Do Single-Sex Classes Affect Academic Achievement?
An Experiment in a Coeducational University

Alison L. Booth, Lina Cardona-Sosa, and Patrick Nolen[1]

Revised August 2018

Abstract

We examine the effect of single-sex classes on the educational attainment of students within a coeducational university. Before students arrived on campus, we randomly assigned them to all-female, all-male, and coed classes, and thereby avoid the selection issues present in earlier studies on single-sex education of students in primary and secondary school. We find that one hour a week of single-sex classes benefits women: females score a quarter of standard deviation better overall and are 7.7% more likely to pass their first year course. Furthermore, women assigned to all-females classes in their first year are roughly 57% less likely to drop out of university and are 61% more likely to get a top ranked degree under the UK system. There is evidence that single-sex classes cause women to adopt behaviors associated with better academic outcomes, such as attending more classes and doing optional assignments. However, these behavioral changes cannot explain much of the all-female effect.

JEL Classification: C90, C93, I23, J16

Keywords: field experiment, single-sex classes, coeducation, higher education, gender, student achievement

[1] Corresponding author: Patrick Nolen, Department of Economics, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, United Kingdom. Email: pjnolen@essex.ac.uk. The experiment reported in the paper was conducted at the University of Essex, whose Ethics Committee provided approval. An earlier version of this paper appeared as an IZA Discussion Paper No 7207.

1. Introduction

More women than men attend university, yet women are under-represented in technical fields like mathematics, the physical sciences, and engineering (National Academy of Sciences, 2006). They are also under-represented in economics, the discipline on which we focus in this paper.

Some universities and academics have tried to reduce gender gaps in economics by targeting females who study these subjects. For example, Harvard Business School has implemented policies such as special training for women aimed at decreasing gender disparities that develop after women begin their studies, and the American Economic Association (AEA) has introduced a mentoring and network program for female economists. While both approaches have been shown to increase female performance and retention in their areas, policy makers have not adopted programs of this type at a national level. More recently, a new initiative to increase the number of women in economics has been launched from Harvard University: The Undergraduate Women in Economics Challenge.[2] In our paper we report an additional approach to those that have been put forward thus far: the use of single-sex classes within a coeducational environment as a potential means of improving female outcomes. Thus, we contribute an important piece of evidence to inform this initiative, and to the broader issue of gender in the classroom.

Single-sex education can be provided at either the school-level (in which case all classes within the school are single-sex), or at the class-level (in which case certain classes within a school are single-sex but the overall school environment is coeducational). The treatment discussed in our paper is of the latter type – some single-sex classes within a coed university. However most of the studies within the literature on single-sex education fall into the former type – single-sex classes within a single-sex school, and the vast majority of these are directed at primary and secondary education.

US policy-makers have, over the past decade or so, allowed the expansion of single-sex education in primary and secondary schools, with the idea that the effects on females will carry on into later life. According to the National Association for Single Sex Public Education (NASSPE),

[2] See http://scholar.harvard.edu/goldin/UWE for further information.

in 2002 there were only about a dozen US public schools offering single-sex classes. In 2010 there were 540, of which 91 were all-girl or all-boy schools (in other words, schools in which all classes were single-sex).[3] This expansion in single-sex education, however, has occurred despite the lack of much conclusive evidence showing benefits of single-sex education – be it in classes or schools. Indeed, a 2005 US Department of Education systematic review found 'minimal' evidence supporting single-sex education; Smithers and Robinson (2006) argue that observed benefits of single-sex education are due to student selection into schools; Halpmen et al. (2011) state "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."

However, despite the criticisms of single-sex education, there are a number of potential psychological reasons for expecting a positive effect for young women in single-sex classes. For example, females in all-female classes may feel more confident and gain higher levels of self-efficacy (Gist and Mitchell (1992)); they may experience a reduction in stereotype threat (as discussed in Spencer et al. (1999), Steele (1997), or Steele et al. (2002)) or a reduction in psychological threat (as discussed in Cohen et al (2006)).

Recently there have been some important new studies aiming to estimate the effect of single-sex education on various educational and economic outcomes. These typically rely on comparing students attending different types of primary or secondary schools, or looking at the effect within a school where *ex ante* the student knew she was very likely to attend single-sex classes. For instance, Park, Behrman and Choi (2013) compare students in coed high schools that are primarily publicly-funded with students in gender-segregated schools that are primarily privately-funded. Eisenkopf, Hessami, Fischbacher, and Ursprung (2014) examine students who chose to attend a school where 85% of the population is female, meaning it is very likely that many students will be in single-sex classes. Booth and Nolen (2012a; 2012b) look at publicly-

[3] This represents a small proportion (less than 1%) of the nearly 100,000 public schools in the US at that time (see https://nces.ed.gov/fastfacts/display.asp?id=84). In the private sector there are proportionately 10 times more fully single-gendered schools (10% according to Morrison (2014) or almost 10% using weighted averages presented in Long and Conger (2013)). That the proportion is higher in the private sector – and is growing (Morrison (2014)) - might be viewed as indicating a demand for single-sex schools. In the UK, of independent private schools, 27% are single-sex (see ISC (2014))

funded selective single-sex schools which they compare with the academic stream of publicly-funded coeducational schools. In these studies, the *ex ante* choice of a primarily-female school may bias the results, as argued by Jackson (2012), who shows that females who benefit from single-sex education are those with a preference to attend single-sex schools. It was for this reason that we chose in the present study to assign students randomly to single-sex and coed classes in a field experiment, and to follow them over time.

While as noted there are few studies estimating the effects of single-sex education – be it in classes or in schools - there is a related literature showing that the *proportion* of females in a classroom or local environment has an effect on the educational and economic outcomes of students at primary and secondary level. For example, Lavy and Schlosser (2011) find that, as the proportion of females in a classroom increases, the cognitive outcomes of both males and females improve. Other studies look either directly or indirectly at the subsequent choices of technical majors for students in secondary schools and find that these choices are affected by the proportion of females in a classroom.[4] That is not the focus of our present paper; instead, we investigate the educational outcomes of university students who have already chosen a course of study.

In view of this literature and the policy-importance of the topic, we designed a field experiment to examine the effects of single-sex classes on the performance of first-year university students in a highly-ranked and publicly-funded economics department. There are several novel features of our approach. First, we *randomly* assigned students into a single-sex environment, and hence selection is not an issue. As noted above, few other studies have estimated the effects of single-sex schooling with random assignment. Specifically, we randomly assigned first-year economics students into all-female, all-male, or coeducational classes, and we then examined the effect on first-year scores and pass rates, and on scores in subsequent years of the degree program.

[4] For example, Schneeweis and Zweimüller (2012) show that a female studying in a class with a higher proportion of females is more likely to choose to study in a technical school later on. Annelli and Peri (2013) find that a woman in a high school class with a larger share of female peers will have a higher wage; this is primarily because the higher proportion of same-sex peers increases the likelihood a woman will choose majors associated with high earning jobs such as economics, business, medicine, or engineering.

The second novel feature of our approach is that we implemented this random assignment to single-sex classes in the economics faculty of a *coeducational university* – the University of Essex.[5] This is in contrast to earlier studies that have primarily focused on the effects of fully sex-segregated primary and secondary schools. Our results are therefore of particular relevance to the many coeducational colleges and universities interested in reducing existing gender inequalities in higher education.

The remainder of the paper is set out as follows. The next section lays out the experimental design, context, and subjects involved, while the third section reports descriptive information about the data and the predetermined variables. The fourth section presents and discusses the main results. The fifth section presents the behavioral changes and discusses the overall findings. The final section concludes.

2. Experiment Design

Below we discuss the educational environment, the randomization, and the student evaluation.

2.1 *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 students 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. Introductory courses run over 20 weeks and have the same structure: each week a senior professor from the economics department gives a two-hour lecture and a graduate teaching assistant (TA) gives a one-hour class. The lecture takes place in a large auditorium while the one-hour class is taught by a TA in small classroom that can hold no more than 30 students.

We emphasized, when training the TAs, that they were not to discuss with students any of the details – or objectives – of the experiment. In classes the TAs discuss problem-sets with

[5] While the important study of Oosterbeek and van Ewijk (2014) looks at varying the proportion female in university students' coeducational study groups, they did not consider fully sex-segregated groups, which are the focus of our interest.

students; these problem-sets relate directly to the material taught in that week's *Introduction to Economics* lectures, and are designed by the professor in charge of that course. The professor also tells the TAs what material should be covered in each class.[6]

During their first year of study students receive 240 hours of instruction, 60 hours per course. The focus of our experiment was the one-hour of class time per week taught by TAs in the introductory economics course and its impact on educational outcomes. That amounts to 20 hours of instruction, or 8.3% of the total instruction received by a student in the first year.

2.2 *Random assignment to classes*

We begin this subsection by providing a brief background to the institutional framework behind UK undergraduate admissions, before describing in detail University of Essex admissions and how the randomization of students to classes was achieved.

All students (both home and international) wishing to study for an undergraduate degree in the UK must apply by recording their interest - well in advance of the start of the academic year - with the Universities and Colleges Admissions Service (UCAS). This is a UK-based central organization, whose main role is processing applications for entry into UK universities. At the same time, secondary school students are also required to rank up to five preferred universities. This is done before the final secondary school exams.

Students who apply to Essex University are given a university-specific registration number as soon as their application is complete. Students subsequently confirm whether Essex is their top choice or a backup. Registration numbers for students who do not pick Essex as a top choice are then reused for other applicants.

Since each applying student can – and usually does – list up to five schools, there is little certainty that a student who has listed Essex University as one of her preferred places to study will actually end up there. Indeed, fewer than 80% of students who initially register an interest in Essex end up attending (because of the one in five aspect). Therefore, many of the earlier

[6] The problem sets are the same across all classes in EC100 and in EC111 (though they differ between EC100 and EC111).

Essex registration numbers are often reassigned as students get conditional offers elsewhere. Given the substantial reuse and reordering of registration numbers, there is considerable randomness associated with their final arrangement. Some low numbers will belong to early applicants, while others will be late applicants who did not get in to the university of their choice.

We now describe how students were allocated to classes. Before incoming first-year students arrived on campus, the timetabling-office randomly assigned them to a single-gender or coeducational class for an introductory economics course. The randomization was conducted based on the number of classes, the unique identifying registration number given to each student as described above, and gender.

Students were ordered by registration number and then assigned to classes one by one until all classes in a course were filled. For example, in EC111 (where there were three all-female classes, four all-male, and 10 coed), the student with the lowest registration number was assigned to the first single-sex class (either all-female or all-male depending on the student's gender), the student with the second-lowest registration number was then assigned to the next appropriate single-sex class, and so on until there was one student in each of the three all-female classes or in each of the four all-male classes. Then the next student, with the gender where the respective single-sex classes were filled, was placed into the first coed class. Once all single-sex classes had one student, the next students with the lowest registration numbers were placed in the remaining coed classes until all classes had one student. The process then started again and continued until each class had two students; then three students, and so on. The procedure continued until all students were assigned to a class. Thus lower registration numbers are equally as likely in single-sex and coed classes.[7]

[7] Note that this method is akin to a matched pair design and is similar to the method used in Miguel and Kremer (2004:165), who ordered schools alphabetically and assigned every third school to a project group. This 'ordered list' method - used in our approach as well as in Miguel and Kremer (2004) - is in principle inferior to a 'random list' method. For example, we might have instead flipped a coin each assignment round to see if we would fill the single-sex or coed classrooms first, or alternatively we might have randomly ordered how the classrooms are filled within each assignment round. However, because student registration number is so weakly related to our outcome, we chose to allocate by student registration number as described in the text.

As in other years, our cohort was, roughly, 35% female and 65% male. We assigned these students into a total of 37 classes: of these, 20 were 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.[8] We chose to create this number of all-male and all-female classes because it kept the gender distribution in coed classes at roughly what it would have been without the experiment; i.e. in the 19 coed classes, each class was, on average, 30% female and 70% male.

Once assigned to a class, students were not allowed to change their class and attendance at the assigned class was enforced.[9] The procedures regarding class-assignment and attendance are the same used 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 lectures. Furthermore, the assignment to classes in other courses was independent of the introductory economics class assignment. At no stage were the students told the purpose of the class assignment nor did we have any enquiries. All students enrolled in the course are supposed to attend the classes and do the compulsory exercises. Lectures and classes began immediately after student arrival at campus. During the first course meeting, students took an IQ test, filled out a demographic questionnaire, and participated in a risk and competition experiment. The IQ test was a modified 20-minute version of Raven's Matrices appropriate for university-aged students and different languages. Our sample consists of 570 first year students.

2.3 Student Grades

It is important to understand the grading details, as these affect students' incentives to exert effort during the year. The grade for introductory economics is based on assignments, tests, and an end-of-year written examination. The academic year comprises three terms. At the end of

[8] Given that treatment was at the class level, we cluster standard errors at the class level and also look at class level regressions in the results section.

[9] 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 include the provision that they must attend lectures and classes.

the first term, students are given a take-home assignment that they are required to do on their own. They are also given a one-hour test in a lecture hall while being supervised. At the end of the second term, students again do a take-home assignment and take a one-hour test. Each term, the assignments and tests are marked by the class teacher, who can see the name of the student.[10] 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 are no discrepancies across TAs. During the summer term (the third and final term for the academic year), students are not presented with any new material but are invited to several revision lectures and - towards the middle of this term - take a mandatory two-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. Furthermore, an external examiner – a senior professor from another UK university – afterwards reviews the exams to ensure they are of a particular standard and that the marks are appropriate. No curve is forced on final exam marks.[11]

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 mark will be based on 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 manner.[12] 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.

[10] 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.

[11] One potential outcome measure might be the double-blind exam scores, as at first blush these appear to be completely "clean" measures of achievement. However, students' effort in studying for the exam would be influenced by what they already know about their homework grades. Students with better homework grades might work harder for the exams because they have learnt they have a talent for that subject, or they might slack off for the same reason. Therefore it is important to focus on all components of the students' grades and their overall score.

[12] Note that the max-rule is applied by the economics department automatically, and students do not themselves make any choice about its application *ex post*. However they might choose to do no coursework; indeed at least 15% of our sample did not do all the coursework.

Degrees in the UK are classified into one of four categories: a first-class degree, which is the best degree classification; an upper second (also referred to as a 2.1), which is the second-best classification; a lower second degree (also referred to as a 2.2), which is the third-best; or a third-class degree, 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. In order to continue into the second year of study, all a student has to do is 'pass' all courses in the first year, i.e. get a score of 40 percent or higher in all four courses. Thus, when looking at the effect of single-sex classes in the first year, we will focus not only on scores but also on the margin that will matter the most for students, whether or not they passed their course.

If a student does not pass the first year (where passing is defined as getting 40 percent or more in *all* courses), the student nearly always has to take the *entire* year again (unless there are extenuating circumstances such as illness). A non-trivial number of students fail their first year (nearly 20% in our sample).

2.4 Allocation of TAs

Lectures for the courses were given on Monday and Tuesday. Then classes were taught from after the last lecture on Tuesday until Friday, so that the material from lecture that week was discussed in the classes that week. Classes were only held for part of the day Wednesday because the afternoon is reserved for sports matches and club meetings. Therefore, the majority of classes are on Thursday and Friday. TAs were PhD students who are assigned to classes by the director of graduate studies. Each TA usually teaches three classes and the director aims to assign that person classes all on the same day. Therefore, once the TAs were assigned, we made sure that each TA taught both coed and single-sex classes. That means that, roughly, there were the same number of single-sex classes and coed classes on each day. We tried to balance classes between morning and afternoon sessions but with only 37 classes that was not always possible. However, the regression results show that the time of day and day of the week of classes do not predict the treatment status of a class (either single-sex or coed), and therefore it suggests that our assignment mechanism worked.

3. Descriptive Statistics

In this section we report descriptive statistics for the predetermined characteristics, followed by density figures showing differences between the treatment and control groups across men and women combined, as well as disaggregated by gender. We report these densities across several measures of academic performance in introductory economics. We use the sample of all students taking the introductory economics course.

Since class type was randomly assigned, we expect that predetermined characteristics should not differ by class type. Summary statistics and differences in means between individuals attending single-sex and coed classes are reported in Table 1. Panel A shows the nineteen variables at the individual level while Panel B reports those variables at the class-level. At the bottom of each panel we present the relevant p-values for the joint test of whether the predetermined variables predict assignment to single-sex classes at either the individual or class level. Under the table we report the number of cases for each panel, plus some further information including how p-values were calculated.

[Insert Table 1 here]

Our particular interest in Table 1 is in checking the balance of pre-determined characteristics across coed and single-sex classes. As with all randomizations, some differences are likely to occur. Perusal of the table reveals that the only variables whose difference across class-types is statistically significant are age and number of sisters for women, and for men ethnicity. From Panel B, it can be seen that there are no statistically significant differences in the four class-level variables across class-type for either gender. At neither the individual or class level, can the variables, jointly, explain assignment to a single-sex class for males or females.

Our main results can be seen descriptively in Figure 1, which displays treatment and control group densities for exam grades, coursework grades, and overall grades (which can include for any individual *either* only the exam mark, *or* a 50:50 mix of exam and coursework, depending on which is higher). The solid line is for men and women in single-sex classes, while the dashed line is for men and women in coed classes. The horizontal axis ranges from a possible

score of zero percent to a maximum of 100 percent (never awarded in this type of course). The vertical dashed line at 40 percent shows the pass mark, while the second vertical dashed line gives the coed mean grade. (Note that in the male figures the coed mean is for men in a coed class, while in the female figures the coed mean is for women in a coed class).

[Insert Figure 1 here]

First, consider the distributions of scores on the exam, which was mandatory and double-blind graded. For women in all-female classes, the distribution lies to the right of the more dispersed distribution for women in coed classes and the distributions are significantly different (Mann-Whitney p-value = 0.068). This is not the case for men. Notice, for men in all-male classes, the distribution has a slightly lower variance than for men in coed classes, though, they are not significantly different (Mann-Whitney p-value = 0.648). Next we turn to coursework grades, given in the middle of Figure 1. For men, there is little difference in the distributions of coursework grades for the treatment and control groups. This contrasts with the situation for women: here the distribution for women in the treatment group is more right-skewed than it is for the control group.

Finally, we consider the overall end-of-year grades for men and women presented at the bottom of Figure 1. Again we find that there is no significant difference in the distributions for men in coed and single-sex classes (Mann-Whitney p-value = 0.543), while the distribution for women in all-female classes is significantly different from that for women in coed classes (Mann-Whitney p-value = 0.069). Furthermore, we see that, for women in coed classes in particular, the distribution is fatter in the left tail than it is for women in single-sex classes. Thus inspection of Figure 1 reveals that women in the control group (coed) are more likely to be found performing poorly than are both women in single-sex classes and all men. [13]

The distributions for women in Figure 1 suggest that there may have been a heterogeneous impact of the treatment for women: it appears that females at the lower part of

[13] While the means were the same for students in coed and single-sex classes in Table 1, one might worry that they differed at different parts of the distribution and that this could have led to the effects seen in Figure 1. To examine this we check the balance across the entire distribution of predicted outcomes in Figure A1 and for IQ in particular in Figure A2 in the Appendix. There was no significant difference in the distributions of the predicted outcome or the distribution of IQ scores between coed and single-sex classes for men or women.

the distribution benefited more from being assigned to an all-female class. To examine if there were differential effects, we follow Barrera-Osorio et al. (2011) and look at predicted versus actual outcomes for students in coed and single-sex classes. To predict outcomes, we used the 19 individual variables in Table 1 and regressed the exam score on those variables for students in coed classes. We did this separately for males and females. Using the estimated equation, we then predicted the exam score for each student and plotted the actual exam score against the predicted score the student achieved. Figure 2 shows local polynomial plots of actual versus predicted outcomes. In the graphs, if the predicted and actual score were the same, they would line up on the 45-degree line.

[Insert Figure 2 here]

For males, the predicted and actual score line up fairly well for men in coed or all-male classes between the grades 38 and 61 (over 60% of our sample has grades that fall in this range). Note, crucially, though, that the relationship between predicted and actual exam scores for men in all-male classes does not appear to differ from the relationship between predicted and actual exam scores for men in coed classes. This is not the case for females.

In the top right panel of Figure 2, we plot the predicted and actual exam scores for women in coed and all-female classes. For women in all-female classes who are predicted to score below 50, the actual scores are much higher than the predicted scores. However, for women in all-female classes predicted to score 50 or above, the predicted and actual scores line up in roughly the same fashion as they do for the coed classes. This suggests that women who might be expected to perform below average on the exam are benefiting from being in an all-female class.

Similarly, for women in all-female classes, the actual coursework scores are above predicted scores for women expected to be in the lower part of the distribution. This suggests that women who might have been in the bottom half of the coursework distribution are more likely to benefit from the treatment while there is no effect (differential or otherwise) for men.

Finally, we look at the aggregate grades (which can be either *only* the exam mark or 50:50 coursework and exam), shown at the bottom of Figure 2. Here we see some evidence that men may be benefiting from all-male classes at the lower part of the predicted distribution. For

women, there is again the clear pattern that we have seen throughout: women at the lower part of the predicted distribution have much higher actual scores than what is predicted if they are assigned to all-female classes. In particular, women who might not have passed Introduction to Economics are benefiting more from the treatment.

4. MAIN REGRESSION RESULTS

Our goal here is to see whether or not the treatment – being in a single-sex class – increases the academic performance of students relative to the control of being in a coeducational class. In subsection, 4.1, we look at the impact of the treatment on a number of continuous dependent variables measuring first-year academic performance. However, recall that the marks a student receives in the first year do not affect the type of degree she earns upon graduating. Moreover, a student need only get 40% or higher in all first year courses in order to continue her studies. For these reasons we also report, in subsection 4.2 below, estimates of the effect of single-sex classes on pass rates alone. Indeed, given the heterogeneous effect of the treatment for women suggested by Figure 2, we might expect that the effects on the pass rate may be larger than for the mean of the distribution. Finally, in subsection 4.3, we report estimates of the treatment on longer-run outcomes – those in the second and third years, where obtaining scores above the pass mark actually matter for the student's final degree qualification.

4.1 Introductory Economics Scores

Regressions estimating the impact of the treatment on each of our three dependent variables (exam, coursework, aggregate grade) are reported in Table 2. All scores examined in Table 2 were normalized by the coed distribution for that outcome, so the estimated coefficients are expressed in terms of standard deviations. These values are reported at the bottom of each column. The notes at the bottom of Table 2 give further details. Standard errors, clustered at the class level, are shown in square brackets. For the OLS regressions, p-values for the variables of interest were also calculated using the Cameron, Gelbach, and Miller (2008) wild bootstrap method since there are only 37 clusters.

[Insert Table 2 here]

The first three columns of Table 2 present OLS results for aggregate score. Column [1] reports estimates from a specification with just two explanatory variables – a single-sex treatment variable taking the value one if the individual is in an all-female or an all-male class, and zero otherwise, and another dummy variable taking the value one for females and zero otherwise. The coefficient of the single-sex treatment dummy is 0.140 and it is statistically significant at the 10 percent level; aggregate scores were higher, on average, in the single-sex classes than the coed classes.

Column [2] reports estimates from a specification that differs from Column [1] in that we now disaggregate the single-sex-class variable into separate all-female and all-male dummies. The estimated coefficients show that the impact of the treatment on aggregate scores is working through the all-female-class type, where the coefficient is 0.252 and is statistically significant at the 10 percent level. This is interpreted as follows: being assigned to an all-female class leads to a 0.252 of a standard deviation increase in aggregate score.[14] The all-male coefficient is 0.08 and is not statistically significant. Analogously the female coefficient has also diminished in magnitude and lost statistical significance (in comparison to the estimate in column [1]).

Column [3] is an expanded specification that we ran to explore the robustness of the treatment-variable estimates. It includes the full set of explanatory variables – all the individual and class level variables listed in Table 1. The estimated coefficient on the all-female treatment variable changes little across the specifications in Columns [2] and [3]. The all-male and female coefficients remain statistically insignificant estimated in Column [3], just as they were in Column [2] – our preferred specification.

Columns [4] and [5] of Table 2 present the OLS results for exam score. The specification in Column [4] is the analogue to Column [2], while Column [5] is the analogue to Column [3]. Across both specifications, the all-female coefficients are of similar magnitude and statistically

[14] That represents an 8.3% increase in the overall score for females in an all-female class. Also, as shown in Table A2, there was no significant difference in the all-female effect for students in EC100 and EC111 and if we control for GTA fixed effects the results are the same.

significant at the 10 percent level. The all-male and female coefficients remain very imprecisely estimated across these specifications too. Being in an all-female class has a positive and significant effect for females.

Next, in Columns [6] and [7], we report estimates of similar specifications for coursework score. The estimated all-female coefficient in Column [6] is roughly the same in the analogous specifications in columns [2] and [4]. However, in Column [7] – which includes all predetermined variables – the estimated coefficient for all-female is no longer significant and the point estimate has decreased by more than half. Since the gender and identity of the student is known by the TA marking the coursework, and the student may choose not to turn in the assignments, the grade given by TAs could be influenced by this knowledge and behavior. Furthermore, not all students take the coursework seriously as indicated by the fact that many do not do the complete set of four pieces of coursework. These could be possible reasons for the lack of robustness of the all-female estimate on coursework. To explore some reasons, we next look at the probability of doing all coursework and the average score if one did all coursework.

In Column [8], we see that women assigned to an all-female class are more likely to do all the coursework assignments. When all predetermined variables are added in Column [9], the all-female coefficient stays significant and positive and is only slightly smaller in magnitude: women in all-female classes are, roughly, 8 percentage points, or 10%, more likely to do all coursework. Since we do not have a robust effect on the coursework score, all we can say is that, for women, the treatment is associated with a greater probability of doing all the coursework.[15]

To summarize, Table 2 shows that students in single-sex classes did better than their counterparts assigned to coed classes; the result is driven by women assigned to all-female classes. Females in all-female classes did a quarter of a standard deviation better in their end-of-the-year grade than females in coed classes. This effect is primarily driven by the 0.246 standard deviation increase in the exam scores for women in all-female classes, as there is no robust effect on coursework scores. Being assigned to an all-female class did cause females to

[15] In the last two columns of Table A2 we look at the effect on coursework scores for students that did all the coursework. This is a selected sample and we find no effect being assigned to a single-sex class.

do more of the coursework assignments, though; they were 10% more likely to have done all assignments.

4.2 First Year Pass Rates

As already noted, a student needs to get 40% or higher in first year courses in order to continue her studies. For this reason, and given the heterogeneous results suggested by the descriptive analysis above, we also estimate the single-sex-class effect on *pass rates* in introductory economics, since this is the margin that matters for continuation. In this sub-section, our main dependent variable of interest is whether or not a student passed the introductory economics course. These results are shown in Table 3.

[Insert Table 3 here]

Columns [1]–[4] provide strong evidence of a positive effect of single-sex classes for females: 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 women in all-female classes are 6.7 percentage points more likely to pass their introductory economics course than those assigned to coed classes. For all students in coed classes in the introductory courses, the percentage passing was 87%, so the increase of 6.7 percentage points represents a 7.7% increase in the pass rate for females. That is a large effect, given that no additional resources were needed. When we add, in Column [2], predetermined individual-level and class-level controls from Table 1, the point estimate changes little.

Since our field experiment takes place at the class-level, we also report class-level regressions in Columns [3] and [4] of the percentage of students in a class passing. Here the number of observation is 37. Column [3] reports the parsimonious specification while Column [4] adds predetermined class-level controls. Column [3] shows that the pass rate for all-female classes was 6.3 percentage points higher than for coed classes. When the class-level variables are included, there is only a small increase in the magnitude of this coefficient. Indeed, the point estimates are very similar across these four columns, showing that women in all-female classes are between 6.0 and 6.7 percentage points more likely to pass introductory economics than their

counterparts assigned to coed classes. The similarity in the magnitude of the estimated effect of being in an all-female class across all four specifications shows we have a robust estimate of the effect of being in a single-sex class for females.

The pass mark of 40 percent is at the lower end of the distribution. Figure 2 suggested that this was where assignment to an all-female class had the biggest effect. Nonetheless we wished to further explore, using regression techniques, what was happening towards the top of the marks distribution. We therefore estimated the effect of class-type on the probability of obtaining a first-class mark in introductory economics. (In the UK a first-class mark is 70% and over, and in our sample 110 students, or 19%, achieved this.) These results are displayed in Columns [5] and [6] of Table 3. Both specifications support what was suggested in Figure 2, namely that single-sex classes did not affect the top of the distribution.

Finally, we investigated if the treatment in Introductory Economics had a spillover effect on our subjects' first year performance *overall*, proxied by its effects on passing the entire year. These estimates are reported in Columns [7] and [8] of Table 3. Both specifications show that women in an all-female class for introductory economics had a significantly greater probability of passing the entire year, relative to women in coed classes. No similar effect was found for men in all-male classes.[16] A student must pass *all* her courses in the first year to progress into the second year, and being assigned to an all-female class increases a woman's chance of passing her Introductory Economics class.[17] Hence, perhaps unsurprisingly, a female assigned to an all-female class is more likely to pass the entire year and be allowed to progress into the second

[16] Columns [1]-[3] have a positive insignificant point estimate for the all-male coefficient but in Column [4] it is significant. The change in significance is driven by an increase in the point estimate, not by an increase in precision. Therefore, we could lack the power to identify a small, positive effect of single-sex education for males (Park et al. (2013) shows a positive effect at the secondary school level). Figure 2, shows that low skilled males may benefit (but that high skilled males might be harmed). Therefore, with only one significant estimate in Table 3 and mixed suggestions in Figure 2, we cannot say we have conclusive evidence that single-sex classes are beneficial for males.

[17] Introduction to Economics is the core course for Economics students and one of the two core courses for Business students. All other courses in the first year (e.g. Methods of Economic Analysis) – which are all required and prescribed by the departments – build on the material in EC100 and EC111. Therefore, Introduction to economics is the main gateway course and if a student cannot pass it, then it is unlikely she will pass the other courses. This is one reason that the point estimate on all-females in column [7] is potentially larger than for introduction to economics course itself; the material learned helps students in other classes. However, with all controls (our preferred specification), as shown in column [8], the point estimate for passing the year is the same as for Introduction to Economics itself.

year. This is important because it means a student is not allocating effort towards introduction to economics and away from other courses.

Overall, the results in Tables 2 and 3 provide a compelling picture of the effect of single-sex classes on first year outcomes. Females in all-female classes for introductory economics are more likely to pass their first year course than females in coed classes; the effect of being in an all-female class is much stronger at the lower end of the distribution; and the effects are present even in the double-blind marked exams that are reviewed by external examiners. Thus there is strong evidence that women studying introductory economics in a coeducational university can benefit from being in an all-female class. But are these short-run effects that last only for one year, or do they have a longer-run effect? To this issue we now turn.

4.3 Longer-Run Academic Effects

We begin with a descriptive, graphical analysis of students' academic performance after the 1st year. Figure 3 displays the distributions of the students' final grades on graduation. The final degree grade is determined by a weighting of 40% of the second year average and 60% of the third year average grade. The distributions, shown separately for men and women, are disaggregated by treatment and control groups. The grades for the overall degree form the basis for the degree classification that each student receives at graduation.

The distributions for women in single-sex classes are less left-skewed than for their counterparts in coed classes. This is in contrast to the male distributions.[18] Recall that these distributions are only for students who actually continued in the university, therefore the results are to be interpreted with caution since the treatment also had an effect on the probability of a female continuing with her studies. As this selection is likely negative, the results suggest that the effects of treatment carry over throughout the second and third year.

[18] Figure A3 in the appendix shows the average grade distributions for students in the second and third year of university. In the second year, females assigned to all-female classes in their first year Introductory Economics classes (and who are now only in coed classes and lectures) were doing better than those who were assigned to coed classes. The difference in the distributions is significant (Mann-Whitney p-value = 0.003). This is not the case for males. In the third year, the distribution for females assigned to all-female classes still lies clearly to the right of that for females assigned to coed classes, although the statistical significance is reduced to the 10 percent level (Mann-Whitney p-value = 0.101).

[Insert Figure 3 near here]

Figure 3 appears to show that the treatment is having an effect on the overall degree score. However, the results in the figure cannot be interpreted as causal because the sample is changing from year to year: only 505 of the 570 students are left in the second year and only 487 are still enrolled and earn a grade in the third year; and the treatment had an effect on the probability of a female continuing. However, the bias introduced in the sample selection is very likely working *against* us finding an effect. This is because women assigned to all-female classes who now continue in their studies are the marginal ones who scored lower in the first year.[19] To examine the causal effect of our treatment on later outcomes, though, we will look at degree receipt.

We have information on whether a student earned a degree and, if she did, the classification of her degree. An advantage of looking at the degree classification of students is that we can use all observations that were originally treated; that is, we have complete *outcome* data for all randomized subjects. Thus we can get at the *causal* effect of being assigned to a single-sex class on one's degree outcome. In the UK, degree *class* is the most important higher educational outcome, with degrees classified into one of the following categories: a first class degree (with an overall average of 70%+); an upper second (60-69%); a lower second (50-59%); or a third (40-49%). Scores below 40% are a failure. 'Good' degrees are regarded as those that are upper-second and above. With such a degree, students can carry on to postgraduate work and get a good job.[20]

[Insert Figure 4 near here]

The degree outcomes for our 2010 cohort of economics students are illustrated in Figure 4, which presents differences in degree classification for single-sex and coed classes. Once more the raw data suggest that single-sex classes have benefitted women, who are more likely to get

[19] Table A1 in the appendix shows regression results that further explore the differences in the distributions shown in figure 3.

[20] In the UK the primary hurdle to getting a graduate job is earning a degree classified as an upper-second or first (see Feng and Graetz (2017) or Smith et al. (2000)). A 'good degree' also means that a student will earn much more over her lifetime in the labor force. Getting a lower-second degree means a student can continue studies in some MSc programs or enter some graduate positions but her options will be limited.

a good degree than their coed counterparts. Moreover, they are less likely to get no degree.

In Table 4, we estimate the *causal* effect of being assigned to a single-sex class in one's first year on one's degree outcome.[21] Table 4 presents the results from regressing gender and class-type on the degree classification variables; the standard errors are once more clustered at the class level.[22]

[Insert Table 4 near here]

In columns [1] and [2] we report estimates of the probability of obtaining a first class degree. In columns [3] and [4] we report estimates of the probability of obtaining either a first class or an upper-second class degree; here the dependent variable takes the value one for a first-class or upper-second class degree, and zero otherwise. We see that women in all-female classes were significantly more likely to be awarded a 'good degree' (classified as either a first or a second class, first division) than their counterparts in the coed classes. These estimates confirm the findings from the unconditional distributions of Figure 4. For men, the all-male class effect is statistically insignificant.

The estimated effect of being assigned to an all-female class in first year is a 28 percentage point increase in the likelihood of getting a 'good degree.' That is a 61% increase for females. This means that the intervention had a very large impact on the employability of females graduating from Essex.

Finally, columns [5] and [6] of Table 4 give marginal effects from the estimation of the probability of obtaining a third-class degree or a fail, while the probability of failing is shown in the last column. In both probit models, women in all-female classes are significantly *less* likely to be in that degree classification than their counterparts in the coed classes. Indeed, being in an all-female class means a woman is 10.4 percentage points less likely to fail out of university; that

[21] We have 566 observations for our degree receipt analysis. We lose four students who had "extenuating circumstances" and were allowed additional time to study towards their degrees. Those students, therefore, cannot be classified into achieving a specific degree classification or no degree because they are technically still enrolled at the university or could choose to re-enroll after the "extenuating circumstances" have been resolved. We did run the regressions where we allocated the four students to having each type of degree and our results did not change.
[22] We also estimated class-level regressions where the dependent variable was the proportion in class who got the particular degree classification and found the same results as those reported in this paper.

is a 57% decrease in the likelihood that she does not get a degree.

In summary, we find that the effects of assignment to an all-female class carried on well into final degree grades and final degree classification. Women assigned to all-female classes are more likely to get 'good degrees' and are less likely to fail out of university. Since access to postgraduate study and to good jobs in the UK is affected by degree class, the effects are likely to be even longer-run than simply affecting degree classification.

5. BEHAVIORAL CHANGES AND POTENTIAL CHANNELS

In this section we explore the mechanisms that might be driving the effects of single-sex classes on females. In particular, we focus on whether the treatment was associated with potentially grade-enhancing changes in student behavior – such as attending classes or doing all the assignments – and we also examine if the TAs might have taught women assigned to all-female classes differently.

[Insert Table 5 here]

Columns [1] and [2] in Table 5 examine if there is evidence that TAs might have changed their teaching style in single-sex classes. At the end of the academic year, before students take their exams or know their overall grades, students rate their TAs on a scale from 1 to 5. A score of 5 means a student strongly agrees that "Overall the classes were taught well" and a score of 1 means a student strongly disagrees with the statement. Column [1] shows that TAs teaching all-female classes did not get a higher score than TAs who taught coed or all-male classes.[23] However, since no TA only taught single-sex classes, we can include TA fixed effects to see if a TA was given a higher score when teaching a single-sex class rather than a coed class. This is done

[23] The scores are normalized by the coed class mean shown at the bottom of the columns. Thus, as before, the coefficients are in terms of standard errors of that distribution.

in column [2]; TAs did not receive any higher score when they taught a single-sex class (either all-female or all-male).[24]

Recall that students in the first year are given the opportunity to do coursework assignments which can help their grade through the 'max rule' as described earlier. In column [3] we replicate the results from column [8] of Table 2 and show that females assigned to an all-female class are 9.3 percentage points more likely to do all the coursework (an 11% increase). Besides coursework, though, attendance in the classes taught by TAs could help a student learn the material better and be better prepared for the exam. As we see in column [4], being in an all-female class increases attendance by 0.243 standard deviations of the coed mean attendance. Thus an obvious question is: could the single-sex effect for females be explained by the choice to do more optional coursework and the increase in attendance?

To see how the changed behavior of a student is associated with passing, we include both doing all coursework and attendance in Column [5].[25] The all-female coefficient is just about half of what it is in our preferred specification from Table 3. This means, even given the biases associated with controlling for attendance and doing all the coursework (both of which were positively affected by the treatment), only half the effect of the all-female coefficient can be explained away. Finally, we use only the coed population in column [6] and estimate the effect of doing all coursework and attending classes on the likelihood of passing. Using these estimates we also find that the behavioral changes can explain only about half of the estimated effect of being assigned to an all-female class.[26] This suggests that female student effort may be an

[24] One may question if the TA score is correlated with the quality of instruction. We explored this and found that students in classes where TAs received higher scores were more likely to pass the course and that the all-female coefficient stays the same.

[25] Table 5 focuses on the pass rate. Table A3 in the appendix looks at how doing all coursework and attending classes also affect exam scores and degree classifications. As with the pass rates, the effect of intermediate outcomes on aggregate and exam scores in the first year can only explain, at most, half of the all-female effect. With regards to degree classification, the intermediate outcomes can explain none of the effect of the all-female impact.

[26] Column [6] shows doing all coursework is associated with a 0.362 increase in the pass rate and that increasing attendance by one standard deviation is associated with a 0.032 increase in the pass rate. Being assigned to an all-female class increased the likelihood of doing all coursework by 0.093 and of attending classes by 0.243 standard deviations. That suggests the behavioral changes account for 0.362*0.093+0.032*0.243=0.04 which is less than 60% of the effect of our preferred specification in Table 3.

important mechanism driving the effect of all-female classes, but that a pure effort explanation must go beyond simply turning in homework and showing up for class. [27]

6. Conclusions

This paper reports the results of a field experiment designed to examine the effects of single-sex classes on the performance of first-year university students in a highly-ranked and publicly-funded economics department. The novel features of our design are twofold. First, we *randomly* assigned students into a single-sex or a coed class environment, and hence selection is not an issue. Second, we implemented this random assignment to single-sex classes in the economics faculty of a *coeducational university.*

We examine if single-sex classes within a coeducational environment have an effect on female test scores, pass rates, and continuing to study at university, as well as longer-run outcomes such as degree class. We find that females assigned to all-female classes score a quarter of a standard deviation higher in their introductory economics class and are 7.3% more likely to pass the course. Furthermore, the benefits to females of single-sex classes appear to carry over past the first year and lead to 57% decrease in the likelihood that a female does not graduate with a degree in economics and a 61% increase in the likelihood that she graduates with a 'good degree.'

These results all occurred with no additional expenditure on the part of the university. Typically, programs that increase female pass rates or improve student performance are much more costly. For example, a study on the Achievement Awards demonstration in Israel by Angrist and Lavy (2009) found that females were more likely to earn a high school matriculation certificate if offered $1500 for completion but that men were not any more likely to earn the qualification. While comparing the costs and benefits of programs based on different outcome variables is not easy, a program like ours that causes a female to be 7.7% more likely to pass a

[27] Also, the coefficient on all-male is significant now (at the 10% level). In this case it is due to the fact that the estimate has become more precise. Given that being in an all-male class has no effect on doing all coursework or attendance then controlling for those two aspects is equivalent to saturating the regression and could be giving us more power to estimate the effect of single-sex education on males. However, this is the only case where the all-male point estimate was similar to other specifications (as in columns [6] and [7]) and became significant when we controlled for more observable factors.

course, score a standard deviation better overall in classes, and leads to a 57% decrease in the likelihood she doesn't graduate with an economics based degree at no additional cost is rather extraordinary.

Our study focuses on only a few aspects of single-sex classes – their effects on first-year pass rates, grades across years, and degree class. While these outcomes are likely to influence wages and job prospects for students, other factors that we are not examining – such as socialization – could also play a role in labor market outcomes.[28] Therefore, while we find large positive effects of single-sex first year classes for females studying economics, we wish to emphasize that more research is needed. For instance, it is not clear if more than one hour of single-sex education in a coeducational institution is going to be more beneficial or if fully gender-segregated education would produce better results. [29]

We also emphasize that our analysis conditions on women who selected to study economics at a top-rated university department, and it is possible that selectivity of subject area and/or institution might be a bigger margin driving gender gaps than the one we have explored. Nonetheless, the margin that we are examining in our paper shows that persistence is still important even among students at a selective university, and this is of considerable policy relevance. For example, Arcidiacono, Aucejo, and Hotz (2016) show, using data for minorities in STEM fields at University of California campuses, that minorities entering STEM programs have lower persistence rates despite the fact that they would have graduated had they attended different schools.

The research on schools summarized by Halpern et al. (2011) suggests that having *all* instruction done in a single-sex environment has no positive effect or could be harmful. In contrast, we have shown that single-sex classes within a broader coeducational environment

[28] Booth, Cardona and Nolen (2014), using a different experimental subsample, report how risk attitudes evolved across the first eight weeks of the Introductory Economics course, and showed how women in single-sex classes modified their risk preferences over time.

[29] Our results are not directly comparable with those of Oosterbeek and van Ewijk (2014), who conducted a randomized experiment for first year undergraduate students of economics and business at Amsterdam University. Their treatment was the share of females in the classroom (which they manipulated to be between 0.14 and 0.51). They do not find any effect of proportion female on individuals' outcomes. In contrast, we have a female class-ratio of either 1 or 0. When the ratio is 1 (single-sex class) we observe a positive effect on females.

benefit females in a subject area where women are in the minority. However, will there be benefits for women studying in subject areas where females are in the majority? Will the benefits of single-sex classes be even greater when women are even more of a minority? Will men benefit from single-sex classes if they are in the minority? Clearly more research is necessary to establish if there is an optimal amount of education in single-sex classes that is beneficial for students, and in what other contexts females or males might benefit.

Since we do not yet fully understand the mechanism driving our findings, there is a worry that expansion of our treatment could lead to unintended consequences – as found in a quite different context by the intervention reported in Carrell, Sacerdote, and West (2013). For instance, while students did not notice the few single-sex classes in this study, if the experiment were expanded and all classes in Introductory Economics were single-sex, they would surely notice this expanded policy. Nevertheless, any extension of our treatment should further explore the underlying mechanisms, and we hope that further studies will replicate and expand on our research in order to shed more light on this important issue.

Acknowledgements
For helpful comments on an earlier version of this paper, we thank the editor Jonah Rockoff, two anonymous referees, and seminar participants at the economics departments of the Paris School of Economics, the Australian National University, Glasgow University, the University of Reading, Heidelberg University, and the University of Essex.

References

[1] Angrist, J., Lavy, V. (2009) "The Effects of High Stakes High School Achievement Awards: Evidence from a Randomized Trial" American Economic Review, vol. 99, no. 4, pp. 301-331

[2] Annelli, M. and Giovanni, P. (2013) "The Long-Run Effects of High-School Class Gender Composition" NBER Working Paper 18744

[3] Arcidiacono, Peter, Esteban M. Aucejo, and V. Joseph Hotz (2016) "University Difference in the Graduation of Minorities in STEM Fields: Evidence from California" *American Economic Review*, vol. 106, no. 3, pp. 525-562

[4] Barrera-Osorio, Felipe, Marianne Bertrand, Leigh L. Linden, and Francisco Perez-Calle. (2011) "Improving the Design of Conditional Transfer Programs: Evidence from a Randomized Education Experiment in Colombia." *American Economic Journal: Applied Economics*, 3(2): 167-95.

[5] Booth, A.L., Cardona-Sosa, L., and Nolen, P.J. (2014) "Gender differences in Risk Aversion: Do Single-Sex Environments Affect Their Development?" Journal of Economic Behavior & Organization, vol. 99, pp. 126-154

[6] 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

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

[8] Cameron, Colin, Jonah Gelbach and Douglas Miller (2008) ""Bootstrap-Based Improvements for Inference with Clustered Errors", *Review of Economics and Statistics*, August 2008, Vol. 90, 414-427.

[9] Carrell, Scott and West, James (2010) "Does Professor Quality Matter? Evidence from Random Assignment of Students to Professors," Journal of Political Economy, vol. 118, no. 3, pp. 409-432

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

[11] Eisenkopf, G., Hessami, Z., Fischbacher, U., Ursprung, H. (2014) "Academic Performance and Single-Sex Schooling: Evidence from a Natural Experiment in Switzerland" Journal of Economic Behavior & Organization *forthcoming*

[12] Feng, Andy and Georg Graetz (2017) "A Question of Degree: The Effects of Degree Class on Labor Market Outcomes" *Economics of Education Review,* forthcoming

[13] Gist, M.E., Mitchell, T.R. (1992). "Self-Efficacy: A Theoretical Analysis of Its Determinants and Malleability" *The Academy of Management Review,* Vol. 17, No. 2 (Apr., 1992), 183-211

[14] 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

[15] 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

[16] Independent Schools Council (ISC) (2014) "ISC 2014 Census Report"

[17] 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

[18] 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.

[19] 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

[20] Long, MC and Conger, D. (2013) "Gender sorting across K-12 schools in the United States" American Journal of Education, vol. 119, no. 3, pp. 349-372

[21] Miguel, T. and Kremer, M. (2004) "Worms: Identifying impacts on education and health in the presence of treatment externalities," Econometrica, vol. 72, no. 1, pp.159-217

[22] Morrison, N. (2014) "Single-Sex Education Belongs in the 21st century" Forbes, April 30th (https://www.forbes.com/sites/nickmorrison/2014/04/30/single-sex-education-belongs-in-the-21st-century/#1c41fcd030bd)

[23] 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.

[24] Oosterbeek, H., van Ewijk, R (2014) "Gender peer effects in university: Evidence from a randomized experiment" Economics of Education Review

[25] Park, H., Behrman, J., Choi, J. (2013) "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*

[26] 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

[27] Schneeweis, N., and Zweimüller, M. (2012) "Girls, girls, girls: Gender composition and female school choice" Economics of Education Review, Vol. 31, vol. 4, pp. 482-500

[28] Smith, Jeremy, Abigail McKnight, and Robin Naylor (2000) "Graduate Employability: Policy and Performance in Higher Education in the UK" *The Economic Journal,* vol. 110, no. 464, pp. F382-F411

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

[30] 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

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

[32] 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