and, if assigned to a single-sex class, are only in that environment for one hour a week.[5] The randomization allows for straightforward identification of the effect of single-sex education on pass rates and course grades. This is an advantage over other recent work trying to identify the effects of single-sex education because our students attend the same school, get the same instruction, and did not choose to attend the school because there was a high probability they would be assigned to single-sex classes. Furthermore, we are able to follow students over time and can examine what courses they chose to take in the second year of study[6] and whether or not their grades in second

5. First-year students receive twelve hours of instruction per week. Eight hours are lectures by economics faculty members and four hours are additional classes taught by graduate teaching assistants. This field experiment deals with one of the classes taught by graduate teaching assistants.

6. In the first year nearly all economics students take the same set of courses. There is the possibility for

year courses are affected by the first year class allocation. Thus the experiment is able to examine two of the key claims made by single-sex advocates: does being in an all-female class cause females to do better? And does exposure to all-female classes cause females to take more technical courses? We are also able to look at whether there is a longer term effect of single-sex education by seeing how class allocation affects the grades of students more than a year after they stopped being taught in single-sex classes.

Our sample is different than those used in previous studies; most of the research to date has focused on students from primary and secondary schools. Our students spend less time in their single-sex classes – one hour per week – than students in most other studies. Furthermore, university students are older and arguably less likely to be influenced by their environment than students who are in primary or secondary school. Despite these factors, which could be seen as biasing us towards finding no effect, we find a considerable impact of single-sex schooling: females in all-female classes are over 7% more likely to pass their introductory economics course, score 8% higher on the course grade, and score 10% higher in their required courses a year after being assigned to a single-sex class than their counterparts assigned to coed classes. Moreover, this increase in female attainment occurs with no additional expenditure.

We are focusing on one aspect of single-sex education – its effect on pass rates, grades and course selection – which can influence wages and job prospects for students when they leave university. However other factors, such as socialization, are not being examined and could also have an effect on labor market outcomes. Therefore, while we find a strong positive effect on pass rates and second year grades for females, we do not wish to advocate for single-sex education based only on this study; more research is needed. For instance, we show that one hour of single-sex instruction is beneficial for a student who spends the rest of her time in a mixed gender environment; however would two hours still be beneficial? Or three hours? The research summarized by Halpern et. al. (2011) suggests that having all instruction done in a single-sex environment has no positive effect or could even be harmful. Therefore, more research is necessary to establish the optimal amount of single-sex education that is beneficial for students, how the effects might vary across subject areas, and the best setting for that education (classes only, laboratories only, or lectures and classes).

We designed our study to see if single-sex education could have an effect within an overall coeducational environment, and also to look at two potential mechanisms through which single-sex education might have an effect. The first mechanism is related to tracking, as discussed in Duflo et al. (2011), and the second is related to the psychological effects of stereotype threat, as in Steele et. al. (2002). Tracking by ability allows teachers to be

a student to substitute one option depending on the student's background or degree scheme.

more effective because it reduces the variance in ability present in the classroom. If the variance in ability differs by gender, then segregating classes by gender could cause females in single-sex classes to do better. More precisely, if the variance of the distribution of ability for females is smaller than the variance of the distribution for males, then segregating by gender means that females in single-sex classes are all closer to the median.[7] The teacher can then teach to the median student rather than spending time teaching to different parts of the distribution. When looking at stereotype threat, one would also expect a benefit from single-sex environments. If students are exposed to the stereotype that females are worse at economics than males, then females may fear speaking up in classes in order to avoid embarrassment or being perceived as different. Furthermore, females may face added anxiety when taking tests because they have been conditioned to believe they are worse at economics. Therefore, a female enrolled in an all-female class should be more willing to participate in discussion and feel less pressure when taking exams. Consequently, she should do better in courses given the absence of the negative stereotype caused by having males present.

The next section lays out the conjectures that our field experiment was designed to test. The third section discusses the experimental design, context, and subjects involved. The fourth section presents the results and the fifth section concludes.

# II. Conjectures

This section lays out two sets of conjectures. The first deals with the effects of single-sex education on pass rates, course selection, and second year course grades. The second subsection deals with conjectures about the mechanisms driving the effects.

## *II.A. Pass Rate, Course Selection, and Exam Marks*

Systematic reviews of the literature on single-sex education have shown that observed benefits of single-sex environments are primarily due to selection into schools.[8] However, the economics of education literature has shown that there is an effect of having more females

7. As shown by Ellison and Swanson (2010), at the top of the distribution, there is a large and significant gap in mathematics test scores. For example, there is a 2.1 to 1 male-female ratio among students scoring 800 on the math. Therefore, if men are more likely to be in the top tail of the distribution, single-sex classes could reduce the variance in ability for females.

8. See US Department of Education (2005), Smithers and Robinson (2006), and Tompson and Ungerleider (2004). Doris et. al. (2012), when looking at single-sex schooling in Ireland, show that there is no effect of single-sex education. Jackson (2012) shows that most students do not benefit from single-sex education and that females at single-sex schools take fewer science courses.

in a class on cognitive outcomes and student choices regarding which type of classes to take. Therefore, to examine the effect of having a fully single-sex class, we designed a field experiment, a randomized control trial, to test the following conjectures.

Conjecture 1. Females in single-sex classes are more likely to pass their introductory economics course than females in coeducational classes.

Supporters of single-sex education argue that females may face pressure to conform to gender stereotypes in order to fit in socially. For instance, a female may be exposed to the stereotype that females are worse at economics than males. If this is the case, then, a female who does well in an economics course, may be ostracized or teased by her peers. Psychologists have shown that the framing of tasks and cultural stereotypes does affect the performance of individuals (see Steele et al. 2002 and the references there in). Therefore, by educating females and males separately, females will not face pressure to conform to negative stereotypes and instead will perform better on tests and thus be more likely to pass the courses in which they are enrolled. We focus on pass rates and not on the final mark because of the weaker incentives first-year students face.[9]

Conjecture 2. Females who studied in single-sex classes during their first-year are more likely to take the technical courses required for a BSc during their second year of study than females who studied in coeducational classes during their first year.

If females enrolled in single-sex classes do better and feel more comfortable studying economics, then they may be more likely to take more technical, mathematically intensive courses in their second year of study. This is an argument often made by single-sex advocates. While not looking at fully gender-segregated classes, Schneeweis and Zweimüller (2012) examine the college choices of students in Austria based on the type of high school they attended. Using variation in the proportion of females in adjacent cohorts within schools, they find that females studying with a higher proportion of females are more likely to enter a technical college when they continue their education. Therefore, educating females in single-sex class may cause them to choose to enroll in technical course in their second year. We focus on choices in the second year because nearly all students take the exact same set of courses in their first year.[10] However, in the second year, they can chose to take more

9. UK degrees are awarded an overall classification – a first class degree, an upper-second class degree (2.1), a lower-second class degree (2.2), or a third class degree – and the marks received in the first year of study usually do not affect the degree classification. In our setting, students only have the incentive to get a mark of 40 or higher. This will be discussed further in the results section.

10. All students in economics have to take the same three year-long courses in the first year. The fourth course will depend on the degree scheme the student enrolled in when coming to the university. Conditional on the degree scheme, students have no options in the first year unless they seek special permission to be exempt from requirements.

mathematical courses from a set of available options.

Conjecture 3. Females who were in single-sex classes during their first year earn higher grades in their required second year courses than females who were in coeducational classes.

Females who were randomly assigned to single-sex classes during their first-year may gain long term benefits from only briefly being educated separately from males. Cohen et. al. (2009) show that even subtle interventions to lessen the psychological threat of being negatively stereotyped can have long lasting benefits. Cohen et. al. (2006) found that African-Americans who completed a series of writing assignments designed to provide self-affirmation immediately earned higher GPAs and that two years later (in the follow-up survey Cohen et. al., 2009) they found that those students earned higher marks as well. European-American students did not benefit from the intervention immediately or in the two year follow-up. If females in single-sex classes do better because stereotype threat is being reduced, then they may earn higher marks in the second year because they continue to benefit from the first-year intervention.

## *II.B. Mechanism Conjectures*

To understand how single-sex education affected female outcomes we will look at two general hypotheses. The first is based on tracking literature and the second is based on stereotype threat.

Conjecture 4. The distribution of ability in single-sex classes does not differ from the distribution of ability in coeducational classes.

The literature on gender gaps in test scores shows that males are more likely to be in the top part of the distribution than females. Therefore, when females are segregated from males, the distribution of ability in all-female classes may be smaller than in coeducational classes; there may be a form of tracking taking place. Recent work by Duflo et. al. (2011) shows that tracking by ability can benefit both high and low-performing students. In examining the effects of tracking, Duflo et. al (2011) lay out a model that is useful to consider the mechanisms by which tracking could benefit students. In our situation, graduate teaching assistants (GTAs) are incentivized to have students do well (e.g. get good marks) or they will be removed from their teaching position.[11] If the GTA is teaching at a level too far away

11. Classes are taught by Graduate Teaching Assistant (GTAs) from the economics department, who are PhD students hired in a competitive hiring process from the pool of PhD applicants. If first-year students

from a student's ability, the student will not benefit. Thus, if there is a large distribution of ability, the GTA has to teach to each part of the distribution to allow students to learn and minimize complaints. If the distribution of ability is small, the GTA can teach to the median level and all students will benefit. Given this framework, if single-sex classes have a lower distribution in ability than coeducational classes, females could benefit due to being in a class with a decreased variance in ability and not because of a psychological factor such as a reduction in stereotype threat. Before any instruction took place we had students take an IQ test and will use that as a measure of ability to examine if there is evidence supporting conjecture four.

Conjecture 5. Females of lower ability benefited more from being assigned to single-sex classes than females of higher ability.

Stereotype threat only emerges when problems are difficult (see Spencer et. al (1999), Steele (1997), Steele et al. (2002)). The argument for this is that the frustration of doing a difficult task is heightened by the added anxiety of stereotype threat; therefore making a difficult task even harder to successfully complete than if stereotype threat was not present. Thus a student who is struggling to do well on a test will suffer more from the presence of stereotype threat than a student who finds the exam relatively easy. Students with a higher IQ will find the economics exam easier and thus score a higher grade. If this is correct, then females with a relatively lower IQ will find the economics exam relatively more difficult and, consequently, benefit more from being assigned to an all-female class. Given this we predict that females with a relatively lower IQ are likely to benefit more from being assigned to an all-female class.[12] To see if the data provide support for this hypothesis, we will examine if there is a differential effect of single-sex classes on females based on IQ scores.

## III. Experimental Design

Our experiment was designed to test the conjectures listed above. Given the issues of selection raised when the conclusions of other studies were examined,[13] we used random

file a legitimate complaint about their Graduate Teaching Assistant (GTA), or if the professor taking the lectures sees problems when reviewing marks from the first term test, a GTA can be relieved of their teaching duties.

12. Cohen et. al. (2009) found this type of effect when looking at African-American students. They found lower performing African-American students benefited more than higher performing African-American students when the psychological threat of being negatively stereotyped was reduced. They used pre-intervention performance to determine high and low ability students, though, and we will use IQ scores.

13. Recent work has aimed to get at the effect of single-sex education but has been forced to rely on comparing students attending different types of schools or looking at the effect within a school where ex ante

assignment to class type in order to identify the effect of the educational environment. Below we discuss the randomization, the experimental set-up and look at the descriptive statistics. Other studies have aimed to look at the effect of single-sex education but have been forced to do so by comparing the effect.

## *III.A. Subjects and Educational Environment*

The students in our experiment arrived at the University of Essex in October 2010 and they all took a year-long introductory economics course, either EC111 or EC100. EC111 is primarily for economics students and EC100 is primarily for student in the business school. The introductory economics course was one of the four required year-long courses that students had to take in their first year. Both introductory courses run over 20 weeks and have the same structure: each week a professor from the economics department gives a two hour lecture and a graduate teaching assistant (GTA) gives a one hour class. The lecture takes place in a large auditorium designed to hold over 1000 students. The one hour class is taught by a GTA in small classroom that can hold no more than 30 students.[14]

During their first year of study students receive 240 hours of instruction, 60 hours per course. The focus of our experiment will be the one hour of class time per week taught by GTAs in the introductory economics course. That amounts to 20 hours of instruction, or 8.3% of the total instruction received by a student in the first year.

## *III.B. Experimental Design*

Students were randomly assigned to a single-sex or coeducational class. Class assignments were made randomly by the timetabling office before students arrived on campus. Once

the student knew she was very likely to attend single-sex classes. For instance Park et. al. (2012a, 2012b) try to use a natural experiment regarding the allocation of students into high schools in Korea to get at the effects of single-sex education but the single-sex and coed schools differ in how they are funded and teacher allocation making it hard to attribute the effect solely to the gender make-up of the schools. Eisenkopf et. al. (2011) look at students who chose to attend a school that prepares them to be teachers. Since teaching is a primary female occupation roughly 85% of their sample is female. The minimal male presence means when students are assigned to classes some classes end up being all-female but the students knew this ex ante and their school choice is related to the probability of being in an all-female environment. The ex ante selection is likely to bias the results because, as shown by Jackson (2012), the few females that benefit from single-sex education are those with a preference to attend single-sex schools.

14. For this experiment, GTAs were told not to discuss with students any of the details – or objectives – of the experiments. (We control for the gender of GTAs in our regressions.) In classes the instructors discuss with students problem-sets that relate directly to the material taught in that week's Introduction to Economics lectures. The Professor of the course tells the GTAs what material should be covered in each class. The problem sets are the same across all classes in EC100 and in EC111 (though they differ between EC100 and EC111).

assigned to a class, students were not allowed to change their class and attendance at the assigned class was enforced.[15] The procedures regarding class-assignment and attendance are the same used at Essex each year. There was no change in the way students were assigned to classes in any of the other year-long courses and courses were scheduled by timetabling so that there was no conflict between class times and any lecturers. Furthermore, the assignment to classes in other courses was independent of the introductory economics class assignment.

There were 20 classes in EC100 and 17 classes in EC111. Of the 20 classes in EC100, 4 were all-female, 7 were all-male, and 9 were coed. Of the 17 classes in EC111, 3 were all-female, 4 were all-male, and 10 were coed. That means – at the class level – we have 7 all-female classes, 11 all-male classes, and 19 coed classes. As in other years our sample was, roughly, 35% female and 65% male. We chose to create this number of all-male and all-female classes because it kept the gender distribution roughly what it would have been without the experiment taking place; i.e. in the 19 coed classes each class was, on average, 30% female and 70% male.

During the first class meeting, GTAs had students take an IQ test, fill out a demographic questionnaire, and participate in a risk and competition experiment. The IQ test was a modified 20-minute version of Raven's Matrices appropriate for university-aged students.

The grade for introductory economics is based on assignments, tests, and an exam. At the end of the first term, students have a take-home assignment that they have to do on their own. They will also take a one-hour test in a lecture hall while being supervised. At the end of the second term, students will again do a take-home assignment and take a one-hour test. The assignments and tests are marked by the class teacher who can see the name of the student.[16] Generally no curve is forced on the assignment or test marks; the course professor gives a detailed outline of what marks should be awarded based on potential answers. However, the professor responsible for the course does look at marks to make sure there is no discrepancy across GTAs. During the summer term (the third and final term for the academic year), students take a 2 hour exam. As is the standard procedure in the UK, that exam is double-blind marked by two members of the economics department and neither marker knows the name or gender of the student. No curve is forced on final exam marks.

A student's overall grade in the course is based on a max-rule. The scores from assignments and tests are averaged. If the average is above the exam mark, the student's final

15. The university is required to take attendance so that it can provide evidence that international students who are in the UK on student visas are actually attending classes. Indeed the visa requirements for most international students includes the provision that they must attend lectures and classes.

16. As shown in Lavy (2008) when teachers are aware of the gender of the student prejudices can occur in marking that could lead to gender differences in test scores that are not there when objective measures are used.

mark will be 50% of the coursework (assignments and tests) mark plus 50% of the exam mark. If the coursework average is below the exam score, then the student's final mark will be the exam mark. Students know this rule from the beginning – it is explained to them during the first lecture – and all courses in the economics department are graded in the same way. Given that coursework may not count towards the final score a number of students choose not to do it. However, all students *must* take the exam.

Degrees in the UK are classified into one of four categories: a first class degree, which is the best degree classification; an upper second (i.e. 2.1) which is the second best classification; a lower second degree (i.e. 2.2.) which is the third best; or a third which is the lowest classification. At Essex University (and at many other British universities), the degree classification is based on the scores students get in their second and third years only. All a student has to do is 'pass' all courses in the first year, i.e. get a score of 40 or higher in all four courses, in order to continue into the second year of study. Thus, when looking at the effect of single-sex education in the first year we will focus on the margin that will matter the most for students, whether they 'passed' their course.

In the second year of study students in the economics department have to take 5 full-year courses (one of the courses can be composed of two half-year courses). Two courses are required, microeconomics, EC201, and macroeconomics, EC202, and the rest can be chosen from a list of courses, though, there is less flexibility for some students because of degree type. All economics students are offered two mathematically intensive courses: EC251, Mathematical Methods in Economics; and EC252, Introduction to Econometric Methods. Students who take both of these courses and pass can graduate with a BSc rather than a BA. Students in the business school who were required to take EC100 in their first year also have to take 5 full-year courses in their second year. All of them must take BE111, management accounting. If a student in the business school wants to graduate with a BSc the student can choose from a set of more technical courses depending on the type of degree. Therefore, when looking at the effect of single-sex classes on the choice of a student to take technical courses, we will focus on whether the student chose to take the requirements for the BSc. When looking at grades in the second year, we will look at EC201 and BE111, because they are both required and taught in the autumn term (the first term of the academic year).[17]

17. Even though the courses are taught in the autumn term, the exam for the courses is given during the summer term (the third and final term of the academic year). The exams are mandatory and double-blind marked.

## *III.C. Descriptive Statistics*

To examine the effect of single-sex classes on pass rates we will look at the sample of all students who took introductory economics. To look at the effect of single-sex classes on the probability of taking the technical courses required for a BSc in the second year or second year marks we will focus on students enrolled in the economics department or the business school.[18] Since class type was randomly assigned we expect that the observables should not differ by class type. Table 1 shows the summary statistics by class type.

[Insert Table 1 Here]

The first seven variables in table one are outcome variables that could have been affected by assigning students to single-sex classes. The summary statistics show that, for a female, being assigned to an all-female class is associated with a higher probability of passing the first year course, with a higher overall exam score in the course, and with a higher average coursework mark. In the second year, females who had been assigned to all-female classes in their first year, scored higher on their required course despite no longer being in all-female classes for over 12 months. None of these effects are present for males. However, there is no effect on the probability of males or females enrolling in the BSc program due to being assigned to a same-gender class.

These summary statistics provide initial evidence for conjectures one and three; assignment to all-female classes is associated with females being more likely to pass their introductory course and that the effects of being assigned to an all-female class lasts for at least a year. However, based on the summary statistics, there is no evidence for conjecture two, that girls assigned to all-female classes are more likely to enroll in a BSc program. We will examine these initial effects more closely in the regression analysis below.

The last 14 variables presented in Table 1 allow us to examine how well the randomization worked. For the majority of the variables, there is no difference between the treatment (single-sex) and control (coed) groups. Boys in all-male classes have a marginally lower IQ score than those assigned to coed classes and there are slightly more black students in all-male classes. Females in all-female classes are marginally more likely to be native English speakers, have more female siblings, and are less likely to have a father with a university degree. As with all randomizations, some differences are likely to occur. To see if these differences matter, we will control for observed differences in the regression analysis.

18. Students enrolled in joint degree schemes such as government with economics or sociology with economics have to take the introductory economics course. However those degrees are not overseen by the economics department or the business school and, consequently, the students do have the same required courses in their second year. Therefore, we do not focus on these students when looking at second year course choices or grades.

The two major differences between coed and single-sex classes that can be seen in the descriptive statistics are that females in the all-female group are younger than their classmates in coed classes, and that class sizes in both all-female and all-male classes are larger than the coed classes. The age difference occurs because the few students, less than 4% of the female sample, over the age of 21 were randomly assigned to coed classes. To examine if the age distribution had a difference we will control for age in the regression analysis and see if it affects the estimate of our variable of interest. The class size difference occurred because of the assignment to classes and could have an effect on the learning of students in classes.

For introductory economics, the number of classes is typically based on predicted enrollment and classes were capped at 20 students in EC111 and 22 in EC100. To minimize the number of classes needed, as many classes as possible are filled to the maximum number of students and then, if needed, more are created. In the year of our experiment, two more coed classes in EC111 and EC100 were created and each had fewer than the maximum number of students. When those additional classes are excluded from the sample, the enrollment between coed and single-sex classes at the beginning of the year is the same. However, over the year, some students drop out or switch to different degree schemes (e.g. switch from being an economics student to a government student). A higher proportion of students from coed courses switched courses or dropped out. This means that, by the end of the year, there were fewer students on average in coed classes than in single-sex classes. Based on the education literature (Angrist and Lavy, 1999; Hoxby, 2000; and Krueger and Whitmore 2001), we know that smaller class sizes benefit students. Therefore, if anything, this difference in class size will drive us to not find an effect of single-sex classes. Furthermore, we can also examine this empirically by controlling for class size differences and seeing the effect it has on our estimate of interest.

# IV. Results

In this section we present the results. We examine if there is any evidence to support the conjectures on the main effects – pass rates, second year marks, and technical course choices – or the mechanism conjectures laid out on the second section.

## *IV.A. Pass Rate, Course Selection, and Exam Marks*

Table 2 shows the regression results that allow us to examine whether there is evidence supporting conjectures one, two, or three.

[Insert Table 2 here]

Columns [1]-[4] in table two show that there is strong evidence for conjecture one, that females assigned to all-female classes are more likely to pass the introductory economics course than females assigned to coed classes. Column [1] shows that girls in all-female classes are 6.7 percentage points more likely to pass their introductory course than females assigned to coed classes. The pass rate for the introductory course is 89% so the increase of 6.7 percentage points represents a 7.5% increase. That is a large effect given that no additional resources were needed. When we control for age and class size the estimated effect doesn't change, as seen in column [2]. This suggests that the randomization worked because inclusion of these variables does not affect the estimate.[19] Furthermore, controlling for all observed differences from table one does not change the point estimate, as seen in column [3].[20] Again this suggests that it is unlikely that the results are being driven by unexplained heterogeneity. Since this experiment takes place at the class level, we also run class-level regressions as shown in column [4]. Column [4] shows that the pass rate of all-female classes were 6.3 percentage points higher than coed and all-male classes. The consistency of the estimated effect of being in an all-female class across all four specifications shows that the randomization likely worked well and that the estimate is getting at the causal effect of being assigned to an all-female class.

Our dependent variable of interest is whether or not a student passed the introductory economics course. We chose this because the marks a student receives in the first year do not affect the type of degree a student earns upon graduating. In the first year of study at Essex, a student need only get a 40% or higher in all courses in order to continue in her studies. However, if we look at the impact of being assigned to an all-female class on the average grade received by a student, we see that there is a positive effect. In column [5], we see that females in all-female classes score 4.5 points higher overall than their female counterparts in coed classes. But if we look at the 15th quantile, we see the estimated effect is more than double and much more significant.[21] This shows that being assigned to an all-female class is having a large effect for those students most at risk of not passing. In fact, at the 15th

19. The class sizes in EC100 are, by design, larger than those in EC111. Thus we have to interact class size with the EC111 dummy variable since, because of the design, class size and the EC111 dummy will be correlated.

20. We control for whether a student is a native English speaker (=1), black (=1), the number of sisters, and if the student's father had a university degree (=1). We initially do not control for raw IQ score, and there are two reasons: first, we do not have IQ scores for all 570 students, and in our initial estimates we want to work with as large a sample size as possible. Second, IQ relates to one of the mechanism conjectures that we investigate later in the paper.

21. We chose the 15% quantile because, on average over the years, 15% of students get a mark below 40 in the introductory economics course.

quantile, being assigned to an all-female class increases the score of a female by over 22%.

Columns [7]-[9] of Table 2 allow us to examine if there is any evidence in support of conjecture two: that females who studied in single-sex classes during their first-year are more likely to take the technical course required for a BSc during their second year of study than females who studied in coeducational classes. In the economics department and business school, students have to take a certain block of courses in their first year. In their second year, students take different courses based on their degree scheme. Students pick a provisional degree scheme upon entering the University but can change it at any time as long as they have the required courses for the scheme. One major choice a student faces in the second year is whether to take technical courses required to earn a BSc instead of a BA degree.[22] We therefore examine if being assigned to an all-female class causes a female to be more likely to take the BSc required courses.[23] As can be seen in column [7], there is no evidence supporting the second conjecture. When all controls are added in column [8], or when the class level effect is examined in column [9], there is still no significant effect. There are a few reasons we may not find an effect. In our case, students choose a provisional degree program when entering the university. Therefore the students would actually have to change their degree scheme if they were move from a BA degree to a BSc degree. If students view changing as sufficiently painful, or have a preference for the status-quo, then this could explain the lack of observed effect. In Schneeweis and Zweimüller (2012) these confounding effects do not exist.

Next we examine how long the effect of being in an all-female class lasts. To do this, we look at whether there is any evidence to support conjecture three, that females who were in single-sex classes during their first year earn higher marks in the second year than females who were in coeducational classes. In the second year, students in both the business school and economics department have to take core courses. We examine the effect of first year class assignment on second year grades in these required core courses. Column [10] shows that females who studied in all-female classes do, on average, 5 points better in their second year than females who studied in coed classes. That represents a 10% increase in the average second year grade. When controls for classes and age are added, as in column [11], the result becomes more significant but the point estimate stays roughly the same. When all initial differences are control for, as in column [12], the results do not go away. This consistency of

22. An economics student would have to take two EC251, Mathematical Methods in Economics and EC252, Introduction to Econometric Methods, to earn a BSc and a Business student would have to take additional accounting courses.

23. In economics students have to take EC201, intermediate macroeconomics, and in the business school students have to take BE111, management accounting. Both courses are taught in the autumn term but the exams take place in the summer term.

the point estimate shows the robustness of the result. Finally when examining the effect in the second year at the first year class level we see the result is still present (see column [13]).

## *IV.B. Mechanism Conjectures*

We have found a large, significant effect of educating females in all-female classes for one hour a week while they are attending a coeducational university. This suggests that some amount of single-sex education, at least for females, could be beneficial. However, the literature, as summarized by Halpmen et. al. (2011) shows that there could be negative effects from fully segregating students by gender. Therefore, we want to understand the mechanism driving our estimated effects, and to consider if another type of educational policy could replicate these effects.

By changing the gender composition of classes, student peer-groups could be affected. One potential effect is that the average ability of a student's classmates could be changed due to the creation of single-sex classes. Since other studies have shown that females are less likely to be high performers in mathematically-based subjects, it could be the case that females in all-female classes may have lower ability classmates than females in coed classes. Duflo et. al. (2011) show that tracking students by ability can cause students to benefit. The mechanism that they outline is that, as the variance in ability in a class decreases, a teacher can teach more effectively. Thus, if separating classes by gender causes the variance in ability in classes to decrease, students may be benefiting owing to a tracking effect and not to single-sex instruction. To test if a tracking effect is driving the result, we see if our evidence supports conjecture four, that the distribution of ability in single-sex classes does not differ from the distribution of ability in coeducational classes.

To do this, we look at the distribution of ability. Before taking the course, we had students complete an IQ test. If we look at the distribution of ability in figure one – as measured by the raw score on the IQ test – we see that there is little difference between the ability of females and males. Both follow a normal distribution and we cannot reject the null-hypothesis that the two distributions (male and female ability) are the same using the Kolmogrov-Smirnov test (p-value = 0.29).

[Insert Figure 1 Here]

Furthermore, we calculated the standard deviation of ability by class. The standard deviation in the raw IQ score for coed classes was 3.00 points, for all-female classes 3.21, and for all-male classes 2.86. Using a t-test for differences in means, one cannot reject the hypothesis that the standard deviation for coed and all-female classes is the same (p-value

= 0.38). Therefore, there is no evidence that single-sex education is causing the deviation in ability to change at the class level. Therefore we cannot reject conjecture four. Our conclusion is that we have no evidence that the effects we found from single-sex education can be attributed to the same mechanism that causes tracking to benefit students.

Another potential reason that females in all-female classes may do well is because of reduction in psychological threats caused by studying with males. For instance if females are stereotyped to be bad at economics then they face not only the anxiety of doing well in the course but also the anxiety of being different and having to overcome a stereotype. Of course, if a female is low ability and she would not face the prospect of breaking the stereotype and, consequentially, she would be less likely to benefit from single-sex classes because she did not face stereotype threat to begin with. Therefore, to examine if there is evidence that the reduction in stereotype threat may be the reason we observe large benefits of single-sex education we examine if there is evidence in support of conjecture five, that females of higher ability benefited more from being assigned to single-sex classes than females of lower ability.

[Insert Table 3 here]

Column [1] in Table 3 shows the main regression specification for our reduced sample. Since not all students took the IQ test, our sample drops from 375 to 313 observations. When standard errors are clustered at the class level, our coefficient on being in an all-female group is now insignificant at the 5 percent level. However, the point estimate is roughly the same as in columns [10] – [12] of Table 2 suggesting that the estimated effect is equivalent and issues of sample selection are not likely to be a concern. Column [2] in Table 3 shows the effect of IQ and how the effect of being in an all-female group differs by one's IQ z-score. As suspected, students with higher IQs do better; a student in the top 5% of the distribution will score 6 points better than an average student. This means that students with lower IQs are likely to face more difficulty when doing the coursework and answering questions on the exam. As predicted by stereotype threat, the effect of being in an all-female class is more beneficial for low IQ females: for females in the bottom part of the distribution the effect of being in an all-female class is larger; females with an IQ z-score of 1.96 (those in the top 5% of the distribution) will score 2 points higher in their mandatory second year course due to being in an all-female class, while those with an IQ z-score of -1.96 (those in the bottom 5% of the distribution) will score 8.13 points higher due to being in an all-female class.[24]

24. We arrive at 8.11 for the top 5% of the distribution because 1.96*1.474+5.22=8.11. The same calculations are done for the bottom 5% of the sample, i.e. -1.96*1.474+5.22=2.33.

This provides some support for conjecture five.[25] To examine if there is other evidence that psychological threats are decreased by being in an all-female group we will look at one other measure of 'comfort': attendance at class.

Column [3] shows that students in all-female classes were 5.9 percentage points more likely to attend classes – that is a 10% increase in the base rate of attendance, 58.1%. This suggests that students were more likely to attend class if they were in an all-female group; this does not hold for all-male groups. One may suspect that this is because GTAs taught all-female groups differently however, as shown in columns [4] and [5] this does not appear to be the case. Each term students are allowed to provide feedback on GTA teaching; they score GTAs on a scale from 1 to 5 with 5 representing that the GTA was of high quality. Of the 37 classes, GTAs in all-female classes did not score any higher than GTAs teaching all-male or coed classes. Since no GTA taught only all-female or all-male groups, we can even use GTA fixed effects to see if, within GTAs, students felt GTAs taught their all-female classes better than the coed classes. As can be seen in column [5] this is not the case. Therefore, being assigned to an all-female class causes a female to attend more classes and it does not appear to be because the GTA is teaching any differently.

Column [6] shows that attendance in classes has a large positive effect on whether a student passes their first year course; a student who attends all classes is 28 percentage points more likely to pass. Could it be that if we design a mechanism that causes females to attend class we can mimic the effect of all-female classes, i.e. is the all-female effect due to the change in attendance? To consider this, let us take the point estimate on attendance in column [6] seriously even though attendance is endogenous and, consequently the estimate is likely to biased upwards. Being in an all-female class increases attendance by 5.9 percentage points, which will translate into a 0.281*0.059=0.016 percentage increase in the pass rate for females in all-female groups. The 1.6 percentage point increase is less than one-third of the 6.7 percentage point effect on the pass rate from being in an all-female class estimated in column [1] of Table 2. When looking at the effect of being in an all-female class on one's second year score, the change in a student's attendance during the first year potentially explains even less of the effect. A student who attends all classes in the first year will score 14.9 points higher in the second year course, as shown in column [8] of Table 3. The change in attendance caused by being assigned to an all-female class can only explain a 0.059*14.9=0.87 point increase in the second year score. That is not even 20% of the 5 point

25. We do not look at the effect of All-Female interacted with IQ z-scores for pass rates because the effect is only at the one part of the distribution; around the cut-off of 40. In order to test if there is a differential effect of the treatment we need there to be an incentive to do better at each part of the distribution, so that there is the potential for the treatment to effect all parts of the distribution differently. This is the case with grades in the second year since they count towards the overall degree classification a student receives.

increase caused by being assigned to an all-female class that is estimated in column [10] of Table 2. In fact, the estimate on being in an all-female class in column [8] of Table 3 is practically the same as the estimate in the preferred specification in column [10] of Table 2. That means that even controlling for attendance does not take away the effect of being assigned to an all-female class on one's second year grades.

Finally, to also examine if the main mechanism driving the all-female class is attendance, we instrument for attendance using the time of the class interacted with the day of the week. Students are much more likely to attend a class if it is held later in the day. Furthermore, classes earlier in the week – Monday, Tuesday, or Wednesday – have higher attendance than those held on Thursdays or Fridays. Therefore, we predict attendance at a class based on day of the week dummy variables interacted with time.[26] Column [7] shows that IV linear probability model (LPM) results for the effect of attendance on the pass rate. There is no estimated effect of attendance on passing the one's first year course and the estimated effect of being in an all-female class, while not significant, is roughly the same as the estimate in the main specification shown in column [1] of table two. Column [9] shows the IV estimates for the effects of attendance on the aggregate grade for the required second year course. The point estimate is double that of the one in column [8] but the standard error is three times as large. The point estimate and significance of the effect of being in an all-female class does not disappear and is still roughly what is was in the main specification shown in column [10] of Table 2.

The results in Table 3 show that there is evidence that all-female classes have a direct effect on pass rates and course scores as predicted by the reduction of stereotype threat; females with lower IQ z-scores have higher estimated benefits from being in all-female classes, females are more likely to attend classes despite the lack of evidence that GTAs are adjusting their teaching, and the main effect of being assigned to an all-female class cannot be attributed solely to the change in attendance.

Given that the evidence suggests that it is the reduction of stereotype threat causing females to do better, one could explore what policies or interventions – other than teaching in gender-segregated classes – could reduce stereotype threat and cause females to do better. For instance, writing self-affirming essays or framing tasks differently have been examined by psychologists (see Cohen et. al. (2006) or Steele et. al. (2002) and the references therein for examples of interventions) and found to reduce stereotype threat. However, these interventions take up class time and have an added cost whereas our experiment was able to increase the first year pass rate of females by 7.5% and the average score of females by 10%

26. Classes are taught on Tuesdays, Wednesdays, Thursdays, and Fridays and at 09:00, 10:00, 11:00, 12:00, 14:00, 15:00, 16:00, and 17:00 each day. First stage regression results are shown in the appendix.

in the second year with no added expenditure.

# V. Conclusions

Advocates of single-sex education have been pushing to expand single-sex education in the hopes of addressing existing gender gaps in test scores and choices to study technical degrees. However, there has been mixed evidence on the effect of single-sex education and a lack of "well-designed research showing that single-sex education improves students' academic performance."[27]

To examine the effects of single-sex education, we randomly assign students to all-female, all-male, and coed classes. The instruction students receive does not vary with class type and students do not choose to attend the course or university because they would like to be educated in a single-sex environment. Therefore, the issues of selection highlighted by critiques of previous studies are not present.

We examine if single-sex education in a coeducational environment has an effect on female test scores, pass rates, and choice of what courses to study. We find that females assigned to all-female classes earn higher grades in their classes immediately and that the effect is present for at least one year. However, we find no effect on the likelihood that a female will chose to study a technical subject.

The positive benefits of single-sex education is only present for females and seems to be driven by the reduction of stereotype threat. Females in all female courses are more likely to attend classes and females who are more likely to do poorly, those with lower IQs, benefit more from being assigned to an all-female class. Therefore, other mechanisms that also aim to reduce the effect of stereotype threat to females studying economics in a coeducational institution may be able to produce the same positive effect as single-sex education for females.

27. Quote from Halpmen et. al. (2011)

# References


Angrist, J., Lavy, V. (1999) "Using Maimonides' Rule to Estimate the effect of class size on scholastic achievement" *Quarterly Journal of Economics*, vol. 112, no. 2, pp. 533-575

Booth, A.L., and Nolen, P.J. (2012a) "Choosing to compete: How different are girls and boys?" *Journal of Economic Behavior & Organization*, vol. 81, pp. 542-555

Booth, A.L., and Nolen, P.J. (2012b) "Gender Differences in Rick Behaviour: Does Nurture Mattter?," *The Economic Journal*, vol. 122, pp. F56-F78

Cohen, G., Garcia, J., Apfel, N., Master, A. (2006) "Reducing the Racial Achievement Gap: A Social-Psychological Intervention" *Science*, vol. 313, pp. 1307-1310

Cohen, G., Garcia, J., Purdie-Vaughns, V., Apfel, N., Brzustoski, P. (2009) "Recursive Processes in Self-Affirmation: Intervening to Close the Minority Achievement Gap" *Science*, vol. 324, pp. 400-403

Doris, A., O'Neill, D.,Sweetman, O. (2012) "Gender, Single0Sex Schooling and Maths Achievement" IZA Working paper #6917

Duflo, E., Dupas, P., Kremer, M. (2011) "Peer Effects and the Impacts of Tracking: Evidence from a Randomized Evaluation in Kenya" *American Economic Review*, vol. 101, no. 5, pp. 1739-1774

Eisenkopf, G., Hessami, Z., Fischbacher, U., Ursprung, H. (2011) "Academic Performance and Single-Sex Schooling: Evidence from a Natural Experiment in Switzerland" University of Konstanz Working Paper #2011-35

Ellison, G., Ashley Swanson (2010) "The Gender Gap in Secondary School Mathematics at High Achievement Levels: Evidence from the American Mathematics Competitions" *Journal of Economic Perspectives*, vol. 24, no. 2, Spring, pp. 109-128

Else-Quest, Nicole M., Janet Hyde, Marcia Linn (2010) "Cross-National Patterns of Gender Differences in Mathematics: A Meta-Analysis" *Psychological Bulletin*, vol. 136, no. 1, pp. 103-127

Fryer, R., Steven Levitt (2010) "An Empirical Analysis of the Gender Gap in Mathematics" *American Economic Journal: Applied Economics*, Vol. 2, No. 2, April, pp. 210-240

Gneezy, U., Leonard, K., and List, J. (2009) "Gender differences in competition: evidence from a matrilineal and a patriarchal society," *Econometrica*, vol. 77, pp. 1637–64

Goldin, Claudia, Lawrence F. Katz, Ilyana Kuziemko (2006) "The Homecoming of American College Women: The Reversal of the College Gender Gap" *Journal of Economic Perspectives*, vol. 20, no. 4, pp. 133-156

Guiso, Luigi, Ferdinando Monte, Paolo Sapienza, Luigi Zingales (2008) "Culture, Gender, and Math" Science, vol. 320, May 30

Halpmen, D., Eliot, L., Bigler, R., Fabes, R., Hanish, L., Hyde, J., Liben, L., Martin, C. (2011) "The Psuedoscience of Single-Sex Schooling," *Science*, vol. 333, 23 September, pp. 1706-1707

Hoxby, C. (2000) "The Effects of Class size on student achievement: New evidence from population Variation" *Quarterly Journal of Economics*, vol. 115, no. 4, pp. 1239-1285

Hyde, J., Sara Lindberg, Marcia Linn, Amy Ellis, Caroline Williams (2008) "Gender Similarities Characterize Math Performance" *Science*, vol. 321, July 25

Jackson, C. K. (2012) "Single-sex schools, student achievement, and course selection: Evidence from rule-based student assignments in Trinidad and Tobago" *Journal of Public Economics*, vol. 96, pp. 173-187

Krueger, A., Whitmore, D. (2001) "Th effect of attending a small class in the early grades on college-test taking and middle school test results: evidence from project STAR" *Economic Journal*, vol. 111, no. 468, pp 1-28

Lavy, V. (2008) "Do gender stereotypes reduce girls' or boys' human capital outcomes? Evidence from a natural experiment" *Journal of Public Economics*, vol. 92, pp. 2083-2105.

Lavy, V., and Schlosser, A. (2011) "Mechanisms and impacts of gender peer effects at school", *American Economic Journal: Applied Economics*, vol. 3, pp. 1–33

Lydall, Ross (01 May 2012) "St. Paul's head calls for more all-girl schools" The Evening Standard

National Academy of Sciences (2006) "Beyond Bias and Barriers: Fulfilling the Potential of Women in Academic Science and Engineering" National Academies Press: Washington, D.C.

Niederle, M., Lise Vesterlund (2010) "Explaining the Gender Gap in Math Test Scores: The Role of Competition" *Journal of Economic Perspectives*, Vol. 24, No. 2, pp. 129-144

Park, H., Behrman, J., Choi, J. (2012a) "Causal Effects of Single-Sex Schools on College Entrance Exams and College Attendance: Random Assignment in Seoul High Schools" PSC Working Paper #1-25-2012 – forthcoming in Demography

Park, H., Behrman, J., Choi, J. (2012b) "Do Single-Sex Schools Enhance Students' STEM (Science, Technology, Engineering, and Mathematics) Outcomes?" PIER Working Paper 12-038

Schneeweis, N., and Zweimüller, M. "Girls, girls, girls: Gender composition and female school choice" *Economics of Education Review*, Vol. 31, vol. 4, pp. 482-500

Smithers., A., and Robinson, P. (2006) "The Paradox of Single-Sex and Co-Educational Schooling" University of Buckingham: CEER publication.

Spencer, S.J., Steele, C.M., Quinn, D.M. (1999) "Stereotype threat and women's math performance" *Journal of Experimental Social Psychology*, vol. 35, pp. 4-28

Steele, C.M. (1997) "A threat in the air: How stereotypes shape intellectual identity and performance" *American Psychologist*, vol. 52, pp. 613-629

Steele, C., Spencer, S., Aronson, J. (2002) "Contending with group image: The psychology of stereotype and social identity threat" *Advances in Experimental Social Psychology* vol. 34, pp. 379-440

Thompson, T., and Ungerleider, C. (2004) "Single-Sex Schooling: Final Report" Canadian Centre for Knowledge Mobilisation: Waterloo, Ontario.

**Table 1: Descriptive Statistics**

| | *Females* | | | | *Males* | | | |
|---|---|---|---|---|---|---|---|---|
| VARIABLES | Coed | All-Female | Difference | S.E of Diff | Coed | All-Male | Difference | S.E of Diff |
| Student Passed First-Year Course (=1) | 0.853 | 0.932 | 0.079* | [0.047] | 0.865 | 0.903 | 0.038 | [0.032] |
| Score for Introductory Economics | 53.085 | 57.541 | 4.456* | [2.448] | 51.998 | 53.445 | 1.447 | [1.739] |
| Score for Introductory Economics Exam | 50.126 | 54.874 | 4.748* | [2.663] | 49.139 | 50.785 | 1.646 | [1.881] |
| Avg Coursework Score for Introductory Economics | 52.715 | 57.142 | 4.427* | [2.527] | 52.143 | 50.698 | -1.445 | [1.759] |
| Score for Second Year Course | 50.961 | 55.71 | 4.749* | [2.557] | 55.520 | 54.277 | -1.243 | [1.821] |
| Student Enrolled in BSC degree (=1) | 0.626 | 0.525 | -0.101 | [0.081] | 0.615 | 0.538 | -0.077 | [0.063] |
| Percent of Introductory Economics Classes Attended | 0.604 | 0.662 | 0.058* | [0.032] | 0.592 | 0.6 | 0.008 | [0.026] |
| Age | 19.671 | 18.858 | -0.813*** | [0.186] | 19.612 | 19.435 | -0.177 | [0.195] |
| IQ Raw Score | 11.736 | 11.233 | -0.503 | [0.491] | 12.174 | 11.486 | -0.688* | [0.379] |
| Class Size at End of Year | 16.250 | 18.541 | 2.291*** | [0.315] | 15.873 | 16.799 | 0.926*** | [0.245] |
| Native Langauge English (=1) | 0.279 | 0.406 | 0.127* | [0.067] | 0.386 | 0.403 | 0.017 | [0.056] |
| Number of Siblings | 1.359 | 1.503 | 0.144 | [0.190] | 1.896 | 1.635 | -0.261 | [0.172] |
| Number of Sisters | 0.706 | 0.994 | 0.288* | [0.153] | 1.085 | 0.977 | -0.108 | [0.124] |
| Student is Eldest Child (=1) | 0.525 | 0.595 | 0.070 | [0.072] | 0.421 | 0.413 | -0.008 | [0.053] |
| White (=1) | 0.483 | 0.504 | 0.021 | [0.076] | 0.553 | 0.46 | -0.093 | [0.057] |
| Black (=1) | 0.308 | 0.237 | -0.071 | [0.064] | 0.207 | 0.306 | 0.099** | [0.048] |
| Mother has University Degree (=1) | 0.430 | 0.309 | -0.121 | [0.075] | 0.413 | 0.372 | -0.041 | [0.056] |
| Father has University Degree (=1) | 0.527 | 0.395 | -0.132* | [0.075] | 0.488 | 0.505 | 0.017 | [0.057] |
| Born in UK (=1) | 0.230 | 0.333 | 0.103 | [0.063] | 0.354 | 0.345 | -0.009 | [0.052] |
| Born in EU (=1) | 0.326 | 0.269 | -0.057 | [0.068] | 0.247 | 0.241 | -0.006 | [0.045] |
| Born in China (=1) | 0.234 | 0.177 | -0.057 | [0.053] | 0.161 | 0.218 | 0.057 | [0.038] |

*When looking at outcomes from the first-year there are 570 students in the sample of which 202 are female,123 are in all-female classes, and 173 are in all-male classes. When looking at the BSC degree choice there are 400 students in the sample of which 145 are female, 95 are in all-female classes, and 118 are in all-male classes. When looking at second year score outcomes the sample consists of 375 studnets of which 140 are female, 95 are in all-female classes, and 113 are in all-male classes. *** p<0.01, ** p<0.05, * p<0.1*

**Table 2: Effect on Pass Rates, BSc Enrollment, and Second Year Marks**

| | Dependent Variable (=1) if Student passed Course | | | | Aggregate Grade for Course | | Dependent Variable (=1) if Enrolled in BSc | | | Aggregate Grade for Required Second Year Course | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| VARIABLES | [1] | [2] | [3] | [4] | [5] | [6] | [7] | [8] | [9] | [10] | [11] | [12] | [13] |
| Female (=1) | -0.019 | -0.027 | -0.042 | | 0.817 | -5.000 | 0.127* | 0.068 | | -2.615 | -3.289 | -3.775 | |
| | [0.040] | [0.037] | [0.038] | | [2.536] | [3.651] | [0.074] | [0.079] | | [2.732] | [2.607] | [2.700] | |
| All-Female (=1) | 0.067** | 0.064*** | 0.063*** | 0.063** | 4.447* | 9.000** | -0.077 | -0.054 | 0.040 | 5.011* | 5.438** | 5.982** | 4.539* |
| | [0.031] | [0.025] | [0.021] | [0.027] | [2.587] | [3.946] | [0.090] | [0.100] | [0.086] | [2.835] | [2.106] | [2.528] | [2.343] |
| All-Male (=1) | 0.034 | 0.043 | 0.035 | 0.050 | 1.411 | 3.000 | -0.056 | -0.068 | -0.089 | -0.924 | -1.106 | -0.621 | 1.307 |
| | [0.029] | [0.026] | [0.024] | [0.031] | [1.478] | [2.777] | [0.062] | [0.085] | [0.053] | [1.880] | [1.884] | [2.232] | [2.101] |
| EC111 (=1) | 0.005 | -0.068* | -0.035 | 0.009 | 2.863** | -0.000 | 0.025 | 0.026 | 0.007 | 5.000*** | 5.058*** | 1.146 | 3.353 |
| | [0.025] | [0.041] | [0.049] | [0.026] | [1.362] | [2.270] | [0.057] | [0.050] | [0.053] | [1.480] | [1.409] | [1.226] | [2.152] |
| Age Controls | No | Yes | Yes | No | No | No | No | Yes | No | No | Yes | Yes | No |
| Class Size x EC111 Controls | No | Yes | Yes | No | No | No | No | Yes | No | No | Yes | Yes | No |
| Observable Differences | No | No | Yes | No | No | No | No | Yes | No | No | No | Yes | No |
| Regression Type | Probit | Probit | Probit | OLS | OLS | 15th Quantile | Probit | Probit | OLS | OLS | OLS | OLS | OLS |
| Constant | | | | 0.859*** | 52.111*** | 40.000*** | | | 0.591*** | 54.615*** | 55.321*** | 61.859*** | 53.154*** |
| | | | | [0.023] | [1.226] | [2.180] | | | [0.049] | [1.671] | [2.085] | [2.234] | [1.834] |
| Observations | 570 | 570 | 570 | 37 | 570 | 570 | 400 | 400 | 37 | 375 | 375 | 375 | 37 |
| R-squared | | | | 0.116 | 0.021 | | | | 0.097 | 0.046 | 0.107 | 0.129 | 0.119 |

*Standard errors are clustered at the class level and shown in brackets. Marignal Effects are reported for the probit regressions. The controls for observable differences are the set of controls in table one for which there were significant differences: whether someone was a native English speaker (=1); number of sisters; whether a father had a university degree (=1); and wether a student was black (=1). *** p<0.01, ** p<0.05, * p<0.1*

Figure 1

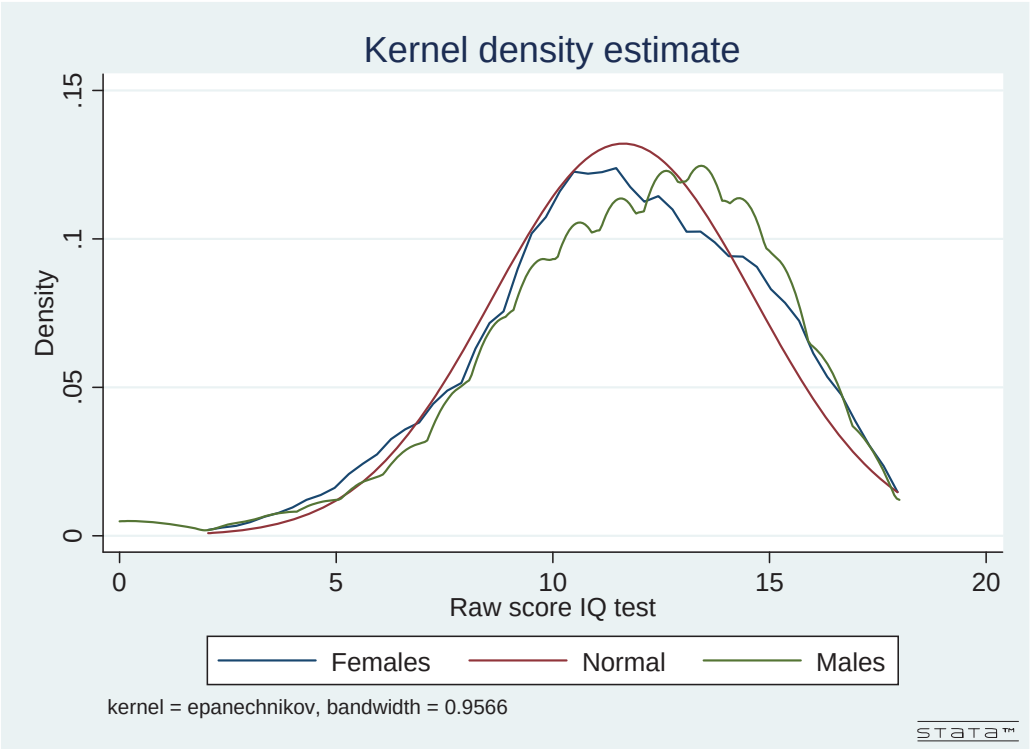
Kernel density estimate
Density
0
.05
.1
.15
0
5
10
15
20
Raw score IQ test
Females
Normal
Males
kernel = epanechnikov, bandwidth = 0.9566
Stata™

**Table 3: Mechanisms**

| VARIABLES | Aggregate Grade for Required Second Year Course [1] | [2] | Attendance (0-1) [3] | GTA Score [4] | [5] | Dependent Variable (=1) if Student passed Course [6] | [7] | Aggregate Grade for Required Second Year Course [8] | [9] |
|---|---|---|---|---|---|---|---|---|---|
| Female (=1) | -2.913 | -2.402 | 0.039 | | | -0.033 | -0.02 | -3.287 | -3.93 |
| | [3.090] | [2.877] | [0.029] | | | [0.041] | [0.047] | [2.534] | [2.565] |
| Z-Score IQ | | 3.025** | | | | | | | |
| | | [1.476] | | | | | | | |
| All-Female (=1) | 5.058 | 5.111 | 0.059** | 0.053 | -0.138 | 0.045 | 0.08 | 4.656* | 4.32* |
| | [3.289] | [3.168] | [0.028] | [0.182] | [0.138] | [0.032] | [0.050] | [2.490] | [2.245] |
| All-Female * Z-Score IQ | | -1.545 | | | | | | | |
| | | [2.169] | | | | | | | |
| All-Male (=1) | -2.584 | -2.207 | 0.012 | 0.035 | -0.043 | 0.035 | 0.04 | -1.094 | -1.26 |
| | [2.026] | [2.054] | [0.033] | [0.116] | [0.139] | [0.026] | [0.030] | [1.716] | [1.662] |
| All-Male * Z-Score IQ | | -3.028 | | | | | | | |
| | | [1.860] | | | | | | | |
| EC111 (=1) | 5.295*** | 5.082*** | -0.009 | 0.180 | 0.452 | 0.005 | 0.00 | 4.958*** | 4.92*** |
| | [1.557] | [1.562] | [0.023] | [0.110] | [0.516] | [0.023] | [0.025] | [1.315] | [1.229] |
| Percent of Classes Attended (0-1) | | | | | | 0.281*** | -0.04 | 14.868*** | 29.06** |
| | | | | | | [0.048] | [0.312] | [3.408] | [12.774] |
| GTA Fixed Effects | No | No | No | No | Yes | No | No | No | No |
| Regression type | OLS | OLS | OLS | OLS | OLS | Probit | IV LPM | OLS | IV |
| Constant | 55.522*** | 55.543*** | 0.581*** | 4.127*** | 4.286*** | | 0.89*** | 45.436*** | 36.67*** |
| | [2.019] | [2.029] | [0.023] | [0.081] | [0.559] | | [0.190] | [3.201] | [8.666] |
| Observations | 313 | 313 | 570 | 37 | 37 | 570 | 570 | 375 | 375 |
| R-squared | 0.061 | 0.063 | 0.025 | 0.074 | 0.602 | | -0.007 | 0.101 | 0.051 |
| F-Stat First Stage Regression | | | | | | | 14.62 | | 10.43 |

*Standard errors are clustered at the class level and shown in brackets. Marignal Effects are reported for the probit regressions. *** p<0.01, ** p<0.05, * p<0.1*