A Preliminary Analysis of SACMEQ III South Africa

NIC SPAULL

Stellenbosch Economic Working Papers: 11/11

KEYWORDS: SACMEQ, SOUTH AFRICA, PRIMARY EDUCATION, EDUCATION, EDUCATION PRODUCTION FUNCTION, EDUCATION POLICY, ECONOMICS OF EDUCATION
JEL: I20, I21, I28

NIC SPAULL
DEPARTMENT OF ECONOMICS
UNIVERSITY OF STELLENBOSCH
PRIVATE BAG X1, 7602
MATIELAND, SOUTH AFRICA
E-MAIL: SPAULL@SUN.AC.ZA

UNIVERSITEIT
STELLENBOSCH
UNIVERSITY

A WORKING PAPER OF THE DEPARTMENT OF ECONOMICS AND THE
BUREAU FOR ECONOMIC RESEARCH AT THE UNIVERSITY OF STELLENBOSCH

A Preliminary Analysis of SACMEQ III South Africa[1]

NIC SPAULL

ABSTRACT

The many and varied links between student socio-economic status and educational outcomes have been well documented in the South African economics of education literature. The strong legacy of apartheid and the consequent correlation between education and wealth have meant that, generally speaking, poorer students perform worse academically. The present study uses the recent Southern and East African Consortium for Monitoring Educational Quality (SACMEQ III) dataset for South Africa to identify those factors that have a significant effect on student maths and reading performance in Grade 6. The research confirms previous findings that socio-economic status, and particularly school socioeconomic status, is important when understanding student success or failure. Other factors which contribute significantly to student performance are homework frequency, preschool education, and the availability of reading textbooks. In contrast, teacher-subject knowledge was found to have only a modest impact on Grade 6 student performance. Policy interventions are also highlighted. The study concludes that South Africa is still a tale of two schools: one which is wealthy, functional and able to educate students, while the other is poor, dysfunctional, and unable to equip students with the necessary numeracy and literacy skills they should be acquiring in primary school. Nevertheless, it suggests that there are some options available to policy-makers which are expected to have a positive effect on student performance.

Keywords: SACMEQ, South Africa, primary education, education, education production function, education policy, economics of education
JEL codes: I20, I21, I28

[1] This document was produced within the Social Policy Research Group in the Department of Economics at Stellenbosch University with the financial assistance of the PSPPD (Programme to Support Pro-Poor Policy Development in South Africa), a partnership programme between The Presidency, Republic of South Africa, and the European Union (EU). The contents do not necessarily reflect the position of The Presidency or the EU.

1. Introduction

The many and varied links between student socio-economic status and educational outcomes have been well documented in the South African economics of education literature[2]. The strong legacy of apartheid and the consequent correlation between education and wealth have meant that, generally speaking, poorer students perform worse academically. Unfortunately little has changed. The links between affluence and educational quality can partially explain this outcome since the poor receive a far inferior quality of education when compared to their wealthier counterparts. This is troubling for two reasons:

1) The received wisdom in economics dictates that an individual's labour-market prospects are directly correlated with their stock of human capital, which itself is correlated with the quality and duration of schooling. Offering an inferior quality of education to the poor disadvantages them in the labour-market and entrenches their poverty. The inter-generational effects of this inadequate education mean that children of impoverished parents are likely to be poor themselves.
2) Given the racial dimension of poverty, and that the poor are more likely to be black, one can go further and say that on average, black students receive an inferior quality of education to their white peers. That this is the reality 17 years on from apartheid is particularly disconcerting.

It is therefore necessary to improve the quality of education provided to the poor if these cycles of poverty are to be broken. While it is easy to understand why affluent schools outperform poor schools, it is less clear why certain poor schools succeed where other, equally poor schools, fail. Is this difference due to variations in school management, socio-economic status, and the provision of textbooks? Or perhaps differences in teacher quality, parental education, and preschool education? Qualitative analysis is unable to reveal the answers to these questions since it cannot isolate specifically which of these factors, or a myriad of others, is the determinant of student success or failure. While not without its own problems, quantitative analysis is more able to identify these factors, and in addition, to quantify the size of their impact.

A serious problem which plagues quantitative analysis is limitations of the data. What question can one pose to capture 'school quality' or 'teacher motivation'? Or perhaps the 'dedication of a

[2] See Donaldson (1992); Crouch & Mabogoane (1998); Lam (1999); Anderson *et al* (2001); Van der Berg (2001); Van der Berg & Burger (2003); Ross & Zuze (2004); Fiske & Ladd (2004); Gustaffson & Patel (2006); Gustaffson (2007); Van der Berg & Louw (2006); Bhorat & Oosthuizen (2006); Van der Berg (2007); Van der Berg (2008); Taylor & Yu (2009).

principal to the success of his students'? These are all likely to be highly influential factors which can go some way to explain why two equally-resourced schools perform very differently. These data limitations mean that we are not able to capture, as well as we would like, those factors which we know are important. Since the quality of the analysis is capped by quality of the data, one can only infer so much from limited data. Nevertheless, the importance of quantitative educational research in South Africa warrants our research attention in spite of these difficulties.

The recently conducted Southern and East African Consortium for Monitoring Educational Quality (hereafter SACMEQ[3]) survey in 2007 provides the research community with new data on primary education in South Africa. The survey asked a representative sample of Grade 6 students various demographic and schooling questions, and tested their academic ability in reading and maths. By including these demographic and schooling variables in a regression explaining student test scores, we are able to identify those factors associated with differences in student performance.

The aim of this research is therefore to analyse this new dataset using education production functions to determine which of the included variables have the greatest impact on student performance, and thus, which should receive policy priority.

1.1 Background to SACMEQ III

SACMEQ is a consortium of education ministries, policy-makers and researchers who, in conjunction with UNESCO's International Institute for Educational Planning (IIEP), aim to improve the research capacity and technical skills of educational planners (Moloi & Strauss, 2005: 12). By generating information from school surveys SACMEQ enables decision-makers to monitor general conditions of schooling and the quality of basic education. SACMEQ III, which is the most recent SACMEQ survey, was conducted in 2007 in South Africa. It is also the latest survey data on primary-school student performance in the country.

Murimba (2005) discusses the origins, vision, mission, goals, structure, context, implementation and methodology of SACMEQ. Additionally, SACMEQ (2010:1) provides an overview of the three SACMEQ projects: SACMEQ I, II and III. This is included verbatim in Appendix C for those unfamiliar with the three SACMEQ surveys.

The South African SACMEQ II data has been the basis of numerous research papers over the past 6 years. These include: analysing the education system using Hierarchical Linear Modelling (HLM) (Van

[3] The countries that participated in the third SACMEQ survey were: Botswana, Kenya, Lesotho, Malawi, Mauritius, Mozambique, Namibia, Seychelles, South Africa, Swaziland, Tanzania (Mainland), Tanzania (Zanzibar), Uganda, Zambia, and Zimbabwe

der Berg & Louw, 2006; Gustaffson, 2007), developing indicators for quality, access and equity (Strauss, 2005), investigating cross-country differences in conjunction with SACMEQ's West African equivalent PASEC (Programme d'Analyse des Systèmes Educatifs de la CONFEMEN) (Fehrler *et al.*, 2009), understanding the relationship between socioeconomic status and student performance (Van der Berg, 2008; Taylor & Yu, 2009) as well as a comprehensive country report by Moloi and Strauss (2005).

2. Data

The SACMEQ III dataset is a useful addition to existing primary school data in South Africa. The high quality and depth of the SACMEQ education data is rare in developing countries, of which South Africa is no exception. The SACMEQ team has assembled a dataset that has been collected in a technically rigorous way, creating a quality dataset which is likely to yield much insight.

The SACMEQ III survey conducted in 2007 surveyed 9083 Grade 6 students and 1488 teachers from 392 schools across South Africa. Students completed three tests – Maths, Reading and Health – and, in addition, gave extensive demographic and home-background information. The teachers included 498 reading teachers, 498 maths teachers, and 492 health teachers. Each completed the Health test, with maths and reading teachers also completing subject-specific tests for their respective disciplines. The school head was also surveyed and asked numerous questions relating to the school.

This was the first nation-wide education survey in South Africa where teachers were tested in addition to students[4]. This provides a valuable opportunity to better understand the impact of teacher-knowledge on student performance.

The SACMEQ tests and questionnaires were only available in two languages: English and Afrikaans. Consequently, it is almost certain that students who do not speak English or Afrikaans as a first language would be at a disadvantage. Given that isiZulu and isiXhosa are the two most spoken languages in South Africa, this disadvantage is likely to extend to the majority of students. Furthermore, language, race and socio-economic status are highly correlated. Those students who speak English as a first language are more likely to be white or Indian and affluent. The correlation between Afrikaans and SES is less strong but still relevant. One must therefore be cautious in

[4] Although the SACMEQ II questionnaire did contain a teacher-test, due to South African teacher-union objections, South Africa was one of the few SACMEQ countries that did not complete the teacher-test section of the SACMEQ II survey. This being said, in SACMEQ III teachers were allowed to refuse to write the tests, which some of them did. This is discussed in greater detail in section 4 'Regressions' below.

attributing the entire difference between wealthier students and poorer students to SES, since some of this difference may arise due to linguistic advantage.

The sample was stratified both by province (explicit strata) and school size (implicit strata). The 'province' stratification was accomplished by separating the sampling frame into provincial lists before undertaking the sample, while the 'school size' stratification used the number of Grade 6 students in each school. The sampling method of *probability proportional to size* (PPS) was used to select schools within strata and simple random sampling was used to select students within schools (SACMEQ, 2010: 4).

All questions in the survey were multiple-choice, with 55 reading questions, 49 maths questions and 86 HIV/AIDS questions. The level of the Reading-test questions ranged from *Level 1: Pre Reading* to *Level 8: Critical Reading*; the Maths-test questions ranged from *Level 1: Pre Numeracy* to *Level 8: Abstract Problem Solving*. For a more detailed exposition of what each level entails and a comprehensive breakdown of summary statistics by performance level, see SACMEQ (2010: 6). The Health test consisted of 86 true-or-false questions regarding HIV/AIDS. The results from all three tests were standardised by SACMEQ to have a mean of 500 and a standard deviation of 100 in the first survey that any of these tests was introduced, and results from subsequent waves were converted to the same metric[5].

The teacher Reading and Maths tests were very similar to the student tests, with only a few more challenging questions included. SACMEQ used Rasch scaling to transform these scores into values which are directly comparable with student test-scores. Teachers wrote the same Health test as the students.

2.1 Variables created

In SACMEQ III, as is the case with most surveys which target children, it is not possible to get an accurate representation of the monetary value of family income. Consequently, socio-economic status (SES) was inferred from a series of possession questions. In SACMEQ III, students were asked whether or not each of 31[6] items was found in the place where they stayed during the school week. To construct an SES variable, all 31 items were used in a Multiple Correspondence Analysis (MCA) forming the SES index. The SES variable was transformed to be the negative of the MCA index to ensure that the largest positive value of MCA was assigned to the wealthiest student for ease of

[5] These figures are for all SACMEQ-participating countries, thus 500 is the mean of the data when all participating countries are seen together, and have been standardised to the SACMEQ II mean. South Africa has a mean of 498 for reading and a standard deviation of 115, thus South Africa is marginally below the SACMEQ II average reading score (500) with greater variation than the SACMEQ average standard deviation (100).

interpretation. The SES variable in the regressions is normal with a mean of -0.05, a standard deviation of 1.008, while the minimum and maximum values are -2.22 and 2.38 respectively. A MCA was used rather than Principal Component Analysis (PCA) since MCA is more appropriate for categorical variables (Booysen *et al.,* 2008:2).

Although some researchers have included parental education in the SES variable, for the purposes of this analysis parental education has been treated separately with 2 variables: dummies for both '*mother has matric'*, and '*at least 1 parent has a degree*'. Taylor and Yu (2009: 62) provide a thorough analysis regarding the use of parent's education in deriving an index for SES, discussing the methodological complications and outlining the various ways of including parental education in an SES index.

In addition to the SES variable, a '*school-SES'* variable was created by taking the average of all the students SES scores in that specific school and assigning this average to each student as school-SES.

Two index variables, '*school building'* and '*school equipment'*, were created and reflect the total number of items a school had in each of these categories. The school equipment index included items such as a telephone, clock, photo-copy machine etc., with a maximum of 18 items. The school building index included items such as a library, staff-room, store room etc., with a maximum of seven items. All other variables used in the regressions are self-explanatory.

3. Descriptive analysis

3.1 Socio-economic Status

Decomposing student performance by socio-economic status shows that 17 years on from apartheid, South Africa's education system is still a tale of two schools (Figures 1 and 2 below[7]). The reading and maths distributions show that students from the uppermost quintile of SES far outperform students from the lower four quintiles. When decomposed by quintile, the distribution is bi-modal by top quintile and bottom-four quintiles, suggesting that there may be two data-generating processes at work. It would seem that student performance does not improve evenly across the various SES quintiles.

[6] These 31 items were: daily newspaper, weekly or monthly magazine, clock, piped water, bore hole, table to write on, bed, private study area, bicycle, donkey/horse cart, car, motorcycle, tractor, electricity (mains, generator, solar), refrigerator/freezer, air-conditioner, electric fan, washing machine, vacuum cleaner, computer, internet, radio, TV, VCR player, DVD player, CD player, audio-cassette player, camera, digital camera, video camera, telephone/cell-phone (from Question 14 in Student Questionnaire).

[7] All kernel density curves presented use the Gaussian kernel function with no additional smoothing.

The South African reality that certain provinces are wealthier than others means that these socio-economic differentials necessarily extend to geographic differentials as well (Figures 5 and 6 below). The wealthiest two provinces, the Western Cape and Gauteng, have a different distribution of student maths and reading scores as compared to the other seven provinces.

The strong correlations between student test scores, SES, and the wealthiest two provinces may indicate some form of causation between wealth and performance. However, it should be stressed that SES is not necessarily the main reason why the wealthier provinces perform better academically. It may be that parents with higher than average ability are more likely to work and live in the commercial hubs of Gauteng and the Western Cape and thus students in these provinces may have higher than average ability. Or it could be that the education departments and schools in these two provinces are better managed and more able to create environments where students excel. Or it could also be that wealthier parents place a higher premium on education as compared to poorer parents and thus spend more time, energy and money on improving their children's education. More than likely, it is a combination of all of the above effects. In each of these cases, SES still plays an indirect role in improving student performance.

To further understand the links between socio-economic status and student performance, SES was used as the explanatory variable on student maths and reading performance in a locally weighted regression (Lowess). These are shown in Figure 7 below. The shapes of both curves suggest that student SES is only significantly positively related to student performance at higher levels of SES. The kernel density of SES was superimposed on the Lowess curves to show that only a relatively small sub-sample of students fall in the range where additional SES is beneficial. To be specific, the fourth quintile of SES begins at an SES value of 0.19 and the fifth quintile at 0.89. Thus, 60% of students lie to the left of an SES value of 0.19 in the graph below, i.e. the flat portion of the Lowess curves. It is interesting to note that this threshold area corresponds to a student reading/maths score of 500, which is approximately the SACMEQ mean.

A complete list of summary statistics for all variables used in the paper is reported in Appendix A. The statistics included are: number of observations, mean, standard deviation, minimum, and maximum.

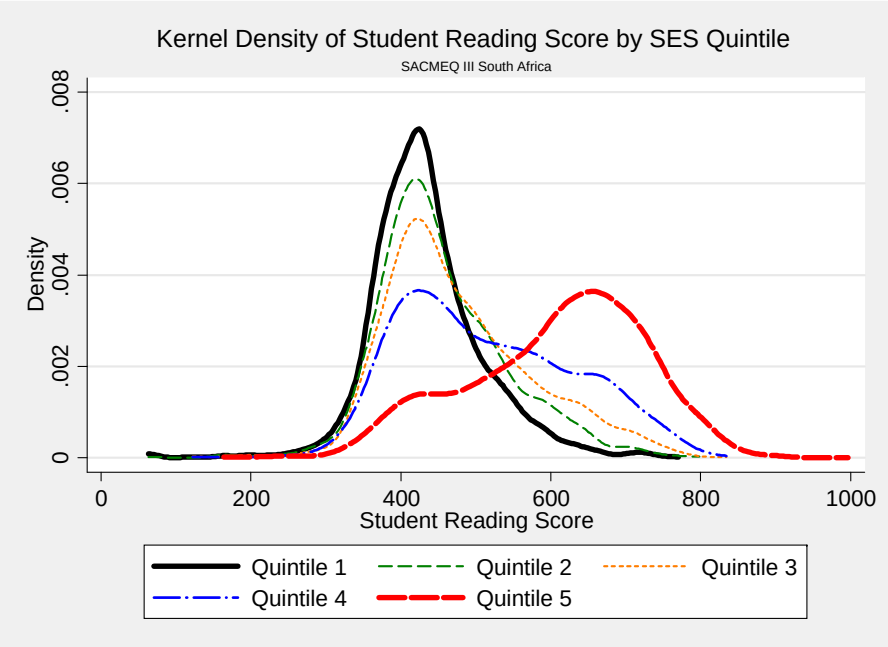

Figure 1

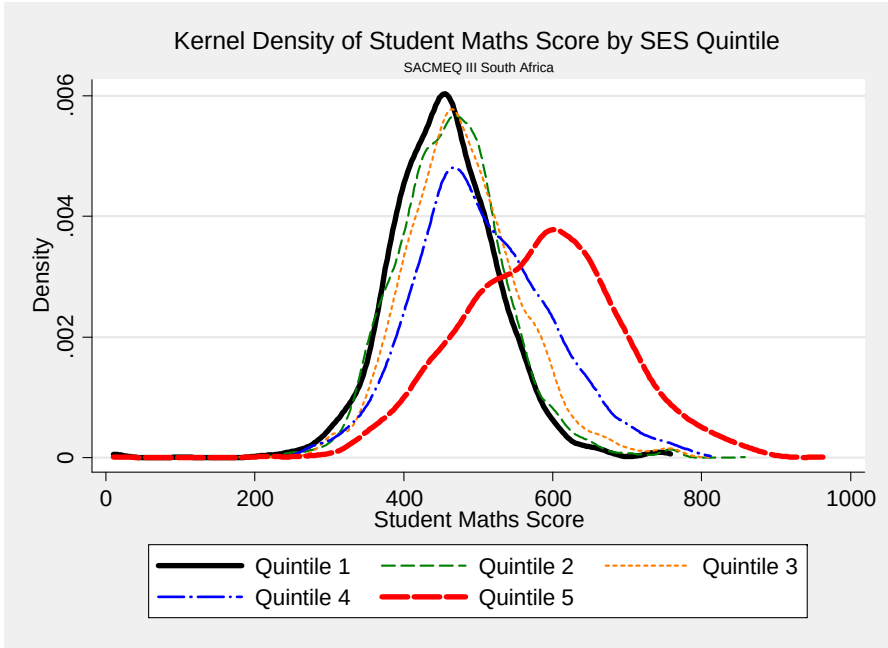

Figure 2

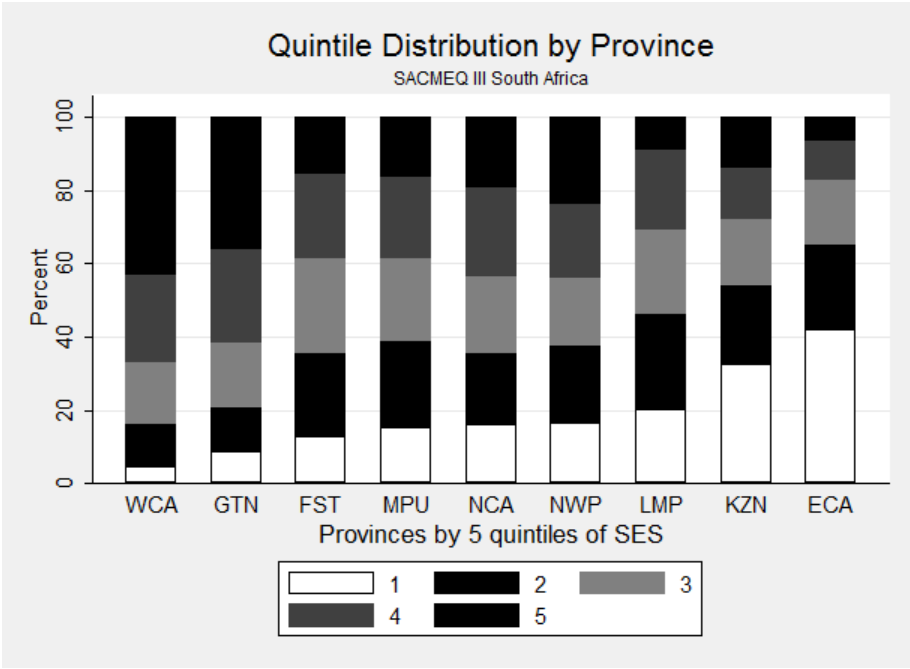

Figure 3

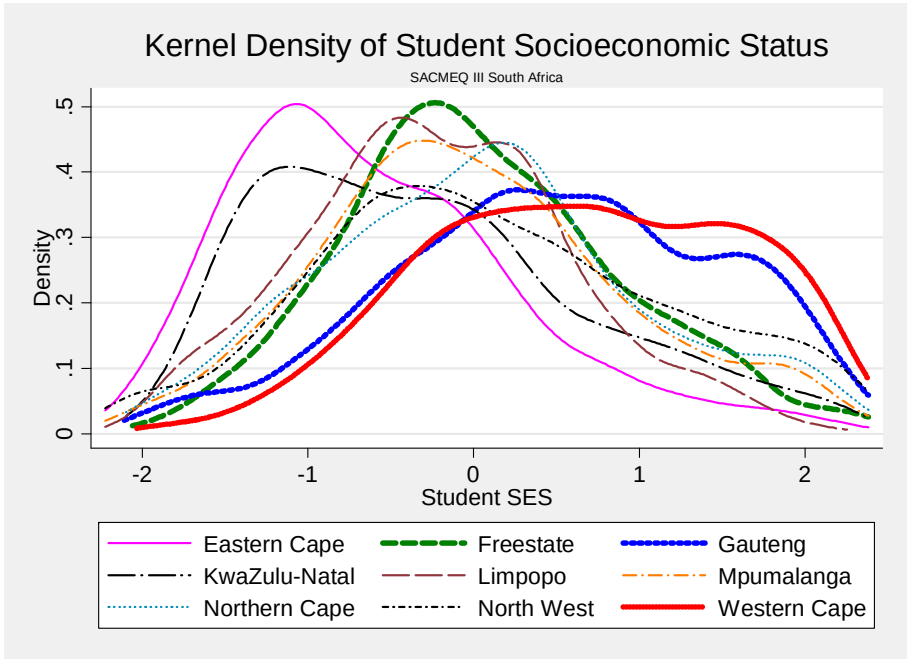

Figure 4

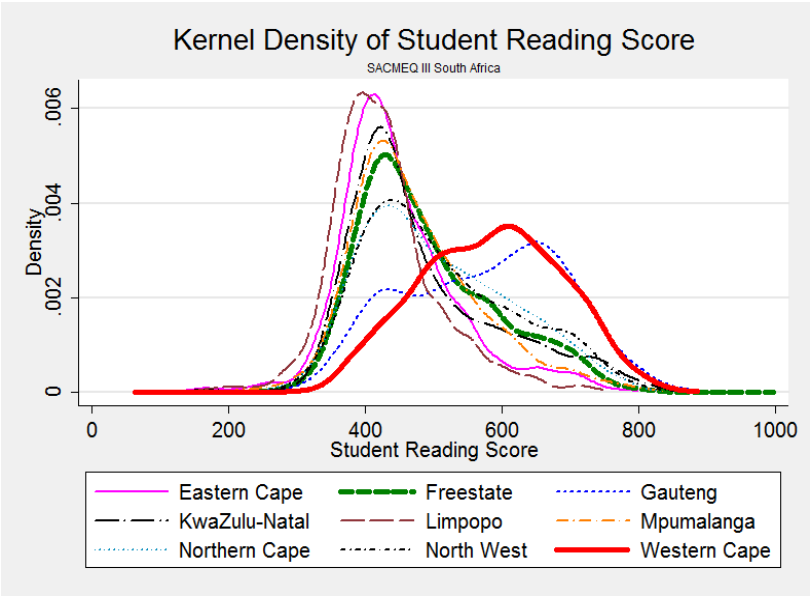

Figure 5

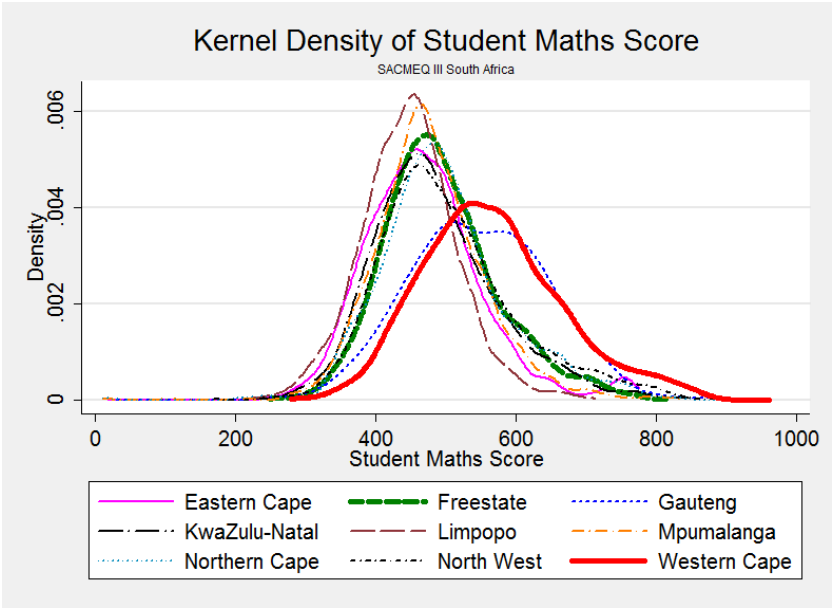

Figure 6

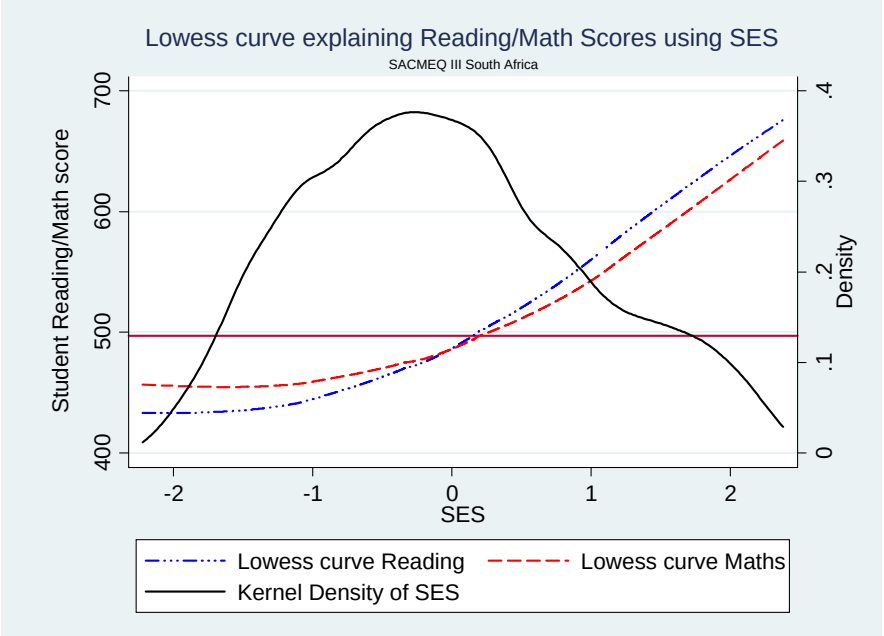

Figure 7

4. Regression analysis[8]

In the introduction to this paper I outlined the benefits of quantitative research, specifically its ability to identify and isolate the impact of different variables. However, this process of isolation is not without its own complications as many of the variables are dependent on each other, sometimes in complex ways. For example, it is difficult to disentangle the multi-directional causation between socio-economic status, geographical location, school-quality and student performance. Do good schools make students clever, and it just so happens that those students are wealthy? Or are wealthier students academically advantaged from a conducive home-background, such that already-clever students go to wealthier schools? In reality, it is likely to be a combination of the two. It is possible to ameliorate the effects of these complications if the correct interpretations are applied. For example, if the variable "Curriculum statement was present in class" had a very large and significant coefficient, one should question whether that curriculum statement was the *cause* of the positive impact on student performance. Perhaps it is signalling a motivated teacher or a generally well-organised school. If so, the large coefficient is probably more attributable to high teacher motivation or school organisation. This highlights the importance of interpretation, and also the complementarity of qualitative studies which often reveal such nuances.

To investigate the causal impacts of student and school characteristics on student performance, a selection of variables was regressed on student reading score, student maths score, and student health score. Although some coefficients are more sensitive to the model specification than others, the *relative* size and significance of coefficients is surprisingly consistent, irrespective of the specification. This is encouraging since the purpose of these education production functions is to show which of the included variables have the greatest impact on student performance and thus, which should receive policy priority. Following from this premise, the interpretations below aim to highlight which factors have a significant influence on student performance.

To account for the survey design of SACMEQ, STATA's built-in '*svy*' command was used in all the student regressions, with the primary sampling unit (PSU) being the school. This accounts for sample-stratification and the clustering of errors while also weighting the observations in order to be representative of the population.

The regression results can be found in Appendix B, and include: 'Student reading, maths and health regressions' (Table B1), 'Variations in student reading regressions' (Table B2), 'Variations in student

[8] Should researchers wish to verify these results or the variable-creation process, they can contact the author for his STATA do-files and original data, once these data have been released.

maths regressions' (Table B3), and 'Variations in student health regressions' (Table B4). Since the Health test was entirely 'True or False'-type questions, this introduces 'noise' into the data due to guessing. Consequently, the coefficients in the health regressions are likely to be far less accurate than those of the reading and maths regressions. As such, most of the interpretation below does not focus on the health regressions, but they were included in the appendix for the sake of completeness.

4.1 Model variations

For each of the three student tests (reading, maths and health), four different regressions were run. The first two used the entire sample with the only difference being the exclusion of the 'teacher test score' variable in one of the regressions. The third and fourth regressions limited the sample to the top quintile of SES, and the bottom four quintiles of SES respectively (i.e. the wealthiest 20% of students and the poorest 80% of students). The reasons for these regressions are as follows:

4.1.1 Sample-selection issues involving 'teacher test score'

Although teachers were asked to complete the teacher test, they were allowed to refuse to write it. Subsequently of the 498 reading teachers, 83 did not write the reading-teacher's test (16.7%), of the 498 maths teachers, 97 did not write the maths teacher's test (19.5%), and of the 492 health teachers, 65 teachers did not write the health test (13.2%). This creates a problem if one wishes to include the teacher test score variable since doing so reduces the sample size by approximately the same percentage as the proportion of those teachers that did not write the test: roughly 15%. This is because it is only possible to include those students in the sample whose teachers wrote the test, and thus have non-missing values for this variable. Since there is likely to be a sample selection issue at play, with weaker teachers refusing to take the test, it is possible that limiting the sample could bias the results. If the missing values are not missing-at-random (MAR) their exclusion will necessarily bias the coefficients. The question is therefore the severity of that bias, and not the presence of absence of it.

Although it may be possible to impute teacher-test scores, the fact that the selection process likely depends on the same variable as that which would be imputed (i.e. teacher knowledge) means that any imputation method would have its own complications. As such, teacher test scores were not imputed. In order to see whether limiting the sample would change coefficients in a material way, two regressions were run for each student regression; one including teacher test score (and thus a

smaller sample size), and one excluding teacher test score (with the full sample).[9] These results are reported in Tables B2, B3, and B4 in Appendix B.

Apart from a few relatively minor variables becoming significant or insignificant where previously they were not, the coefficients on most variables did not change much. Thus, while it is unfortunate that we do not have teacher test-scores for around 15% of students and that there is no easy way to impute this variable, the opportunity to include a teacher subject-knowledge variable is valuable enough to warrant limiting the sample. Therefore, teacher test-scores were included in all other regressions.

4.1.2 Splitting the student sample by SES quintile

Given that the student distributions of reading and maths scores are very different for the top quintile as compared to the bottom four, I ran two regressions: one on students from the top SES quintile, and the other on students from the bottom four quintiles of SES. The aim was to determine if the same factors are equally important for each of these sub-sets of students – i.e. whether they share the same data-generating process.

It is important to remember that regression coefficients for dummy variables are calculated with reference to a base category *in that sample*. Therefore, one must take care when comparing coefficients between these two variations due to the drastically different samples. For example, in both the reading and maths regressions, the coefficient on 'extra tuition' is almost three times as large for the top quintile regression as compared to the bottom-four-quintile regression. Since we expect the quality of education offered to the richest 20% of students to be much higher than that offered to the poorest 80%, students in the top quintile will only attend extra tuition if they are performing particularly badly (i.e. they are weak students). Thus this variable is most probably indicating which students are underperforming and therefore attending extra classes. This may not be the case for poorer students. Average students, not only underperforming students, may attend extra lessons due the lower quality of education provided to poorer students in their normal school hours. Consequently, average students attending extra lessons may moderate the signalling effect of this variable.

[9] The alternative of adjusting for sample selection bias by using a Heckman two-step model, which first models the selection equation and then the variable of interest, was not pursued in this case because of the difficulty of finding an appropriate exclusion restriction, i.e. one or more variables linked to the selection/participation process but not affecting the variable of interest, student performance.

4.2 Model fit

The R-squared output of three standard regressions (Table B1) show that the standard model specified in this paper is best able to explain Reading-scores (0.599), then Maths-scores (0.491) and least able to explain Health-scores (0.345). This difference in explanatory power, at least between the Reading and Maths scores, has been found elsewhere in the literature for similar data (Van der Berg, 2008: 27).

When the samples are limited to the top SES quintile and the bottom-four SES quintiles, it is interesting to see that the included variables are able to explain more variation in wealthy-student performance (60.7% for reading and 55.2% for maths (Table B2 & B3) than poorer student performance (45.6% for reading and 31.3% for maths (Table B2 & B3). The most likely cause of this difference is that variables that are important for understanding poor students' performance have been excluded from the model. Variables such as school management and teacher quality are thought to be extremely important in understanding why some poor schools perform better than others. If the variation in school management and teacher quality is greater between poor schools than between wealthy schools, as we expect to be the case, then the exclusion of these variables will affect the bottom-four-quintile regression more than the top-quintile regression. Capturing these variables in a survey questionnaire is a difficult task, but necessary if we are to explain why some poor schools perform well in spite of their disadvantaged background.

It must also be noted that there is greater variation among student test scores in the top quintile compared to the bottom four. In addition to the above explanation, this also contributes to the higher R-squared for the top quintile regressions.

4.3 Frequency of English spoken at home

Students who spoke English 'always' in the home environment scored 37.5 points higher on the literacy test than those that did not, while students who spoke English 'sometimes' at home scored 19 points higher on average *ceteris paribus*. The impact of either speaking English 'sometimes' or 'always' at home was also large and significant for the numeracy score and the health score.

These positive returns to speaking English at home can partially be explained by the fact that the SACMEQ III tests were only conducted in English and Afrikaans. As argued earlier, given that isiZulu and isiXhosa are the two most spoken languages in South Africa, these students are likely to be at a disadvantage to their native English-speaking counterparts. Hence, one would expect those who spoke English 'sometimes' or 'often' in the home environment to better understand the literacy and numeracy tests, and thus to perform better.

4.4 Whether a student has used a computer before

The coefficient on the variable 'Used a computer before' is large and significant across all three regressions. Clearly those students who have used a computer before are more academically able than those that have not. While this relationship may be causal, i.e. computer use increases student numeracy or literacy, it is also possible that this variable could simply be distinguishing between already better performing students and weaker students through socio-economic status. On a technical note, it is prudent to ask whether this variable introduces multicollinearity (with SES) in the regression. Consequently the regressions were run with and without the 'Used PC' variable. Since SES did not change substantially in size or significance, 'Used PC' was kept as a separate variable.

4.5 Frequency of grade repetition

The large negative impact from grade repetition can be seen across all three regressions with the negative effect increasing as the number of grade-repetitions increases. Considering that grade repetition is meant to bring students up to the required level by holding them back a year, it is disconcerting that even after repeating a grade, these students perform consistently worse than those that did not repeat. What is all the more worrying is that grade repetition is more prevalent amongst the poorer quintiles (Table D1).

While it is tempting to conclude that grade repetition is not helping students, or worse, is harming students, one cannot make such conclusions from these coefficients. This is because we expect weaker students to repeat grades more often than stronger students. While repeating the grade may or may not help, these students are still likely to be on the lower end of the performance distribution after repeating the grade. If this is the case, these variables are also signalling which students are weaker to begin with, making it difficult to draw strong conclusions about the usefulness, or lack thereof, of grade repetition.

4.6 Frequency of homework

If students received homework either once or twice a week or most days of the week, they performed significantly better than students who received homework less frequently. There was no discernable effect for students who only received homework once or twice a month or not at all. Students in the lowest two quintiles of SES received the least homework overall. Between 12% and 15% of students from the bottom four quintiles received homework only once or twice a month or not at all. This is compared to only 6% of top-quintile students (see Table D2). Clearly students who receive homework frequently are more literate, more numerate and more knowledgeable about HIV/AIDS. The fact that the positive impact of homework is highly significant and stable across all

three standard regressions lends credibility to the notion that homework is important for student performance. Given that this is a relatively easy and almost cost-free policy option, teachers should be encouraged to give regular homework to students.

According to the received wisdom in the pedagogical literature, practice is imperative for student learning. Therefore, it seems logical that the benefits of homework are derived because students are practising to read and practising mathematical problems and that this process improves learning. However, due to differing home-backgrounds, homework frequency may not have the same impact for poor students compared to more affluent students. This is because wealthier students are more likely to be encouraged to complete their homework by their parents or caregivers. This is in stark contrast to poorer students who often have many chores to do after school, may have no access to electricity, and little private time to work. In addition, poorer students may have to work after school hours in order to supplement the low household income. Consequently, innovative solutions such as "after-school home-work clubs" or similar initiatives may be necessary if poorer students are to reap the benefits of increased homework.

4.7 Impact of socioeconomic status:

Both the literature and the preceding descriptive analysis suggest that socio-economic status is an important variable in understanding student performance, and indeed this is the case. Of all the variables included in the model, school socio-economic status has the largest impact on student performance. The Lowess curves in Figure 7 above indicate that the effect of socio-economic status on student performance is non-linear. Put differently, wealth has a greater effect on student performance at higher levels of wealth. While Figure 7 is for *individual* socio-economic status, the Lowess curves for *school* socio-economic status show a similar pattern. Consequently, individual and school SES were included in the model in a quadratic form. Both individual SES and school SES are jointly significant in almost all models presented in the Appendix B.

One of the most interesting findings of this research project was how dominant school SES was as a determinant of student performance. Across all the regressions school SES remains particularly large, surprisingly stable and highly significant. The relative sizes of the coefficients on individual SES and school SES indicate that a school's overall socio-economic status has a greater impact on learner performance than does a child's individual status. This means that placing a poor child in a wealthy school is likely to more than compensate for any negative effects of a poor home background. The sheer size of the school SES coefficients, and the fact that school SES ranges from -1.8 to +1.8, means

that the contribution of school SES to student performance for wealthy students dwarfs any other single variable in the model.

Plotting the school SES and school SES squared coefficients from the standard reading and maths regressions (coefficients from Table B1), shows both the non-linearity and the combined magnitude of these coefficients. For students with high school SES, the return to being in that wealthy school can exceed 100 points – approximately one standard deviation for both reading and maths.

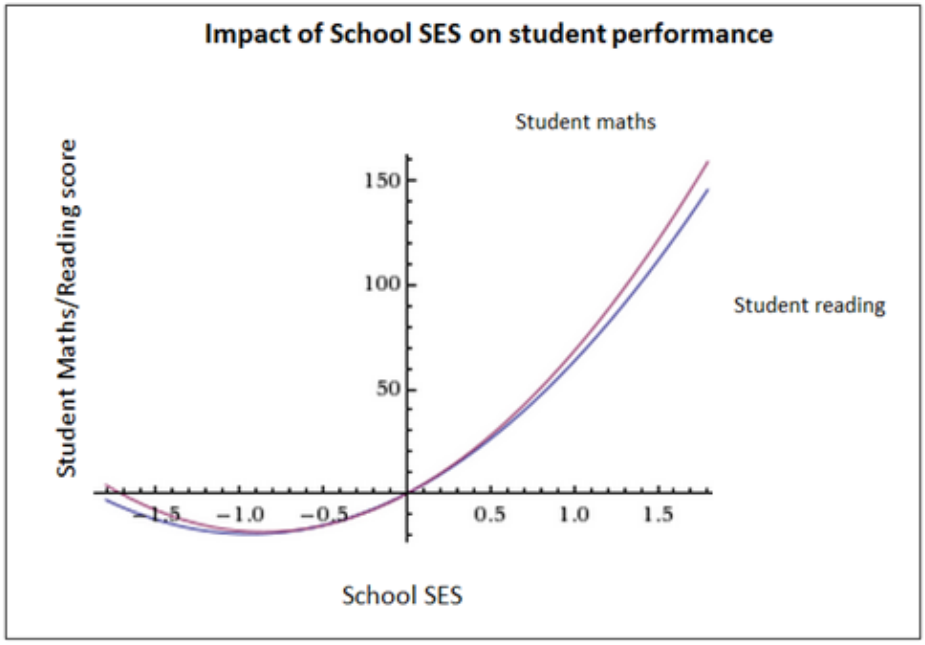

Figure 8

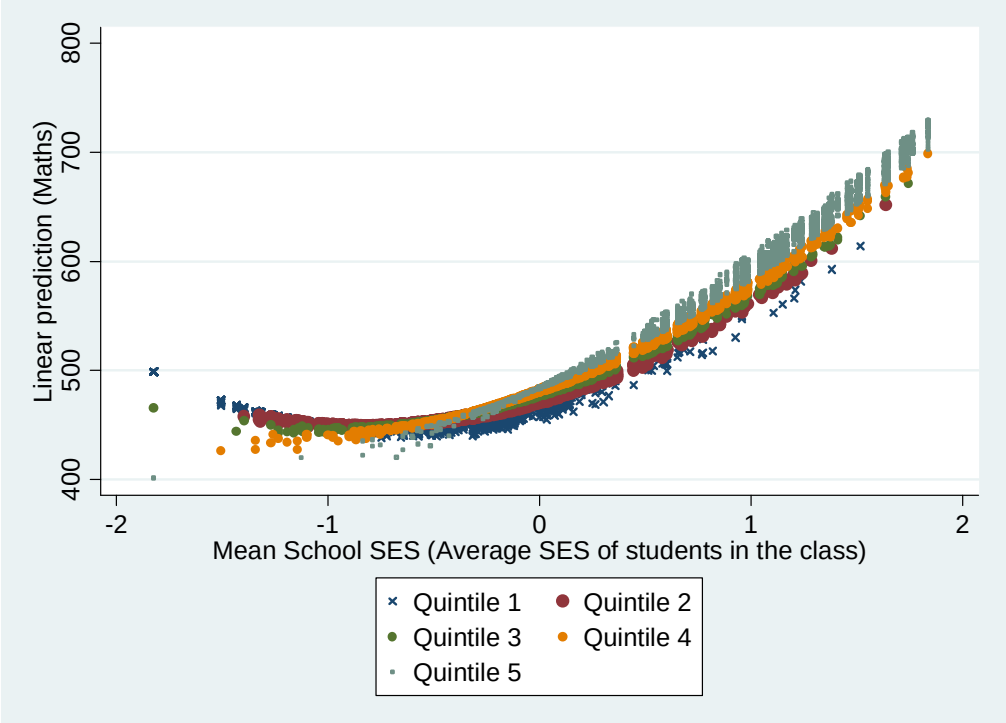

Figure 9

While the above graph and analysis have shown that *school* SES is more important than *student* SES in determining student performance, it is unclear whether there is an interaction effect between these two variables. Since both were included in a quadratic form, interacting all four variables makes the interpretation of this combined effect cumbersome. Instead of including interaction effects, I predicted the student maths score conditional on student SES, school SES, and their quadratics. By plotting these predicted values on school SES (Figure 9) it become possible to see the variation in student maths performance due to *individual* SES for each level of *school* SES. One important observation is that poor students, the crosses in Figure 9, do not significantly underperform relative to their richer counterparts *for a given level of school SES.* Those poorer students who attended wealthy schools experienced gains to school SES similar to wealthy students. This is compared to those students of a low personal SES but who attended poorer schools. In Figure 9, this can be seen by the fact that the crosses (poorest students) rise in a similar fashion to wealthier students as school SES increases.

It is highly likely that the school SES variable included in the regressions is capturing elements of school quality. Affluent schools are more likely to exhibit those characteristics that we know to be important for student success. These include better school management, greater parental and governing-body involvement, sufficient school discipline, little teacher absenteeism, high teacher quality and motivation, and generally a more functional school environment, all of which aid student learning and thus performance

4.8 Impact of preschool education:

Since SACMEQ III was the first of the SACMEQ surveys to ask students about their preschool education, it is of particular interest to see the extent that such education impacts literacy and numeracy. The regression results in Table B1 show that preschool education has a marked impact on Grade 6 academic achievement, and this is true across all three subject-types. It is particularly strong for reading, as can be seen by the large and highly significant coefficients for all possible reading regression specifications (Table B2).

All of the pre-school dummy variables in the model have 'no preschool' as the reference category. While there is no significant relationship between a few months of preschool education and academic performance, there is a strong, highly significant relationship for the other levels of preschool education. Regarding reading and maths scores, it is interesting to note that there is not a large difference between one year and three years of preschool education, as statistical tests also show.

Preschool education is also of particular interest to policy makers since it could well be a mechanism through which social mobility is improved. There is a large body of international literature indicating the importance of pre-school education, and conversely, the negative effects of forgoing early education (see Gustafsson (2010) for a recent discussion of preschool education in South Africa). A lack of preschool education could place economically disadvantaged students at an educational disadvantage[10] which further increases the number of hurdles students must overcome if they are to succeed at school and later in life.

Given the above, it is worrying to see the strong correlation between preschool education and wealth, as measured by SES (see Table D4). As one would expect, poorer quintiles have less preschool education and higher quintiles have more preschool education. Almost 40% of students in the poorest quintile receive no preschool education whatsoever[11]. Since the difference in performance between one year and more than one year of preschool education is small, the point of emphasis should be on those students who received no preschool education.

4.9 Impact of textbook availability:

Previous education production function studies have found that the educational returns to textbooks are large and significant in South Africa (Van der Berg & Louw, 2006; Gustafsson, 2007) and Sub-Saharan Africa (Fehrler *et al*, 2009). The student reading regressions show that students who have their own reading textbook, or share with not more than one student, perform significantly better than students who have to share their textbooks with more than one student. There is no discernable impact of maths textbooks on student maths performance.

Similar to the trends seen in grade repetition and homework frequency, richer students are far more likely to have access to reading textbooks than their poorer counterparts. Amongst the poorest 20% of students, 36.8% either do not have a reading-textbook or must share with two or more students. The figure for the richest 20% of students is only 15.3%. Given that the reading-performance gains to reading textbooks are only evident when students either have their own textbook or share with not more than one other, policy should focus on ensuring that no student need share with more than one student. Given the well-defined, and relatively low costs of this policy option, it would seem that providing reading textbooks where they are in short supply – particularly in poor schools – is the low hanging fruit of the South African primary education system.

[10] Over and above the negative cognitive impact of no preschool education, it is highly likely that there are social and emotional skills developed in preschool which help the student in later school life.

[11] It should be noted, however, that the situation has improved since 2007, with increased access to pre-school education seen across the board, and specifically on the part of poorer learners (Gustafsson, 2010).

4.10 Orphans, orphanages and children's homes:

One of the many problems brought about by the HIV/AIDS epidemic in South Africa is the tragic increase in the number of orphans. Included in all the regressions were variables on orphan-status and whether the student was in an orphanage or children's home. The variable 'orphan' takes a value of one if the student indicated that both of their parents were deceased, and zero otherwise. The 'orphanage' variable takes a value of one if the students indicated that they lived in an orphanage or children's home, and zero otherwise.

A particularly startling finding from this preliminary analysis of the SACMEQ III data is the large negative effect of being in an orphanage or children's home. The regression results indicate that those students who lived in an orphanage or children's home fared substantially worse across all three subject tests. What is all the more remarkable is how highly stable and uniformly significant the 'orphanage' variable is across the three specifications (Table B1).

Initially, one would be prudent in thinking that these results could be driven by only a few students who live in the same orphanage and all attend a few underperforming schools. However, upon closer inspection of the data it becomes clear that this is not the case. The 58[12] students (0.67% of the total) who indicated that they lived in an orphanage or children's home each attended one of 46 different schools in 35 different districts. Since these 58 students were distributed across such a large number of schools and districts, one would expect that the orphanage dummy variable is not picking up school-level factors.

Alternatively, one could perhaps argue that orphanages or children's homes send students to under-performing schools (perhaps due to resource constraints), in which case the 'orphanage' dummy might simply be capturing poorly performing schools[13]. To ensure that this was not the case, the regressions were re-run including a school-level dummy variable, which took a value of 1 if there was a student who lived in an orphanage or children's home in the school and zero otherwise. This variable was negative and significant at conventional levels, and took the following values: reading (-12.35), maths (-10.59) and health (-11.38). The orphanage variable was still significant and took the following values: reading (-25.1), maths (-28.98), and health (-27.35), all significant at the 5% level.

[12]Given that South Africa has relatively few orphanages, it is possible that children living in child-headed households selected this option – i.e. the phrase 'children's home' may have been misunderstood. This is important from a policy perspective.

[13] It should be noted that it is extremely unlikely that orphanages have a special knack for selecting poorly performing schools since the regression already controls for a myriad of factors, perhaps most importantly school SES. Hence, it is unlikely that orphanages systematically selected underperforming schools.

This leads one to conclude that students who live in orphanages attend below-average schools (after accounting for numerous factors in the regression), but more importantly, that they perform substantially worse than their classmates and substantially worse than the average Grade 6 South African student, and all of this in an already below-average school. Therefore, it is almost certain that this orphanage variable is capturing the adverse economic, social, psychological and emotional impacts of staying in an orphanage, rather than simply accommodation. Policy makers, principals and teachers should all be aware of the multi-faceted problems faced by those living in orphanages and children's homes.

4.11 Student HIV/AIDS knowledge

The third student-level regression in Table B1 aims to explain the variation in the student health test-score. It shows that student HIV/AIDS knowledge is affected by many of the same factors as numeracy and literacy, since the coefficients on many variables are strikingly similar across the regressions.

It soon becomes evident that the health-test score may simply be testing academic ability, which could explain the high degree of similarity across the largely academic variables[14]. However, it is more likely that academic performance is highly correlated with factors which positively affect health knowledge: more educated parents, better schools and more capable health-teachers. Consequently, it is not that the health-test is simply a different way of testing academic ability (as literacy and numeracy do), but rather that student HIV/AIDS knowledge is highly correlated with those factors that improve academic ability. What is also surprising is the similarity of the size and significance of the coefficients considering that the health-test was simply 86 true or false questions regarding HIV and AIDS.

It should be noted that underlying the natural interpretation of health-score results (and even in the construction of the health-test itself) is the assumption that a student or teacher with more knowledge about HIV/AIDS is less at risk than one who has less knowledge about HIV/AIDS. This is not necessarily true since unprotected sexual activity may be influenced primarily by factors other than knowledge about HIV/AIDS. These other factors could be social norms, peer-pressure or even rape – all of which are unlikely to be influenced significantly by knowledge or information.

[14] These similar variables include whether a student has used a computer before, the frequency of English spoken at home, parental education, preschool education, grade repetition and homework frequency (Appendix B, Table B1).

4.12 Impact of teacher knowledge on student test scores

The inclusion of a teacher test-score variable in the SACMEQ III survey is particularly useful in determining the impact of teacher knowledge on student scores. By including the reading-teacher's reading-score in the student reading-score regression, one can draw out this relationship and begin to answer the question: '*Do more knowledgeable teachers produce more knowledgeable students?*'

The standard-regression results indicate that teacher knowledge is statistically significant, with reading teacher knowledge having a greater impact on student reading performance than maths or health teacher knowledge has on each of their student's performance. However, it must be recognised that while this relationship exists, it is exceedingly small. The coefficients on the various teacher test score variables are 0.0704 (reading-teacher), 0.0482 (maths-teacher), and 0.0653 (health teacher). Thus, a 100 point increase in teacher test-score, which is in the same order of magnitude as one standard deviation in each of the teacher test-score distributions[15], would only raise student reading-scores by 7.1 points, student maths-scores by 4.8 points and student health scores by 6.5 points. When seen in light of the size of some of the other coefficients in the student regressions, clearly teacher-knowledge is not a significant determinant of student test performance. This is in stark contrast to the initial assumptions of most education researchers who would expect teacher knowledge to have a large impact on student performance.

Importantly, the impact of teacher knowledge on student performance is much smaller for students from poorer backgrounds. The split-SES regressions (Tables B2, B3 and B4) show that for the poorest 80% of students, the impact of teacher knowledge is almost half that of the impact of teacher knowledge in wealthier schools. The coefficient on teacher maths-score in the mathematics regression (Table B3) is no longer significant in the 'bottom-4-SES quintiles' regression. It thus appears highly probable that students and teachers in poorer schools face multiple constraints which overshadow the impact of teacher knowledge.

To stress the small size of the teacher test-score coefficients, it is revealing to consider how student performance would change if all teachers performed satisfactorily in the teacher tests. Since the teacher tests contained many of the same questions as the student tests[16], one would expect all teachers to score almost full marks on the teacher test. This was certainly not the case.

[15] The precise standard deviations for each of the teacher test scores are maths-teacher maths test (111.35), reading teacher reading test (81.3), and health teacher health test score (100.52).
[16] Rasch scaling used these overlap-questions to convert the teacher test scores into figures comparable with the student test scores.

Comparing the lowest performing decile of teachers with the best performing decile of teachers shows that there are large discrepancies between teacher knowledge: Reading from 641 to 931, and Maths from 612 to 991. Applying the coefficients from the standard regressions to these differences shows how little teacher knowledge impacts on student performance. Thus the student reading gain from raising the weakest performing 10% of teachers (with a score of 641) to be equivalent to the best-performing 10% of teachers (with a score of 931), is only 20.4 points. For Maths the equivalent figure is 18.3. These figures are comparable with the impact size of far less dramatic changes such as increasing student homework frequency to most days of the week.

This does not mean, however, that teachers do not matter, only that teacher knowledge is not as strongly correlated with teacher quality as one might have expected. Factors such as teacher motivation and the ability of the teacher to convey their subject-knowledge may better capture what makes a 'good' teacher. Thus, it would seem that the ability to teach students well at the Grade 6 level is not very dependent on subject-knowledge, but perhaps more on the teacher's ability to convey that subject-knowledge.

4.13 Placing South Africa in regional context

Although the SACMEQ data for most countries has not been made publicly available, SACMEQ has released the mean scores for each country split by important sub-groups (see SACMEQ, 2010: 12, Figures D1 to D7 and Tables D6 & D7 in Appendix D). These include the mean reading and maths scores split by boys and girls, urban and rural, as well as the poorest 25% of students and the wealthiest 25% of students. This provides a broad overview of where countries lie in the regional distribution. Although each country wrote the same maths and reading tests and thus the scores are comparable across countries, some distinctions may be country-specific. For example, the rural-urban distinction is somewhat subjective and open to interpretation by the school head. This being said, it is unlikely to lead to large discrepancies across countries. Another important point to consider is that the various sub-groups are not homogenous across countries, for example, Malawi is likely to have a higher proportion of schools in rural areas than the Seychelles or Mauritius, and this is likely to affect the 'rural student' averages in each of these countries.

Graphing the mean reading and maths scores for each sub-group is particularly revealing (Figures D1-D6). Given that South Africa has more qualified teachers, lower pupil-to-teacher-ratios and better access to resources, one would expect that South African students would perform at the top of the regional distribution. Unfortunately this is not the case. In a league table of student

performance, South Africa ranks 10th out of the 15 SACMEQ countries for student reading performance and 8th out of 15 for student maths performance (Figure D1).

Comparing average student scores across countries ignores large variations both within and between countries. Consequently, these simplistic rankings should not be used to draw detailed conclusions or to unequivocally prescribe the policies of the 'best' countries. They may also be influenced by factors that are not immediately visible such as Gross Enrolment Rates. South Africa has almost 100% enrolment in Grade 6 while some other SACMEQ countries have lower GER's. This gives rise to a sample selection issue since we expect that those students who are not in the schooling system are more likely to be poor and rural, and thus below average students. Their exclusion therefore increases the mean score of that country. This is less of a problem in primary school, as compared to secondary school, since most SACMEQ countries have high rates of primary enrolment. Notwithstanding the above concerns, comparing the averages of important sub-groups, such as by gender, location and socio-economic status can provide interesting and useful insights.

When ranked by the performance of the wealthiest 25% of students, South Africa ranks 4th out of 15 for reading. However, when ranked by the performance of the poorest 25% of students, South Africa ranks 14th out of 15 for reading! For maths the figures are 6th out of 15 for wealthy students and 12th out of 15 for poor students. To put this in perspective, the average 'poor' South African student performs *worse* at reading than the average 'poor' Malawian or Mozambican student (Figures D3 and D4). This is in spite of the fact that the average 'poor' South African student is significantly wealthier than the average 'poor' Malawian or Mozambican student (Figure D7).

Similarly, South Africa ranks 13th out of 15 in the performance of rural students reading scores, and 9th out of 15 in the performance of urban students reading scores (Figures D5 & D6). For maths the figures are 8th out of 15 for urban students and 12th out of 15 for rural students.

These figures are no doubt startling to both South African policy-makers and the average South African citizen. South Africa's abysmal regional performance should provoke introspection and instigate change in educational policy. Clearly the South African primary school system is significantly underperforming relative to its regional counterparts given its large relative advantage in material resources.

Following the release of the rest of the SACMEQ III data, there is likely to be much fruitful cross-country analysis. This is eagerly anticipated.

5. Policy Recommendations

The foregoing has highlighted those areas that are significant determinants of student performance and thus which areas should receive policy priority. The following summarises the focus areas.

5.1 Pre-school education

Providing at least one year of quality pre-school education to all students is likely to improve student performance. This is especially true for poorer students who would otherwise start primary school at a disadvantage, and a disadvantage that is unlikely to diminish throughout their schooling career. Improving the *quality* of preschool education offered to the poor is also necessary if the full benefit of this policy intervention is to be felt.

5.2 Access to reading textbooks

Students from low-income households are less likely to have direct access to textbooks. Since there is a strong positive correlation between reading-textbook access and reading performance, targeting policies and funds towards reading-textbook provision will have an impact on student performance. This is especially true for learners from a disadvantaged socio-economic background.

5.3 Homework frequency

The research shows performance gains associated with those students who received homework either once or twice a week, or most days of the week. Practical policies that encourage teachers to prescribe homework, and enable students to complete that homework, should be explored and implemented. These policies are likely to be inexpensive, but yield significant gains in student performance.

5.4. School quality

The particularly large and highly significant coefficients on school socio-economic status indicate that wealthy schools are better able to help students reach their potential. However, it is only partially true that wealth can buy results. Yes, one can employ more and better teachers and provide adequate educational resources, but many of the factors that determine success in wealthy schools, such as management, discipline, and parental involvement, are not dependent on wealth. Policy-makers should identify ways and means of ensuring that poorer schools are better managed.

5.5 Teacher knowledge and quality

Teachers' subject expertise has a very small positive impact on learner performance at a Grade 6 level. While improving teacher subject-knowledge is likely to provide modest gains, policy should focus rather on helping teachers convey the subject material to their students.

5.6 South Africa's regional performance

Given South Africa's status as a middle-income country, education policy makers should ask how it is possible that primary schooling systems in neighbouring low-income countries are able to outperform South Africa, when South Africa has a clear resource advantage. A poignant example is how Tanzania outperforms South Africa for every sub-population (rural-urban, rich-poor, male-female, overall) when South Africa's GDP per capita is more than ten times higher than that of Tanzania[17].

6. Conclusion

The motif that runs through much of the analysis above is that South Africa is still a tale of two schools: One which is functional, wealthy, and able to educate students; with the other being poor, dysfunctional, and unable to equip students with the necessary numeracy and literacy skills they should be acquiring in primary school. While the constitution promises equal *access* to education, it cannot promise an equal *quality* of education. Until such a time as the primary education system in South Africa is able to offer a quality education to all students, not only the wealthy, the existing levels of educational inequality will remain. This has consequences for the labour market, poverty and hereditary poverty.

Although parts of the primary education system in South Africa are dysfunctional by regional standards, specifically those accessed by the rural and the poor, there are still policy interventions that can help. In brief, these are 1) ensure all learners have access to at least one year of quality preschool-education, 2) provide adequate access to reading textbooks, 3) increase the frequency of homework in poorer schools, 4) improve school management and discipline, 5) improve the ability of teachers to convey their subject-knowledge, and 6) learn from other African countries who produce better results with fewer resources. These interventions are likely to improve the performance of primary-school students, particularly so for those from poorer backgrounds.

[17] According to the World Bank Development Indicators (2009), Tanzania's GDP per capita was $509 while South Africa's was $5786.

7. References

Anderson, K. G., Case, A. and Lam, D. (2001). **Causes and Consequences of Schooling Outcomes in South Africa: Evidence from Survey Data.** *Social Dynamics.* 27(1): 1-23. Available: http://casr.ou.edu/pubs/KA_Causes__Consequences.pdf [September 2010]

Bhorat, H. & Oosthuizen, M. (2006). **Determinants of Grade 12 pass rates in the post-apartheid South African schooling system.** Dakar: Sisera. Online. Available: http://www.idrc.ca/uploads/user-S/114787836912006_6_DPRU_SAGA_Bhorat_Determinant.pdf [October 2010].

Booysen, F., Van der Berg, S., Burger, R., von Maltitz, M., & Du Rand, G. (2008). **Using an Asset Index to Assess Trends in Poverty in Seven Sub-Saharan African Countries.** *World Development*. 36(6): 1113-1130.

Crouch, L. & Mabogoane, T. (1998). **No magic bullets, just tracer bullets: The role of learning resources, social advantage, and education management in improving the performance of South African schools.** Research Triangle Park: Research Triangle Institute. Online. Available: http://www.rti.org/ddsp/documents/no_magic.pdf [October 2010].

Donaldson, A. R. (1992). **Content, quality and flexibility: The economics of education system change**. Spotlight 5/92. *South African Institute of Race Relations*: Johannesburg

Fehrler, S. Michaelowa, K & Wechtler, A. (2009). **The Effectiveness of Inputs in Primary Education: Insights from Recent Student Surveys for Sub-Saharan Africa**. *Journal of Development Studies.* Vol. 45 (9) pp 1545-1578.

Fiske, E.B. & Ladd, H.F. (2004). **Elusive Equity: Education Reform in Post-Apartheid South Africa.** Washington D.C.: Brookings Institution Press.

Gustafsson, M.(2007). **Using the hierarchical linear model to understand school production in South Africa**. *South African Journal of Economics* 75(1), March: 84-98. Available: http://www.ekon.sun.ac.za/wpapers/2007 [September 2010]

Gustafsson, M. (2010). **Policy note on pre-primary schooling: An empirical contribution to the 2009 Medium Term Strategic Framework**. Stellenbosch Working Paper 05/1020. Online. Available: http://www.ekon.sun.ac.za/wpapers/2010/wp052010 [12 December 2010]

Gustafsson, M. & Patel, F. (2006). **Undoing the apartheid legacy: Pro-poor spending shifts in the South African public school system.** *Perspectives in Education*, 24(2): 65-77.

Lam, D. (1999). **Generating Extreme Inequality: Schooling, Earnings, and Intergenerational Transmission of Human Capital in South Africa and Brazil**. *Population Studies Center, University of Michigan*. Research Report No. 99-439.

Moloi, M. & Strauss, J. (2005). **The SACMEQ II Project in South Africa: A Study of the Conditions of Schooling and the Quality of Education – South Africa Working Report.** SACMEQ. Online. Available: http://www.sacmeq.org/education-south-africa.htm [September 2010]

Murimba, S. (2005). **The Southern and Eastern Africa Consortium for Monitoring Educational Quality: Mission, Approach and Projects**. *Prospects: Quarterly Review of Comparative Education*. Vol. XXXV, no.1

Ross, K. N. Zuze, L. (2004). **Traditional and alternative views of school system performance**. *Educational Quality*. IIEP Newsletter, October-December, 8-9. Online. Available: http://www.sacmeq.org/downloads/newsletter/NleOct2004Sacmeq%20article.pdf [September 2010]

SACMEQ (2010). Contributors: Hungi, N., Makuwa, D., Ross, K., Saito, M., Dolata, S., van Capelle, F., Paviot, L., & Vellien, J. **SACMEQ III Project Results: Pupil achievement levels in reading and mathematics.** Southern and East African Consortium for Monitoring Educational Quality. Online. Available: http://www.sacmeq.org/downloads/sacmeqIII/WD01_SACMEQ_III_Results_Pupil_Achievement.pdf [January 2011]

Strauss, J. (2005**). A Model for Evaluating South Africa's Education System Based on SACMEQ Research Data.** Paper presented to the International Invitational Educational Policy Research Conference, Paris, France, 28 September 2 October, 2005. Online. Available: http://www.sacmeq.org/education-south-africa.htm [23 June 2010]

Taylor, S. & Yu, D. (2009). **The importance of socio-economic status in determining educational achievement in South Africa**. Stellenbosch: University of Stellenbosch. Available from: http://ideas.repec.org/p/sza/wpaper/wpapers73.html [September 2010].

Van Capelle, F. (2010). **StatPlanet: Interactive Data Visualization and Mapping Software**. (Online) Available: http://www.sacmeq.org/statplanet [September 2010]

Van der Berg, S. (2001). **Social Policy to Address Poverty**. In, Bhorat, H., Leibbrandt, M., Maziya, M., Van der Berg, S. and Woolard, I (eds). 2001. *Fighting Poverty: Labour markets and inequality in South Africa.* UCT Press: Cape Town.

Van der Berg, S. & Burger, R. (2003). **Education and socio-economic differentials: A study of school performance in the Western Cape**. Cape Town: University of Cape Town. Available: http://www.commerce.uct.ac.za/Research_Units/dpru/WorkingPapers/PDF_Files/wp73.pdff [October 2010].

Van der Berg, S & Louw, M. (2006). **Unravelling the mystery: Understanding South African Schooling Outcomes in Regional Context.** Paper to the conference of the Centre for the Study of Africa Economies, Oxford University, March 2006. Available: http://www.jet.org.za/events/conferences/School%20quality%20research%20seminar%202/VanderBerg-Louw%20SA%20schooling%20outcomes.pdf [September 2010]

Van der Berg, S. (2007). **Apartheid's enduring legacy: Inequality in education**. *Journal of African Economies*, 16(5): 849-880.

Van der Berg, S. (2008). **How effective are poor schools? Poverty and educational outcomes in South Africa**. *Studies in Educational Evaluation*, 34(3): 145-154. Available:

https://www.expertenkulturen.uni-goettingen.de/de/document/download/2dd028dc312af2f890efbd5b984d1856.pdf/69_Berg.pdf [September 2010]

World Bank. (2009). **World Bank Development Indicators**. Online. Available: http://data.worldbank.org/data-catalog/world-development-indicators?cid=GPD_WDI [18 March 2011]

8. Appendices

Appendix A:

Summary statistics

Variable	Number of observations	Mean	Std. Dev.	Min	Max
Student regression variables					
Student Reading score	9071	495.096	116.1771	62.94938	996.5053
Student Maths score	9063	495.4149	96.69129	10.33382	962.909
Student Health score	9063	502.9555	98.54655	26.73984	1018.472
Reading-teacher Reading score	8094	757.7259	81.69571	289.8527	1090.261
Maths-teacher Maths score	7884	763.6243	108.8479	469.2937	1204.372
Health-teacher Health score	8121	763.2316	104.8102	526.0151	1018.472
Under 12	9083	0.427898	0.494801	0	1
Over 12	9083	0.0875	0.282582	0	1
Male	9083	0.492434	0.49997	0	1
Urban area (large city)	9083	0.31858	0.465951	0	1
Orphan (double-orphan)	9083	0.090253	0.28656	0	1
Orphanage or children's home	9083	0.006709	0.08164	0	1
Lived with parents	9083	0.729696	0.444141	0	1
3 or more siblings	9083	0.560154	0.496396	0	1
Less than 3 meals per day[18]	9083	0.061122	0.239567	0	1
More than 10 books at home	9083	0.341737	0.474318	0	1
Used a computer before	9083	0.5047892	0.5000046	0	1
No. of days absent[19]	9083	0.984972	2.328556	0	26
Speak English at home sometimes	9037	0.611331	0.487475	0	1
Speak English at home always	9037	0.153481	0.36047	0	1
Mother has matric	9083	0.372178	0.483412	0	1
At least one parent has degree	9083	0.127889	0.333985	0	1
SES	9083	-0.04991	1.007677	-2.22339	2.381679
SES squared	9083	1.017793	1.175427	9.54E-08	5.672396
School SES	9083	-0.04991	0.746817	-1.82459	1.836909
School SES squared	9083	0.560165	0.690549	4.85E-09	3.374235
School building index	8991	3.155051	2.256223	0	7
School equipment index	8713	11.10456	5.582146	0	18

[18] This dummy variable took a value of one if the student indicated that they normally miss one morning meal, and one lunch meal, and one evening meal in a week.

[19] This is a student-answered question asking how many days they were absent in the last month of full time schooling.

Preschool - months	9083	0.04847	0.214769	0	1
Preschool - 1 year	9083	0.328723	0.469775	0	1
Preschool - 2 years	9083	0.153553	0.36054	0	1
Preschool - 3 years or more	9083	0.201501	0.401143	0	1
Repeated a grade once[20]	9083	0.203473	0.402603	0	1
Repeated a grade twice	9083	0.050355	0.218688	0	1
Repeated a grade 3 or more times	9083	0.031091	0.173573	0	1
Homework - 1 or 2 times a month	9083	0.087108	0.282008	0	1
Homework - 1 or 2 times a week	9083	0.31545	0.46472	0	1
Homework - Most days	9083	0.558578	0.496584	0	1
Read. Textbook -Teacher only	9083	0.063997	0.24476	0	1
Read. Textbook - Share 2+	9083	0.161068	0.367614	0	1
Read. Textbook - Share with 1	9083	0.281021	0.449523	0	1
Read. Textbook - own textbook	9083	0.447617	0.497276	0	1
Extra English tuition	9083	0.095506	0.293928	0	1
Math Textbook -Teacher only	9083	0.17213	0.377514	0	1
Math Textbook -Share 2+	9083	0.117446	0.321968	0	1
Math Textbook -Share with 1	9083	0.236797	0.42514	0	1
Math Textbook - own textbook	9083	0.361557	0.480478	0	1
Extra Math tuition	9083	0.098229	0.297641	0	1
NorthWest[21]	9 083	0.601601	0.2377963	0	1
Eastern Cape	9083	0.16285	0.369249	0	1
Free State	9083	0.049548	0.217021	0	1
Gauteng	9083	0.17221	0.377584	0	1
KwaZulu-Natal	9083	0.236756	0.425114	0	1
Limpopo	9083	0.134316	0.34101	0	1
Mpumalanga	9083	0.085412	0.279509	0	1
Northern Cape	9083	0.02048	0.141644	0	1
Western Cape	9083	0.078268	0.268608	0	1

[20]The question from which the 'repeated' variables are derived asked: "How many times have you repeated a grade since you started school including Grade 6?"

[21] The North West province was used as the reference category for the 'Province' dummy variables.

Appendix B:

Student regressions[22]

Table B1 – Student reading, maths and health regressions

	Student reading	Student math	Student health
Under 12	-18.4682***	-14.9951***	-5.4437*
Over 12	-6.9541	-6.6306*	-4.2368
Male	-12.7782***	0.2822	-10.6057***
Urban area (large city)	-11.3319**	-9.8852*	-8.0273
Orphan (double-orphan)	-7.0686*	-4.6329	-4.9614
Orphanage or children's home	-36.0692***	-37.4391***	-36.1027***
Lived with parents	-5.3555**	-3.7669	-1.9748
3 or more siblings	-10.6353***	-7.3624***	-8.3194***
Less than 3 meals per day	-12.3229***	-8.1409	-6.557
More than 10 books at home	9.6724***	5.5628**	6.5805**
Used a computer before	21.9527***	15.3216***	14.2765***
No. of days absent	1.1213	0.4921	1.8729
Speak Eng. at home sometimes	19.0144***	15.5334***	19.1453***
Speak Eng. at home always	37.4995***	15.3195***	29.8589***
Mother has matric	13.8079***	9.8831***	16.8851***
At least one parent has degree	14.0110***	14.7947***	8.1425**
SES	2.7922*	0.6305	0.8834
SES squared	1.6120*	3.4595***	-0.3433
School SES	41.2664***	42.9921***	24.7332***
School SES squared	21.9729***	25.1229***	8.8551*
School building index	3.3918	-0.6604	1.33
School equipment index	0.5135	0.8054	1.0204
Preschool - months	5.1781	1.9331	-1.001
Preschool - 1 year	10.8312***	7.6765*	8.7963**
Preschool - 2 years	18.9847***	5.1315	11.5173**
Preschool - 3 years or more	11.8728***	9.6649***	14.6994***
Repeated a grade once	-19.1497***	-11.4923***	-15.5356***
Repeated a grade twice	-25.0088***	-15.7366***	-23.0066***
Repeated a grade three or more	-44.1622***	-27.6076***	-43.4977***
Homework - 1 or 2 times a month	2.9002	5.352	22.6319***
Homework - 1 or 2 times a week	19.0363***	19.2336***	39.1989***
Homework - Most days	19.4584***	22.2710***	41.6106***
R/M Textbook -Teacher only	-0.7394	-4.8909	
R/M Textbook -Share 2+	0.5809	-15.6056	
R/M Textbook -Share with 1	20.3619***	0.7174	
R/M Textbook - own textbook	18.3743***	-2.0984	

[22] Any coefficient that is significant at conventional levels (i.e. at the 10% level or lower) has been greyed out for ease of identification. However, traditional 'stars' to denote significance have also been included with their usual connotations (* for 10%, ** for 5%, and *** for 1%).

Extra English tuition	-16.6696***		
Extra Math tuition		-12.3972**	
Reading-teacher Reading score	0.0704**		
Maths-teacher Maths score		0.0482**	
Health-teacher Health score			0.0653***
Eastern Cape[23]	14.0399	25.4944*	0.8404
Free State	7.2751	9.4735	-26.9115***
Gauteng	27.1634***	16.7485**	2.8013
KwaZuluNatal	13.6635*	12.5188*	14.4944
Limpopo	-31.2522***	-17.9163***	-32.9000**
Mpumalanga	0.6254	6.9153	-5.871
Northern Cape	19.4718**	10.5318	-19.4695**
Western Cape	25.1349***	25.3479***	-4.3864
Constant	362.8019***	400.4745***	381.9382***
N	7724	7494	7712
F-stat	84.36911	45.80037	28.39298
Prob > F	0	0	0
R-squared	0.59945	0.49138	0.34504
Testing for differences between coefficients:			
Preschool1=2	0.04576	0.56005	0.53166
Preschool1=3	0.78058	0.60841	0.13368
Preschool2=3	0.03214	0.1872	0.47031
Repeat1=2	0.14391	0.33137	0.12477
Repeat1=3	0.00006	0.01302	0.00152
Repeat2=3	0.00954	0.13976	0.02977
Homework1=2	0.0003	0.00132	0.00467
Homework1=3	0.00039	0.00009	0.00118
Homework2=3	0.90435	0.35871	0.58257
SES SESsq joint significance	0.04999	0.0016	0.79432
SSES SSESsq joint significance	0	0	0.00001
Textbookown=share1	0.62625		

* p<0.1 ** p<0.05 ***p<0.01

[23] The North West province was used as the reference category for the 'Province' dummy variables.

Table B2 – Variations in student reading regressions

	Student reading	Student reading, no teacher-scores	Student reading, top quintile only	Student reading, bottom 4 quintiles only
Under 12	-18.4682***	-18.8480***	-23.5384***	-17.3265***
Over 12	-6.9541	-8.6300**	-3.4861	-6.3972
Male	-12.7782***	-11.8188***	-9.8124**	-12.8693***
Urban area (large city)	-11.3319**	-8.5119*	-8.7347	-12.1803**
Orphan (double-orphan)	-7.0686*	-5.2419	-27.2601**	-4.7395
Orphanage or children's home	-36.0692***	-38.4046***	-0.071	-40.1924***
Lived with parents	-5.3555**	-4.5554*	-3.3002	-5.3690**
3 or more siblings	-10.6353***	-11.6370***	-7.4246	-10.6041***
Less than 3 meals per day	-12.3229***	-13.0053***	-19.4697*	-11.0435**
More than 10 books at home	9.6724***	9.5082***	15.7560***	7.5028***
Used a computer before	21.9527***	21.6609***	12.37	21.8216***
No. of days absent	1.1213	1.0868	0.4542	1.0489
Speak Eng. at home sometimes	19.0144***	19.3987***	15.4398*	20.1082***
Speak Eng. at home always	37.4995***	38.3178***	41.0539***	30.6389***
Mother has matric	13.8079***	14.7102***	11.4799**	13.1540***
At least one parent has degree	14.0110***	16.7647***	10.8544**	16.5528***
SES	2.7922*	2.7340**	85.3557*	4.4706
SES squared	1.6120*	1.7471*	-22.7984	1.4328
School SES	41.2664***	41.2500***	46.8748***	46.3439***
School SES squared	21.9729***	23.7798***	1.1526	29.0022***
School building index	3.3918	4.0820*	8.0427***	2.2508
School equipment index	0.5135	0.5548	1.0064	0.5242
Preschool - months	5.1781	5.5249	11.7773	5.216
Preschool - 1 year	10.8312***	10.5252***	19.2327**	9.8075**
Preschool - 2 years	18.9847***	18.3344***	27.6487***	18.5509***
Preschool - 3 years or more	11.8728***	13.2678***	29.4235***	7.3979**
Repeated a grade once	-19.1497***	-19.4476***	-24.0369***	-18.2677***
Repeated a grade twice	-25.0088***	-25.4895***	-26.0744**	-24.5108***
Repeated a grade 3 or more times	-44.1622***	-44.1726***	-43.8954***	-41.3855***
Homework - 1 or 2 times a month	2.9002	1.5976	16.8851	2.7656
Homework - 1 or 2 times a week	19.0363***	18.4801***	22.726	20.0216***
Homework - Most days	19.4584***	20.3871***	23.185	19.2998***
R/M Textbook -Teacher only	-0.7394	4.1468	24.2590*	-6.13
R/M Textbook -Share 2+	0.5809	3.3225	-16.5707	1.7399
R/M Textbook -Share with 1	20.3619***	20.2501***	18.919	20.3356***
R/M Textbook - own textbook	18.3743***	15.2153***	26.3938**	15.1924***
Extra English tuition	-16.6696***	-17.3116***	-35.6491***	-11.5047*
Reading-teacher Reading score	0.0704**		0.0914***	0.0548*
Eastern Cape	14.0399	15.8082	13.2863	9.1594
Free State	7.2751	5.046	12.886	3.9168

Gauteng	27.1634***	27.2487***	29.8651***	21.7647**
KwaZuluNatal	13.6635*	14.9029**	29.3778***	6.2532
Limpopo	-31.2522***	-30.7637***	-46.6106***	-33.201***
Mpumalanga	0.6254	2.4408	0.9479	-1.8568
Northern Cape	19.4718**	16.2224**	29.7928**	14.1662*
Western Cape	25.1349***	29.6712***	11.0254	31.0188***
Constant	362.8019***	410.5610***	251.3250***	382.3219***
N	7724	8591	1559	6165
F-stat	84.36911	93.68868	50.39694	48.82521
Prob > F	0	0	0	0
R-squared	0.59945	0.60134	0.60661	0.4562
Testing for differences between coefficients:				
Preschool1=2	0.04576	0.04211	0.20392	0.06557
Preschool1=3	0.78058	0.4331	0.07387	0.57132
Preschool2=3	0.03214	0.12429	0.75578	0.00497
Repeat1=2	0.14391	0.11362	0.87753	0.13946
Repeat1=3	0.00006	0.00001	0.1874	0.00026
Repeat2=3	0.00954	0.0048	0.27698	0.02736
Homework1=2	0.0003	0.00006	0.65737	0.00012
Homework1=3	0.00039	0.00003	0.6247	0.00035
Homework2=3	0.90435	0.58393	0.94181	0.84466
SES SESsq joint significance	0.04999	0.02522	0.02307	0.23764
SSES SSESsq joint significance	0	0	0	0
Textbookown=share1	0.62625	0.19802	0.18044	0.22914

* p<0.1 ** p<0.05 ***p<0.01

Table B3 – Variations in student maths regressions

	Student math	Student math, no teacher-scores	Student math, top quintile only	Student math, bottom 4 quintiles only
Under 12	-14.9951***	-15.7817***	-25.9308***	-12.8508***
Over 12	-6.6306*	-4.773	-1.494	-6.9614
Male	0.2822	0.95	2.8811	-0.1579
Urban area (large city)	-9.8852*	-7.442	-10.7663	-9.8972*
Orphan (double-orphan)	-4.6329	-6.3031	-15.8533	-3.736
Orphanage or children's home	-37.4391***	-33.2328***	5.6995	-43.6218***
Lived with parents	-3.7669	-4.4086**	-4.9521	-3.632
3 or more siblings	-7.3624***	-6.9467***	-11.8617***	-6.0520***
Less than 3 meals per day	-8.1409	-9.8784**	1.0745	-9.3085*
More than 10 books at home	5.5628**	6.7218***	17.1200***	2.5304
Used a computer before	15.3216***	15.6602***	20.6284**	15.4363***
No. of days absent	0.4921	0.4862	-0.2754	0.6039
Speak Eng. at home sometimes	15.5334***	15.5507***	7.4698	16.8416***
Speak Eng. at home always	15.3195***	16.5256***	4.0598	16.7740***
Mother has matric	9.8831***	10.1898***	7.8587	9.8535***
At least one parent has degree	14.7947***	14.5775***	16.5155***	13.9542***
SES	0.6305	0.4563	6.0918	3.7916
SES squared	3.4595***	3.1671***	2.3298	4.8778**
School SES	42.9921***	42.9073***	22.3440**	44.2171***
School SES squared	25.1229***	26.7511***	27.7259***	24.6848***
School building index	-0.6604	-0.1868	7.0075***	-1.5183
School equipment index	0.8054	0.8498	-0.7442	0.9119
Preschool - months	1.9331	2.6672	12.6499	0.7186
Preschool - 1 year	7.6765*	6.7048*	16.6846**	7.0222
Preschool - 2 years	5.1315	5.1454	10.8079	5.3808
Preschool - 3 years or more	9.6649***	8.8394***	26.4924***	4.6508
Repeated a grade once	-11.4923***	-12.0539***	-13.2442*	-10.7934***
Repeated a grade twice	-15.7366***	-13.6430***	-25.4886**	-14.9027***
Repeated a grade 3 or more times	-27.6076***	-26.5436***	-7.2938	-28.0978***
Homework - 1 or 2 times a month	5.352	-1.8623	18.6484	6.0509
Homework - 1 or 2 times a week	19.2336***	14.4378**	28.9902**	19.9021***
Homework - Most days	22.2710***	18.3064***	25.8779**	23.3082***
R/M Textbook -Teacher only	-4.8909	-5.1804	7.3814	-5.6461
R/M Textbook -Share 2+	-15.6056	-14.6506	5.7458	-17.8242
R/M Textbook -Share with 1	0.7174	-0.2621	10.9067	-1.0626
R/M Textbook - own textbook	-2.0984	-2.6214	14.2184	-5.5041
Extra Math tuition	-12.3972**	-12.6322**	-30.4693***	-5.2962
Maths-teacher Maths score	0.0482**		0.0766***	0.036
Eastern Cape	25.4944*	25.3213**	-0.4005	25.2607
Free State	9.4735	10.5015	9.383	10.0022

Gauteng	16.7485**	15.5196**	24.2014**	15.0757*
KwaZuluNatal	12.5188*	12.7330**	29.9663***	9.9852
Limpopo	-17.9163***	-17.5225***	-14.9856	-18.5964***
Mpumalanga	6.9153	2.7056	10.3606	4.8136
Northern Cape	10.5318	12.3191**	20.4999*	8.0706
Western Cape	25.3479***	28.6175***	30.0918**	25.3166***
Constant	400.4745***	438.7964***	336.3101***	412.8011***
N	7494	8582	1433	6061
F-stat	45.80037	50.1653	36.53574	27.3103
Prob > F	0	0	0	0
R-squared	0.49138	0.5019	0.55219	0.31326
Testing for differences between coefficients:				
Preschool1=2	0.56005	0.69598	0.43435	0.73801
Preschool1=3	0.60841	0.53518	0.0824	0.60404
Preschool2=3	0.1872	0.24662	0.0063	0.85012
Repeat1=2	0.33137	0.69286	0.32564	0.37366
Repeat1=3	0.01302	0.01456	0.73077	0.00802
Repeat2=3	0.13976	0.07698	0.3462	0.1186
Homework1=2	0.00132	0.00006	0.45409	0.00143
Homework1=3	0.00009	0	0.60483	0.00004
Homework2=3	0.35871	0.20036	0.61957	0.34442
SES SESsq joint significance	0.0016	0.00125	0.10361	0.05628
SSES SSESsq joint significance	0	0	0	0

* p<0.1 ** p<0.05 ***p<0.01

Table B4 – Variations in student health regressions

	Student health	Student health, no teacher-scores	Student health, top quintile only	Student health, bottom 4 quintiles only
Under 12	-5.4437*	-6.3193**	-2.2412	-5.4926
Over 12	-4.2368	-5.3771	-9.5142	-2.6482
Male	-10.6057***	-11.3309***	-13.2324***	-10.2130***
Urban area (large city)	-8.0273	-11.7591**	1.5958	-10.5440*
Orphan (double-orphan)	-4.9614	-3.9951	-9.8246	-4.3312
Orphanage or children's home	-36.1027***	-46.5856***	-32.9950**	-31.9357***
Lived with parents	-1.9748	-2.0461	-6.0782	-1.1784
3 or more siblings	-8.3194***	-9.4634***	-9.7910*	-7.7068***
Less than 3 meals per day	-6.557	-6.7599	-11.0491	-5.2629
More than 10 books at home	6.5805**	5.8958*	10.5435*	5.1794
Used a computer before	14.2765***	15.0349***	26.4235***	13.0270***
No. of days absent	1.8729	1.7202	0.7119	1.8644
Speak Eng. at home sometimes	19.1453***	17.8172***	7.848	20.4897***
Speak Eng. Aa home always	29.8589***	30.6077***	26.0720***	27.6516***
Mother has matric	16.8851***	16.6623***	20.3706***	15.7614***
At least one parent has degree	8.1425**	10.2437***	10.9257**	6.2276
SES	0.8834	0.5435	-13.8261	2.8184
SES squared	-0.3433	-0.2756	4.5169	0.4459
School SES	24.7332***	24.5946***	28.8892**	28.4392***
School SES squared	8.8551*	7.1991	-3.1272	11.9052*
School building index	1.33	1.2836	4.7719*	0.6098
School equipment index	1.0204	1.0041	0.4533	1.1788
Preschool – months	-1.001	-2.6569	8.4842	-1.7196
Preschool - 1 year	8.7963**	8.0283**	10.4643	8.6041**
Preschool - 2 years	11.5173**	12.4604***	16.9076*	10.5125*
Preschool - 3 years or more	14.6994***	13.0003***	24.3450***	11.5218**
Repeated a grade once	-15.5356***	-15.5836***	-15.2680*	-15.8238***
Repeated a grade twice	-23.0066***	-21.6764***	-9.6114	-23.8530***
Repeated a grade 3 or more times	-43.4977***	-46.3383***	-33.1405	-42.5254***
Homework - 1 or 2 times a month	22.6319***	26.1158***	7.7443	24.4332***
Homework - 1 or 2 times a week	39.1989***	43.4626***	29.4623	39.9087***
Homework - Most days	41.6106***	48.3760***	18.3947	44.0820***
Health-teacher Health score	0.0653***		0.0824**	0.0608**
Eastern Cape	0.8404	6.5495	-32.4463*	3.6979
Free State	-26.9115***	-24.9392***	-20.1981	-28.2361***
Gauteng	2.8013	8.0068	-4.0913	3.8488
KwaZuluNatal	14.4944	15.9984*	10.4714	14.4796
Limpopo	-32.9000**	-32.3498***	-30.2415	-32.7795**
Mpumalanga	-5.871	-4.6926	-3.2603	-6.5124
Northern Cape	-19.4695**	-20.5319**	5.0345	-24.6931**

Western Cape	-4.3864	0.9141	-13.307	-2.1835
Constant	381.9382***	428.7283***	386.6096***	384.2282***
N	7712	8584	1535	6177
F-stat	28.39298	27.54519	14.58234	20.45519
Prob > F	0	0	0	0
R-squared	0.34504	0.33093	0.31408	0.2613
Testing for differences between coefficients:				
Preschool1=2	0.53166	0.28589	0.41378	0.71403
Preschool1=3	0.13368	0.16474	0.06759	0.5434
Preschool2=3	0.47031	0.89754	0.29115	0.85385
Repeat1=2	0.12477	0.20335	0.68708	0.12003
Repeat1=3	0.00152	0.00009	0.49348	0.00169
Repeat2=3	0.02977	0.00368	0.38978	0.04495
Homework1=2	0.00467	0.00141	0.07737	0.01501
Homework1=3	0.00118	0.00004	0.38512	0.00172
Homework2=3	0.58257	0.23548	0.14176	0.38312
SES SESsq joint significance	0.79432	0.8928	0.95501	0.44475
SSES SSESsq joint significance	0.00001	0.00002	0.02053	0.00001

* p<0.1 ** p<0.05 ***p<0.01

Appendix C (SACMEQ, 2010: 1):

Background to SACMEQ Projects I, II, and III (SACMEQ, 2010: 1)

a) SACMEQ I Project (1995-1998)

This was the first educational policy research project conducted by SACMEQ. It commenced in 1995 and was completed in 1998. Seven Ministries of Education participated in the project (Kenya, Malawi, Mauritius, Namibia, Tanzania (Zanzibar), Zambia, and Zimbabwe), and each one of them prepared a national educational policy report. These reports have set down agendas for government action by using national surveys to explore issues related to: baseline indicators for educational inputs, the general conditions of schooling, equity assessments for human and material resource allocations, and the literacy levels of Grade 6 pupils. Around 20,000 pupils from 1,000 primary schools were involved in the SACMEQ I project.

b) SACMEQ II Project (1998-2004)

This was SACMEQ's second educational policy research project. It started in 1999 and was completed in 2004. Fourteen Ministries of Education (Botswana, Kenya, Lesotho, Malawi, Mauritius, Mozambique, Namibia, Seychelles, South Africa, Swaziland, Tanzania (Mainland), Tanzania (Zanzibar), Uganda, and Zambia) completed the SACMEQ II Project. SACMEQ II Project national reports provided measures of change in the conditions of schooling and the quality of education between 1995 and 2000 for six SACMEQ countries. The project involved around 40,000 students, 5,300 teachers and 2,000 school heads from 2000 primary schools.

c) ***SACMEQ III Project (2005-2010)***

This is SACMEQ's third educational policy research project. All fifteen SACMEQ Ministries of Education (Botswana, Kenya, Lesotho, Malawi, Mauritius, Mozambique, Namibia, Seychelles, South Africa, Swaziland, Tanzania (Mainland), Tanzania (Zanzibar), Uganda, Zambia, and Zimbabwe) participated. The SACMEQ III Project will assist Ministries of Education to track changes in the general conditions of schooling and pupil achievement levels between 1995 and 2000 (for 6 Ministries of Education), and between 2000 and 2007 (for 14 Ministries of Education). The SACMEQ III Project will also provide Ministries of Education with information about the knowledge levels of pupils and their teachers in matters relating to HIV and AIDS...Data for the SACMEQ III Project were collected during the last quarter of 2007 from 61,396 pupils, 8,026 teachers, and 2,779 schools.

Appendix D:

Additional graphs and tables

Table D1 - Grade repetition

Quintile	Never repeated	Repeated once	Repeated twice	Repeated three times or more	Total
1	1,133	466	141	77	1,817
2	1,246	404	109	58	1,817
3	1,242	411	105	58	1,816
4	1,311	382	73	51	1,817
5	1,490	252	45	29	1,816
Total	**6,422**	**1,915**	**473**	**273**	**9,083**

Table D2 – Homework frequency

Quintiles	No homework	1-2 times /month	1-2 times/week	Most days	Total
1	101	197	640	870	1,808
	6%	11%	35%	48%	100%
2	71	201	630	911	1,813
	4%	11%	35%	51%	100%
3	64	169	623	954	1,810
	4%	9%	34%	53%	100%
4	59	154	589	1,009	1,811
	3%	9%	33%	56%	100%
5	27	83	395	1,306	1,811
	1%	5%	22%	72%	100%
Total	**322**	**804**	**2,877**	**5,050**	**9,053**
	3.56%	**8.88%**	**31.78%**	**55.78%**	**100**

Table D3 – Textbook availability

Reading textbooks

Quintile	No textbooks	Only teacher	Share with 2+	Share with 1	Own textbook	Total
1	6.8%	8.1%	21.9%	27.7%	35.6%	100%
2	4.0%	7.5%	20.0%	31.1%	37.3%	100%
3	3.5%	5.5%	16.8%	30.7%	43.4%	100%
4	2.9%	5.4%	14.6%	27.7%	49.4%	100%
5	3.6%	5.4%	6.3%	23.8%	60.8%	100%
Total	**4.2%**	**6.4%**	**16.2%**	**28.2%**	**45.0%**	**100%**

Maths textbooks

Quintile	No textbooks	Only teacher	Share with 2+	Share with 1	Own textbook	Total
1	13.9%	16.5%	16.0%	24.2%	29.4%	100%
2	10.2%	17.9%	14.8%	26.0%	31.2%	100%
3	11.0%	17.9%	12.6%	24.6%	33.9%	100%

4	10.0%	17.9%	9.1%	23.8%	39.1%	1000%
5	7.4%	16.6%	5.7%	20.4%	49.9%	100%
Total	**10.6%**	**17.3%**	**11.8%**	**23.9%**	**36.4%**	**100%**

Table D4 – Preschool education and SES

	Years of preschool education					
Quintiles	**None**	**Few months**	**1 year**	**2 years**	**3+ years**	**Total**
1	**39%**	4%	35%	11%	10%	**100%**
2	**32%**	6%	37%	12%	13%	**100%**
3	**28%**	5%	35%	15%	17%	**100%**
4	20%	5%	32%	18%	25%	**100%**
5	11%	4%	25%	22%	38%	**100%**
Total	**26%**	**5%**	**33%**	**15%**	**20%**	**100%**

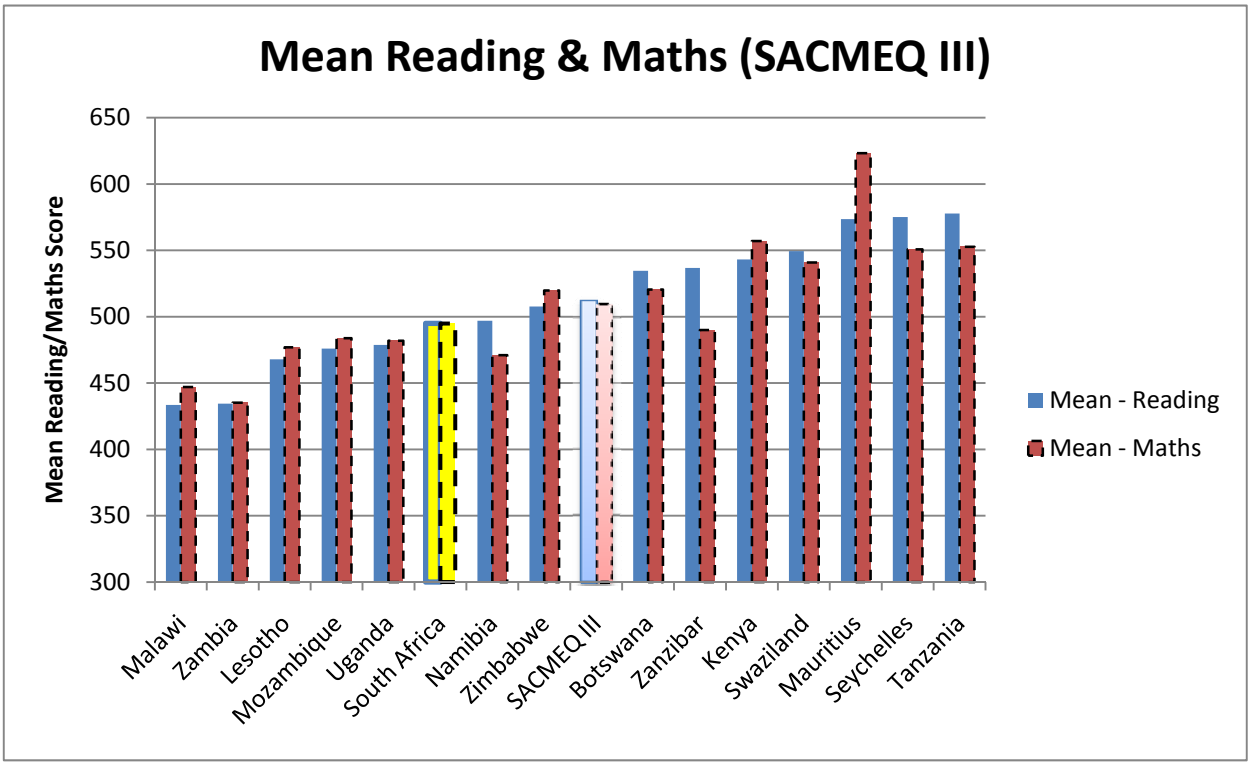

Figure D1

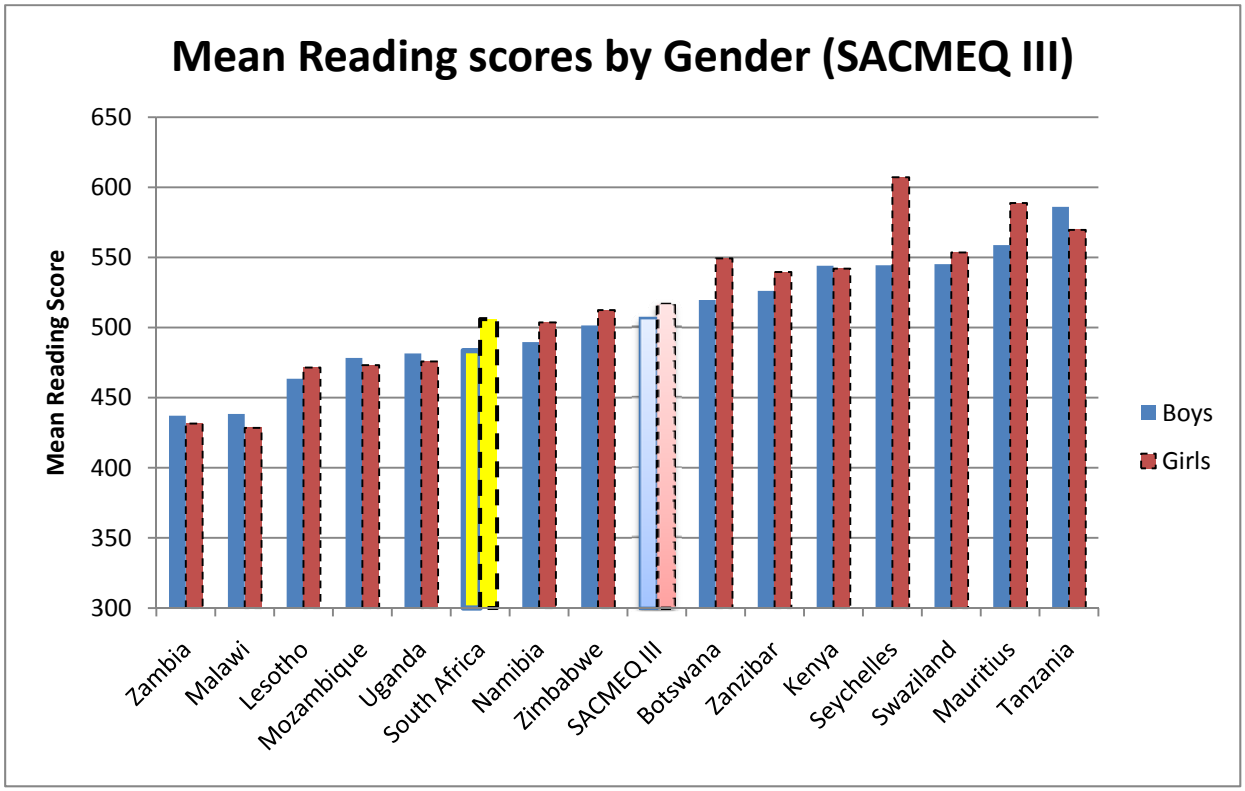

Figure D2

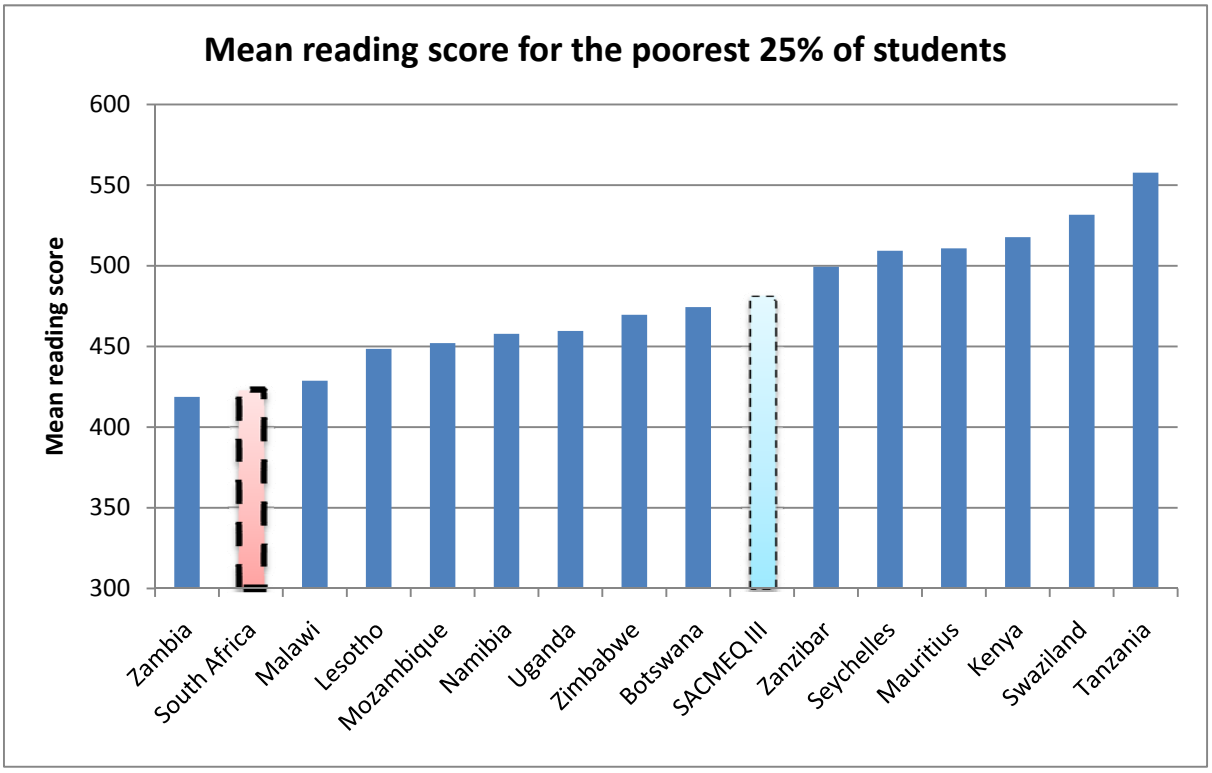

Figure D3

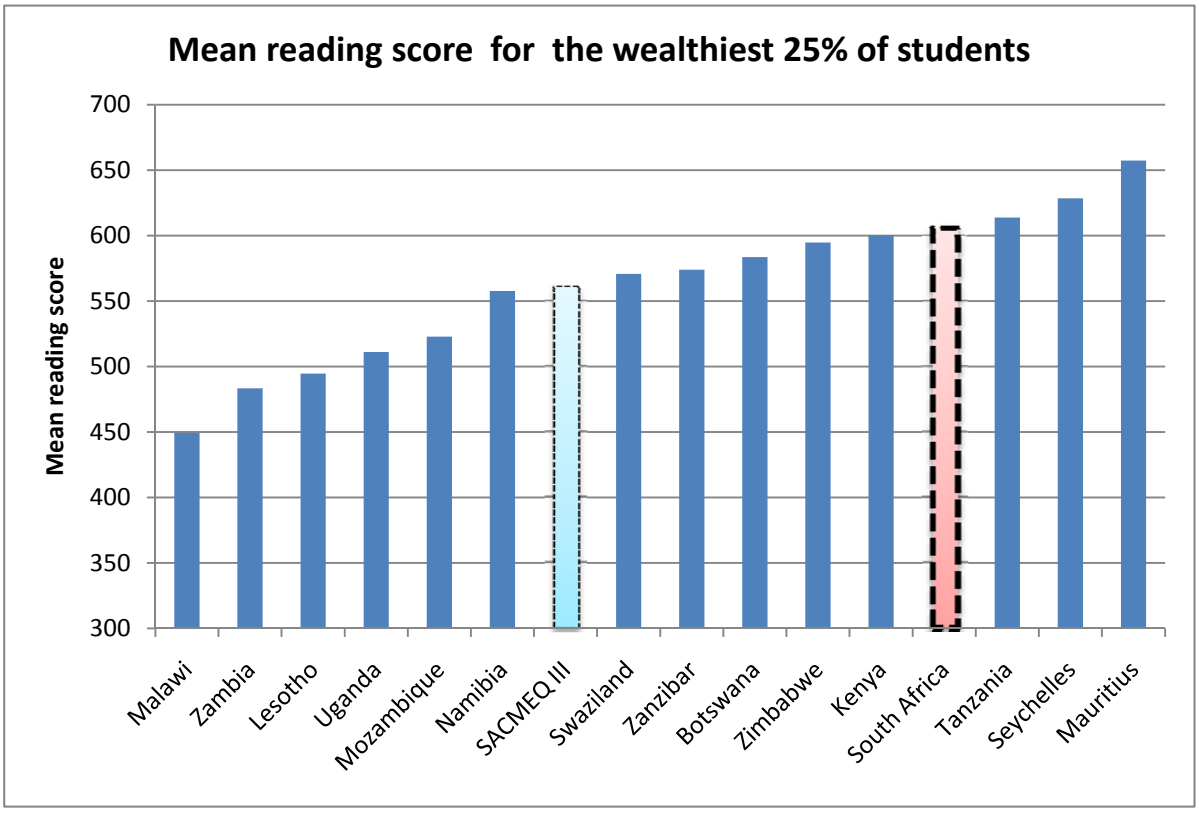

Figure D4

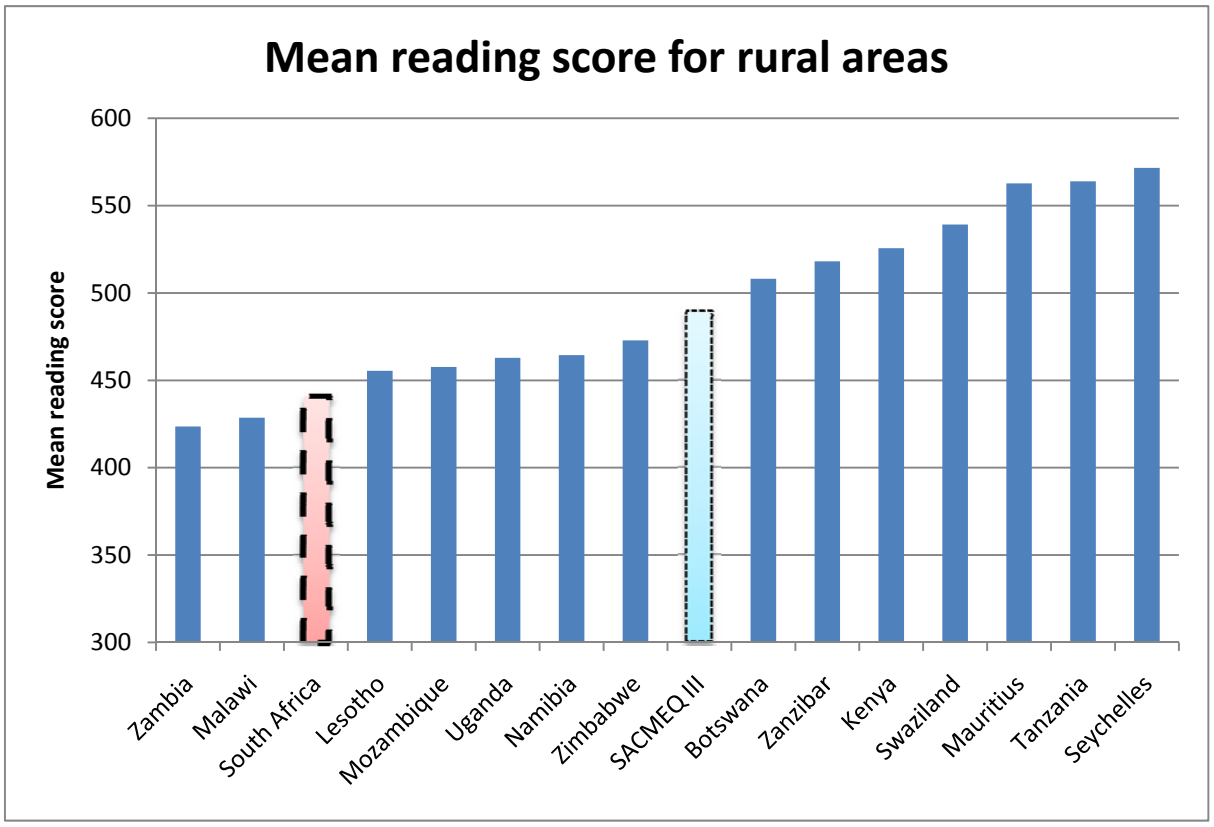

Figure D5

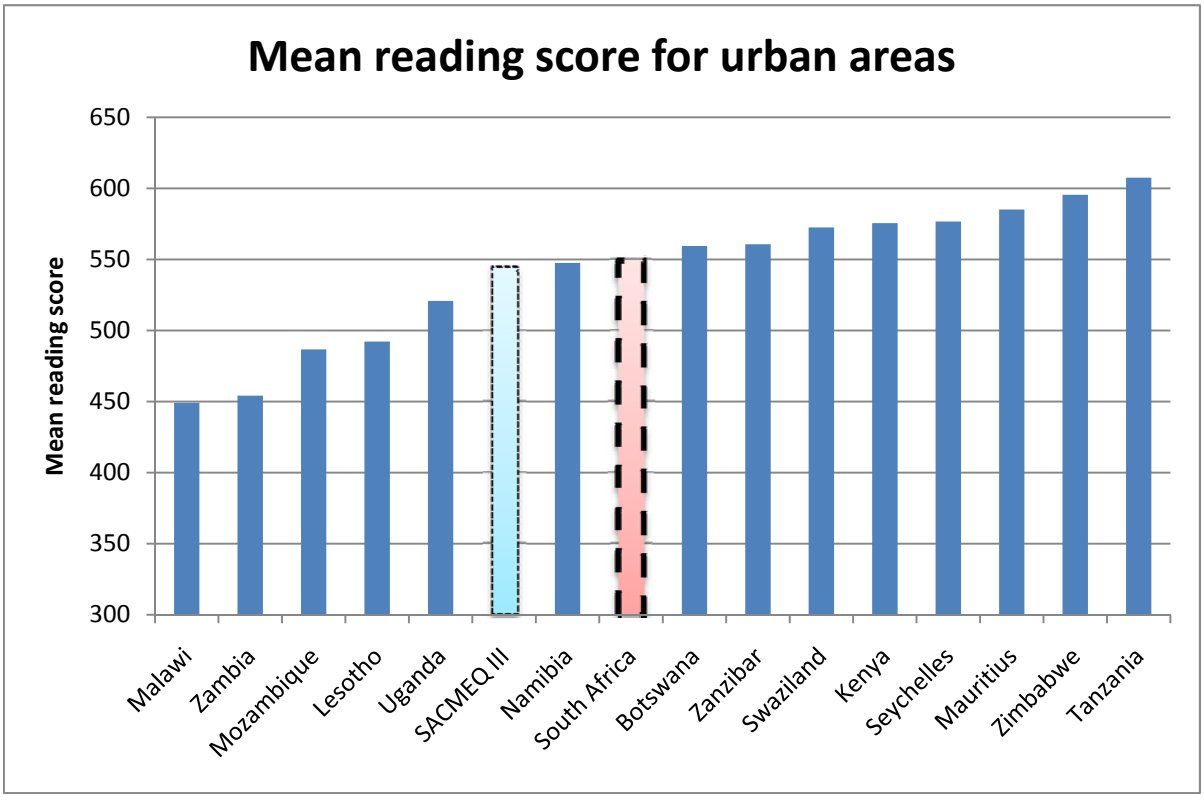

Figure D6

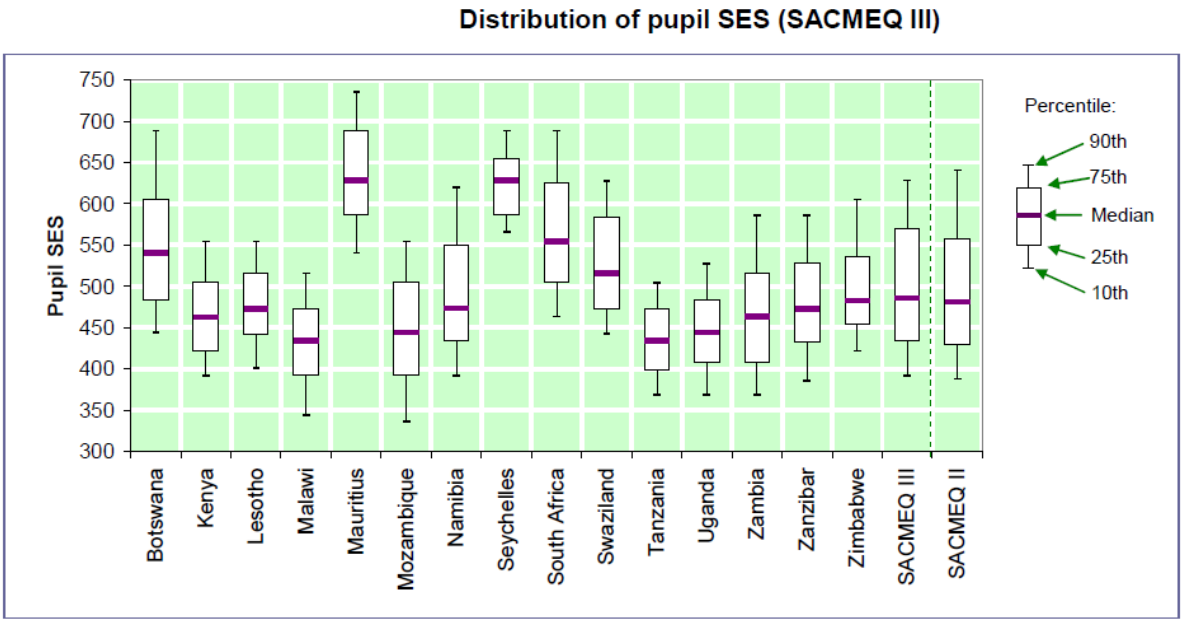

Figure D7

Source: SACMEQ (2010)

Table D5: SACMEQ III Student Maths Scores

	Boys	Girls	Rural	Urban	Low SES (Bot25%)	High SES (Top 25%)	Overall
Botswana	517.5	523.6	501.1	538.8	479	553.1	520.5
Kenya	567.6	546	544.5	580	540.9	595.8	557
Lesotho	477.1	476.8	469.3	492	460.2	498.3	476.9
Malawi	452.7	441.1	443.7	457.6	444.7	454.4	447
Mauritius	616.1	630.7	613.2	634.1	554.2	719.2	623.3
Mozambique	488.2	478.6	477.6	487.5	470.8	510.8	483.8
Namibia	472	470.1	448.5	506.1	443.7	513.5	471
Seychelles	535.2	566.7	550.2	550.9	498.7	593.6	550.7
South Africa	491.2	498.4	456.7	533.1	446.2	578.6	494.8
Swaziland	545.5	536.2	535.6	552.9	533.4	552.4	540.8
Tanzania	568.5	537.5	542.1	575.7	540.4	579.4	552.7
Uganda	486.7	477.2	470.8	511.5	465.4	504.2	481.9
Zambia	440.8	429.2	428.6	447.2	424.5	463.1	435.2
Zanzibar	489.3	483.9	477.8	500.5	471.1	510	489.9
Zimbabwe	520.8	519	492.1	589.6	487.8	588.8	519.8
SACMEQ III	511.9	507.6	493.9	533.2	488.7	541.7	509.7

(Source: SACMEQ, 2010)

Table D6: SACMEQ III Student Reading Scores

	Boys	Girls	Rural	Urban	Low SES (Bot25%)	High SES (Top 25%)	Overall
Botswana	519.7	549.4	508.1	559.5	474.4	583.6	534.6
Kenya	544.1	542.1	525.6	575.6	517.8	600.2	543.1
Lesotho	463.5	471.5	455.5	492.3	448.5	494.6	467.9
Malawi	438.4	428.5	428.6	449.1	428.8	449.3	433.5
Mauritius	558.8	588.9	562.7	585.2	510.8	657.3	573.5
Mozambique	478.4	473.2	457.7	486.7	452.1	522.8	476
Namibia	489.6	503.7	464.4	547.5	457.8	557.7	496.9
Seychelles	544.4	607.2	571.6	576.7	509.3	628.5	575.1
South Africa	483.5	506	440.8	549.2	423.2	605.6	494.9
Swaziland	545.2	553.6	539.2	572.6	531.6	570.7	549.4
Tanzania	586.1	569.7	563.9	607.6	557.7	613.8	577.8
Uganda	481.5	475.9	462.9	520.9	459.6	511.1	478.7
Zambia	437.1	431.5	423.6	454.2	418.8	483.4	434.4
Zanzibar	526.2	539.6	518.1	560.7	499.4	573.9	536.8
Zimbabwe	501.5	512.5	472.9	595.5	469.6	594.7	507.7
SACMEQ III	506.8	517.1	489.9	544.8	481.3	561.2	512

(Source: SACMEQ, 2010)