DO THE MATHS:

AN ANALYSIS OF THE GENDER GAP IN MATHEMATICS IN AFRICA

Andy Dickerson, Steven McIntosh and Christine Valente

Department of Economics

University of Sheffield

Preliminary Draft

September 2011

Abstract

This paper uses data on 19 African countries to examine the gender difference in maths test scores amongst primary school children. In almost all countries studied, there is a small, but significant, difference in maths test scores in favour of boys. This difference cannot be explained in terms of a wide range of pupil, family and school level characteristics, including both school heterogeneity and within-school variation in access to educational inputs. In country-level analysis, however, differences in female sexual and reproductive status account for about half of the cross-country variation.

Keywords: Cognitive maths skills; gender differentials; developing countries

JEL Codes: O15, I20

Corresponding author: Christine Valente E: christine.valente@sheffield.ac.uk; T: +44 (0)114 222 3412; F: +44(0)114 222 3458

Do the maths: An analysis of the gender gap in mathematics in Africa

1. Introduction

There is widespread evidence of the existence of a female disadvantage in performance in mathematics tests in high- and middle-income countries (for examples relying on large-scale, arguably representative surveys, see Hedges and Nowell 1995; Guiso et al. 2008; Fryer and Levitt 2010). The causes behind these differences have been the subject of much debate. Natural scientists have pursued explanations based on evidence of genetic differences between genders, e.g. that male performance in quantitative and spatial ability tests is more dispersed than that of women, so that larger numbers of men have unusually high scores. Social scientists and economists have, on the contrary, found explanations based on cultural values. In 39 middle- to high-income countries taking part in the 2003 Programme for International Student Assessment (PISA), Guiso et al. (2008) show that up to a third of the variation in the female maths gap can be accounted for by differences in GDP per capita and in indicators of gender inequality. However, Fryer and Levitt (2010) show that this finding does not hold in all the countries from the 2003 TIMMS[1] evaluation exercise, and point to a potential explanation based on the presence of a few outliers who are characterised by the dominance of single-sex schooling, and who exhibit both high female performance in maths and low levels of female empowerment. In the United States, Fryer and Levitt (2010) find that, while there is no difference between girls and boys when they enter school, by the end of the fifth year, boys have moved 0.2 standard deviations ahead of girls in maths. Their results suggest that the gap is not caused by less investment in maths by girls, by low parental expectations of girls in maths, or by testing mechanisms that are biased towards girls. Interestingly, the gap they observe by the end of the fifth year of schooling is very similar to the gap observed by Goldin et al. (2006) in a nationally-representative sample of twelfth graders in 1992 (0.17 s.d.). Pope and Sydnor (2010) however emphasise that this small average difference hides much wider discrepancies at the top of the distribution, as in the same sample, they calculate a 1.40 male-to-female ratio in the top 5% of maths scores. In a similar vein to the study by Guiso et al. (2008), Pope and Sydnor (2010) find that 40% of the variation across states in stereotypical gender differences in 8th-grade test scores can be

[1] Trends in International Mathematics and Science Study.

accounted for by differences in survey responses to a question about gender roles.[2] Goldin et al. (2006) document the increase in the ratio of female-to-male US college graduates during the second half of the 20th century, which they attribute to changes in attitudes towards female labour force participation, marriage, motherhood, and access to effective contraception; and show that the female progress in college education was accompanied by an increase in the uptake of- and performance in maths and science courses in high school.

In developing countries, and especially in Sub-Saharan Africa, the main concern of researchers and policy makers with respect to the education of girls has been to find ways to allow them to get one at all. Even at primary schooling age, African net enrolment rates (NER) are far from unity in many countries, and especially so for girls. In 2003, the ratio of female to male enrolment in primary education in Sub-Saharan Africa was just 68% (World Bank Databank, Education Statistics). While in several countries included in the present analysis, enrolment rates where near unity and achieved gender parity around the time of the analysis (e.g., the primary NER was 99% for boys and 98% for girls in 1999 Cameroon), in many other countries male enrolment rates were below 70% and those of girls substantially below that (e.g., Benin, Chad and Guinea) (UNESCO 2007). The gender gap is also marked when considering primary *completion* rates. Lewis and Lockheed (2008) report that in 2005 the gender gap in primary completion rates was highest in Sub-Saharan Africa at 10 percentage points (56% versus 66%) than in any other cluster of developing countries. This state of affairs has led to the United Nations setting as a Millennium Development Goal the elimination of gender inequality in enrolment in primary and secondary education, and female education enrolment rates have indeed been rising in a number of developing countries.

Both demand-side and supply-side arguments have been put forward to explain why girls have lower enrolment or achievement in developing countries. On the demand side, different schooling investments may arise for boys and girls if costs, benefits, or preferences are gender-specific. Differences in direct costs should be minimal in co-educational settings, but may be important where single-sex schooling is the norm. Opportunity costs are likely to vary with gender in many developing countries, e.g., if a daughter is of greater use than a son

[2] Pope and Sydnor (2010) define a stereotypical adherence index which is equal to the average male-female ratios in math and science and the female-male ratio in reading for the top 5 percent of students. The attitudinal question used in their analysis comes from the General Social Survey, which asks: “Is it much better for everyone involved if the man is the achiever outside the home and the woman takes care of the home and family?”.

performing domestic and caring duties around the home. Benefits may be lower if girls expect lower earnings in adulthood (as found to be the case in India by Rosenzweig and Schultz 1982). This could be because of outright discrimination and lower wages for the same work, crowding into lower paid work, or less time spent expected in paid work if traditional gender roles about child-rearing and gender roles within families are continued. Furthermore, there may be gender-specific parental benefits from given earnings if tradition dictates that only offspring of a certain gender is expected to look after their parents in their old age. Or cultural norms may lead to a gender-specific, subjective valuation of educational attainment. On the supply side, education provision may vary by gender both in terms of quantity and quality. In the Sub-Saharan African countries studied here, almost all schools are co-educational, so that widespread differences in school quantity would only arise if girls were rationed out of schools in theory opened to both genders. However, gender differences in schooling quality can arise within school. Teachers could have lower expectations of girls, or simply be prejudiced against them, and so give them less attention or access to fewer resources in schools. Coclough et al. (2000), for instance, relate case studies in Ethiopia and Guinea in which girls were found to carry out more non-school activities during school hours, such as cleaning the classroom, offices and latrines, and fetching water.

A large number of studies have estimated the role of demand-side factors on gender differences in enrolment, and consistently found that the enrolment of girls is more sensitive to changes in household income than that of boys (e.g., Glick and Sahn 2000 for Guinea; Bjorkman 2006 for Uganda; Sackey 2007 for Ghana). Fewer studies have considered the role of supply-side factors, but recent randomized controlled trials show that the addition of female teachers (Banerjee et al. 2005), and the introduction of community schools (Burde and Linden 2010) may both contribute to bridging the enrolment gap.

Recent evidence has showed that cognitive ability matters more for individual labour market outcomes and macroeconomic growth than years of education completed (Hanushek and Woessman 2008). Estimates of the impact of cognitive skills on earnings are scarce outside developed countries. However, returns to cognitive skills are generally estimated to be large, and where maths and reading skills are considered separately, maths skills appear to matter more for income (Glewwe 1996 and Jolliffe 1998 for Ghana, Moll 1998 for South Africa). As gender gaps in school enrolment decrease, it therefore becomes important to consider inequality in the acquisition of cognitive skills during the time spent at school, and especially so for maths. The literature considering gender differences in test scores in developing countries is sporadic. To the best of our knowledge, there are only two exceptions in the economics literature. Appleton (1995) studies gender differences in pass rates in the examination at the end of primary schooling in Cote d'Ivoire and shows that the female pass rate is lower and much more sensitive to household expenditure than that of boys. At very high levels of household consumption, girls actually outperform boys. Alderman *et al.* (1996) explore the gender gap in cognitive skills amongst individuals aged 10-25 in Pakistan. Conditional on having started school, females obtain lower numeracy scores, but higher literacy scores. Analysing the gender gap in cognitive ability in their entire sample, Alderman *et al.* (1996) find that the dominant explanation lies in differences in the demand, by parents, for primary education between the genders, including the larger sensitivity of demand for female schooling to local availability of schools and distance to the nearest school. In the education literature, Nguyen (2006) finds that, in Vietnam, girls aged 6-18 have higher overall grades when enrolled in school, which she argues is due to a selection effect.

The direction, if any, of gender differences in maths tests in Sub-Saharan African schools is unclear *a priori*. On the one hand, lower levels of female empowerment would lead to an expectation of wider gaps on the basis of the existing literature focussing on richer countries. On the other hand, gender biased enrolment rates would tend to lead to more positive selection into schooling for girls than for boys, which could thus lead to smaller gaps in maths scores.

In this paper, we document systematically the existence of a gender gap in mathematics for primary school pupils in 19 Sub-Saharan countries. We test a wide range of hypotheses which could account for the gender gap in maths: gender differences in schooling quality (between- and within-class), in observable parental educational inputs, school attendance, coeducational schooling, and cross-country differences in female status in public and in private life. Child-specific educational inputs cannot explain the poorer performance of girls. Single-sex schools eliminate the female disadvantage in some but not all countries. Female empowerment in the economic and political spheres does not improve the relative performance of girls in maths. On the contrary, female status in the private sphere, as proxied by the total fertility rate and by the prevalence of sexual harassment of pupils, explains up to half the gender gap in maths.

The next section describes the two African data sets that we use, and the empirical methodology that we employ. The results of the micro-analysis are presented in Section 3, while Section 4 presents the results of the cross-sectional analysis. Section 5 concludes.

2. Data and Methodology

Our data come from two pan-African surveys of educational quality, the Southern and Eastern African Consortium for Monitoring Educational Quality (SACMEQ)[3] and the Program for the Analysis of Education Systems (PASEC), which, taken together, cover 19 Western, Southern and Eastern African nations.

(i) SACMEQ

SACMEQ is a consortium of 15 Ministries of Education in Southern and Eastern Africa.[4] We utilise data for 12 countries from the second survey (SACMEQ II) which gathered data from 14 countries in Southern and Eastern Africa between 2000 and 2002.[5] The focus of the survey is on assessing the conditions of schools and the quality of primary education. Information is gathered from pupils and teachers using survey questionnaires together with tests of their numeracy and literacy. Characteristics of their schools are also recorded. In total, more than 40,000 pupils were surveyed in over 6,100 classes spread between almost 2,300 schools.

Three separate survey questionnaires are administered to Standard Grade 6 pupils, their teachers and the school headteacher. Information on the school's facilities as well as on the education and experience of the headteacher is gathered from the headteacher questionnaire. The teacher questionnaire provides details on the class teacher's education and experience, the classroom characteristics and equipment available, as well as the methods used to teach literacy and numeracy. Finally, the pupil questionnaire asks about their home environment, the input and involvement that others have in their education both within and outside school, and also on the school equipment that they have at their disposal. The same questionnaire is administered across all countries.

To obtain the test score results, a maths test was given to the pupil respondents. For SACMEQ II mathematics, some items were based on TIMMS while other items were newly

[3] We use version 6.0 of the SACMEQ II data (dated October 2009): see Ross *et al.* (2004) for details.
[4] Botswana, Kenya, Lesotho, Malawi, Mauritius, Mozambique, Namibia, Seychelles, South Africa, Swaziland, Tanzania (Mainland), Uganda, Zambia, Tanzania (Zanzibar) and Zimbabwe.
[5] We exclude Zanzibar to focus on mainland Tanzania only, and the Seychelles, on account of the much higher standard of living in this country.

written by the individual country National Research Coordinators (NRCs). In each country, prior to the data collection, a committee of specialists and Grade 6 teachers was formed in order to review each of the test items and establish the sub-set of items considered to be 'essential items' to be mastered if the pupils were to be able to undertake a successful programme of study at Grade 7.

(ii) PASEC

PASEC is the evaluation tool used by CONFEMEN, a network of Education Ministers from 41 Francophone countries. The programme consists of a number of national surveys using a combination of headmaster, teacher, and pupil questionnaires, as well as unified maths and (French) literacy tests. The questionnaires used cover the same themes as in SACMEQ, albeit in less detail. They are adapted to each country by a national team (e.g., in terms of the household items which the pupils are asked to say whether they have one in their home), but most questions are the same across countries.

Similar to SACMEQ, test results are comparable across countries within PASEC, but not comparable to SACMEQ results.[6] Indeed, test items are based on common aspects of the curricula across PASEC countries. PASEC data are collected by national teams, with technical support from the CONFEMEN central office and a scientific committee.

Compared to SACMEQ, PASEC surveys have the advantage of testing both children who are in Grade 2 and those in Grade 5 at the time of the survey, and of administering two tests at each level, one at the beginning of the year, and another one at the end of the year, which allows for an analysis of value-added.

PASEC surveys have been carried out since 1995, but data quality is only considered high for surveys carried out from 2003 onwards (http://www.confemen.org/spip.php?article272). We therefore focus on these data. More specifically, we use data from seven countries surveyed between 2003 and 2006, leading to a sample of nearly (over) 14,000 Grade 5 (Grade 2) pupils in nearly 1000 schools and as many classes.[7]

These sources of data have been used in previous research estimating the effectiveness of inputs in primary education. Fehrler *et al.* (2009) use both SACMEQ and earlier PASEC

[6] Future PASEC surveys will include a number of items comparable to SACMEQ (personal communication from PASEC).

[7] The seven countries in question are: Benin, Cameroon, Chad, Gabon, Guinea, Madagascar, and Mauritania.

data, while Michaelowa (2001) uses earlier PASEC only, to estimate education production functions, pooled across countries within each survey. Both studies apply a hierarchical linear model (HLM) framework to their data to consider the effect of individual as well as school- or class-level characteristics on test scores, and, in the case of Michaelowa (2001), national characteristics. The two models differ in that Michaelowa (2001) includes country *random effects* but Fehrler et al. (2009) include country *fixed effects*. Contrary to the present analysis using school fixed effects, however, neither model allows for omitted variables correlated with school fixed effects, or, in case of Michaelowa (2001), with country fixed effects. With this caveat in mind, Michaelowa (2001) finds no gender differences in combined maths and reading scores, but show a positive correlation between combined test scores and, among others, (i) having literate parents and coming from households owning books and more household items such as a radio and television, and eating meals regularly, (ii) teachers of higher quality, better attendance and self-reported motivation, (iii) schools with books and more basic equipment and *larger* class sizes (up to 62 pupils), (iv) countries with larger education expenditure per student and higher current literacy rates. Fehrler et al. (2009) find positive correlations between test scores and availability of textbooks, but not teacher education and training (at least in PASEC countries), while a negative correlation is observed between test scores and teaching in shifts.

Since the tests taken in the SACMEQ countries and the PASEC countries are different, and to enable comparison of results between the two countries, the dependent variable is measured as the percentage of questions answered correctly.

3. Results

(i) Descriptive Statistics

Table 1 (a and b) contains descriptive statistics on all of the variables used in the analysis, separately for the PASEC (Year 5) and SACMEQ samples respectively, and in both cases separately for boys and girls. The first rows in each case show the scores on the maths test for boys and girls. In the PASEC results, boys score more highly than girls at both the beginning and the end of year 5, by 2.1 and 0.5 percentage points respectively, only the former being statistically significant. In SACMEQ, the difference is a statistically significant 1.4 percentage points.

The remaining rows in Tables 1a and 1b contain information on the explanatory variables used in the analysis, and shows how they differ between boys and girls. Looking first at the PASEC (Year 5) results in Table 1a, we can see that the boys in the sample are slightly, but statistically significantly, older than the girls. In terms of household wealth, there is no statistically significant difference between the genders, but most of the other indicators of family background work in favour of girls. Thus, girls are more likely than boys to have a literate father, a literate mother, to receive regular daily meals, to receive parental help and help from a tutor with school work. Family background is therefore generally more conducive towards good outcomes for girls than for boys. Thus, when we control for family background in the multivariate analysis that follows, we should expect to see the gap in test scores between the genders increase. This should not be a surprise, given some of the previous research discussed in the Introduction above. Studies in developing countries have repeatedly found that measures of family income have a positive effect on girls' enrolment in school, but no effect on boys enrolment. Thus, most boys participate in education whatever their background, while for girls they are more likely to participate if they come from a richer family background. This gender-differentiated propensity to be enrolled in school implies that the girls we observe in our sample are more likely to have better-off parents. As a consequence, the female lag in maths scores observed in the raw data are likely to underestimate the true extent of the gender gap, and should tend to increase as we control for school fixed effects and family background characteristics.

The following rows in Table 1a show that girls are much more likely to be involved in domestic work at home, as expected, which works against their success in school. On the other hand, boys are more likely to report that working affects school attendance, though the difference between the genders in this case is quite small, but nevertheless statistically significant.

The final rows in Table 1a report various school variables, to investigate whether there is any difference in the quality of education between the genders, in terms of observable variables at least. The data show that there is in fact little difference in resources between the genders, in terms of the use of a maths book, or the standard of other equipment. Female pupils are more likely to have a female teacher and a female headteacher.

The raw data from SACMEQ paint a similar, albeit more detailed picture. The main advantage of SACMEQ compared to PASEC is its wealth of variables on the quality of the pupil's learning environment, both at home and at school. In particular, the availability of pupil-specific learning resources within the school allow us to control for potential discrimination in access to school inputs, be it in terms of access to physical inputs (chair, desk, teaching materials) or in terms of attention received from the teacher. Compared to the boys, the girls in the sample come from wealthier families, have a higher average number of meals per day, and have better educated fathers and mothers. Girls are less likely to be absent from school for work reasons, and in this case no more likely to be absent for family reasons. In terms of help received, girls are significantly more likely to receive more help with their homework, and have their homework and schoolwork checked. Within schools, girls are significantly more likely to report that their teacher always checks their work, more likely to have their own text book, and more likely to be in school with more equipment, a lower pupil-teacher ratio and with electricity. All of these factors are working in favour of girls, for the same reason as discussed above, that it is only the more advantaged girls who are in education. It is unclear from the raw data whether the female advantage in access to school inputs is driven by school heterogeneity or by better access to these inputs within a given school. Our school fixed-effects estimates shed more light on this.

(ii) Pooled Regression Results

Tables 2a and 2b report the results from a multivariate analysis of maths test scores, for the PASEC Year 5 final scores and the SACMEQ scores respectively. Moving from left to right in the table, the estimated equations are gradually built up in terms of the explanatory variables that they include, the aim being to investigate whether any of the explanatory variables can explain the gender difference in maths test scores.

Looking first at the PASEC (Year 5) scores in Table 2a, column 1 includes only the gender variable identifying female pupils, and country level fixed effects. Looking within countries, therefore, girls score on average 1.6 percentage points lower on the maths test than boys, a difference that is strongly statistically significant. The country fixed effects variables show that test scores are highest in Madagascar and lowest in Mauritania.

Column 2 adds controls for whether the pupil is taught in a single-sex school, and an interaction of this variable with the female indicator. The results show that a single-sex school lowers test scores for boys, but raises them for girls, the interaction coefficient outweighing the base effect of the single-sex school. This effect offsets the general negative effect of being female on maths test scores.

Column 3 shows the female coefficient when we move from country level fixed effects to school fixed effects. As can be seen, this increases the size of the female disadvantage in maths test scores, to just over 2 percentage points. Column 4 adds controls for pupils' age, and also whether pupils receive help from teachers or have a maths text book which they can take home with them. The idea for including the latter two variables is to examine whether girls are disadvantaged relative to boys within the same school, in terms of the attention or resources they receive. Both variables, however, attract statistically insignificant results, and their inclusion has no effect on the female coefficient.

Column 5 adds control for family background. Recall from above that, due to selection into school for girls on the basis of household characteristics, the inclusion of these variables might be expected actually to increase the conditional gender gap in maths test scores. In fact, the inclusion of the household variables has no effect on the gender coefficient, this being due to most of the added variables themselves having a statistically insignificant effect on test scores. It might seem a little surprising that household wealth and literacy variables have no effect on children's test scores. A plausible explanation for these results has to do with the fact that all of the analysis is conditional upon enrolment in education. Most of the wealth effects are likely to affect enrolment itself, according to the previous literature reviewed above, and once in schools, wealth plays a negligible role on attainment. In addition, recall that the estimated equation in column 5 includes school fixed effects. Therefore, not only does the equation condition on enrolment, but also enrolment within a particular school, so that wealth effects in terms of access to higher quality schools in richer neighbourhoods are also conditioned out. Thus, the inclusion of the household variables is not intended to provide an estimate of the total effect of household wealth on children's test scores, but only to control for any remaining differences between boys and girls within the same schools, in order to obtain an accurate estimate of the gender effect.

As a robustness check, Column 6 in Table 2a restricts the sample to observations with no imputed values for any variables. As can be seen, the coefficients are robust to this change, suggesting that imputed values were not affecting the results.

The conclusion so far is that none of the control variables in the PASEC dataset can explain the gender difference in maths test scores, which remains around 2 percentage points no matter which variables are added to the specification. A similar outcome is observed with the SACMEQ equations in Table 2b, despite being able to control for a much wider set of pupil-specific schooling characteristics and home environment. In this case, the raw difference between the genders, controlling only for country fixed effects, is around 1 percentage point in favour of boys. The highest test scores are observed in Mauritius and the lowest in Malawi. The single-sex school story (column 2) is slightly different in SACMEQ countries, compared to the PASEC countries. In SACMEQ, single-sex schools benefit both genders, by about 3 percentage points, with no significant difference between the genders in terms of the benefit received.[8] Controlling for school fixed effects does not change the gender differential in maths scores, which remains around 1 percentage point (column 3). Adding child-specific schooling characteristics (within school, column 4) increases the size of the gender gap to 1.66 percentage points since. This suggests that the female advantage in schooling quality noticed in the raw data applies *within-school*, suggesting that the gender gap in tests scores is not due to anti-female discrimination in access to a wide range of measureable inputs. Similarly for household characteristics, the favourable position of the girls who we observe enrolled in school means that controlling for household characteristics widens the gender gap, to almost 2 percentage points (column 5). Note that in the case of the SACMEQ countries, in which enrolment rates are overall higher, household wealth and paternal education do have their expected positive and statistically significant effect on children's test scores. This is consistent with the interpretation that, in PASEC countries, the correlation between household characteristics and test scores captures more selection into enrolment.

Returning to Table 2a, one column that has not been commented upon yet is the final column of results in column 7. Recall from the data description above that in PASEC, the respondents are tested both at the beginning of Year 5 and at the end. A measure of improvement in test scores over the year, or 'value added' can therefore be derived, and its determinants investigated, by controlling for initial maths score at the beginning of the year. The results show that over the course of Year 5, girls improve by just over 1 percentage point less than boys. Thus, amongst pupils who are observed at both the start and the end of Year 5, the gap in performance on maths tests between boys and girls is widening. Not only do girls start Year 5 at a lower level than boys on maths test, they finish Year 5 even further behind.

If we were to repeat the full analysis of Table 2a for the PASEC sample, but using initial test scores at the start of Year 5 as the dependent variable, then the gap between girls and boys is

[8] However, the number of children in single sex schools is again very low as seen from Table 1b.

observed to be even wider than at the beginning of the year than at the end (see the results in Table A1 in the appendix). In column 1, the raw gap (controlling only for country fixed effects) is estimated to be 3 percentage points, and this expands to 3.5 percentage points in the full specification in column 6. Thus the gender gap is wider at the beginning of Year 5 than at the end. This presents something of a puzzle, since it was just discussed in the previous paragraph that the value added results in Table 2a showed girls falling further behind boys over the course of Year 5. The answer to the puzzle is that the value added results are based only on individual pupils who are observed at both the beginning and the end of Year 5. The results presented in Table A1 are for all individuals observed at the beginning of Year 5. However, drop out from school during the course of a school year can and frequently does occur. Furthermore drop-out from school is non-random, as shown in Figure 1 and Figure 2. These diagrams show the distribution of marks (as usual, in terms of percentage correct scores) at the beginning of Year 5, separately for those who remain in education until the end of Year 5, and those who drop-out. Figure 1 is for boys and Figure 2 is for girls. As can be seen, those who drop-out during Year 5 perform less well at the beginning of Year 5, as might be expected. However, the difference in initial test scores between stayers and drop-outs is greater for girls than for boys.[9] Therefore, girls lose more of their lower ability members through drop-out during the course of Year 5, so that their apparent position improves relative to boys (comparing Table A1 to Table 2a). However, when we focus just on those girls present at the beginning and the end of Year 5, in a value added analysis, we observe girls actually falling further behind during the course of the year.

Before leaving the pooled analysis, recall that the PASEC data set also survey pupils in Year 2, so that the same analysis as above can be undertaken for the younger age group. The results are displayed in Table 3. These makes clear that the girls lag boys in maths test scores in Year 2 as well, with the gap actually being larger. In column 1, where the only controls are country fixed effects, the gender gap is just over 3 percentage points. The gap remains similar in size when school fixed effects, pupil-specific schooling characteristics and household characteristics are added in subsequent columns. The final column, looking at value added during Year 2, shows that girls achieve lower value added and so are falling behind boys over the course of the year. The gender gap is also present when looking at

[9] The Kolmogorov-Smirnov test statistic for equality of distribution functions between stayers and drop-outs is 0.042 for boys (p-value: 0.087) and 0.080 for girls (p-value: 0.000).

initial maths test scores at the start of Year 2[10], and so is present at all points of the school system surveyed in PASEC.

(iii) Individual Country Analysis

As well as the pooled analyses discussed in the previous sub-section, we also ran the same set of specifications for each country individually. The results are displayed in Tables 4a and 4b, for PASEC and SACMEQ countries respectively. The specifications mirror those estimated in the previous sub-section, with the obvious exception that country level fixed effects analyses are no longer estimated. For the sake of brevity, only the coefficients on the female dummy variable are reported in these tables.

Looking first at the PASEC (Year 5, final) results in Table 4a, the results in column 1 do not condition for any other variables, and so are essentially raw gender differentials. The results show that girls perform worse than boys in six of the seven PASEC countries (in three cases the gap being statistically significant), the exception being Mauritania. The magnitude of the gender gap in these countries varies in size, from 3.6 percentage points in Guinea, to just 0.3 percentage points in Cameroon. Once school fixed effects are entered into the equation, in column 2, so that girls and boys within the same school are compared, girls perform worse than boys in final Year 5 maths tests in all seven countries, with all gaps being statistically significant with the exception of that in Cameroon. The addition of further control variables cannot close the gender gap at all, and often widens it. The final results, in column 7, report the results from the value added specification. In all countries except Cameroon, the female coefficient is also negative in this specification, suggesting that the gap between boys and girls widens during Year 5, and by a statistically significant amount in four of the countries.

Turning to the SACMEQ countries in Table 4b, the raw gender differentials in column 1 are negative, revealing lower scores for girls, in eight of the twelve countries, with the same number of negative coefficients observed in the school fixed effects specification in column 2. Of these eight, all but one are statistically significant, whereas two of the four positive coefficients are statistically insignificant. Only in Botswana and Mauritius do girls report significantly higher maths test scores, controlling for school fixed effects, than boys. Furthermore, in the case of Mauritius, the female advantage disappears once further controls are added in subsequent columns for pupil, household and pupil-specific schooling

[10] Estimating the same set of specifications as previously, the gender gap is between 2 and 3 percentage points in all specifications (full results available from the authors on request).

characteristics. Of the negative raw gender gaps, most remain of a similar size or even widen when further control variables are added to the specification, thus revealing the now familiar story that the relatively low performance of girls in maths test cannot be explained by household or pupil-specific schooling characteristics.

4. Cross-Country Analysis

What, then, explains the gender difference in test scores? One possibility, given that the gender gap is observed as early as the beginning of Year 2, is that girls are biologically less able in the area of maths. Such a possibility cannot be explored with the data available to us here, though such a hypothesis would need to explain why girls then do not perform worse than boys in all countries of the world and in all time periods. In this section, we first test whether the explanation put forward by Guiso et al. (2008) for the OECD countries included in PISA, but showed by Fryer and Levitt (2010) to be sensitive to the inclusion of other countries, holds in Africa. We then suggest an alternative cultural explanation which accounts for between 44 and 51% of the cross-country variation in the female maths disadvantage in our African dataset.

In Figure 3, the raw gender gap in each country is plotted against GDP per capita in 2000. There is a clearly upward-sloping relationship, though this significant correlation is driven by countries with GDP per capita greater than 1000 dollars. Figure 4 plots the raw gender gap against the 2007 gender gap index introduced by the World Economic Forum in 2006. This index is the mean value of four sub-indexes measuring gender inequality in terms of economic participation and opportunity, educational attainment, health and survival, and political empowerment.[11] An index value closer to unity indicates more *equality*. Contrary to what is observed by Guiso et al. (2008) in richer, non-African PISA countries, there is no clear correlation between female empowerment as measured by the gender gap index and the maths gender gap. Figure 5 shows that there is no systematic pattern either when concentrating on political empowerment, again contrary to the findings in Guiso et al. (2008) for PISA countries. Results of OLS regressions presented in Table 5a confirm this graphical analysis, and further explore the correlation between the maths gender gap and specific

[11] Economic participation and opportunity is measured as a weighted mean of 5 ratios, namely: ratio of female over male labour force participation; wage; earned income; legislators, senior officials and managers; professional and technical workers. Educational attainment is measured as the weighted mean of four ratios: literacy rate; net enrolment in: primary; secondary; and tertiary schooling. Health and survival is calculated as the weighted mean of female over male healthy life expectancy and female over male sex ratio at birth. Political empowerment is the weighted average of the women-to-men ratio of seats in parliament; ministers; and female-over-male years as head of state during the last 50 years.

dimensions of female status summarised in the index. It is interesting to note that difference in GDP per capita explain about 30 percent of the variation in relative female performance in our sample (Column (1)). Similar to Guiso et al. (2008), we control for GDP per capita in the other regressions as a general control for standards of living. Columns (3) to (6) regress the maths gender gap on (GDP and) each of the four sub-indexes, while Column (7) focuses on the relationship between the maths gender gap and gender inequality in primary schooling enrolment (see also Figure 6). Higher enrolment parity may indicate greater opportunities for girls in education, but could also increase the gender gap by bringing lower ability girls into education. The inclusion of each of the sub-indexes substantially increases the R-squared, thus suggesting that they may be part of the explanation. However, there is too little variation in the health subindex (which only varies between .95 and .98 in this sample) to produce a reliable estimate. Despite more variation, the educational sub-index or its more specific dimension analysed here (i.e., primary enrolment parity index) do not have a statistically significant effect in this small sample. Finally, both political empowerment and economic participation are correlated with larger *disadvantages* in maths, and in the case of economic participation, this result is significant at the 95% level.

We then explore the effect on the maths gender gap of an alternative dimension of female empowerment, namely woman status in terms of sexual and reproductive health. In the SACMEQ survey, head teachers were asked whether the school had to deal with sexual harassment of pupils by teachers, and could answer "Never", "Sometimes", or "Often". Figure 7 plots the raw gender gap against the percentage of head teachers who answer that they have to deal with this type of behaviour at least "Sometimes". There is a clear negative relationship between the extent of the problem at the country level and the female-to-male difference in maths performance. A regression of the maths gender gap on the share of head teachers reporting sexual harassment accounts for 51% of the cross-country variation in the maths performance gap (Table 5b, Column (2)). This result is highly significant, but it becomes non-distinguishable from zero when controlling for GDP per capita in this small sample (Column (1)). Interestingly, 85% of the cross-country variation in the prevalence of sexual harassment of pupils by teachers is explained by differences in the total fertility rate (Column (7)).[12] When regressing the maths gender gap on the total fertility rate, which is available for all 19 countries and not just for the countries included in SACMEQ, we see that

[12] The R-squared obtained when regressing the harassment prevalence rate on the total fertility rate without controlling for GDP per capita is 0.847.

variations in fertility significantly explain 44% of the variation in differential maths performance (Column (4)). It is striking that, once fertility is controlled for, GDP does not explain any of the remaining variation in the maths gender gap. In Column (6), we present a robustness check in which we test the sensitivity of the effect of the fertility rate to controlling for genetic factors that might be correlated both with fertility and gender differences in cognitive ability. In order to do so, we expand the dataset into a corresponding set of country dyads and use the measure of genetic distance between the population of each pair of countries compiled by Spolaore and Wacziarg (2009). The estimate of the effect of the pair-wise difference in fertility rates on the pair-wise difference in relative female performance is virtually unchanged by controlling for genetic distance.

OLS regressions of the total fertility rate on the four different sub-indexes comprising the gender gap index all achieve R-squared statistics of less than 0.06.[13] Harassment and fertility thus appear to capture similar cultural traits that affect female relative performance in maths but that are not correlated with female equality as measured by the gender gap index.

5. Conclusion

This paper has used data on a number of Sub-Saharan African countries, to calculate and analyse determinants of the lower performance of girls in maths test scores in primary school. In the pooled specifications, and in most of the individual country specifications, girls perform worse than boys on maths tests, by around 2 or 3 percentage points, where the test score is the percentage of correct answers on the test. This gap is present at different points in time, at the beginning and end of Year 2, and at the beginning and end of Year 5, for the PASEC countries where these multiple observations are taken. It therefore seems to be a permanent feature of girls' performance. Furthermore, the 'value added' specifications for the PASEC countries show that girls' performance deteriorates relative to that of boys during Year 2 and also during Year 5, suggesting that girls cannot close that achievement gap, and if anything it gets wider as they move through school.

Girls appear to fare better in single-sex schooling in some of the countries covered by our study, but this finding does not apply in general. Caution should prevail in interpreting this result, however, since fewer than 2% of children in our sample are educated in single-sex

[13] Results available upon request.

schools. These results add to those of Fryer and Levitt (2010) in suggesting that this is an area that merits further research.

The two surveys used for the analysis contain information on a wide range of characteristics, about the individual pupils, their family background and their schools. However, controlling for these characteristics has close to no effect on the estimated gender coefficients. These results suggest that the explanation for girls' poorer performance in maths test is not that their families or schools have characteristics that are less conducive to cognitive development and achievement. If anything, the girls that we observe in education in these African countries actually have family or pupil-specific schooling characteristics that tend to *improve* their test scores.

Turning to cross-country analysis, we found that girls perform relatively better in countries where there are fewer reports of sexual harassment by teachers and where fertility rates are lower, irrespective of the status of women on the labour market, in terms of educational participation measured by enrolment, of their relative life expectancy and the degree of sex-selective abortion if any, or in terms of political representation. While lack of data precludes a causal interpretation of this finding or an analysis of the exact pathways through which the observed correlation arises, our results rule out a substantial number of explanations. Indeed, they show that, in sub-Saharan Africa, (i) any intrinsic gender difference in maths ability cannot explain more than about half the female deficit in maths test scores and (ii) cultural values about the role of women in the private but not in the public sphere account for a large share of the variation in the female gap in maths. Furthermore, the analysis carried out at the pupil level suggests that differences in observable resource allocation at home (including absenteeism) and at school are unlikely to be important pathways through which the role of women in the private sphere translates into poorer maths scores in the 19 countries analysed here. The main remaining candidate explanation left by these results therefore appears to be related to the psychological effect of stereotypes.

We tested whether national differences in fertility rates could also explain differences in relative female performance in the set of countries who took part in the 2003 TIMMS evaluation, and where Fryer and Levitt (2009) have shown that the gender gap index fails to significantly account for cross-country differences in female relative maths performance. However, the relationship plotted in Figure 9 is not statistically significant. It therefore increasingly seems that cultural differences between countries are likely to also affect what cultural difference matters for the gender gap in mathematics.

References

Abu-Ghaida and Klasen, S. (2004). 'The costs of missing the Millennium Development Goal on gender equity', *World Development*, 32, 1075-1107.

Alderman, H., Behrman, J., Ross, D. and Sabot, R. (1996). 'Decomposing the gender gap in cognitive skills in a poor rural economy', *Journal of Human Resources*, 31, 229-254.

Appleton, S. (1995). 'The interaction between poverty and gender in human capital accumulation: The case of the primary leaving exam in the Cote d'Ivoire', *Journal of African Economies*, 4, 192–224

Banerjee, Abhijit, Suraj Jacob, and Michael Kremer with Jenny Lanjouw and Peter Lanjouw (2005) "Promoting School Participation in Rural Rajasthan: Results from Some Prospective Trials," mimeo, MIT.

Becker, G. (1981). *A Treatise on the Family*. Cambridge, Massachusetts: Harvard University Press.

Benhabib, J. and Spiegel, M. (1994). 'The role of human capital in economic development: Evidence from aggregate cross-country data." *Journal of Monetary Economics*, 34,143-174.

Behrman, J. and Knowles, J. (1999). ' Household income and child schooling in Vietnam', *World Bank Economic Review*, 13, 211-256.

Bjorkman, M. (2006). 'Income shocks and gender gaps in education: evidence from Uganda', Institute for International Economic Studies, Stockholm University, Seminar Paper No. 744.

Burder, Dana and Leigh L. Linden. (2010) The Effect of Village-Based Schools: Evidence from a RCT in Afghanistan. Mimeograph.

Card, D. (1999). 'The Causal effect of education on earnings', in Ashenfelter, O. and Card, D. (eds) *Handbook of Labor Economics, Volume 3a*. Amsterdam; New York and Oxford: Elsevier Science, North-Holland.

Coclough, C., Rose, P. And Tembon, M. 2000. Gender inequalities in primary schooling: The roles of poverty and adverse cultural practice. *International Journal of Educational Development* 20, pp 5-27.

Fehrler, S., Michaelowa, K. and Wechtler, A. (2009). 'The effectiveness of inputs in primary education: insights from recent student surveys in Sub-Saharan Africa', *Journal of Development Studies*, 45, 1545-1578.

Fryer, R. and Levitt, S. (2009). 'An empirical analysis of the gender gap in mathematics', mimeo, Harvard University and University of Chicago.

Glick, P. and Sahn, D. (2000). 'Schooling of girls and boys in a West African country: the effects of parental education, income, and household structure', *Economics of Education Review*, 19, 63-87.

Glewwe, P. (1996). The Relevance of Standard Estimates of Rates of Return to Schooling for Education Policy: A Critical Assessment ," *Journal of Development Economics* 51:2, pp. 267–90.

Goldin, C., Katz, L., and Kuziemko, I. (2006). The homecoming of American college women: The reversal of the college gender gap. *Journal of Economic Perspectives*, 20(4), pp.133-156.

Guiso, L., Monte, F., Sapienza, P., and Zingales, L. (2008). Culture, Gender and Math, *Science*, 320, pp.1164-1165.

Hanushek, E. and Woessmann, L. (2008). 'The role of cognitive skills in economic development', *Journal of Economic Literature*, 46, 607-668.

Harmon, C., Oosterbeek, H. and Walker, I. (2003). 'The returns to education: microeconomics', *Journal of Economic Surveys*, 17, 115–155.

Hedges, L. and Nowell, A. (1995). Sex Differences in Mental Test Scores, Variability, and Numbers of High-Scoring Individuals, *Science*, 269, pp.41-45.

Jolliffe, D. (1998). Skills, Schooling, and Household Income in Ghana, *World Bank Economic Review*, 12(1), pp.81-104.

Kingdon, G. (2002). 'The gender gap in educational attainment in India. How much can be explained?', *Journal of Development Studies*, 39, 25-53.

Klasen, S. (2002). 'Low schooling for girls, slower growth for all? Cross-country evidence on the effect of gender inequality in education on economic development', *World Bank Economic Review*, 16, 345-373.

Knowles, S., Lorgelly, P. and Owen, P. (2002). ‘Are educational gender gaps a brake on economic development? Some cross-country empirical evidence’, *Oxford Economic Papers*, 54, 118-149.

Lewis, M. and Lockheed, M. (2008). ‘Social exclusion and the gender gap in education’, World Bank Policy Research Working Paper No. 4562.

Lucas, R. (1988). ‘On the mechanics of economic development’, *Journal of Monetary Economics*, 22, 3-42.

Machin, S. and McNally, S. (2005). ‘Gender and student achievement in English schools’, *Oxford Review of Economic Policy*, 21, 357-372.

Mankiw, N., Romer, D. and Weil, D. (1992). ‘A contribution to the empirics of economic growth’, *Quarterly Journal of Economics*, 107, 407-437.

Michaelowa, K. (2001). ‘Primary education quality in Francophone Sub-Saharan Africa: determinants of learning achievement and efficiency considerations’, *World Development*, 29, 1699-1716.

Moll, P. (1998). Primary Schooling, Cognitive Skills and Wages in South Africa, *Economica*, 65, pp.263-284.

Nguyen, P. (2006). Effects of social class and school conditions on educational enrolment and achievement of boys and girls in rural Viet Nam. *International Journal of Educational Research*, 45, pp.153-175.

Pope, D. and Sydnor, J. (2010) Geographic Variation in the Gender Differences in Test Scores, *Journal of Economic Perspectives*, 24 (2), pp.95-108.

Psacharopoulos, G. and Patrinos, H. (2004). ‘Returns to investment in education: a further update’, *Education Economics*, 12, 111–134.

Romer, P. (1990). ‘Endogenous technological change’, *Journal of Political Economy*, 99, S71-S102.

Rosenzweig, M. and Schultz, T. (1982). ‘Market opportunities, genetic endowments, and intrafamily resource distribution: child survival in rural India’, *American Economic Review*, 72, 803-815.

Ross, K., Saito, M., Dolata, S., Ikeda, M., & Zuze, L. (2004). *Data Archive for the SACMEQ I and SACMEQ II projects*. Paris: IIEP-UNESCO.

Sackey, H. (2007). ‘The determinants of school attendance and attainment in Ghana: a gender perspective’, African Economic Research Consortium, Research Paper 173.

Song, L., Appleton, S. and Knight, J. (2006). ‘Why do girls in rural China have lower school enrolment?’, *World Development*, 34, 1639-1653.

Figure 1: Male year 5 initial maths score by drop out status

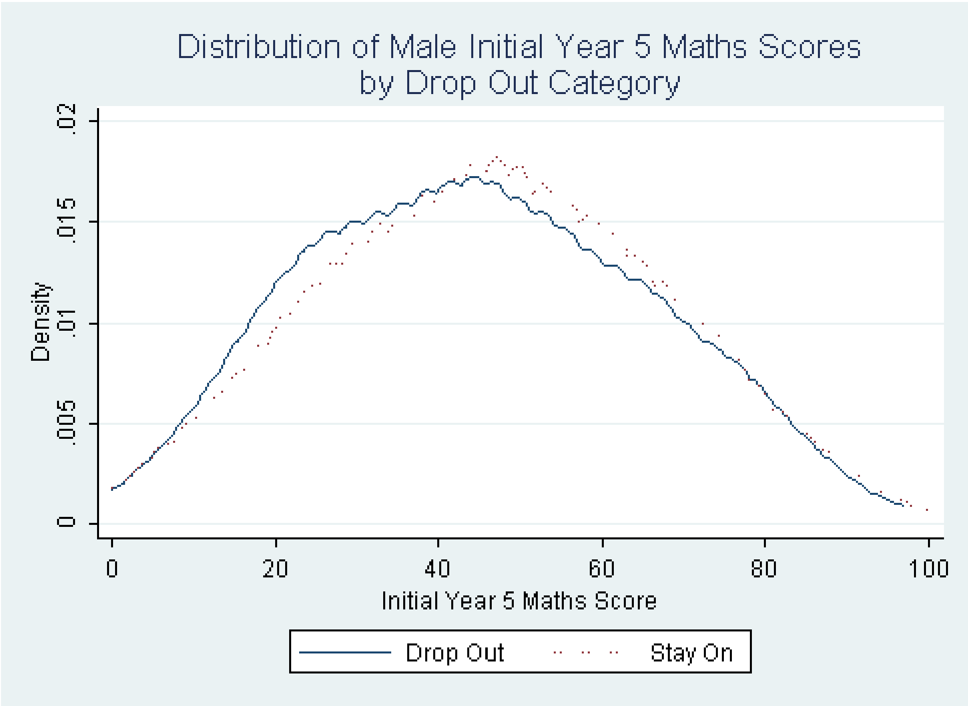


Figure 2: Female year 5 initial maths score by drop out status

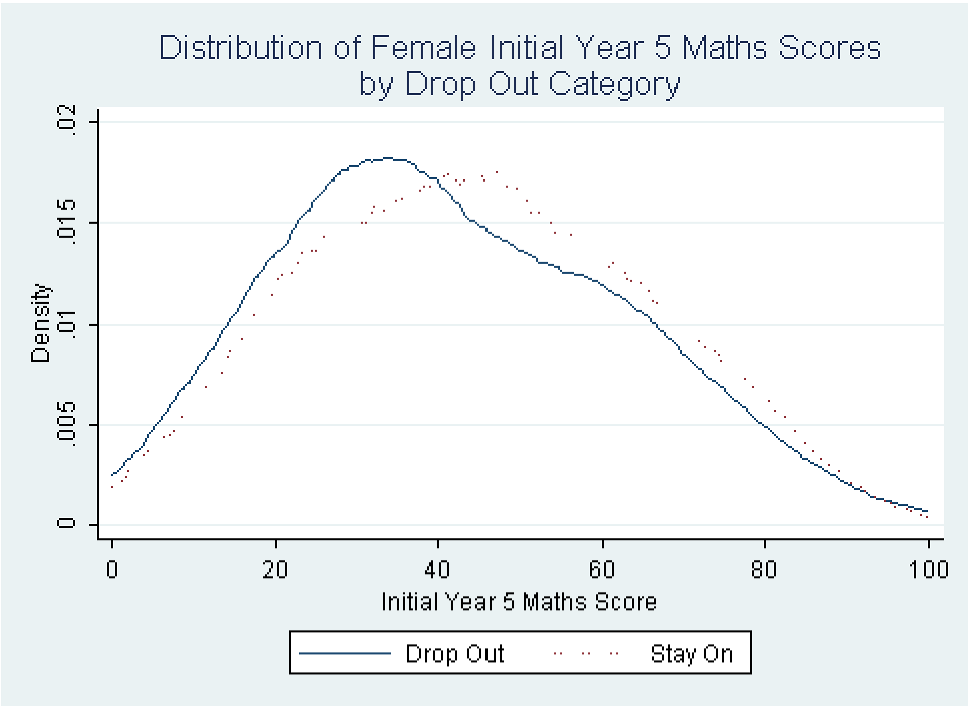
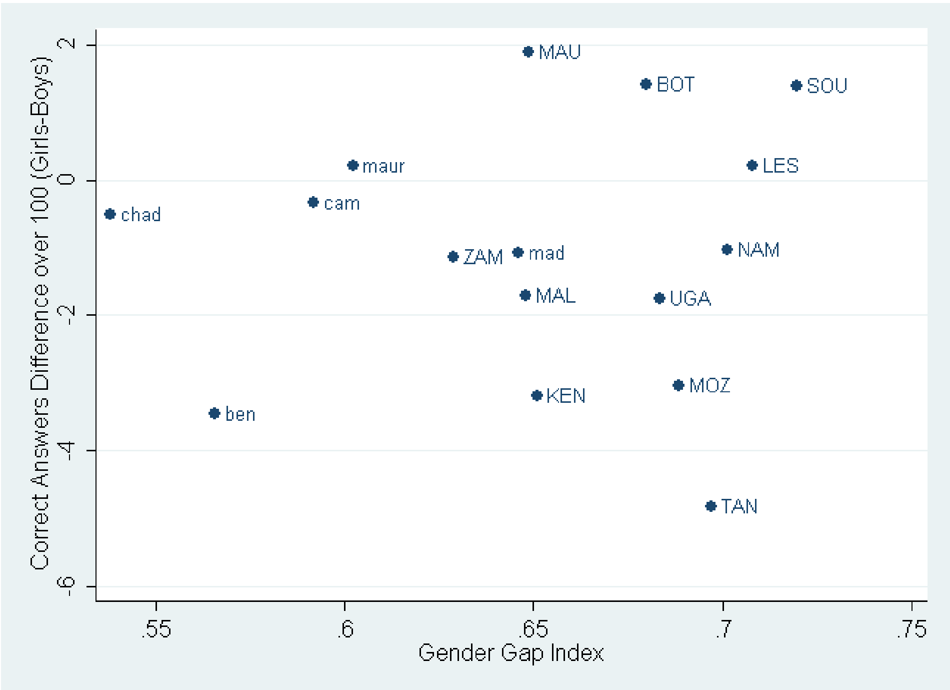


Figure 3: Relationship between gender gap in maths and GDP per capita in 2000

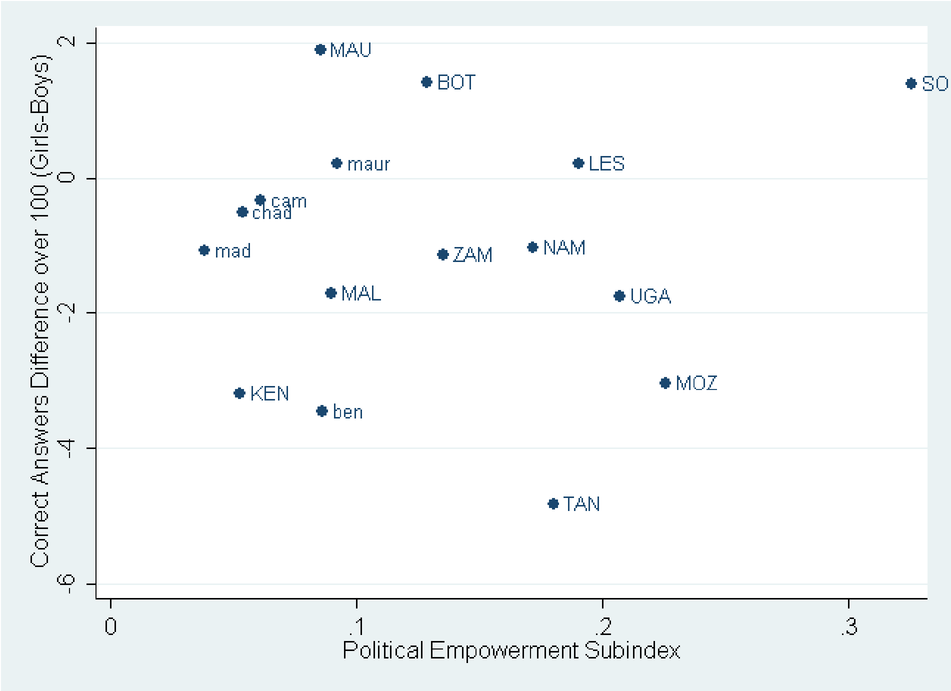
Note: Countries in lower case are from PASEC, and in upper case are from SACMEQ. GDP per capita in 2000 (current dollars) taken from IMF World Economic Outlook Database, April 2010.

Figure 4: Relationship between gender gap in maths and Gender Gap Index



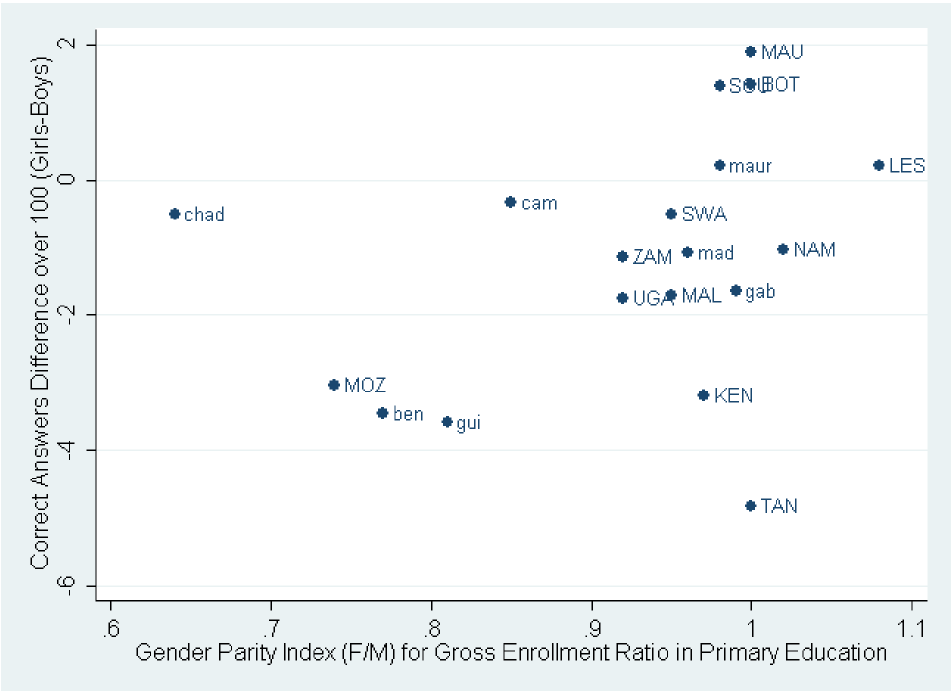
Note: Countries in lower case are from PASEC, and in upper case are from SACMEQ. 2007 Gender Gap Index taken from Hausmann et al. (2007).

Figure 5: Relationship between gender gap in maths and political subindex



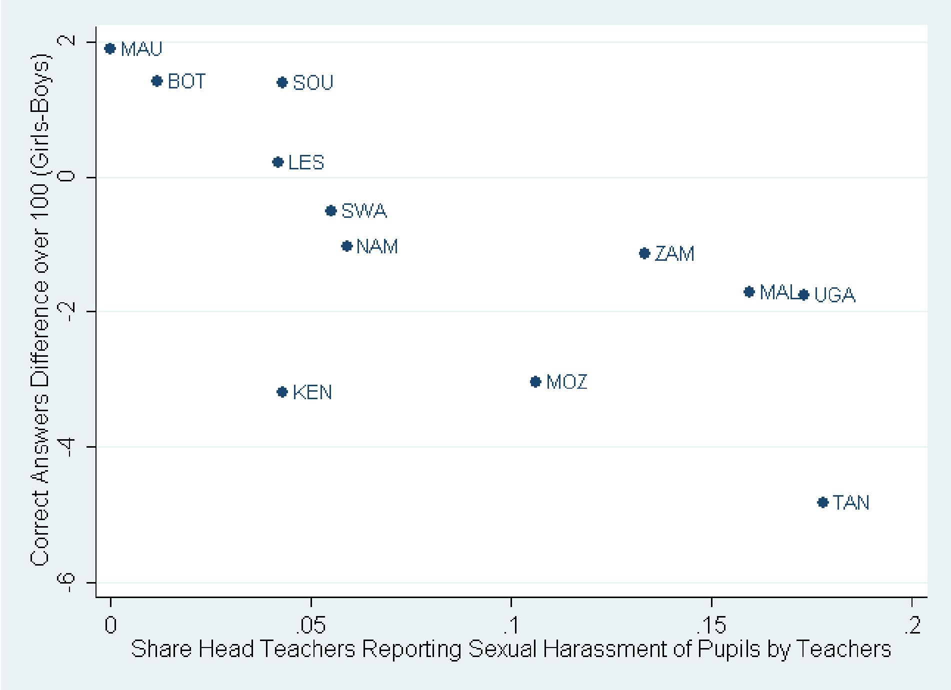
Note: Countries in lower case are from PASEC, and in upper case are from SACMEQ. 2007 Political Empowerment subindex taken from Hausmann et al. (2007).

Figure 6: Relationship between gender gap in maths and female/male enrolment ratio



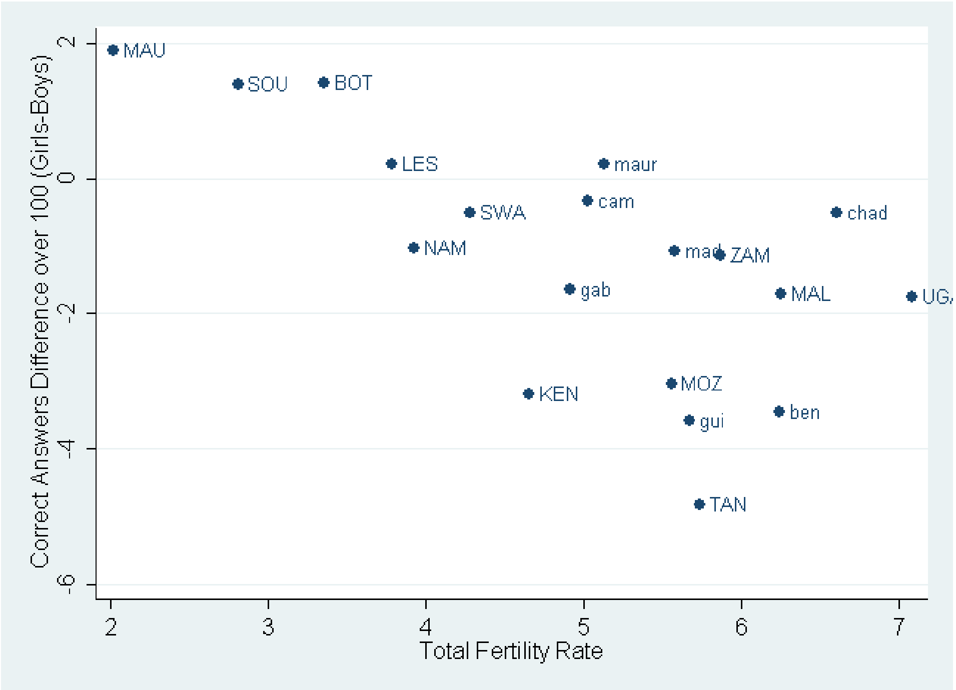
Note: Countries in lower case are from PASEC, and in upper case are from SACMEQ. Female/Male enrolment ratio in primary education for 1999 for SACMEQ countries and in 2004 for PASEC countries taken from UNESCO (2007).

Figure 7: Relationship between gender gap in maths and reported sexual harassment



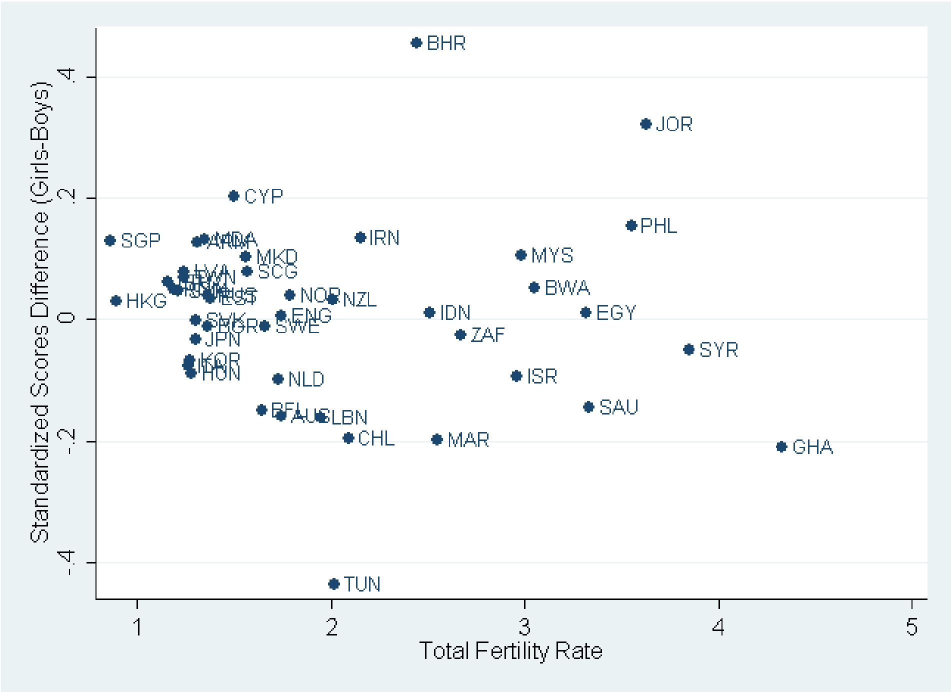
Note: Data only collected in SACMEQ.

Figure 8: Relationship between gender gap in maths and total fertility rate



Note: Countries in lower case are from PASEC, and in upper case are from SACMEQ. Total fertility rate in 2000 taken from US Census Bureau International Database.

Figure 9: Relationship between gender gap in maths and total fertility rate, TIMMS 2003



Note: Countries in lower case are from PASEC, and in upper case are from SACMEQ. Total fertility rate in 2003 taken from US Census Bureau International Database.

Table 1a: Descriptive Statistics – PASEC Sample

	Boys Mean	Girls Mean	Difference	(t-stat)
initial Score Yr5 - Percent	46.452	44.304	2.149***	(-5.98)
final Score Yr5 - Percent	38.995	38.482	0.513	(-1.51)
age	12.183	11.983	0.201***	(-6.51)
middle wealth tercile	0.403	0.396	0.00681	(-0.81)
top wealth tercile	0.309	0.322	-0.0124	(-1.55)
does not live with parents	0.125	0.137	-0.0119**	(-2.07)
single parent	0.252	0.251	0.00176	(-0.24)
father is literate	0.691	0.783	-0.0921***	(-12.14)
mother is literate	0.499	0.600	-0.101***	(-11.90)
Regular daily meals (0 to 3)	2.669	2.745	-0.0757***	(-7.21)
parental help	0.236	0.275	-0.0397***	(-5.31)
sibling help	0.435	0.422	0.0132	(-1.56)
tutor help	0.037	0.043	-0.00579*	(-1.72)
other help	0.047	0.051	-0.00322	(-0.87)
single-sex school	0.018	0.019	-0.00131	(-0.57)
domestic work	0.743	0.906	-0.163***	(-25.14)
works in agriculture or trade	0.636	0.574	0.0621***	(-7.41)
work affects study at home	0.185	0.184	0.000873	(-0.13)
work affects school attendance	0.133	0.123	0.0101*	(-1.75)
work affects schooling because too tired	0.13	0.122	0.00821	(-1.43)
has maths book	0.458	0.462	-0.00347	(-0.41)
teacher help	0.031	0.030	0.000238	(-0.08)
teacher's years of experience	10.10	11.01	-0.910***	(-6.29)
female teacher	0.26	0.305	-0.0453***	(-5.89)
female headteacher	0.196	0.236	-0.0403***	(-5.74)
middle tercile equipment	0.355	0.340	0.0153*	(-1.88)
high tercile equipment	0.293	0.307	-0.0140*	(-1.78)
N	13747			

t statistics in parentheses
* p<0.10; ** p<0.05; *** p<0.01

Table 1b: Descriptive Statistics – SACMEQ Sample

	Boys Mean	Girls Mean	Difference	(t-stat)
maths score - percent	40.273	38.878	1.395***	(8.72)
age	13.653	13.061	0.592***	(28.94)
single-sex school	0.011	0.005	0.00619***	(6.57)
single-sex × female	0	0.005	-0.00529***	(-10.09)
stay with relatives during week	0.114	0.119	-0.00512	(-1.55)
stay boarding school during week	0.052	0.044	0.00816***	(3.71)
stay on own during week	0.031	0.026	0.00563***	(3.29)
lowest wealth tercile	0.343	0.323	0.0198***	(4.08)
middle wealth tercile	0.335	0.341	-0.0054	(-1.11)
top wealth tercile	0.321	0.336	-0.0144***	(-2.98)
number of meals everyday	2.344	2.380	-0.0361***	(-3.97)
number days absent per month	1.774	1.526	0.248***	(9.44)
absent for work reason	0.055	0.040	0.0155***	(7.07)
absent for unpaid fees	0.044	0.044	-0.000256	(-0.12)
if absent for family reason	0.090	0.085	0.00439	(1.51)
number of books at home	22.93	24.88	-1.952***	(-3.54)
father education index	8.965	9.268	-0.304***	(-6.13)
mother education index	8.030	8.541	-0.511***	(-11.33)
gets no homework	0.012	0.011	0.00125	(1.13)
sometimes check homework	0.490	0.469	0.0211***	(4.09)
most times check homework	0.381	0.419	-0.0380***	(-7.54)
sometimes help with homework	0.545	0.541	0.00404	(0.79)
most times help with homework	0.261	0.294	-0.0334***	(-7.25)
sometimes check schoolwork	0.509	0.500	0.00808	(1.57)
most times check schoolwork	0.327	0.350	-0.0227***	(-4.65)
someone does maths sometimes	0.547	0.552	-0.0051	(-1.00)
someone does maths most times	0.242	0.254	-0.0123***	(-2.75)
questions about maths sometimes	0.549	0.549	0.0000932	(0.02)
questions about maths most times	0.263	0.271	-0.00864*	(-1.90)
extra maths tuition	0.364	0.360	0.00385	(0.78)
someone reads sometimes	0.586	0.577	0.00947*	(1.86)
someone reads mosttimes	0.233	0.249	-0.0153***	(-3.47)
questions about reading sometimes	0.560	0.550	0.0105**	(2.06)
questions about reading most times	0.255	0.276	-0.0207***	(-4.55)
extra English tuition	0.344	0.335	0.00927*	(1.90)
no maths text books	0.116	0.116	0.000457	(0.14)
has own math text book	0.479	0.486	-0.00663	(-1.29)
no maths homework	0.055	0.048	0.00697***	(3.05)
teacher corrects maths some times	0.222	0.221	0.00021	(0.05)
teacher corrects maths most times	0.206	0.206	0.000417	(0.10)
teacher corrects maths always	0.477	0.491	-0.0140***	(-2.72)
no reading text books	0.103	0.094	0.00902***	(2.94)

	Boys Mean	Girls Mean	Difference	(t-stat)
has own reading text book	0.477	0.492	-0.0156***	(-3.03)
no reading homework	0.132	0.122	0.00933***	(2.72)
teacher corrects reading some times	0.271	0.258	0.0132***	(2.90)
teacher corrects reading most times	0.201	0.201	-0.000272	(-0.07)
if teacher corrects reading always	0.342	0.370	-0.0282***	(-5.71)
pupil equipment index	5.462	5.563	-0.101***	(-5.23)
write on chair/bench	0.923	0.941	-0.0177***	(-6.83)
write on desk/table	0.895	0.913	-0.0174***	(-5.73)
can borrow library books	0.456	0.453	0.00306	(0.59)
school has electricity	0.443	0.452	-0.00922*	(-1.79)
school has library	0.443	0.437	0.00595	(1.16)
female headteacher	0.735	0.695	0.0398***	(8.53)
pupil to teacher ratio	42.114	41.074	1.040***	(4.16)
male teacher	0.601	0.560	0.0408***	(7.88)
high level teacher education	0.242	0.260	-0.0177***	(-3.95)
middle class equipment tercile	0.338	0.342	-0.00411	(-0.83)
high class equipment tercile	0.277	0.286	-0.00955**	(-2.02)
N	37684			

t statistics in parentheses
* p<0.10; ** p<0.05; *** p<0.01

Table 2a: Pooled Results – PASEC Final Year 5 Maths Test Score

	Column 1 Country FE	Column 2 Country FE	Column 3 School FE	Column 4 School FE	Column 5 School FE	Column 6 School FE: no imputation	Column 7 School FE: Value Added
female	-1.643***	-1.808***	-2.082***	-2.147***	-2.095***	-2.262***	-1.140**
	(0.298)	(0.300)	(0.223)	(0.225)	(0.238)	(0.252)	(0.352)
Benin	33.023***	33.089***					
	(0.395)	(0.395)					
Cameroon	46.880***	47.030***					
	(0.407)	(0.408)					
Chad	32.911***	33.087***					
	(0.469)	(0.471)					
Gabon	44.722***	44.807***					
	(0.444)	(0.444)					
Gguinea	37.770***	37.844***					
	(0.369)	(0.369)					
Madagascar	52.994***	53.065***					
	(0.398)	(0.398)					
Mauritania	23.820***	24.069***					
	(0.442)	(0.452)					
single-sex school		-7.112***					
		(1.583)					
single-sex×female		9.754***					
		(2.238)					
age				-0.334***	-0.286***	-0.282***	-0.365***
				(0.074)	(0.077)	(0.082)	(0.069)
has maths book				0.400	0.315	0.466	(0.264
				(0.317)	(0.330)	(0.354)	(0.296)

	Column 1 Country FE	Column 2 Country FE	Column 3 School FE	Column 4 School FE	Column 5 School FE	Column 6 School FE: no imputation	Column 7 School FE: Value Added
teacher help				-0.214	-0.325	-0.554	0.053
				(0.751)	(0.787)	(0.837)	(0.702)
middle wealth tercile					0.171	0.239	-0.041
					(0.362)	(0.383)	(0.326)
top wealth tercile					0.324	0.155	-0.084
					(0.434)	(0.459)	(0.390)
does not live with parents					1.264***	1.339***	0.936**
					(0.368)	(0.390)	(0.331)
single parent					0.677*	0.713*	0.584*
					(0.296)	(0.320)	(0.266)
father is literate					0.190	0.368	0.174
					(0.314)	(0.330)	(0.282)
mother is literate					0.001	-0.164	0.176
					(0.321)	(0.340)	(0.287)
regular daily meals (0 to 3)					-0.615**	-0.507*	-0.537*
					(0.236)	(0.247)	(0.211)
parental help					0.227	0.173	0.133
					(0.338)	(0.356)	(0.304)
sibling help					0.035	0.111	0.147
					(0.299)	(0.314)	(0.268)
tutor help					0.201	0.498	0.116
					(0.640)	(0.675)	(0.570)
other help					-0.203	0.112	-0.216
					(0.570)	(0.600)	(0.511)
domestic work					-0.502	-0.723	
					(0.366)	(0.393)	
works in agriculture or trade					-0.501	-0.545	

	Column 1	Column 2	Column 3	Column 4	Column 5	Column 6	Column 7
						School FE:	School FE:
						no	Value
	Country FE	Country FE	School FE	School FE	School FE	imputation	Added
					(0.302)	(0.320)	
work affects study at home					0.355	0.350	
					(0.405)	(0.433)	
work affects school attendance					-0.859	-0.824	
					(0.490)	(0.531)	
work affects schooling because too tired					-0.233	-0.146	
					(0.446)	(0.482)	
initial score (/100)							0.352***
							(0.008)
constant			39.714***	43.581***	44.722***	44.460***	28.449***
			(0.146)	(0.924)	(1.263)	(1.337)	(1.158)
no. of children	11904	11904	11904	11761	11147	10066	11386
no. of schools			933	933	930	929	930
R-squared	0.860	0.860	0.008	0.010	0.013	0.015	0.183

Notes:

1. Coefficients are reported together with their standard errors. * p<0.10; ** p<0.05; *** p<0.01.
2. Omitted country category is Benin.
3. ‘help’ refers to help with homework.
4. Wealth terciles based on PCA except for Chad and Guinea.
5. Class equipment terciles based on PCA.

Table 2b Pooled Results – SACMEQ Maths Test Scores

	Column 1 Country FE	Column 2 Country FE	Column 3 School FE	Column 4 School FE	Column 5 School FE
female	-0.981***	-0.968***	-1.178***	-1.666***	-1.901***
	(0.140)	(0.140)	(0.113)	(0.113)	(0.138)
Kenya	8.583***	8.500***			
	(0.331)	(0.332)			
Lesotho	-10.508***	-10.513***			
	(0.335)	(0.335)			
Malawi	-12.731***	-12.796***			
	(0.365)	(0.365)			
Mauritius	10.280***	10.280***			
	(0.343)	(0.343)			
Mozambique	1.625***	1.593***			
	(0.336)	(0.336)			
Namibia	-11.518***	-11.518***			
	(0.302)	(0.302)			
South africa	-5.484***	-5.484***			
	(0.336)	(0.336)			
Swaziland	0.616	0.597			
	(0.335)	(0.336)			
Tanzania	0.807*	0.797*			
	(0.344)	(0.344)			
Uganda	-1.501***	-1.550***			
	(0.352)	(0.353)			
Zambia	-12.695***	-12.743***			
	(0.353)	(0.354)			
single-sex school		2.693**			
		(0.919)			

	Column 1 Country FE	Column 2 Country FE	Column 3 School FE	Column 4 School FE	Column 5 School FE
single-sex×if female		0.324			
		(1.647)			
age				-0.859***	-0.726***
				(0.038)	(0.047)
no maths text books				-0.381	-0.344
				(0.248)	(0.300)
has own math text book				-0.230	-0.231
				(0.195)	(0.245)
no maths homework				-0.349	0.103
				(0.413)	(0.515)
teacher corrects maths some times				1.075**	0.671
				(0.331)	(0.397)
teacher corrects maths most times				1.512***	1.164**
				(0.339)	(0.405)
teacher corrects maths always				1.729***	1.440***
				(0.332)	(0.399)
no reading text books				-0.812**	-0.416
				(0.253)	(0.306)
has own reading text book				-0.057	-0.008
				(0.186)	(0.232)
no reading homework				0.572	0.082
				(0.334)	(0.421)
teacher corrects reading some times				0.500	0.228
				(0.293)	(0.360)
teacher corrects reading most times				0.348	0.089
				(0.305)	(0.373)
teacher corrects reading always				0.494	0.352
				(0.299)	(0.369)
pupil equipment index				0.539***	0.486***

	Column 1 Country FE	Column 2 Country FE	Column 3 School FE	Column 4 School FE	Column 5 School FE
				(0.040)	(0.049)
write on chair/bench				1.704***	1.372*
				(0.458)	(0.535)
write on desk/table				2.193***	2.298***
				(0.342)	(0.407)
can borrow library books				0.527*	0.625*
				(0.208)	(0.256)
middle wealth tercile					0.373*
					(0.183)
top wealth tercile					1.253***
					(0.226)
stay with relatives during week					-1.235***
					(0.231)
stay boarding school during week					-1.849***
					(0.393)
stay on own during week					-1.521***
					(0.445)
number of meals everyday					0.726***
					(0.093)
someone reads sometimes					0.947***
					(0.213)
someone reads mosttimes					0.315
					(0.252)
questions about reading sometimes					0.379
					(0.216)
questions about reading mosttimes					-0.032
					(0.253)
extra English tuition					0.251
					(0.183)

	Column 1 Country FE	Column 2 Country FE	Column 3 School FE	Column 4 School FE	Column 5 School FE
number of books at home					0.003
					(0.001)
father education index					0.139***
					(0.021)
mother education index					0.036
					(0.022)
gets no homework					-2.850***
					(0.850)
sometimes check homework					0.097
					(0.276)
most times check homework					0.655*
					(0.297)
sometimes help with homework					-0.311
					(0.226)
most times help with homework					-1.720***
					(0.259)
sometimes check schoolwork					0.448*
					(0.227)
most times check schoolwork					0.181
					(0.250)
absent for family reason					0.429
					(0.252)
absent for unpaid fees					-0.241
					(0.356)
number days absent per month					-0.287***

	Column 1 Country FE	Column 2 Country FE	Column 3 School FE	Column 4 School FE	Column 5 School FE
					(0.031)
absent for work reason					-1.258***
					(0.342)
constant	42.927***	42.921***	40.166***	43.528***	39.672***
	(0.244)	(0.245)	(0.078)	(0.803)	(1.049)
no. of children	37407	37407	37407	37162	26559
no. of schools			2125	2110	2109
R-squared	0.242	0.243	0.003	0.029	0.052

Notes:

1. Coefficients are reported together with their standard errors. * p<0.10; ** p<0.05; *** p<0.01
2. Omitted country category is Botswana.
3. ‘help’ refers to help with homework.
4. Wealth terciles based on PCA except for Chad and Guinea.

Table 3: Pooled Results – PASEC Final Year 2 Maths Test Scores

	Column 1 Country FE	Column 2 Country FE	Column 3 School FE	Column 4 School FE	Column 5 School FE	Column 6 School FE: no imputation	Column 7 School FE: Value Added
female	-3.124***	-2.979***	-3.109***	-2.675***	-2.679***	-2.572***	-2.779***
	(0.441)	(0.445)	(0.378)	(0.377)	(0.398)	(0.418)	(0.608)
Benin	36.102***	36.060***					
	(0.625)	(0.626)					
Cameroon	56.906***	56.841***					
	(0.614)	(0.615)					
Chad	41.947***	42.007***					
	(0.714)	(0.714)					
Gabon	51.333***	51.259***					
	(0.652)	(0.652)					
Guinea	51.258***	51.190***					
	(0.575)	(0.575)					
Madagascar	57.512***	57.439***					
	(0.557)	(0.558)					
Mauritania	33.073***	33.792***					
	(0.632)	(0.657)					
single-sex school		-4.228					
		(2.812)					
single-sex×female		-4.098					
		(3.452)					
age				2.322***	2.273***	2.245***	0.832***
				(0.136)	(0.142)	(0.149)	(0.124)
has maths book				2.561***	2.379***	2.359***	0.916
				(0.549)	(0.572)	(0.604)	(0.494)

	Column 1 Country FE	Column 2 Country FE	Column 3 School FE	Column 4 School FE	Column 5 School FE	Column 6 School FE: no imputation	Column 7 School FE: Value Added
teacher help				2.558	2.271	0.957	1.468
				(1.651)	(1.699)	(1.866)	(1.460)
middle wealth tercile					0.628	0.243	-0.200
					(0.632)	(0.671)	(0.547)
top wealth tercile					0.288	-0.093	-0.655
					(0.828)	(0.877)	(0.715)
does not live with parents					1.218	1.078	0.441
					(0.673)	(0.708)	(0.582)
single parent					0.456	0.563	-0.063
					(0.527)	(0.573)	(0.457)
father is literate					1.022	1.071	0.730
					(0.537)	(0.565)	(0.464)
mother is literate					1.005	0.811	0.093
					(0.551)	(0.580)	(0.476)
regular daily meals (0 to 3)					-0.368	-0.289	-0.421
					(0.436)	(0.458)	(0.378)
parental help					1.450*	1.414*	0.898
					(0.591)	(0.624)	(0.512)
sibling help					0.352	0.417	-0.644
					(0.540)	(0.567)	(0.467)
tutor help					2.028	1.638	1.221
					(1.312)	(1.376)	(1.134)
other help					4.477***	4.433***	2.598*
					(1.190)	(1.250)	(1.032)
domestic work					-0.232	-0.198	
					(0.561)	(0.593)	
works in agriculture or					-0.092	-0.040	

	Column 1	Column 2	Column 3	Column 4	Column 5	Column 6	Column 7
						School FE: no imputation	School FE: Value Added
	Country FE	Country FE	School FE	School FE	School FE		
trade							
					(0.491)	(0.519)	
work affects study at home					-0.557	-0.669	
					(0.689)	(0.732)	
work affects school attendance					-0.846	-1.728	
					(0.831)	(0.889)	
work affects schooling because too tired					-1.713*	-1.466	
					(0.776)	(0.831)	
Initial score (/100)							0.562***
							(0.010)
Constant			47.679***	26.863***	26.495***	26.986***	11.785***
			(0.252)	(1.202)	(1.834)	(1.923)	(1.587)
No. of Children	12362	12362	12362	12123	11460	10407	11814
No. of Schools			967	964	957	957	962
R-Squared	0.789	0.789	0.006	0.033	0.036	0.036	0.257

Table 4a: Individual Country Equations – PASEC Final Year 5 Maths Test Scores – Female Coefficients

	Column 1 No FE	Column 2 School FE	Column 3 Col (2) + age + pupil-specific school char.	Column 4 Col (3) + household characteristics	Column 5 As (4), no imputed values	Column 6 As (4), Value Added
Benin	-3.456	-3.559	-3.593	-3.532	-3.355	-2.022
P-value	0.000	0.000	0.000	0.000	0.000	0.000
No. of Children	1822	1822	1820	1761	1606	1780
R-Squared	0.010	0.025	0.027	0.044	0.049	0.236
Cameroon	-0.342	-0.428	-0.458	-0.470	-0.534	0.554
P-value	0.677	0.520	0.495	0.505	0.467	0.413
No. of Children	1735	1735	1723	1621	1480	1633
R-Squared	0.000	0.000	0.001	0.014	0.018	0.078
Chad	-0.505	-2.008	-2.070	-1.635	-1.813	-0.580
P-value	0.568	0.002	0.002	0.035	0.032	0.366
No. of Children	1237	1237	1205	1046	915	1126
R-Squared	0.000	0.009	0.012	0.033	0.038	0.164
Gabon	-1.652	-2.265	-2.272	-2.351	-3.079	-0.867
P-value	0.020	0.000	0.000	0.001	0.000	0.145
No. of Children	1491	1491	1465	1393	1191	1413
R-Squared	0.004	0.009	0.017	0.032	0.046	0.244
Guinea	-3.587	-2.787	-2.865	-2.858	-2.819	-1.395

	Column 1	Column 2	Column 3	Column 4	Column 5	Column 6
	No FE	School FE	Col (2) + age + pupil-specific school char.	Col (3) + household characteristics	As (4), no imputed values	As (4), Value Added
P-value	0.000	0.000	0.000	0.000	0.000	0.002
No. of Children	2173	2173	2148	2039	1874	2104
R-Squared	0.012	0.017	0.019	0.025	0.029	0.194
Madagascar	-1.073	-1.306	-1.590	-1.537	-1.847	-1.324
P-value	0.146	0.026	0.007	0.012	0.004	0.011
No. of Children	1964	1964	1930	1857	1721	1876
R-Squared	0.001	0.003	0.011	0.021	0.026	0.270
Mauritania	0.203	-2.158	-2.148	-2.364	-2.432	-1.179
P-value	0.806	0.000	0.000	0.000	0.000	0.031
No. of Children	1482	1482	1470	1430	1279	1454
R-Squared	0.000	0.010	0.019	0.029	0.031	0.216

Table 4b: Individual Country Equations – SACMEQ Maths Test Scores – Female Coefficients

	Column 1 No FE	Column 2 School FE	Column 3 Col (2) + age + pupil-specific school char.	Column 4 Col (3) + household characteristics
Botswana	1.405	1.524	0.454	1.070
P-value	0.002	0.000	0.273	0.036
No. of Children	3321	3321	3321	2248
R-Squared	0.003	0.004	0.063	0.134
Kenya	-3.179	-3.516	-4.378	-4.418
P-value	0.000	0.000	0.000	0.000
No. of Children	3296	3296	3279	2827
R-Squared	0.012	0.022	0.093	0.116
Lesotho	0.198	0.147	-0.456	-0.596
P-value	0.541	0.594	0.109	0.088
No. of Children	3144	3144	3144	2104
R-Squared	0.000	0.000	0.026	0.049
Malawi	-1.706	-1.418	-1.704	-1.765
P-value	0.000	0.000	0.000	0.000
No. of Children	2323	2323	2323	1882
R-Squared	0.012	0.009	0.032	0.046
Mauritius	1.877	1.493	0.520	0.540

	Column 1 No FE	Column 2 School FE	Column 3 Col (2) + age + pupil-specific school char.	Column 4 Col (3) + household characteristics
P-value	0.014	0.026	0.416	0.437
No. of Children	2870	2870	2870	2258
R-Squared	0.002	0.002	0.116	0.217
Mozambique	-3.041	-3.773	-3.807	-3.844
P-value	0.000	0.000	0.000	0.000
No. of Children	3136	3136	2999	2091
R-Squared	0.023	0.042	0.072	0.104
Namibia	-1.027	-0.552	-0.977	-0.949
P-value	0.006	0.022	0.000	0.005
No. of Children	4990	4990	4990	2750
R-Squared	0.002	0.001	0.028	0.056
South Africa	1.383	0.301	-0.326	-0.256
P-value	0.018	0.381	0.353	0.598
No. of Children	3135	3135	3113	1783
R-Squared	0.002	0.000	0.032	0.08
Swaziland	-0.518	-0.499	-1.362	-1.123
P-value	0.189	0.155	0.000	0.014
No. of Children	3138	3138	3138	2054
R-Squared	0.001	0.001	0.037	0.062

	Column 1 No FE	Column 2 School FE	Column 3 Col (2) + age + pupil-specific school char.	Column 4 Col (3) + household characteristics
Tanzania	-4.827	-5.157	-5.413	-5.423
P-value	0.000	0.000	0.000	0.000
No. of Children	2849	2849	2849	2494
R-Squared	0.031	0.048	0.095	0.153
Uganda	-1.758	-2.914	-3.153	-3.096
P-value	0.007	0.000	0.000	0.000
No. of Children	2619	2619	2619	2091
R-Squared	0.003	0.02	0.064	0.091
Zambia	-1.143	-1.478	-1.846	-2.104
P-value	0.004	0.000	0.000	0.000
No. of Children	2586	2586	2517	1977
R-Squared	0.003	0.007	0.049	0.095

Table 5a: Cross-Country Regressions

Explained variable	Relative Female Performance (Difference in correct answers /100)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
GDP per capita	0.001**	0.001***	0.001***	0.001**	0.001***	0.001***	0.001**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Gender Gap Index		-7.506					
		(7.405)					
Economic Subindex			-8.950**				
			(3.911)				
Educational Subindex				0.022			
				(3.116)			
Health Subindex					37.791		
					(46.990)		
Political Subindex						-2.750	
						(4.964)	
Primary GER Parity Index							1.953
							(3.747)
Constant	-2.048***	2.685	3.688	-2.103	-38.853	-1.762**	-3.773
	(0.477)	(4.727)	(2.555)	(2.558)	(45.722)	(0.750)	(3.348)
Countries	19	16	16	16	16	16	19
R-Squared	0.304	0.508	0.622	0.470	0.495	0.482	0.316

Notes: Standard errors in parentheses. Gender gap index data not available for: Gabon, Guinea, Swaziland.

Table 5b: Cross-Country Regressions

	Relative Female Performance (Difference in correct answers /100)				Pairwise Difference in Relative Female Performance		% Harassment
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
GDP per capita	0.001*		0.000				0.000
	(0.000)		(0.000)				(0.000)
% Heads Reporting Sexual Harassment in Schools	-9.003	-23.063***					
	(9.403)	(7.132)					
Total Fertility Rate			-0.772*	-0.910***			0.043***
			(0.378)	(0.249)			(0.010)
Pairwise Difference in Total Fertility Rate					-0.911***	-0.913***	
					(0.127)	(0.128)	
Genetic Distance						0.000	
						(0.000)	
Constant	-1.401	0.901	2.432	3.334**	0.080	0.050	-0.125*
	(1.319)	(0.737)	(2.235)	(1.282)	(0.384)	(0.434)	(0.061)
Countries or Country dyads	12	12	19	19	171	171	12
R-Squared	0.662	0.511	0.448	0.440	0.440	0.440	0.852

Notes: Standard errors in parentheses. Harassment data only collected in SACMEQ. Observations in Columns (4) and (5) are dyads. For these two regressions, standard errors are clustered at the country level. Genetic distance from Spolaore and Wacziarg (2009).

Appendix Table A1: Pooled Results – PASEC Initial Year 5 Maths Test Scores

	Column 1 Country FE	Column 2 Country FE	Column 3 School FE	Column 4 School FE	Column 5 School FE	Column 6 School FE
female	-3.038***	-3.155***	-3.646***	-3.655***	-3.645***	-3.552***
	(0.316)	(0.318)	(0.246)	(0.248)	(0.263)	(0.279)
Benin	46.367***	46.413***				
	(0.419)	(0.418)				
Cameroon	53.090***	53.248***				
	(0.460)	(0.461)				
Chad	37.986***	38.378***				
	(0.473)	(0.476)				
Gabon	51.615***	51.673***				
	(0.436)	(0.436)				
Guinea	44.716***	44.767***				
	(0.398)	(0.398)				
Madagascar	62.087***	62.192***				
	(0.425)	(0.425)				
Mauritania	26.630***	27.171***				
	(0.475)	(0.484)				
single-sex school		-9.833***				
		(1.621)				
single-sex×female		6.946**				
		(2.334)				
age				0.017	0.039	0.064
				(0.080)	(0.084)	(0.090)
has maths book				1.982***	1.871***	1.995***
				(0.345)	(0.359)	(0.387)
teacher help				-0.884	-0.954	-0.921
				(0.833)	(0.879)	(0.945)

	Column 1 Country FE	Column 2 Country FE	Column 3 School FE	Column 4 School FE	Column 5 School FE	Column 6 School FE
middle wealth tercile					0.597	0.314
					(0.397)	(0.422)
top wealth tercile					1.215*	0.770
					(0.476)	(0.505)
does not live with parents					0.998*	1.082*
					(0.405)	(0.432)
single parent					0.390	0.380
					(0.324)	(0.350)
father is literate					-0.064	0.038
					(0.350)	(0.369)
mother is literate					-0.081	-0.165
					(0.356)	(0.379)
regular daily meals (0 to 3)					-0.131	-0.035
					(0.257)	(0.272)
parental help					0.378	0.216
					(0.371)	(0.394)
sibling help					-0.297	-0.277
					(0.330)	(0.349)
tutor help					0.648	0.777
					(0.717)	(0.760)
other help					0.188	0.334
					(0.634)	(0.671)
domestic work					-0.152	-0.317
					(0.402)	(0.433)
works in agriculture or trade					-0.692*	-0.839*
					(0.331)	(0.353)
work affects study at home					-0.152	-0.033
					(0.449)	(0.482)

	Column 1 Country FE	Column 2 Country FE	Column 3 School FE	Column 4 School FE	Column 5 School FE	Column 6 School FE
work affects school attendance					-1.281*	-1.329*
					(0.548)	(0.596)
work affects schooling because too tired					-1.160*	-1.182*
					(0.496)	(0.539)
constant			47.136***	46.047***	46.076***	45.873***
			(0.160)	(1.011)	(1.384)	(1.469)
no. of Children	13747	13747	13747	13573	12763	11453
no. of Schools			956	956	948	947
R-Squared	0.867	0.867	0.017	0.020	0.024	0.023

Notes:

1. Coefficients are reported together with their standard errors. * p<0.10; ** p<0.05; *** p<0.01.
2. Omitted country category is Benin.
3. ‘help’ refers to help with homework.
4. Wealth terciles based on PCA except for Chad and Guinea.
5. Class equipment terciles based on PCA.