Scottish Journal of Political Economy, Vol. 51, No. 2, May 2004

GIRLS ROCK, BOYS ROLL: AN ANALYSIS OF THE AGE 14–16 GENDER GAP IN ENGLISH SCHOOLS

Simon Burgess, Brendon McConnell**, Carol Propper* and Deborah Wilson***

Abstract

We investigate possible explanations for the educational gender gap at age 16. We employ a national dataset of matched exam results of the cohort of pupils who took Key Stage 3 tests in 1999 and GCSEs in 2001. Our key result is the sheer consistency of the gender gap, across both the attainment and the ability distribution, with regard to both raw outcomes and value added. It is primarily driven by performance differentials in English. The generality of the gender gap suggests its source is not within-school practice, which means that policy directed at improving such practice may be misplaced.

I Introduction

The differential educational achievements of boys and girls in the UK have generated debate since the 1970s (OfSTED, 2003). Early work focused on raising the participation of girls, particularly in the traditional boys' subjects of maths, science and technology. The recent focus of this debate, however, has been the perceived underachievement of boys relative to their female peers – the so-called 'gender gap' – as measured in the UK by results attained at the four Key Stages of the National Curriculum. The gender gap at age 16 (when pupils take GCSEs) is seen as cause for particular concern, illustrated by the degree of media coverage given to the annual publication of results. While gender-related differences in performance are also of interest internationally (OECD, 2003; Elwood and Gipps, 1999; Shaw, 2002), we focus here on the gender differentials in English secondary schools in measured outcomes at age 16.

The purpose of this paper is to investigate more fully gender differentiated patterns of performance and to go beyond the 'average' statistics that grab the headlines. We employ a national dataset of the matched exam results of the entire cohort of approximately half a million pupils who sat the compulsory exams at age 14 (known as Key Stage 3 tests) in 1999 and took school leaving exams at age 16 (these are GCSEs or GNVQs) in 2001. We consider the gender

*University of Bristol, CASE, LSE and CEPR
**University of Bristol

gap, at subject level and in aggregate, with respect to three measures of attainment: the percentage of boys and girls gaining at least 5 GCSEs (or equivalent) at grade C or above (%5A*-C); their total GCSE points and a measure of value added between 14 and 16.[1] We investigate patterns of differential performance both across the attainment (performance at age 16 in GCSEs or GNVQs) distribution and across the ability distribution, as proxied by a pupil's average performance at Key Stage 3. We compare differences across different types of schools, in terms of good or poor performance, gender mix, admissions policy and percentage of pupils eligible for free school meals. We also consider whether Local Education Authorities (LEAs) have an impact on the gender gap in schools within their boundaries.

Our key result is that the gender gap is effectively constant across all the ways we cut the data. It is consistent across both the attainment and the ability distributions, whether we use raw GCSE test scores or value added as the outcome measure. It is not affected by school quality, nor by a wide range of observable school characteristics. We show that the gender gap is primarily driven by performance differentials in English. We find that it is negatively related to an increase in eligibility for free school meals (a marker of the poverty of pupils attending a school) and to the proportion of boys in the within-school cohort.

The rest of the paper is structured as follows. Section II briefly reviews the related literature and Section III details our dataset. Section IV presents our results and the final section concludes.

II Previous Literature

First, we review the evidence on the existence of and changes in the gender gap. We then consider both the potential explanations for and the possible strategies that have been put forward to reduce it.

In England in 2001, girls outperformed boys on average at each Key Stage (DfES, 2002a). In all Key Stage 1 tests, taken at age 7, a higher percentage of girls achieved the expected level than boys, with a much larger gap between their respective performances in reading and writing than in mathematics. By the age of 11 (Key Stage 2), girls outperformed boys in English and science, with boys doing slightly better in maths. At Key Stage 3 (age 14) there was no difference in performance in science, but the gender gap was maintained in the other subjects. In 2001 at Key Stage 4 (GCSEs/GNVQs), 44.8% of boys achieved 5 or more passes at grades A*-C compared to 55.4% of girls (DfES, 2002b). An analysis of 1995 data found a similar picture across all Key Stages (Arnot *et al.*, 1998). Stobart *et al.* (1992) found a gender difference in maths GCSE in favour of boys, but noted that this was steadily decreasing. Certainly by 1999, the gender differential in England in GCSE maths had been reversed (Atkinson and Wilson, 2003). Similarly, in Scotland, girls' relative performance in Higher Grade maths was better than boys for the first time in 1995 (Powney, 1996).

[1] This first measure is widely used, by the government and the media, to rank schools according to pupils' performance.

Subject level studies of the gender gap reveal it to be greater in English than in maths or science (Arnot *et al.*, 1998; Gorard *et al.*, 1999; DfES, 2000b; Atkinson and Wilson, 2003).

There is conflicting evidence on whether the gender gap is widening through time across successive cohorts. The gender gap appears to be narrowing among younger pupils (between KS1 and KS2) (OfSTED, 2003). Younger and Warrington (2002) identify a trend of the gap increasing at GCSE between 1988 and 1999, which concurs with the finding of Arnot *et al.* (1998) that girls' performance improved more rapidly than that of boys between 1984 and 1994. In contrast, Gorard *et al.* (1999) use Welsh data and find evidence that the gap, in terms of aggregate measures of achievement at each Key Stage, has decreased since 1992. With regard to within-cohort changes in the gender gap, Atkinson and Wilson (2003) provide some evidence that the gender gap widens – in favour of girls – during Key Stage 4, i.e. between the ages of 14 and 16.

There is also conflicting evidence regarding which type of students may be driving the average performance statistics. Gorard *et al.* (1999) find that the gender gap is not uniformly distributed across the range of attainments; rather it is primarily driven by boys' underachievement at the highest grades in any assessment. In contrast, Boaler *et al.* (2000) suggest that there is under-achievement by girls at the top end of the ability distribution that may be linked to the high pressure, high expectation environments in top-set classes. Younger and Warrington (2002), in a detailed analysis of one co-educational comprehensive school, find that the gender gap in terms of value added is only significant if students have a Cognitive Ability Test (CAT) score below level 4, i.e. towards the lower end of the ability range.

Research into possible explanations for the gender gap has primarily focused on within-school factors, rather than external factors such as ethnicity, social class, etc. (Salisbury *et al.*, 1999). The within-school factors that have been identified as possible (non-mutually exclusive) explanations include: modes of assessment (Arnot *et al.*, 1998); curriculum and question setting (Salisbury *et al.*, 1999); tiering or setting practices (Elwood and Murphy, 2002; Boaler *et al.*, 2000); peer group effects (Hoxby, 2000). More generally, there is a concern about the anti-learning 'laddish culture' that is seen to be pervasive in at least some schools.[2] The relationship between gender and performance is seen as complex and multi-faceted, often largely dependent on local context and conditions.

The possible explanations inform the many strategies that have been put forward to try and reduce the gender gap. Again, these have a predominantly within-school focus, and include: single-sex teaching, both at school and classroom level (Younger and Warrington, 2002; Sukhnandan *et al.*, 2000; Elwood and Gipps, 1999); ensuring gender-neutral modes of assessment, curriculum content and question setting (Arnot *et al.*, 1998); good teaching and classroom management (OfSTED, 2003); mentoring and the use of positive

[2] http://news.bbc.co.uk/1/hi/education/2208596.stm

role models (Sukhnandan *et al.*, 2000). There is no conclusive evidence of the effectiveness of any of these strategies (Salisbury *et al.*, 1999).

III Data

For this analysis we employ one of the national matched exam datasets released by the Department for Education and Skill (DfES). The dataset contains matched examination information for both Key Stage 3 (KS3) and GCSE/GNVQ, generally sat by pupils at the ages of 14 and 16 respectively. Individual exam results can be identified within the dataset, allowing complete information about the subjects studied by pupils, which is complemented by information on pupils' gender. We present results for the 1999–2001 cohort, but also utilise school performance data from the 1997–1999 cohort.

The focus of the analysis is on state maintained secondary schools in England. We omit independent and special schools, as well as other academic centres such as hospital schools and detention centres, from our dataset. The pupil level matched dataset is augmented with data from two other school level datasets. Information on pupil eligibility for free school meals (FSM, a marker of poverty amongst the school population) and school size is obtained from the Annual School Census (ASC),[3] whilst School Performance Table data further supplements the matched dataset with information on schools' admission policy (comprehensive, secondary modern, grammar school or City Technology College (CTC)) as well as funding status: community, foundation, voluntary aided or voluntary controlled (see Atkinson and Wilson (2003) for further details on these funding categories).[4]

We have matched examination data on over half a million pupils in 3,103 schools, 87% of which are co-educational. Summary statistics are presented in Table 1. Comprehensive schools make up the majority (88%) of our dataset, with secondary modern and grammar schools composing 6% and 5% respectively. There are only 15 City Technology Colleges (CTCs), all of which are co-educational. A larger proportion of girls (6.45%) attend single sex schools than boys (4.75%), resulting in boys constituting the larger component (51.6%) of co-educational school pupils. The means and standard deviations of the variables we use in the regression analyses are given in Table 2.

IV Results

We present our results, considering the gender gap across different cuts of our data. In England in 2001 girls outperformed boys by five points at GCSE on average – equivalent to one pass at grade C. We go beyond such averages, by examining gender-related differences in performance by ability (measured by attainment at Key Stage 3), both on aggregate and at subject level; by within-school cohort gender mix; by the performance of the school, and by a measure of deprivation experienced by the families at different schools. Since we have a very large dataset,

[3] Also known as 'Form 7'.

[4] CTCs are distinct from other schools both in terms of funding status and admissions policy.

Table 1
Summary statistics

Admission policy		Comprehensive	CTC*	Secondary modern	Grammar	Total
Boys only	Pupils	15844	0	1917	7483	25244
	Schools	*109*	*0*	*16*	*61*	*186*
Girls only	Pupils	23370	0	3011	7875	34256
	Schools	*145*	*0*	*20*	*62*	*227*
Co-educational	Pupils	444325	2548	19374	5280	471527
	Schools	*2490*	*15*	*144*	*41*	*2690*
Total	Pupils	483539	2548	24302	20638	531027
	Schools	*2744*	*15*	*180*	*164*	*3103*

Note:
*CTC = City Technology College.
Source: Department for Education and Skills.

Table 2
Means and standard deviations

Variable	Mean	Std. dev.
School FSM	0.16	0.14
Own KS3	32.63	6.30
School KS3	32.63	3.06
Own GCSE point score	40.85	18.50
School GCSE point score	40.18	8.86
Own value-added	– 0.60	10.06
School value-added	– 0.62	3.62
Cohort gender mix	0.51	0.17
Cohort size	192.91	64.68

Source: Department for Education and Skills.

and since we approach this problem with no strong priors, we first present the results graphically in order to impose as little structure as possible, and to allow any heterogeneity to appear. Finally, we confirm our findings more formally by regression analysis, at school and then at individual pupil level. We examine two rather different measures of outcomes at age 16: GCSE results in isolation, and value added from age 14. We present results for GCSE points score rather than the percentage of pupils achieving 5 GCSEs at grade C or above. Not surprisingly, as the latter is a summary statistic of the former, the two yield similar results.[5]

The distribution of the gender gap by ability

We do not have measures of exam performance earlier than Key Stage 3 (KS3) so proxy ability by attainment at KS3. We use each student's mean over the

[5] Graphs using %5A*-C as the outcome measure are available from the authors on request.

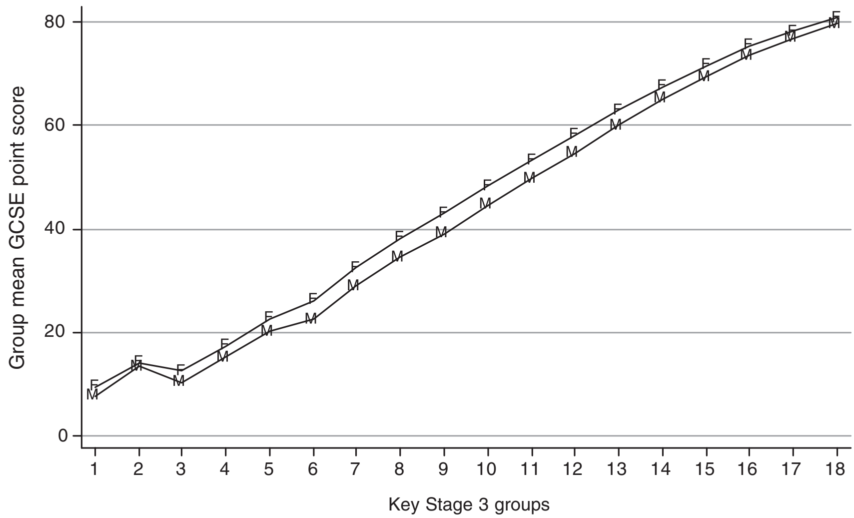

Figure 1. Mean GCSE point score by KS3 group.

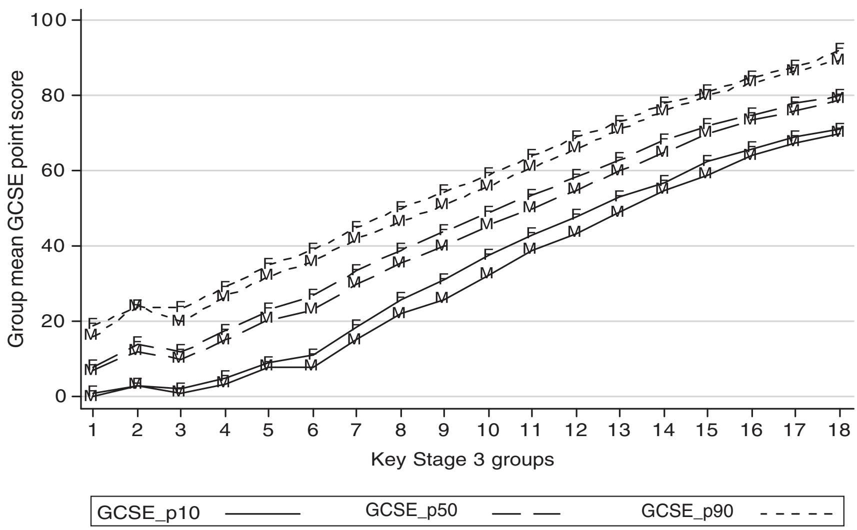

Figure 2. Percentiles of GCSE point score by KS3 group.

three subjects tested (English, maths and science). Figures 1 and 2 present the GCSE points outcome for each mean KS3 group, separately for girls (denoted 'F' in the graph) and boys (denoted 'M'); Figure 1 shows the mean, and Figure 2 some detail of the distribution. Looking at the mean first, we see that there is a clear and consistent gap of about 4 points throughout the distribution. The lines do not cross or touch at any point on the figure, indicating that for each ability group female students score more highly on average than male students. In fact,

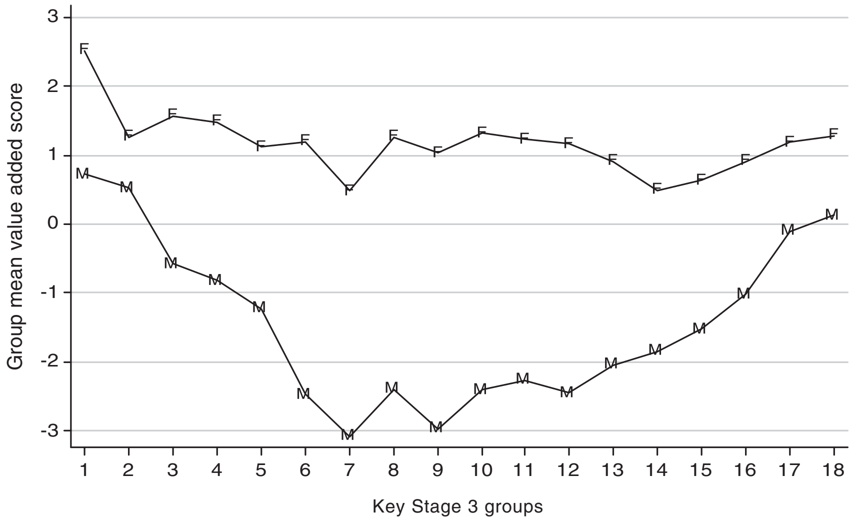

Figure 3. Mean value added by KS3 group.

between KS3 groups 6 and 14, the gap is almost constant. There is, if anything, some narrowing toward the tails of the distribution, a finding confirmed by later regression. Figure 2 shows that this pattern holds for the rest of the outcome distribution. For all but the very lowest two KS3 groups (which account for 1.3% of the distribution of prior attainment), the gender gap is clear and remarkably constant at the 10^{th}, 50^{th} and 90^{th} percentiles of the outcome distribution across the ability distribution. Within each KS3 group, high achieving girls do better than high achieving boys, and low achieving girls do better than low achieving boys, by about 3 and 4 points respectively.

This fact sits alongside the other clear outcome of the figure – namely that ability (prior attainment) has a very substantial impact on GCSE outcomes, and an impact that is much greater than the gender gap. For example, a girl with KS3 score at the 75^{th} percentile of the female distribution achieves a GCSE score better than 66.3% of girls; a boy with the same KS3 mean, scores better than 58.4% of girls. Thus the ‘ability gap’ is far greater than the gender gap, a fact that should not be overlooked in the policy debate on this issue.

Figures 3 and 4 repeat this exercise using as the outcome variable the value-added over the ages 14 to 16. Figure 3 shows the mean and Figure 4 some details of the distribution (Wilson, 2004, discusses how this value-added measure is calculated). Again the pattern is that the gender gap is always in the same direction.[6] It is clear that the overall gap in value-added is driven by the collapse of value-added for boys in the middle of the distribution: boys with average KS3 scores make much less progress over the age range 14–16 than do girls. The detailed distributional picture in Figure 4 backs this up, with the worst

[6] In only 3 of the 90 pairs of data points in Figure 4 do the lines touch, and never cross.

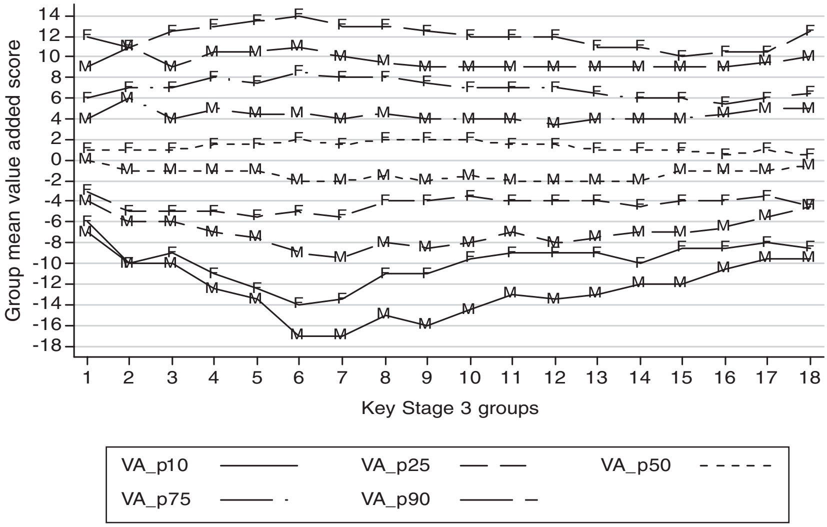

Figure 4. Percentiles of value added by KS3 group.

performing (bottom 10%) boys starting from average KS3 scores doing much worse than the worst performing girls starting from the same prior attainment.

Subject differences in the gender gap

The literature on the gender gap has noted differential performance in different subjects. We therefore analyse the gender gap in GCSE scores by subject and ability (prior attainment). It is important to avoid potential selection effects and so only consider subjects that are compulsory for all pupils to do, so we focus on English, maths and science.[7] Mean GCSE scores are shown by gender and ability in Figures 5 (English), 6 (maths) and 7 (science). It is clear from these that the English scores are driving the overall gender gap, with wide differences between the genders at all levels of the prior attainment distribution, and across the GCSE outcome distribution, though particularly at the lower quartile. However, in maths and science, the picture is rather different. For almost all prior attainment levels, the median boy and the median girl achieve the same maths GCSE score. There are also few differences at the lowest decile. But looking at the high performers (top decile) from each prior attainment level, boys score higher GCSE points in maths than do girls. The average gap is of the order of half a point. The same is true for science: no gap at the median, but boys outperforming girls at the 90th percentile. It is of course received wisdom that 'girls do better in English, boys do better in maths'. These results suggest that there is still some truth in that; whether this is simply the last remaining

[7] We cannot avoid selection effects completely as (some) pupils can choose to do double or triple science, or more than one maths paper, for example.

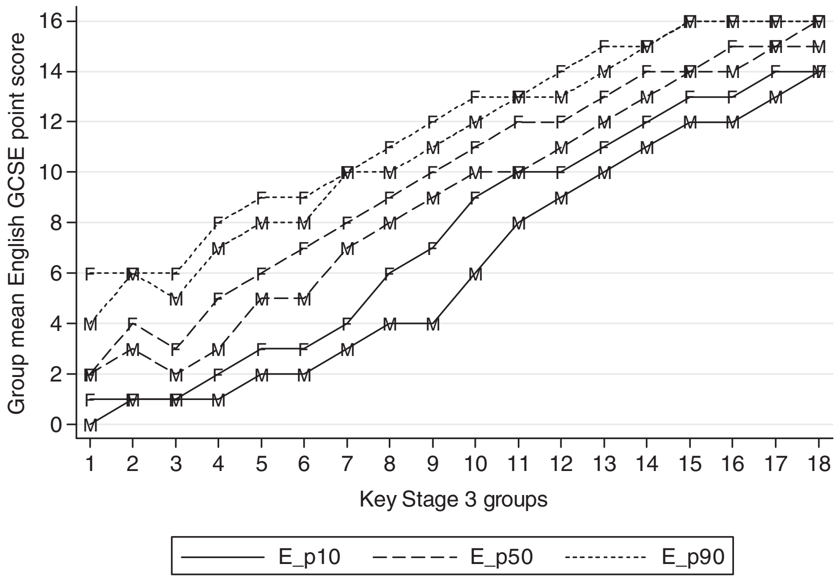

Figure 5. English GCSE score by KS3 group.

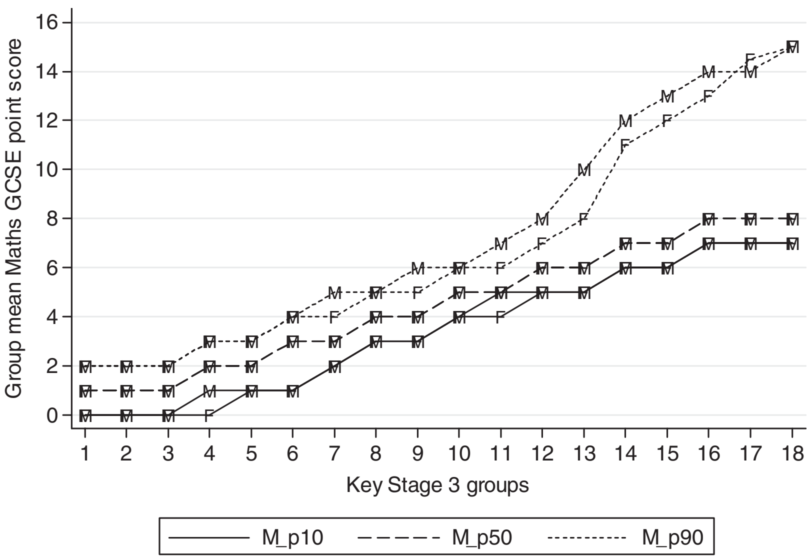

Figure 6. Maths GCSE score by KS3 group.

bastion of male superiority about to fall, or an enduring difference, awaits further data to tell.

Gender mix and the gender gap

We turn now to consider possible correlates of the gap. Given our data and the literature, our focus is on school characteristics. We begin with gender peer

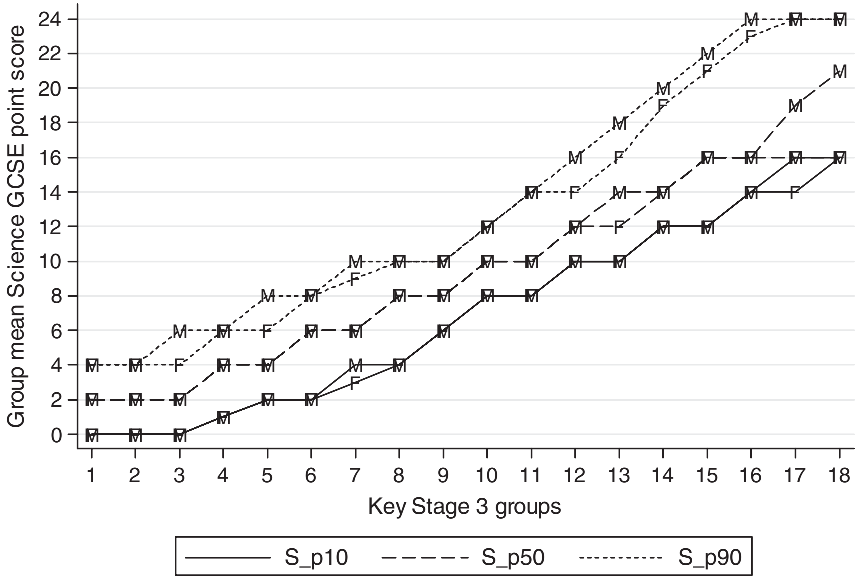

Figure 7. Science GCSE score by KS3 group.

group effects, measuring this as the proportion of a pupil's school-cohort that is female.[8] We separate out pupils in single sex schools (labelled 'G' and 'B' on the figures),[9] and split the remainder of the schools into 5 gender-mix categories, with schools in group 1 having the lowest percentage of boys in the cohort, and group 5 the highest. Note these refer to the gender mix of the 14–16 cohort we are studying in each school, not the school as a whole, so it is immediately relevant to the pupils.[10] We graph two measures of attainment – the school mean GCSE points score (Figure 8) and the school mean value-added (Figure 9).

It is clear that girls outperform boys across both measures of attainment and across the full spectrum of cohort gender mix. There is evidence of different outcomes for both boys and girls in single sex schools, but it is impossible to tell whether this is a 'single sex' effect or a 'grammar school' effect, given the high degree of correlation between single sex schools and grammar schools. Across all co-educational schools, the mean girl GCSE score exceeds the mean boy score. Interestingly, both genders perform less well in GCSE terms the higher the fraction of boys in the school, although the effect disappears in the value-added calculation. This suggests there is a negative peer effect from higher proportions of boys, the impact of which is accounted for by Key Stage 3 (age 14).

[8] See Hoxby (2000) for a discussion of the gender balance as a peer group effect.

[9] Note that very few of the schools in our dataset are single sex (only 4.75% of pupils are in boy-only schools, and only 6.45% in girl-only schools), and that being single-sex is highly correlated with being a grammar school, i.e. one that selects pupils on the basis of ability.

[10] We have repeated this procedure using instead the whole-school gender mix; the results are the same and are not reported here.

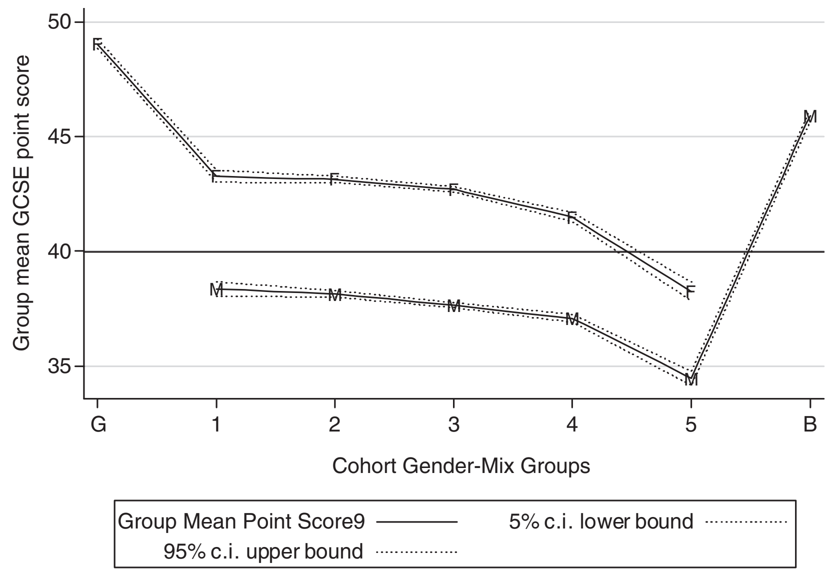

Figure 8. Gender group peer effects – GCSE point score.

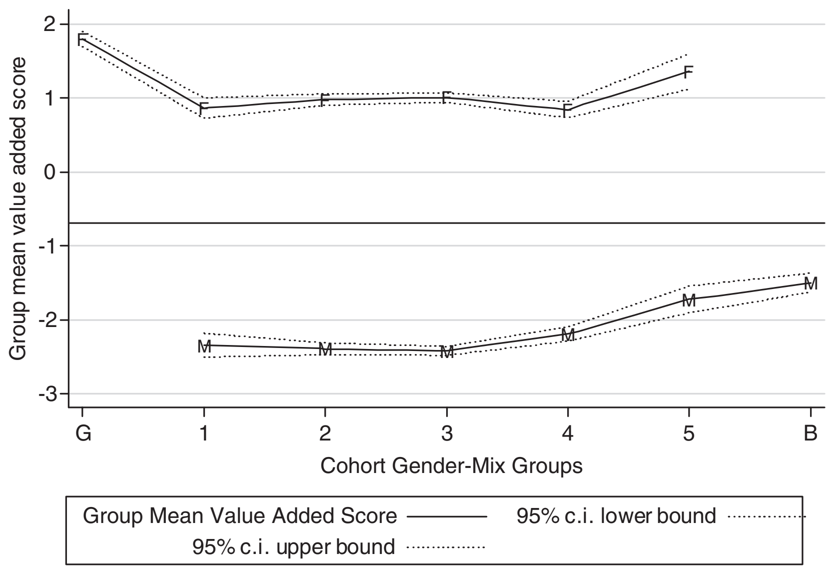

Figure 9. Gender group peer effects – value added.

The gender gap across good schools and poor schools

Is it the case that high performing or highly effective schools are able to reduce the gender gap relative to poorly performing schools? We investigate the extent

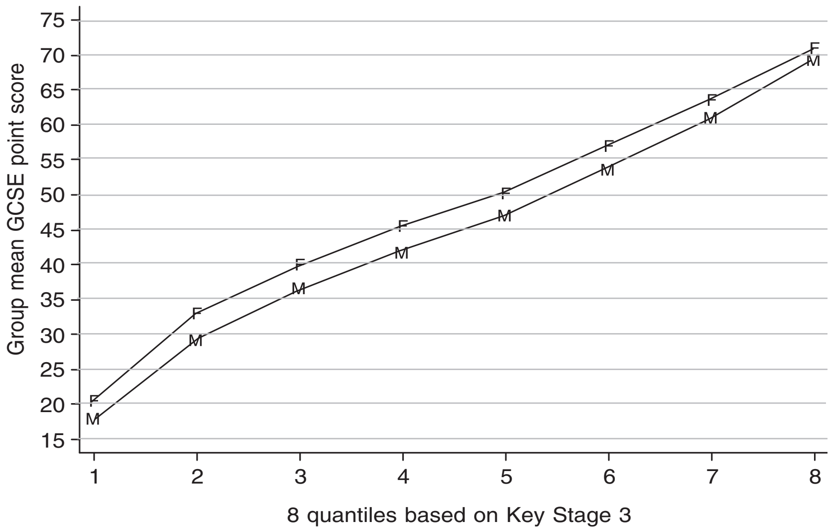

Figure 10. The gender gap in high performing schools.

to which the gender gap is related to observable indicators of school quality, rather than to gender differences *per se*. It may be that good teachers or good school procedures reduce or eliminate the gap: indeed, as our review of the literature illustrates, current thinking regarding strategies to reduce the gap is primarily focused on improving best practice within the school and classroom. We define high (poorly) performing schools as those in the top (bottom) 20% of the performance distribution as measured by the %5A*-C indicator, and highly effective (ineffective) schools as those in the top (bottom) 20% of the value-added distribution. In each case we use lagged performance data taken from the 1997–1999 cohort.

Figures 10 to 13 show that there is no substantial difference in the gender gap across these different schools, and across the distribution of pupils' abilities within those schools. That is to say, the distribution of the gap between girls' GCSE scores and boys' GCSE scores is about the same in schools with highly effective teaching, as it is in schools with ineffective teaching.[11] This is true at the mean, and also across the achievement distribution. Since the schools in the top performance category are different from the schools in the top effectiveness category (only 7.19% of the schools in our sample are in both), this is quite a strong result, and suggests that quality differences between schools have little impact on the gender gap.[12] We pursue this further below. A related question is the impact of school admission policy: whether schools

[11] The same result is obtained when we consider value added as the outcome measure.

[12] Similarly, the degree of overlap between poor performing and ineffective schools is only 7.06%.

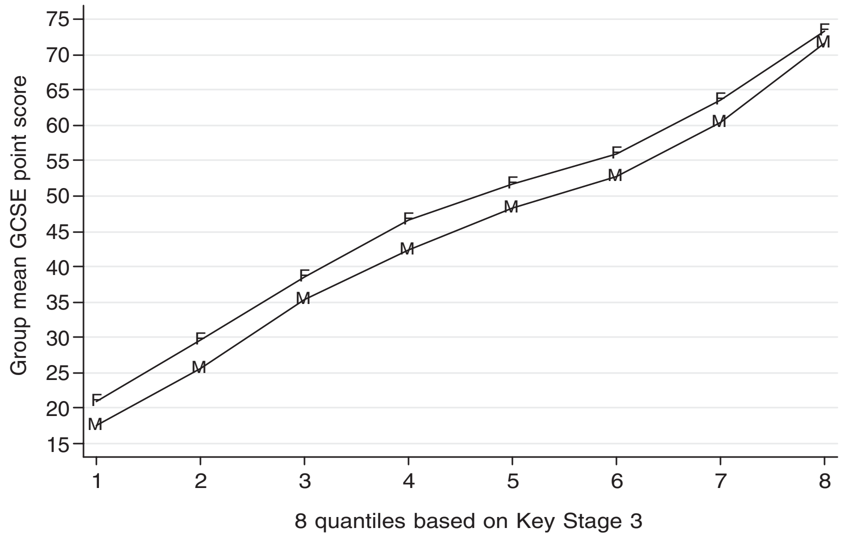

Figure 11. The gender gap in effective schools.

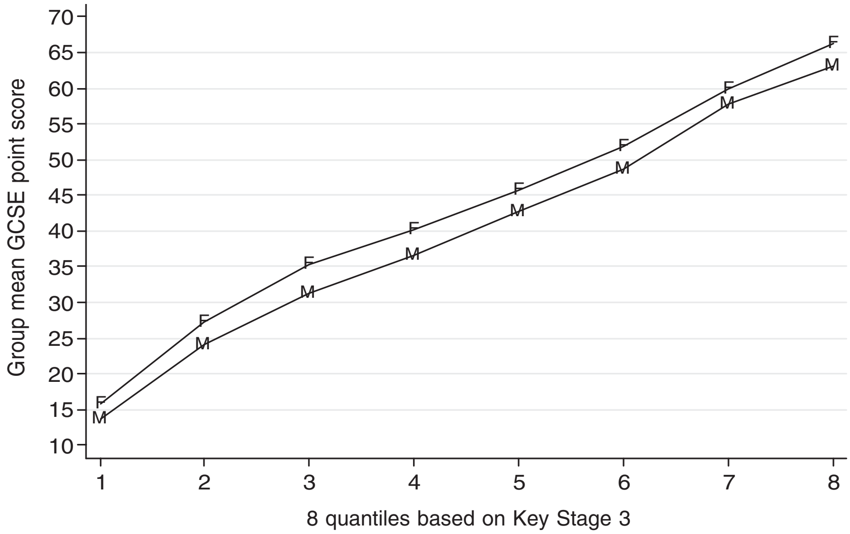

Figure 12. The gender gap in low performing schools.

are non-selective (comprehensive or secondary modern) or do select on ability (grammar schools). Again, our results show little impact on the gender gap of admissions policy (graphs not shown here, but see the regressions below).

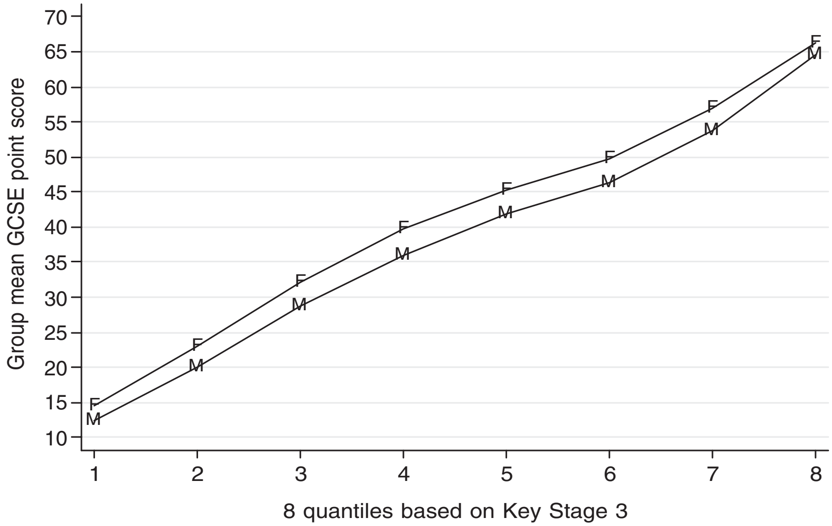

Figure 13. The gender gap in ineffective schools.

The gender gap and free school meals eligibility

It is well known that pupil attainment is to a considerable degree influenced by the home environment, and that poverty at home reduces achievement on average (see Sparkes, 1999, and references therein). The standard measure of this in the UK is the percentage of pupils at a school eligible for free school meals (FSM). We therefore examine whether this is correlated with the gender gap. If true, this would provide some suggestive evidence (no more) that boys and girls respond differently to a deprived home environment.

Figures 14 and 15 show that as expected, both school mean GCSE points and value-added decline as FSM eligibility increases. Note again that, just as with the 'ability gap', the impact of differences in FSM eligibility, the 'poverty gap', is much greater than the gender gap. But the differences in poverty have little impact on the size of the gender gap: the gap is 5 GCSE points in schools in the second decile of FSM eligibility, and is the same in schools in the bottom decile. Similarly, the gender gap in value-added is essentially invariant to FSM eligibility. This suggests that differential response to poverty levels at home is unlikely to be a major determinant of the gender gap.

Regression analysis

We summarise our discussion and confirm our findings more formally by regression analysis, at school and then at individual pupil level. First, we compute the school level gender gap, i.e. the mean difference between girls' and boys' GCSE total point score and regress this on a set of school characteristics. These include: the percentage of pupils eligible for free school meals (school FSM), the percentage for whom English is an additional or second language, the

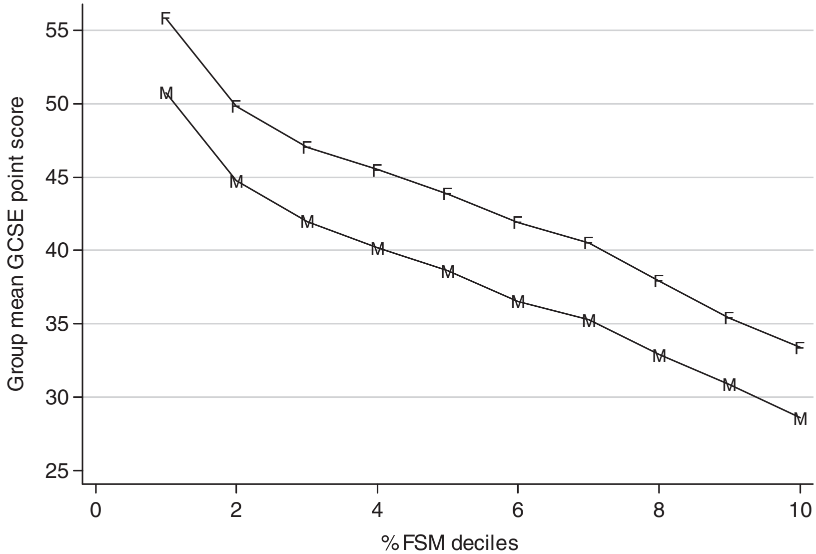

Figure 14. FSM eligibility and the gender gap – GCSE point score.

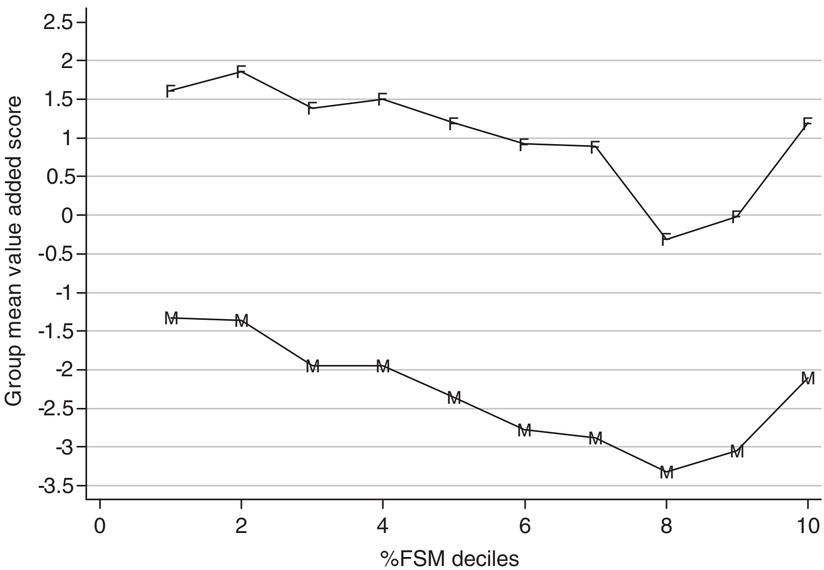

Figure 15. FSM eligibility and the gender gap – value added.

percentage of minority ethnic (non-white) pupils, admissions policy, religious denomination, funding status, number of pupils (school size), school mean Key Stage 3 (KS3) score and within-school cohort gender mix, as described in the sub-section above headed *Gender mix and the gender gap*. In a second regression

Table 3
Regressions of school level gender gap, with and without LEA dummies

Explanatory variables	Without LEA dummies		With LEA dummies	
School FSM	– 3.471***	(0.930)	– 4.838***	(1.293)
English as additional language	0.006	(0.009)	0.009	(0.010)
% ethnic minority pupils	0.789	(0.704)	0.671	(0.810)
Secondary modern	– 0.409	(0.315)	– 0.758	(0.461)
Grammar	– 0.557	(0.901)	– 1.189	(1.003)
Unspecified denomination	– 0.548	(4.010)	– 4.595	(10.46)
Church of England	– 0.043	(0.472)	– 0.026	(0.482)
Christian	0.278	(1.161)	0.544	(1.185)
Roman Catholic	0.005	(0.528)	0.105	(0.539)
Jewish	1.509	(2.015)	1.498	(2.050)
Sikh	– 17.226	(15.49)	– 16.597	(15.34)
7th Day Adventist	0.676	(6.376)	– 0.625	(6.405)
Foundation	– 0.418**	(0.176)	– 0.411*	(0.212)
Voluntary Aided	– 0.082	(0.494)	– 0.121	(0.508)
Voluntary Controlled	0.573	(0.391)	0.515	(0.402)
School size	0.001	(0.001)	0.001	(0.001)
School size squared	– 0.000	(0.000)	– 0.000	(0.000)
School KS3	1.267**	(0.543)	0.973*	(0.587)
School KS3 squared	– 0.019**	(0.008)	– 0.015*	(0.009)
Gender mix2	– 0.123	(0.252)	– 0.208	(0.257)
Gender mix3	– 0.101	(0.244)	– 0.081	(0.249)
Gender mix4	– 0.596**	(0.271)	– 0.568**	(0.277)
Gender mix5	– 0.998***	(0.363)	– 1.157***	(0.387)
Constant	– 15.459*	(8.997)	– 10.570	(9.840)
Observations	2660		2628	
R-squared	0.04		0.12	

Notes:
Standard errors in parentheses.
*Significant at 10%; **significant at 5%; ***significant at 1%.
Proportion of boys in cohort increasing through gender mix2 to gender mix5.

we include all the above plus a set of LEA dummies.[13] We might expect individual LEAs to have an impact on the gender gap. LEAs differ both in terms of education policy and across various other dimensions. Such differences may include: admissions policy; proportion of single sex schools; size; population density; whether predominantly urban or rural. The results of both regressions are shown in Table 3.

The first point to note is the extremely low R^2, equal to 4% in the first regression and rising to just 12% with the inclusion of the LEA dummies. While some variables are significant, the explanatory power is so low that we can say that the gender gap is essentially random across schools relative to the (quite rich) set of characteristics we are able to examine. Second, we turn to the

[13] We omit the 11 LEAs with less than five co-educational schools, which leaves us with 138 dummy variables. These are not included in the results reported here but are available on request from the authors.

Table 4
Pupil level regressions

	GCSE score		Value-added	
	OLS	FE	OLS	FE
Own KS3	2.5683***	2.5178***	– 0.1066***	– 0.1244***
Own KS3 squared	– 0.0017***	– 0.0007*	0.0006	0.0012***
School FSM	0.7422**	(dropped)	0.9372**	(dropped)
Gender	– 3.3956***	– 3.3706***	– 3.3019***	– 3.2810***
Gender*				
Own KS3	– 0.6887***	– 0.7221***	– 0.6930***	– 0.7239***
Own KS3 squared	0.0104***	0.0109***	0.0103***	0.0107***
School FSM	0.8599**	1.1257***	0.7008	1.0471**
School performance	– 0.0985***	– 0.1027***	– 0.1090***	– 0.1024***
School performance squared	0.0014***	0.0014***	0.0016***	0.0014***
School effectiveness	– 0.0264***	– 0.0337***	– 0.0251**	– 0.0324***
School effectiveness squared	– 0.0009	– 0.0002	– 0.0011	– 0.0004
Observations	498864	498864	498864	498864
R^2	0.73	0.71	0.10	0.03
Test of FEs		10.32		10.46

Notes:
*Significant at 10%; **significant at 5%; ***significant at 1%.

Other controls in the OLS regressions were: % ethnic minorities, % English additional language, admission policy, funding policy, religious affiliation, school size, gender mix, pupil age & LEA dummies.

Other interactions in the FE regressions were: % ethnic minorities, % English additional language, admission policy, funding policy, religious affiliation, school size, and pupil age.

significant variables. In the first regression, school KS3 mean score is significant and positive, while its square is significant and negative. In the second regression these are no longer significant but maintain the same signs. This reflects our earlier finding that the gender gap is greatest in the middle of the ability or prior attainment distribution. The FSM eligibility variable, school FSM, is significant and negative in both regressions, showing that controlling for all else, as poverty rises, the gender-related performance differential falls. Finally, again in both regressions, the highest two groups of gender mix are significant and negative: as the proportion of boys in the within-school cohort increases, the gender gap decreases. This could be due either to boys doing relatively better or girls doing relatively worse in a cohort containing a higher proportion of boys. LEA dummies are significant as a group. Little else approaches significance. We conclude from this that the gender gap is not greatly influenced by any of the widely used observable school characteristics.

In Table 4 we present the results of our pupil level regressions with some 0.49 million observations. We regress each pupil's outcome – GCSE score and value-added – on the pupil's gender, age, own prior attainment (KS3), and the set of school characteristics. We also include interactions of gender with own prior attainment and school characteristics to further investigate how the impact of gender may be modified by the learning environment.[14] We also run specifications with school fixed effects instead of school characteristics. Since LEA dummies are included, we drop the three LEAs with less than five schools. We also drop schools with less than 10 pupils in the year, due to the school being the unit of observation in the fixed effects regressions. Missing school level data also led to several thousand observations being dropped. In total, the sample cuts meant that the dataset was reduced from 531,027 to 498,864 observations, a cut of just over 6%.

The coefficient on gender in Table 4 tells us that girls outperform boys on average by approximately 3.4 points at GCSE, conditional on all other variables, including prior attainment.[15] This is a little less than the mean raw gender gap (5 points) because we are controlling for KS3, which also shows a gender difference. The signs on individual (own) KS3 and its square show a concave relationship between prior attainment and GCSE outcome. The percentage of pupils eligible for free school meals shows a surprising positive coefficient. However, this is very small and barely significant. If we drop own prior attainment from the regression, this variable becomes significantly negative as expected and is 16 times larger. This suggests that its impact is accounted for by the time pupils reach KS3.

Consider now the gender interaction terms. The interaction with FSM eligibility confirms the above finding that as the proportion of low income pupils in a school increases, the gender gap falls. The gender gap increases as both school performance and school effectiveness increase; the former at a decreasing rate. Note finally that the R^2 is much higher for the GCSE regressions than for value added, confirming both the predictive power of prior attainment and that pupil progress is more difficult to explain.

Given the size of our sample, it is not surprising that these variables are highly statistically significant. What we are interested in, however, is the *magnitude* of these influences on the gender gap. Consider Table 5. In each row, we show how the gender gap varies as the named variable changes, holding all other influences at their mean values. Note that if all distributions were symmetric (so mean equals median), the values in the 50^{th} percentile column would be equal. The greater gender gap in the middle of the KS3 distribution is again confirmed. The key point, however, is just how small the effects of any of the variables are on the gender gap, even across the extremes of the distributions of own prior attainment and school performance. Even though they are statistically significant, quantitatively they have no real impact. We repeated this analysis at subject level for English, maths and science. The mean size of the gender gaps were -0.55, 0.13 and 0.10 respectively. The only differences in respect to influences on the size of the gender gap were, first, the gender gap in science *widened* towards the tails of the KS3 distribution and, second, an

[14] These interactions are included net of their means so that the coefficient on gender itself is straightforwardly interpretable as the mean conditional gender gap.

[15] Note that girl = 0 and boy = 1 in the gender dummy variable.

Table 5
The impact of school and individual characteristics on the gender gap

	Percentiles of the distribution:				
	10th	25th	50th	75th	90th
Own KS3	– 2.65	– 3.18	– 3.37	– 3.21	– 2.71
School FSM	– 3.52	– 3.49	– 3.42	– 3.30	– 3.13
School performance	– 3.23	– 3.29	– 3.37	– 3.45	– 3.53
School effectiveness	– 3.28	– 3.36	– 3.37	– 3.30	– 3.14

Note:
Each row shows how the gender gap varies as the named variable changes, holding all other variables constant at their mean values.

increase in FSM eligibility *increased* the science gender gap. The results for English mirrored the results for total GCSE score reported above, confirming that it is the gender gap in English that drives the overall performance differential between boys and girls.

These results confirm our finding that the gender gap is not influenced by any of the leading observable characteristics, including measures of school quality. This makes it very hard for policy to be designed to address it. It also suggests that the source of the gap is not related to the behaviour of schools and teachers, but is more generic. Whether this is societal or physiological in nature is beyond the scope of this study.

V Conclusion

The aim of this paper was to examine gender related differences in performance at age 16, both in terms of GCSE results and in terms of the value added between the ages of 14 and 16. We have investigated a number of possible explanations for the underachievement of boys relative to girls. Our striking result is the sheer consistency of this gender gap, across both the attainment and the ability distribution, on aggregate and at subject level, with regard to both raw outcomes and value added. We show that it is not related to whether the school performs well or poorly, or whether it is effective or ineffective. Nor is it affected by a wide range of other observable features of schools such as size, admissions policy, religious denomination or funding status.

There is a clear difference between subjects. The gender gap is primarily driven by performance differentials in English, while boys and girls are still obtaining similar results in maths and science. This may be the result of a slow moving socialisation process, in which case we might expect girls to eventually outperform boys in these 'traditionally male' subjects as well. Or it could be that the different cognitive demands and processes required by the subjects is giving us a clue that the gender gap is rooted in different rates of cognitive maturation between boys and girls, that itself happens at varying rates for different cognitive processes.

Our analysis suggests that the 14–16 gender gap is something very general and is not much affected by any of the leading observable school characteristics. This suggests that the source of the gap is not within-school practice, which in turn means that policy directed at improving such practice may be misplaced. In fact, given our findings regarding the size of both the 'ability' and the 'poverty' gaps relative to the gender gap, focus on the reasons behind these performance differentials may lead to better results in terms of improvements in both boys' and girls' educational attainment. With the release of the national pupil level annual schools census (PLASC), which contains pupil level information on free school meal eligibility and ethnicity as well as gender, we aim to further investigate more fully the patterns of differential performance of different types of pupil.

Acknowledgements

The authors would like to thank the Leverhulme Trust for funding this research through the Leverhulme Centre for Market and Public Organisation, the Department for Education and Skills for supplying the data, and the editors for helpful comments on a previous draft. The views expressed in this paper are the authors' own.

References

Arnot, M., Gray, J., James, M., Rudduck, J. with Duveen, G. (1998). *Recent Research on Gender and Educational Performance*. OFSTED Reviews of Research, London: HMSO.

Atkinson, A. and Wilson, D. (2003). Measuring pupil attainment in English secondary schools: a preliminary analysis. CMPO: University of Bristol, Working Paper 03/063.

Boaler, J., William, D. and Brown, M. (2000). Students' experiences of ability grouping – disaffection, polarisation and the construction of failure. *British Educational Research Journal*, **26**, 5, pp. 631–48.

DfES (2002a). National curriculum assessments of 7, 11 and 14 year olds in England – 2001. http://www.dfes.gov.uk/statistics/DB/SBU/b0333/index.html

DfES (2002b). GCSE/GNVQ and GCE A/AS/VCE/Advanced GNVQ examination results 2000/2001 – England. http://www.dfes.gov.uk/statistics/DB/SBU/b0334/index.html

Elwood, J. and Gipps, C. (1999). *Review of Recent Research on the Achievement of Girls in Single-Sex Schools*. Institute of Education: University of London.

Elwood, J. and Murphy, P. (2002). Tests, tiers and achievement: gender and performance at 16 and 14 in England. *European Journal of Education*, **37**, 4, pp. 395–416.

Gorard, S., Rees, G. and Salisbury, J. (1999). Reappraising the apparent underachievement of boys at school. *Gender and Education*, **11**, 4, pp. 441–54.

Hoxby, C. (2000). Peer effects in the classroom: learning from gender and race variation. NBER Working Paper 7867. Cambridge, MA: NBER.

OECD (2003). Literacy skills for the world of tomorrow: further results from PISA 2000. http://www1.oecd.org/publications/e-book/9603071E.pdf

OfSTED (2003). Boys' achievement in secondary schools. http://www.ofsted.gov.uk/publications/docs/3316.pdf

Powney, J. (1996). *Trends in Gender and Attainment in Scotland*. Scottish Council for Research in Education. http://www.scre.ac.uk/spotlight/spotlight64.html

Salisbury, J., Rees, G. and Gorard, S. (1999). Accounting for the differential attainment of boys and girls at school. School Leadership and Management, **19**, 4, pp. 403–26.

Shaw, L. (2002). Schools' gender gap concern now is boys. *Seattle Times*, 30/09/02. http://seattletimes.nwsource.com/text/134545336_gender30m.html

Sparkes, J. (1999). Schools, Education and Social Exclusion. CASE Working Paper 29. LSE: CASE.

STOBART, G., WHITE, J., ELWOOD, J., HAYDEN, M. and MASON, K. (1992). *Differential Performance at 16+: English and Mathematics*. London: HMSO.

SUKHNANDAN, L., LEE, B. and KELLEHER, S. (2000). *An Investigation into Gender Differences in Achievement, Phase 2: School and Classroom Strategies*. National Foundation for Educational Research. www.nfer.ac.uk

YOUNGER, M. and WARRINGTON, M. (2002). Single-sex teaching in a co-educational comprehensive school in England: an evaluation based upon students' performance and classroom interactions. *British Educational Research Journal*, **28**, 3, pp. 353–74.

WILSON, D. (2004). Which ranking? The impact of a 'value added' measure of secondary school performance. *Public Money and Management*, forthcoming.