

GIRL POWER? AN ANALYSIS OF PEER EFFECTS USING EXOGENOUS CHANGES IN THE GENDER MAKE-UP OF THE PEER GROUP.

PROUD, Steven

Abstract

The effect of a child's peers has long been regarded as an important factor in affecting their educational outcomes. I use exogenous changes in the proportion of girls within English school cohorts to estimate the causal effect of a more female peer group. I find significant negative effects of a more female peer group on boys' outcomes in English, particularly at age 7. Much, but not all, of this deficit is caught up by age 11, but there is still a significant negative effect present. In maths and science, all pupils benefit from a more female peer group in primary schools.

Keywords: gender, education, peer groups

JEL Codes: J1, D1, I21

1. Introduction

Perhaps one of the most influential educational reports of the 20th century, The Coleman report (Coleman et al (1966), 86) introduced the idea that "Attributes of other students account for far more variation in the achievement of minority children than do any attributes of school facilities and slightly more than do attributes of staff." The report was commissioned to investigate the level to which school level integration was progressing following the removal of the segregation laws. The conclusions were that schooling increased the disparity between whites and blacks. This led to an interest in the impact of peer groups on educational outcomes. Early studies include Winkler (1975) who found differential effects of the composition of peer groups on different races, and Summers and Wolfe (1977) who found that both black and non-black pupils benefited from a more balanced mix of black and non-black pupils, and also commented that students who tested at or below grade level were helped by being in a school with high achievers.

However, Manski (1993) identifies problems with trying to measure peer group effects. Manski splits peer effects into three different types; endogenous effects (where the behaviour of one individual affects the behaviour of others and vice versa), exogenous effects (where the characteristics of one individual affect the outcomes of other individuals), and finally correlated effects where there is only correlation between characteristics of one individual and the outcomes of others. The first two present causal effects, which are useful for policy, whereas the final case merely presents a correlation, which is not. However, disentangling whether effects are causal, or merely correlations present their own challenges. Manski discusses the problems inherent with such endogenous peer effects when trying to infer the effects that members of a reference group have on its own members (the reflection problem), and argues that it is not possible to make inferences on effects unless one has prior knowledge of the make-up of the reference group. Furthermore, he argues that studies using apparent random distribution may experience bias to the apparent peer effect if there are unseen family characteristics that are in common with the reference group.

Much of the literature has concentrated on the effect of the quality of the peer group on individuals' outcomes (see, for example Dills (2005)). However, there has been an increasing interest in the effect of other characteristics of the peer group on individuals' outcomes and behaviour. Mora and Orepoulos (2011) find significant effects of a peer group's willingness to leave school early on individuals' decisions. Hoxby (2000)¹ takes advantage of year-on-year changes in the gender make-up of the within school peer group to estimate the effect of the gender and racial make-up of a child's peer group on their educational outcomes. Hoxby's initial strategy utilizes the credibly exogenous variation in the distribution of females across cohorts within a grade, using the raw proportion of girls as a measure of the peer group. Her results suggest that both boys and girls benefit from a more female peer group in all subjects, with a 10 percentage point increase in the proportion of the peer group that is female associated with approximately a 0.02 standard deviation increase in individuals' outcomes. She then combines change in the ability of the peer group and finds an effect too large to be credibly explained by merely changes in the ability of the peer group. Further evidence of the effect of the gender make up of a child's peer group is shown by Lavy and Schlosser (2007) who provide support to Hoxby's results, showing a negative effect of a more male peer group in Israeli schools. This with the test-score gap between girls and boys to estimate the effect of an exogenous. Similarly, Hansen et al (2006) find that female dominant and equally mixed groups perform better than male dominated groups. Furthermore, Whitmore (2005) finds positive effects of a high proportion of girls in kindergarten through to the second grade on the outcomes of both boys and girls. However, Whitmore (2005) also suggests that in the third grade there is evidence that boys do worse in a class with a high fraction of girls. Lavy and Schlosser (2007) suggest that much of this effect can be attributed to the fact that boys provide a more disruptive atmosphere in the classroom. However, Younger et al (1999) and others suggest that teachers give differential attention to boys and girls. Thus, it is possible that when the gender make-up of the class were to change, teachers would also change their behaviour in the classroom, which could have positive or negative effects on children's outcomes.

In this paper, I identify causal effects of a more female peer group on outcomes, taking advantage of the strategy developed by Hoxby (2000) in order to estimate the effect of the gender make-up of a child's peer group in the context of the English schools system. I examine the effect of the within classroom peer group by examining schools with credibly only one class per cohort and further examine the effect of a more –female peer group on the value added that pupils gain from being educated in a more-female peer group. However, it should be noted that these schools may be different than 'standard' primary schools because they are smaller than average. Largely, the results are in line with those suggested by Hoxby (2000) and Lavy and Schlosser (2007), but I also find significant negative effects of a more female peer group on boys in English at all levels of assessment, and significant positive effects of a more female peer group on both boys and girls in maths and science, although these effects largely disappear post age 11. The value added model suggests that although boys are disadvantaged at age 7 by having a more female peer group in English, much of this deficit is caught up by age 11.

¹ Also published in abridged form as Hoxby (2002)

This paper begins by giving an overview of the school system in England, and discusses the pupil level annual school census (PLASC) and the National Pupil Database (NPD) used in this analysis. Section 3 discusses the methodology, as developed by Hoxby (2000), and the innovations used in this research, section 4 discusses the results and finally section 5 offers discussion and conclusions.

2. Data and the English Schools system.

There is a quasi-market in English schools, where parents make a preference of which school within their local authority they would like their child to attend. If the school is over-subscribed, then the school, or local authority, uses admissions criteria, which are usually related to the distance from the family home to the school, whether the child has any siblings at the school, or in some schools their attendance at church or equivalent schools.

In England, it is compulsory² for children to be in education from the September after they turn 4 up until age 16. During their school careers, pupils are taught in different types of schools, depending on their age. While there are several different models of schooling in England, the most common is that pupils attend infant schools from age 4 to age 7, junior schools from age 7 to age 11, and secondary school from age 11 onwards. Often, infant and junior schools are combined to make one single primary school. Similarly, up until the age of 11, the most common model of education has pupils taught with the same peer group for all subjects. In England, secondary schools are usually significantly larger than primary schools (with, on average 184 students per academic year), and many schools stream their pupils into different peer groups for different classes. As such, conclusions of the effect of a more female peer group will be less reliable, as students may be taught with inconsistent peer groups, and will not be taught with their entire peer group either. Because of this, I concentrate only on primary schools³

Table 1 gives summary statistics of cohorts within schools. At age 7 and age 11, pupils are educated in cohorts with a mean size of between 39 and 42. In this paper, I investigate the effect of a “more female” peer group on academic outcomes for children up to the age of 11. Table 1 further shows that, on average, schools at all ages have approximately 49% girls, and the standard deviation in the proportion who are female is approximately 9%.

During their time in primary education, pupils are taught in key stages as set out within the national curriculum. At the end of each key stage, the pupils are assessed by either teacher assessment or in nationally administered tests. These key stages are administered at age 7 (key stage 1), age 11 (key stage 2). At each key stage, pupils are examined in English⁴ and mathematics, and at age 11, pupils are additionally examined in science. Pupils’ achievement at ages 7 and 11 is measured in national levels. In each subject, the national curriculum is separated into strands which assess various skills within the subject, and each level is associated with a certain skill level that needs to be

² This was increased to 18 in 2013, but this does not affect the data collected here.

³ The analysis on secondary schools was carried out, and is available in an earlier version of this paper; Proud (2008)

⁴ Age 7 examinations split English into reading and writing levels.

achieved. In order to analyse this data, the grades are converted into a linear scale, and then normalised by key stage, subject and year, such that the individual level data has a mean 0 and standard deviation of 1.

Table 1 Summary statistics.

	<u>All Schools</u>		<u>Small primary schools</u>	
	Age 7	Age 11	Age 7	Age 11
Mean size of within school cohort	39.00 (19.07)	41.51 (21.52)	20.57 (5.70)	20.66 (5.51)
Mean proportion of within school cohort that is female	0.490 (0.092)	0.492 (0.090)	0.491 (0.118)	0.493 (0.119)
Number of school/cohorts	72457	73520	17675	14400

Notes: One observation per school year. Standard deviations in parentheses. Small schools are defined as schools with 30 or fewer pupils in every observed cohort.

English at age 7 needs to be treated carefully. Pupils are awarded two grades at age 7 for English (reading and writing). As such, in order to have 1 grade per pupil per subject, I consider a single English grade to be the mean of their English related exams at age 7.

I use data from the National Pupil Database (NPD) and the Pupil Level Annual School Census (PLASC), containing data on all pupils in state funded education in England. The National Pupil Database (NPD) is an administrative database covering all pupils in state education in England. This database contains the results of pupils in the key stage assessments, with results available for pupils who sat exams at age 7 between 1998 and 2004, and age 11 between 1996 and 2004.

The vast majority of local authorities in England do not allow explicit selection of pupils by ability. However, there are some local authorities who still do permit this selection into grammar schools. Since the identification strategy will depend on random year-on-year changes in the gender make-up of the peer group, if a school is allowed to explicitly select pupils, then this would bias my results. Furthermore, other non-selective schools in the selective local authority would also bias the results, as they would have the students who hadn't been selected by the grammar schools. Following Atkinson et al (2006), I define a local authority to be selective if 10% or more of the secondary school pupils are in selective schools, and omit these selective local authorities. In addition, all schools within non-selective local authorities which are defined as selective schools are omitted to prevent any similar biases. This leads to 11.2% of schools being dropped. Further to this, there are some schools in PLASC which appear to be unreliable, based on school size. As such, I omit any school with fewer than 10 pupils within the cohort. Similarly, in order to avoid errors with the data, any school which is recorded as having a change in size year-on-year smaller than the 1st or larger than the 99th percentile is omitted. Additionally, since the identification strategy takes advantage of year-on-year changes in the gender make-up of the of the cohort, single sex schools are dropped from the sample, as are schools which are recorded as having a year-on-year change in the proportion of females outside the 1st to 99th percentile. Finally, private schools and schools specifically for students with special educational needs and any school that does not appear in every year of the data are omitted. This leads to 74.68% of all schools included in the data.

3. Methodology

This paper uses the same basic methodology used by Hoxby (2000), but places it in the context of the English schools system. Hoxby's (2000) model suggests that since schools are not changing significantly over time in terms of teachers, facilities, and demographic of intake, then there will be an underlying average outcome μ for females (males) educated in school s in subject j within cohort c . Thus, a female (male) individual, i 's, outcomes, A can be modelled as a function of the underlying average, individual level characteristics, X and the proportion of the cohort who are female, p .

$$A_{i,female, sjc} = \mu_{female, sj} + \gamma_{female} p_{sjc} + \alpha_{female} X_{i, sjc} + \varepsilon_{i, sjc} \quad (1)$$

Following Hoxby (2000), in order to identify the effect of a more *female* peer group on outcomes, I use a first difference methodology to examine the effect of year-on-year changes in the proportion of the within school cohort who are female on the year-on-year changes in the average outcome for the subgroups of female and male students.

$$\Delta \bar{A}_{female, sjc} = \gamma_{female} \Delta p_{sjc} + \alpha \Delta \bar{X}_{female, sjc} + \Delta \varepsilon_{female, sjc} \quad (2)$$

In order to estimate these effects, I estimate equation (2) for *females* and *males* separately using OLS, with year on year changes in the average outcome achieved in national tests as the dependent variable and year on year changes in the proportion of the peer group who are female as an explanatory variable. Further controls include year dummies to control for shocks to the year-on-year change in outcomes which affect all schools, and also dummies for the age at which the examination is sat. In order for γ to be an unbiased estimate of the effect of a more female peer group on individuals outcomes, the change in the proportion of pupils who are female needs to be uncorrelated with changes in any other individual characteristics which are, in turn, correlated with changes in outcomes.

It is unlikely that there will be any omitted variable bias here, since the gender of a child is credibly random and as such, uncorrelated with background characteristics. One worry is that parents with certain characteristics may systematically make different decisions on which school to send their children to, dependent on gender. In order to check for this possibility, I take advantage of a robustness check, as suggested by Hoxby (2000), which for each school regresses the proportion of the within school cohort which is female against a linear and a quadratic time trend, then randomises the order of time, and performs the same regression. In order to obtain a sample where schools do not have a systematic trend over time in the proportion of pupils who are female, the ratio of the R-squareds is calculated, and I consider any school with a ratio less than 1.05 to have credibly random year-on-year changes in the gender make-up.

A necessary requirement for this identification strategy to be useful is for there to be a significant year-on-year variation in the proportion of pupils who are female. As discussed in the previous section, the within school standard deviation of the proportion of the cohort which is female is 9% at ages 7 and 11. As such, we have a relatively large year-on-year variation within school up to the age of 11.

In constructing this model, I have assumed that the proportion of pupils who are female has a linear effect on the individuals' outcomes. One worry is that an increase in the proportion of the peer group that is female will have a different effect depending on the current proportion of the peer group that is female. That is, if we have an all-male cohort, then an additional female pupil *may* have more impact on the peer group than when the cohort is largely female. In order to ensure that the linear model of the peer effects is the correct specification, I use a similar strategy to Hoxby (2000). In order to test the linearity, I construct quartiles for the gender make-up of the peer group. This means we can compare the computed effect of a more female peer group for schools with low proportions of female pupils with schools with large proportions of female pupils. That is, dummies, q_k are created, for the k^{th} quartile of the gender mix distribution, and these are interacted with the year-on-year change in the proportion of the cohort that is female. In order to test the linearity, the year on year change in female (or male) outcomes is regressed against the change in the proportion of the cohort which is female and the interacted terms. Thus, the coefficient δ_k is the change in the marginal effect of a change in the gender make-up of the peer group, depending on how female-dominated it initially was. If any of these coefficients are non-zero, then it suggests there is a non-linear effect.

$$\Delta \bar{A}_{female, sjc} = \gamma \Delta p_{female, sjc} + \sum_{k=2}^4 \delta_k \Delta p_{female, sjc} q_k + \Delta \varepsilon_{female, sjc} \quad (3)$$

I then use an F-test to test that $\delta_k = 0, k = 2, 3, 4$

3.1 Small schools

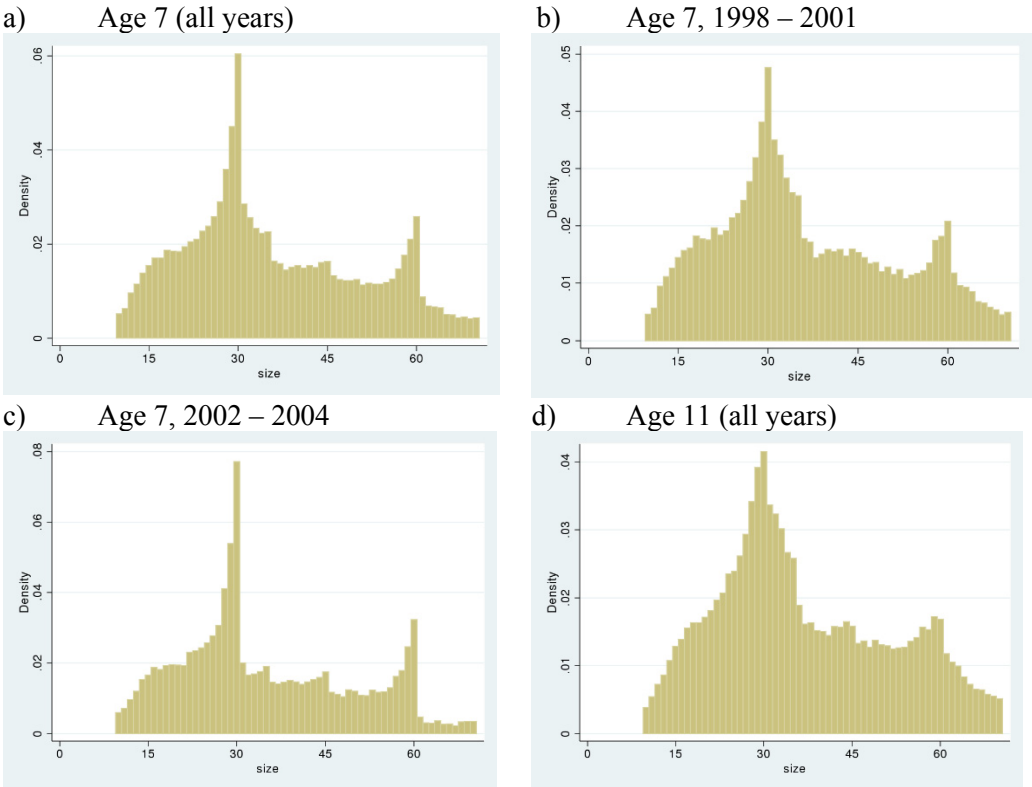
There is a potential problem with this strategy. Since there is no data on classroom level interactions within the school, it is possible that the magnitude of effect could be biased due to measurement error of the true peer group which a pupil experiences. That is there could be non-random allocation of pupils to classrooms within a school, meaning a pupil who attends a large school with a large proportion of girls could still be taught in a male dominated classroom. As such, by examining the year on year changes in the proportion of the cohort who are female, individual children, within the school, may experience a different peer group to the one implied by the gender make-up of the cohort. In order to address this possibility, I use the fact that in infant schools in England (covering pupils from age 4 to age 7) since the start of the 2001/2002 academic year there has been a legal requirement that there should not be more than 30 pupils to a qualified teacher (the [Education \(Infant Class Sizes\) Regulations 1998](#)⁵). Effectively, this means that the maximum class size in infant's school is 30. However, according to statistics from the Department for Children, Schools and Families⁶ in 2009, 310 classes out of 53,160, or 0.6% of all key stage 1 classes, are taught in unlawful classes of 31 or more. Furthermore, 580 classes (or 1.1% of all key stage 1 classes) were in lawfully allowed key stage 1 classes with more than 30 pupils. These lawfully allowed over-size classes

⁵ See <http://www.opsi.gov.uk/si/si1998/19981973.htm> for more details. Accessed 10/8/2009

⁶ See http://www.dcsf.gov.uk/rsgateway/DB/SFR/s000843/SFR08_2009_ClassSizeCommentary.pdf for more details. Accessed 10/8/2009

are approved due to children being allowed entrance following appeals, being moved into a school's area after the start of the school-year or when the LA has placed a child with special needs into a school.

Figure 1 Distribution of school sizes at ages 7 and 11.



Notes: Schools with fewer than 10 pupils per academic year are omitted. Observation is a school-cohort.

If infant schools follow this legislation, then we would expect to see a saw-tooth distribution of school cohort sizes. That is, we would expect to see peaks at sizes of 30, 60, 90 etc. This would indicate schools being filled until they reach the maximum size of 30 for a classroom, and if they have additional capacity, then starting a new class. Figure 1 shows the distribution of school sizes at age 7 and 11.

It is clear from panel a) that there is clustering at and below schools of size 30 and 60 at age 7, indicating that the schools are filling up the available spaces, and then stopping admissions. Panels b) and c) of Figure 1 show the distributions at age 7 before and after the 2002 reform, and while there is a more pronounced drop-off post legislation, there is still evidence that schools are using class sizes of 30 as a general rule pre-reform. Thus, at age, it seems a valid strategy to consider schools with 30 pupils as being schools which primarily teach all of their pupils within the school-year in one class. Similarly, panel d) of figure 1 shows the distribution of school sizes at age 11. There is a similar distribution as that seen in key stage 1 pre-reform, but there are less pronounced falls after school sizes of 30 and 60. However, this evidence is sufficient to make the assumption that schools with 30 or fewer pupils consist of one class. Table 1 shows the distribution of

sizes and gender make-up of school-cohorts in small schools, and as would be expected, there is a larger standard deviation of the proportion of the cohort which is female in these small schools (11.8%).

3.2 Value Added

In order to try and estimate the effect of the gender make-up of the peer group on a child’s progress, rather than just their overall achievement, it is necessary to consider a value added specification. This is constructed by taking the expected performance of a pupil in a given exam based on the median performance for pupils with the same prior attainment. The value added score is thus the difference between the actual score and the expected score⁷. This score is then normalised to a mean 0, standard deviation 1 by subject and exam year. This value added score is then treated in the same way as the results for individual age groups, as described above.

4. Results.

4.1 Results in all schools

Table 2 shows regression results for all schools in English, mathematics and science. The initial specification includes all schools and levels, and I estimate equation (2)

$$\Delta \bar{A}_{female,gjc} = \gamma_{female} \Delta p_{gjc} + \alpha \Delta \bar{X}_{female,gjc} + \Delta \varepsilon_{female,gjc} \tag{2}$$

In English, there is a significant negative effect for male pupils of having a more female peer group, whereas for mathematics and science; girls boys experience a significant positive effect of having a more female peer group. The effect observed for boys suggests that a 10 percentage point increase in the proportion of the peer group that is female would have a 0.0064 standard deviation decrease in boys’ individual attainment at every age group.

For mathematics and science, the effect of a more female peer group is positive for both boys and girls, but of a smaller magnitude than that seen in English. For mathematics, the results for all schools suggest that a 10 percentage point increase in the proportion of the cohort that is female would have a 0.0026 standard deviation increase in boys’ outcomes and a 0.0032 standard deviation increase in girls’ outcomes. For science, this increase is associated with a 0.0041 standard deviation increase in boys’ individual outcomes and 0.0040 standard deviation increase in girls’ outcomes.

Table 2 Results for all schools and levels pooled

	English		Mathematics		Science	
	Male	Female	Male	Female	Male	Female
Proportion female	-0.064*** (0.012)	0.007 (0.011)	0.026** (0.012)	0.032*** (0.012)	0.041** (0.017)	0.040** (0.017)
Observations	126,436	126,436	126,436	126,436	64,330	64,330

⁷ This is the same methodology as is used for creating the value added league tables, as published by the Department for education. See, for instance, http://www.education.gov.uk/performancetables/schools_04/sec3b.shtml

Notes: Dependent variable is the change in mean key stage score within school cohort for male (female) pupils. Robust standard errors in parentheses. *** denotes significance at the 1% level, ** denotes significance at the 5% level, * denotes significance at the 1% level. Method is ordinary least squares. Each cell represents a separate regression. Key stage 1 is not formally assessed for science. Year and exam dummies are also included. Standard errors are clustered at school level.

Table 3 Results for all levels, only considering schools with credibly random year-on-year changes in the proportion of the school cohort that is female

	English		Mathematics		Science	
	Male	Female	Male	Female	Male	Female
Proportion female	-0.086*** (0.020)	-0.043** (0.019)	0.012 (0.021)	-0.003 (0.019)	0.049 (0.031)	0.012 (0.031)
Observations	37,296	37,296	37,296	37,296	18,438	18,438

Notes: Dependent variable is the change in mean key stage score within school cohort for male (female) pupils. Robust standard errors in parentheses. *** denotes significance at the 1% level, ** denotes significance at the 5% level, * denotes significance at the 1% level. Method is ordinary least squares. Each cell represents a separate regression. Key stage 1 is not formally assessed for science. Year and exam dummies are also included. Standard errors are clustered at school level.

In order to examine whether there is any likelihood of selection bias affecting our results, Table 3 shows the results comparing all schools at all levels with schools which meet the strict criteria for random year-on-year changes in the proportion of the peer group that is female, as set out in the methodology. Although approximately two-thirds of the schools in the whole sample have been dropped here, the negative results for English remain insignificantly affected. The results for maths and science, however, become insignificant, although the estimated coefficients for science are the same magnitude as those seen in the entire sample. This suggests that if there is any selection bias, it would only make the results marginally more positive. Since we are most interested in the negative estimated effects for English, this is not a worry for our analysis.

A further worry is that the linear specification may not be appropriate. In order to examine this possibility, I estimate equation (3). If we reject the joint null that all of the coefficients on the interaction terms are non-zero, then it suggests that a non-linear specification may be more appropriate.

Table 4 Testing the linearity of the regressions

	English		Maths		Science	
	Male	Female	Male	Female	Male	Female
Proportion female (1)	-0.080*** (0.014)	0.003 (0.014)	0.012 (0.014)	0.030** (0.014)	0.031 (0.020)	0.037* (0.021)
2 nd quartile (2)	-0.001 (0.007)	-0.007 (0.006)	-0.000 (0.007)	0.002 (0.007)	0.009 (0.010)	0.009 (0.010)
3 rd quartile (3)	-0.011* (0.006)	-0.007 (0.006)	-0.005 (0.006)	-0.002 (0.006)	-0.016* (0.009)	-0.001 (0.009)
4 th quartile (4)	-0.013** (0.006)	-0.004 (0.006)	-0.011* (0.006)	-0.001 (0.006)	-0.004 (0.009)	-0.001 (0.009)
Observations	126,436	126,436	126,436	126,436	64,330	64,330
P-value	0.0669	0.6141	0.1820	0.9457	0.0485	0.7030

Notes: Dependent variable is the change in mean key stage score within school cohort for male (female) pupils. Robust standard errors in parentheses *** denotes significance at the 1% level, ** denotes significance at the 5% level, * denotes significance at the 1% level. Year and exam dummies are also included. Standard errors are clustered at school level. Quartile dummies are based on 4 quantiles based on the proportion of pupils who are female in the previous year. The threshold values are 0.4333, 0.4912, 0.5484. P-value is for the F-test: $H_0: (2)=(3)=(4)=0$

Table 4 presents results for this non-linearity test. The results of the F-tests seem to confirm that largely, the linear specifications appear appropriate for investigating the effect of a more female peer group. However, there is some evidence of a marginal non-linearity for boys in English and science. This does not appear to be a problem for interpretation of the effects in English, as any non-linearity simply means a further reduction in performance when there is initially a more female peer group. That is, in schools which initially have a large proportion of the cohort that is female, a small increase in the proportion who are female has a more negative effect than for schools with a small proportion of the cohort that is female initially. In science, there is a small reduction in the effect of a more female peer group if initially the school has approximately an average make-up of the peer group, but any differences are small. Otherwise, the linear specifications appear appropriate.

Table 5 goes on to estimate the effect of a change in the gender make-up of a within school cohort at ages 7 and 11. The negative effect observed for boys in English remains present at all levels, and remains statistically significant at at least the 10 percent significance level (and generally higher) for all levels. At age 7, a one standard deviation increase in the proportion of the cohort that is female (9.2%) is associated with a 0.0081 standard deviation decrease in individual boys' outcomes. At age 11, the effect of a 1 standard deviation increase (9.0%) is associated with a 0.0036 standard deviation decrease in boys' outcomes. For girls, the only marginally significant effect is observed at age 11, with a 1 standard deviation increase in the proportion of the cohort that is female being associated with a 0.0025 standard deviation increase in individual girls' outcomes.

Table 5 Results broken down by level.

	English		Maths		Science	
	Male	Female	Male	Female	Male	Female
Age 7						
Proportion female	-0.088*** (0.016)	-0.013 (0.015)	0.041** (0.017)	0.039** (0.016)		
Observations	62106	62106	62106	62106		
Age 11						
Proportion female	-0.040** (0.017)	0.028* (0.016)	0.010 (0.016)	0.025 (0.016)	0.041** (0.017)	0.040** (0.017)
Observations	64330	64330	64330	64330	64330	64330

Notes: Dependent variable is the change in mean key stage score within school cohort for male (female) pupils. Proportion female is the proportion of the within school cohort who are female. Robust standard errors in parentheses. *** denotes significance at the 1% level, ** denotes significance at the 5% level, * denotes significance at the 1% level. Method is ordinary least squares. Each cell represents a separate regression. Key stage 1 is not formally assessed for science. Year and exam dummies are also included. Standard errors are clustered at school level.

Table 5 further goes on to examine the effects in mathematics and science. For mathematics, both boys and girls gain statistically significant improvements in their scores, with a 1 standard deviation increase in the proportion of the cohort that is female associated with a 0.0037 standard deviation increase for boys and 0.0035 standard deviation increase for girls in individual outcomes. For science, there are statistically significant effects for both boys and girls of a more female peer group at age 11.

4.2 Results in small schools.

Since these results are based on within school cohorts, many of the pupils are not taught with their implied peer group. As such, results based on primary schools with 30 or fewer pupils in every observed cohort are likely to only have 1 class per cohort, so analysis on this subset will allow us to observe the effect of the peer group a child actually experiences. Table 6 shows the results at ages 7 and age 11 in small schools.

The negative effect observed for all schools for males in English remains present in this estimation, although it is less significant at age 11 than at age 7. The effect at age 7 suggests that a 1 standard deviation increase in the proportion of the class which is female (11.8%⁸) is associated with a 0.015 standard deviation decrease in individual boys outcomes. At age 11, this effect drops to a 0.011 standard deviation decrease. The only other statistically significant effect is observed for girls in mathematics at age 7, with a 1 standard deviation increase in the proportion of girls associated with a 0.009 standard deviation increase in individuals' outcomes.

Table 6 Results by examination level in small schools

	English		Maths		Science	
	Male	Female	Male	Female	Male	Female
Age 7						
Proportion female	-0.132*** (0.029)	0.002 (0.027)	0.037 (0.032)	0.074** (0.029)		
Observations	15150	15150	15150	15150		
Age 11						
Proportion female	-0.091*** (0.035)	0.007 (0.033)	-0.030 (0.035)	0.018 (0.033)	-0.034 (0.035)	0.024 (0.035)
Observations	12600	12600	12600	12600	12600	12600

Notes: Dependent variable is the change in mean key stage score within school cohort for male (female) pupils. Robust standard errors in parentheses. *** denotes significance at the 1% level, ** denotes significance at the 5% level, * denotes significance at the 10% level. Method is ordinary least squares. Each cell represents a separate regression. Key stage 1 is not formally assessed for science. Year and exam dummies are also included. Standard errors are clustered at school level.

4.3 Value added results.

Table 7 shows the results of the value added regressions between ages 7 and 11. Considering the value added for English, the results suggest that both boys and girls

⁸ This is based on the standard deviation of the proportion of the within school cohort that is female in small schools, as shown in **Error! No se encuentra el origen de la referencia.**

achieve a higher value added score when they are in a class with a higher proportion of girls between the age of 7 and 11. Within small schools, there are no significant effects⁹.

Table 7 Value added results

	English		Maths	
	Male	Female	Male	Female
Proportion female	0.073** (0.036)	0.079** (0.034)	0.038 (0.034)	0.053 (0.033)
Observations	18374	18374	18374	18374

Notes: Dependent variable is the change in mean value added score within school cohort for male (female) pupils. Robust standard errors in parentheses. *** denotes significance at the 1% level, ** denotes significance at the 5% level, * denotes significance at the 1% level. Method is ordinary least squares. Each cell represents a separate regression. Key stage 1 is not formally assessed for science. Year and exam dummies are also included. Standard errors are clustered at school level.

The positive effect of a more female peer group on boys value added in English is a surprising one, considering that the effect on the levels achieved by boys is unambiguously negative. However, this may be evidence that boys who had initially been placed in a cohort with a large proportion of girls are catching up with their suffering from less of a disadvantage and are catching up with pupils who had not had the initial disadvantage. This story is consistent with the results presented in Table 5, with a less negative effect observed at age 11 than at age 7 for boys in English.

5. Discussion

In this paper, I have examined the causal effect of seemingly exogenous changes in the gender make-up of a child's within-school peer group using year to year changes in the proportion of girls within the school as an explanatory variable for the outcomes at ages 7, and 11.

The results show statistically significant negative effects of a more female peer group on the level of achievement at all ages for male pupils in English, robust across specifications, and a statistically significant positive effect for both boys and girls in mathematics and science at ages 7 and 11 respectively. Although the effect for boys in English may not be seen as a large effect, if we were to consider moving a boy from a female dominated classroom (with 75% girls) to a male dominated classroom (with 25% girls), this would suggest that for the pooled specification, individual boys outcomes would be raised by 0.033 standard deviations. Considering the effect observed in small schools at age 7, a similar move is associated with a 0.044 standard deviation increase in boys outcomes. These compares favourably with the introduction of the literacy hour in England, which raised childrens' outcomes by 0.08 standard deviations (Machin & McNally, (2008)). The value added results suggest that even though boys are disadvantaged at age 7 by having a more female peer group, those who remain in the same school up to the age of 11 catch up much of the deficit. These negative effects need to be weighed against positive effects for both boys and girls in mathematics and science, although the magnitudes of these effects are generally smaller than those observed for boys in English.

⁹ These results are available in Proud (2008)

Lavy and Schlosser (2007) identify positive effects from an increase in the proportion of girls within the peer group on outcomes, which they argue is an exogenous peer effect, due to the fact that the presence of more female peers lowers classroom violence, whilst improving inter-student and student-teacher relationships. This, they argue is not a result of endogenous changes in behaviour, but rather an exogenous effect, merely due to the gender make-up of the classroom. The results I find for mathematics and science seem to confirm the positive results from Lavy and Schlosser (2007), and also Hoxby (2000). Although the compositional effects suggested by Lavy and Schlosser would provide a convincing explanation to the positive effects observed for mathematics and science, the results for boys in English require a different explanation.

Whilst not a direct peer effect, the gender make-up of classrooms may lead to differential teaching methods within the classroom. Even though teachers may believe that they do not use different methods with girls and boys, Younger et al (1999) find evidence that boys and girls are treated very differently in the classroom. Students claim that boys receive more negative attention than the girls, and there is evidence that teachers have a lower tolerance level to boys' behaviour than to girls, which can "lead to male disillusionment and a negative reaction to learning". (Younger et al 1999, 339) However, they also comment that there is little evidence in observed lessons that boys are given "more support than girls in the teacher-learning process" (Younger et al 1999, 339). Furthermore, Dee (2007) finds that girls taught in a classroom with a female teacher and boys taught with a male teacher tend to perform better than pupils with a teacher of the opposite gender, suggesting that female teachers may direct learning in a way that is more likely to benefit girls rather than boys. This, when combined with Macleod (2005) who comments that only 15.7% of all primary school teachers in England are male and half of 5 to 11 year olds have no contact with male teachers implies that girls are likely to benefit more in education due to the gender of teachers.

A further possibility is related to the significant evidence that boys lag behind their female peers in English, but less so in mathematics and science. Recent studies have included Machin and McNally (2005) and Burgess et al (2004), both of which have found strong evidence of girls outperforming boys in English, but less so with maths and science. If an English class were to consist of a large proportion of (high performing) girls, and a small proportion (lower performing) boys, instruction directed at the average member of the class would be at a higher level than the boys were performing. As such, there is the danger that boys could be "left behind".

Since the methodology used here takes advantage of variation in the proportion of the cohort that is female, it is not possible to make definitive conclusions of the effect of educating pupils in single sex classrooms. However, the results obtained suggest that boys could benefit at all ages from being taught English in English schools with as small a proportion of girls as possible. In mathematics and science, the results shown here tend to imply that both boys and girls benefit from having more girls in the classroom, at least up until age 11. However, since there are only so many girls to go around, it is not possible to increase the proportion of girls that both boys and girls experience. This implies that a mix of the genders is optimal in both mathematics and science.

This research has identified a causal relationship, in England at least, that boys perform worse in English exams when they have a high proportion of girls in their peer group, whereas in mathematics and science, within primary schools, both girls and boys

benefit somewhat from a more female peer group. While Lavy and Schlosser (2007) have offered convincing evidence for the mechanisms in use in mathematics and science, it is not possible to disentangle whether these are exogenous effects, or whether an increase in the proportion of girls exhibiting a certain behaviour leads to changes in both boys' and girls' behaviour, leading to endogenous effects. As such, more research is needed to accurately pinpoint the mechanism driving the detrimental effect of female peers on boys' outcomes in English.

References

- Atkinson, A., P. Gregg and B. McConnell (2006) The Result of 11 Plus Selection: An Investigation into Opportunities and Outcomes for Pupils in Selective LEAs *CMPO Working Paper 06/150*, Centre for Market and Public Organisation, University of Bristol.
- Burgess, S., B. McConnell, C. Propper and D. Wilson (2004) Girls rock, boys roll: An analysis of the age 14-16 gender gap in English schools. *Scottish Journal of Political Economy*, **51**: 2, pp. 209-229.
- Clark, C. (2005) Do peers matter? Evidence from the sixth grade experiment, Mimeo, University of Kentucky).
- Coleman, J. S., E. Q. Campbell, C. J. Hobson, J. McPartland, A. M. Mood, F. D. Weinfeld and R. L. York (1966) Equality of Educational Opportunity, in: E. W. "US Department of Health (ed.) (Washington, DC., Office of Education).
- Dee, T. S. (2007) Teachers and the Gender Gaps in Student Achievement. *Journal of Human Resources*, **42**: 3, pp. 528-54.
- Dills, A. K. (2005) Does cream-skimming curdle the milk? A study of peer effects. *Economics of Education Review*, **24**: 1, pp. 19-28.
- Hansen, Z., H. Owan and J. Pan (2006) The Impact of Group Diversity on Performance and Knowledge Spillover - An Experiment in a College Classroom *NBER Working Paper* National Bureau of Economic Research).
- Hoxby, C. (2000) Peer Effects in the Classroom: Learning from Gender and Race Variation *NBER Working Paper*, National Bureau of Economic Research).
- Hoxby, C. (2002) The Power of Peers. *Education Next*, **2**: 2, pp. 57-63.
- Lavy, V. and A. Schlosser (2007) Mechanisms and Impacts of Gender Peer Effects at School *NBER Working Paper* National Bureau of Economic Research).
- Machin, S. and S. McNally (2005) Gender and student achievement in English schools. *Oxford Review of Economic Policy*, **21**: 3, pp. 357-372.
- Machin, S., and S. McNally (2008) The Literacy Hour, *Journal of Public Economics*, **92**:5-6, 1441-1462.
- Manski, C. F. (1993) Identification of Endogenous Social Effects: The Reflection Problem. *Review of Economic Studies*, **60**: 3, pp. 531-42.
- Mora, T. and P. Oreopoulos (2011) Peer effects on high school aspirations: Evidence from a sample of close and not-so-close friends. *Economics of Education Review*, **30**: 4, pp. 575-581.
- Proud, S. (2008) Girl Power? An Analysis of Peer Effects Using Changes in the Gender Make-Up of the Peer Group., *CMPO Working Papers 08/186*, Centre for Market and Public Organisation
- Summers, A. A. and B. L. Wolfe (1977) Do Schools Make a Difference? *American Economic Review*, **67**: 4, pp. 639-52.
- Whitmore, D. (2005) Resource and Peer Impacts on Girls' Academic Achievement: Evidence from a Randomized Experiment. *American Economic Review*, **95**: 2, pp. 199-203.
- Winkler, D. R. (1975) Educational Achievement and School Peer Group Composition. *Journal of Human Resources*, **10**: 2, pp. 189-204.
- Younger, M., M. Warrington and J. Williams (1999) The gender gap and classroom interactions: reality and rhetoric. *British Journal of Sociology of Education*, **20**: 3, pp. 325-341.