Uncovering indicators of effective school management in South Africa using the National School Effectiveness Study

STEPHEN TAYLOR

Stellenbosch Economic Working Papers: 10/11

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

Uncovering indicators of effective school management in South Africa using the National School Effectiveness Study[1]

STEPHEN TAYLOR

ABSTRACT

For many poor South African children, who are predominantly located in the historically disadvantaged part of the school system, the ongoing low quality of education acts as a poverty trap by precluding them from achieving the level of educational outcomes necessary to be competitive in the labour market. An important question is the extent to which this low quality of education is attributable to poverty itself as opposed to other features of teaching and management that characterise these schools. The literature explaining schooling outcomes in South Africa has reached a consensus that additional educational resources are no guarantee of improved outcomes. While socio-economic status remains the most powerful determinant of educational outcomes, studies have typically struggled to isolate other school and teacher characteristics that consistently predict outcomes, leaving much of the variation in achievement unexplained. Several authors have pointed to an ineffable mix of management efficiency and teacher quality that must surely underlie this unexplained component.

The National School Effectiveness Study (NSES) is the first large-scale panel study of educational achievement in South African primary schools. It examines contextually appropriate features of school management and teacher practice more thoroughly than other large sample surveys previously administered in South Africa. Using the NSES data, this paper identifies specific aspects of school organisation and teacher practice, such as the effective coverage of curriculum and completed exercises, which are associated with literacy and numeracy achievement and with the amount of learning that occurs within a year of schooling. Some suggestions are also made regarding the appropriate way to interpret these results for the purpose of policy-making

---

1. INTRODUCTION

Prior to 1994 education in South Africa was characterised by institutionalised inequality as part of the broader programme of *apartheid*. Schools were governed by separate education departments for each race group.[2] Black, Coloured and Indian schools received considerably less funding and real resources and consequently produced an inferior quality of education in general. Since the transition to democracy a unified Department of Education has been established and considerable progress has been made with regard to improved equity in funding and resource provision. However, inequity in the quality of education has proved a more enduring problem. For many poor children, who are predominantly located in the historically disadvantaged part of the school system, this low quality of education acts as a poverty trap by precluding them from achieving the level of educational outcomes necessary to be competitive in the labour market. An important question is the extent to which this low quality of education is attributable to poverty itself as opposed to other features of teaching and management that characterise these schools.

Analyses of fiscal incidence have demonstrated that massive resource shifts have taken place since the late 1980's, to the extent that government spending on primary and secondary education has become redistributive (e.g. Van der Berg, 2006, Gustafsson and Patel, 2006). Non-personnel funding (including for example, infrastructure and learning support materials), as outlined in the Norms and Standards introduced in 2000, is explicitly pro-poor in design. The level of non-personnel funding received by schools depends on the official "school poverty quintile" into which they are classified. Since 2006 the poorest two quintiles (and more recently also the third quintile) have been classified as "no-fee schools". This means that greater funding is made available to them in compensation for not charging fees. In the mid-1990's pupil-teacher ratios and teacher salaries were made more equitable across the historically different groups of schools. Personnel spending, however, is not strictly pro-poor as better qualified and more experienced teachers who command somewhat higher wages generally choose to work in more affluent schools. Seeing as personnel spending comprises at least 80% of overall government spending on education this limits the extent to which spending can be redistributive.

---

[2] Under the apartheid system there were separate education departments corresponding to the various race groups in South Africa. There were separate departments for white schools (House of Assemblies – HOA), coloured schools (House of Representatives – HOR), Indian schools (House of Delegates – HOD) and black schools (Department of Education and Training – DET) and each of the homelands had an education department.

The substantial increase in resources invested in the historically disadvantaged parts of the school system has unfortunately not produced a commensurate improvement in education quality. This is clearly evident in the test scores of South African students in numerous surveys of educational achievement that have been carried out in recent years.[3] These surveys have unequivocally shown that the overall level of achievement amongst South African children is extremely low. When the data allows for a disaggregation of schools according to the historically different systems a massive disparity is clear. Consequently numerous authors have now described the distribution of educational achievement in South Africa as bimodal (e.g. Fleisch, 2008; Van der Berg, 2008; Taylor and Yu, 2009). By this it is meant that the overall distribution in fact conceals two separate distributions corresponding to two very differently performing parts of the South African school system. Fleisch (2008: 1-2) maintains that there are effectively two education systems within one in South Africa. The difference between the two systems is rooted in the historically separate administration of education for each race group. The majority of South Africa's students (80-85%) are located in the historically disadvantaged system and demonstrate very low proficiency in reading, writing and numeracy. The second system produces educational achievement that is closer to what would be expected in the developed world. This system serves mainly white and Indian children, and increasingly black and coloured middle class children. The vast majority of university entrants are produced by this latter system. Van der Berg (2008: 145) describes these two groups of schools as operating under "separate data generating processes." It is therefore important to be sensitive to this underlying structural aspect when analysing educational achievement data for South Africa.

As alluded to earlier, the quality of education within the historically disadvantaged part of the school system has been largely unresponsive to increased resources. Van der Berg (2008: 153) argues that school resources do not *necessarily* make a difference but that the ability of schools to convert resources into outcomes is the crucial factor, and that this is where the policy attention is required. The ability to convert resources into outcomes is essentially what economists of education call school efficiency. However, this tradition of research has often been unable to illuminate the specific organisational features or teaching practices which promote greater school efficiency. Large-scale sample surveys of educational achievement,

---

[3] The main examples of these are the systemic evaluations, the Trends in International Maths and Science Surveys (TIMSS – 1995, 1999 & 2003), the surveys of the Southern And East African Consortium for the Monitoring of Education Quality (SACMEQ I, II and III) and the Progress in International Reading Literacy Study (PIRLS 2006).

which form the main source of information for education production functions[4] are not always designed for a developing country context and therefore have typically not adequately captured the salient aspects of school management practice in South Africa. Also, responses to questions put to teachers and principals tend to suffer from a systematic bias as respondents are likely to give themselves a more favourable appraisal than would accurately reflect reality. Moreover, behaviour is likely to change upon observation. The result is that aspects of school practice such as time management may not come through significantly in modelling student achievement, even though such factors do indeed matter. School functionality or efficiency remains something of a "black box": resources flow into the box and differential outcomes emerge, yet little is known or can be proven about what occurs within the box to determine the outcomes.

Van der Berg and Burger (2002), in their study of achievement in the Western Cape province, found that approximately two-thirds of the variation in achievement could be explained by socio-economic status (SES), the racial composition of schools and a selection of teacher resource variables. They suggest that the efficiency of school management was probably an important omitted variable. Similarly, Crouch and Mabogoane (1998), combining the unexplained variation in their model with the effect of a dummy variable for historical education department (which they regard as capturing an efficiency dimension because SES was already controlled for in the model), estimated that approximately 50% of the variation in school performance was attributable to the unobserved feature of management efficiency. A production function study by Gustafsson (2007) did manage to identify the correct allocation of management and teaching time as one management level factor that was associated with achievement in South Africa.

Figure 1, which is taken from Van der Berg (2007: 857), can be regarded as suggestive evidence of the influence of unobserved school (dys)functionality that is hindering educational achievement in South Africa. The figure shows lowess regression lines of average school mathematics achievement against school mean SES for South Africa and for the other African countries in the SACMEQ II survey. The asset-based index for SES is comparable across all SACMEQ countries. The lowess line for South Africa lies below that for the other countries across most of the distribution. Only at the most affluent end of the distribution do South

---

[4] Education production functions are a commonly used modelling technique in the economics of education. Production functions model cognitive skills as a function of an individual's personal characteristics that influence their learning efficiency as well as various aspects of school quality that influence skills. Crudely speaking, these models examine how various inputs affect the "production" of cognitive skills.

African schools enjoy a performance advantage. Van der Berg (2007: 857) concludes that poor South African children are performing worse than equally poor children in the other African countries in this sample – this despite favourable characteristics in South Africa in terms of pupil-teacher ratios, the availability of textbooks and teacher qualifications. The figure demonstrates that although SES has a strong influence on achievement in South Africa and elsewhere, there remains room for improvement at given levels of SES. Unobserved aspects of school functionality, management efficiency and teacher behaviour are surely leading candidates to underlie the gap in Figure 1.

Figure 1: Lowess lines for South Africa and other SACMEQ countries
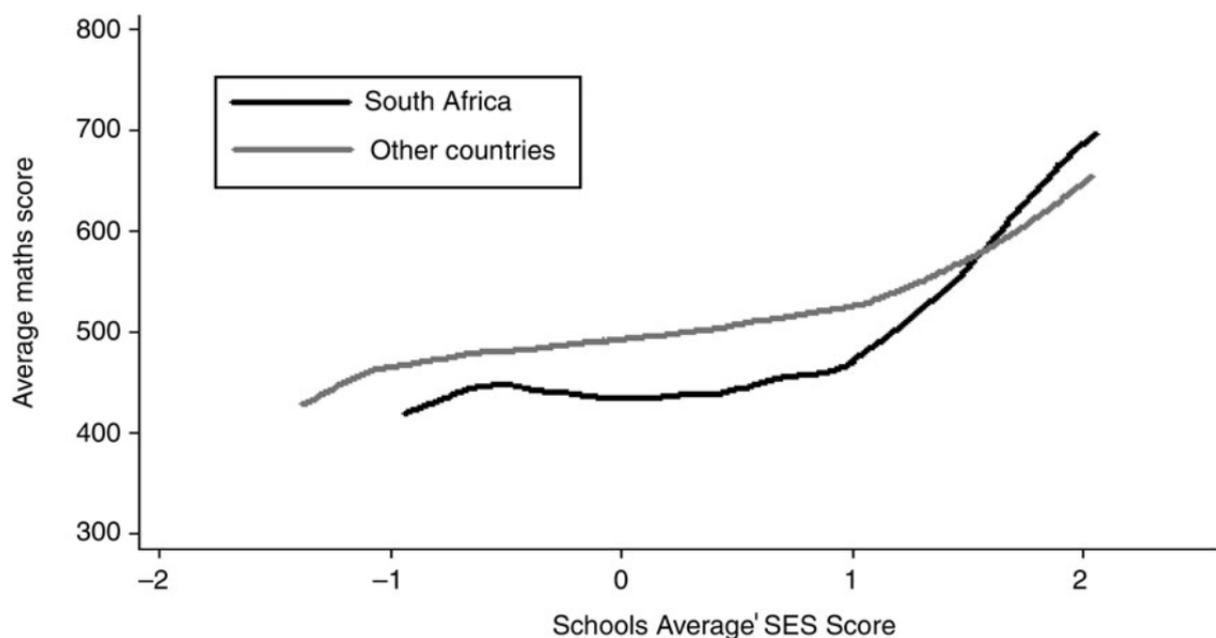
Source: Van der Berg (2007: 857)

Effective school management practice has thus proved hard to observe using large-scale sample surveys. This is evident in the review by Taylor, Muller and Vinjevold (2003) of factors that have been shown to influence student achievement. They split their review into large-scale sample-based studies and small-scale descriptive studies. They group influential factors emerging from large-scale sample studies into the following categories: race, parent education, household income and wealth, settlement type, family structure, gender, language use and language of instruction, teacher qualifications, facilities, pupil-teacher ratios and learning materials. Absent from this list but present under the list of factors described by small-scale studies is management. Taylor *et al* (2003: 61) maintain that the task of management is to

"provide an environment in which teachers can teach and students can learn." It is understandable that case study-type methodologies, which involve extensive observation and open-ended description, are better suited to capturing this management function than sample surveys, which rely mainly on closed-ended questions. The limitation of small studies, however, is that it is not possible to generalise from them conclusions that apply to the school system at large.

The next section argues that the data used in this paper, which comes from the National School Effectiveness Study (NSES), does indeed boast a richer collection of school and teacher variables as well as several other advantages that are unique in the South African context. After introducing the data and describing the overall literacy and numeracy results, Sections 3, 4 and 5 present descriptive analysis of the association of SES with achievement, the perpetuation of the historical ex-department dimension and several indicators of effective teaching and management practice in South African schools, respectively. The predictive power of these indicators is more rigorously analysed in Section 6 using a variety of multivariate regression techniques. It is hoped that these models will take the analysis of educational achievement and in particular the influence of teaching and management practice somewhat further than what has previously been possible in the South African context, due to the unique design of the NSES.

2. THE NATIONAL SCHOOL EFFECTIVENESS STUDY: DATA DESIGN AND BASIC RESULTS

Data for the National School Effectiveness Study (NSES) were collected between 2007 and 2009 on a nationally representative sample of schools in South Africa. The project was managed by JET Education Services and funded by the Royal Netherlands Embassy. Students in 266 schools in eight of the nine provinces of South Africa were tested in literacy and numeracy in 2007 (grade 3), 2008 (grade 4) and 2009 (grade 5).[5] The same individuals were tested in each year thus producing a panel dataset. The same tests were administered each year making the results comparable from one year to the next. In addition to the testing, a wide variety of other information was collected through student questionnaires in 2007, 2008 and 2009, teacher questionnaires in 2008 and 2009 and school principal questionnaires in 2007, 2008 and 2009.[6]

---

[5] Unfortunately the project was blocked from surveying Gauteng due to other testing that was being administered in that province at the same time.
[6] Information on the ex-racial department of schools was imputed from the DoE's Master List of Schools.

At the time of writing this paper, the third wave of data had only recently been cleaned and made available. Therefore, much of the analysis presented here is based on only the first two waves of data from 2007 and 2008, comprising a sample of 11813 students. Some preliminary analysis is included based on all three waves, which due to attrition comprises 8383 students who were surveyed in all three years. This panel nature of the data is distinctly advantageous as it offers the potential to observe the amount of learning that occurs over time rather than a simple cross-sectional snapshot of achievement. Using gain scores as the outcome of analysis means that omitted variable bias, such as that due to innate ability, can at least to some extent be controlled. This is not possible when a single cross-section of achievement is used.

A second advantage of the NSES is the extensiveness with which it covered school management, teacher knowledge and teacher practice issues. A wide variety of issues were surveyed and they were covered with remarkably fine detail for a large-scale sample survey. For example, an extensive document review was carried out including examining the frequency of various types of exercises in student workbooks. English teachers took a short literacy test and mathematics teachers took a short numeracy test, allowing the effects of teacher knowledge on student achievement to be investigated. The SACMEQ II survey had included this in its design, but South African teachers were exempt from taking the test, reportedly due to opposition from teacher unions. The third wave of SACMEQ did test South African teachers and this dataset had only been partially released into the public domain at the time of writing. Spaull (2011) has conducted preliminary analysis of the role of teacher knowledge in South African educational achievement using SACMEQ III.

A further definitional issue relates to the derivation of the overall literacy and numeracy scores. The literacy test consisted of 40 items and the numeracy test 53 items. The scores presented here are percentage scores where each item is given the same weight in the overall score. However, in the most recent dataset with all three waves the literacy percentage scores have been calculated so as to weight up longer items. This is probably the preferable method as it makes sense that an item involving an answer and a sentence to substantiate the answer will provide more information than a multiple choice question, for example, and should have the mark allocation categories of 0, 1, 2 rather than 0, 0.5, 1, as was the case in the unweighted derivation of the scores.

Table 1 summarises the mean scores for literacy and numeracy in each year as well as the gain scores by gender and home language. Note that Table 1 is based on the dataset for the first two waves only and therefore uses the unweighted scores. An initial observation is that the scores

were rather low in general, especially considering that the difficulty level of the test questions ranged from grade 1 level to grade 4 level. The mean achievement in literacy in 2007 (grade 3) was 19.38%, which improved to 27.03% a year later. For numeracy the mean achievement increased from 28.42% in grade 3 to 34.58% in grade 4.

Table 1: Literacy and numeracy results by gender and home language[7]

	Literacy 2007	Literacy 2008	Literacy gain	Numeracy 2007	Numeracy 2008	Numeracy gain
Females	20.39	28.63	8.23	29.42	35.65	6.23
Males	18.27	25.28	7.01	27.33	33.41	6.08
African language	16.93	24.14	7.21	25.08	31.01	5.92
Afrikaans or English	32.75	42.81	10.06	46.62	54.08	7.46
Total	19.38	27.03	7.65	28.42	34.58	6.16

Several other patterns are evident in the table. On all the outcomes female students performed better than male students on average. Students whose home language was Afrikaans or English performed considerably better than those whose home language was one of the other South African languages.[8] Two factors probably drive this difference. Firstly, the tests were administered in English, which would have afforded English speakers an understandable advantage.[9] One would expect this advantage to be reduced in grade 4 as the English ability of students in African language schools improves, but the gap appears to widen in grade 4 as seen in Table 1. Secondly, students who spoke Afrikaans and English came from more affluent homes (as measured by an asset-based index of SES to be introduced in the next section) than African language students, and were located predominantly in historically white and coloured schools. Thus a socio-economic and school system effect also underlies the disparity in achievement by language.

Table 2 shows the mean literacy and numeracy scores for each year, the mean gains from one year to the next and the mean 2-year gain score. Note that in the case of literacy, these scores

[7] A weight was specified to adjust for the sampling design in the analysis in this table, and in the forthcoming analysis when appropriate. The sampling used a one stage stratification design on the basis of province so that weights differed according to province but each student within a given province had the same weight.

[8] Box plots of literacy and numeracy achievement for all 11 home languages are presented in Appendix A.

[9] The decision to administer all three waves of testing in the NSES in English was made because the language of learning and teaching (LOLT) in South African schools changes from the mother tongue in the Foundation phase (grade R-3) to English in grade 4. It therefore made sense to test in English at the grade 4 and 5 levels, and, for the sake of standardising the tests, they were also administered in English in the first wave (grade 3).

are weighted to account for the value of each item. Comparing Tables and 1 and 2, the mean scores are not substantially different, indicating that the weighted version of the literacy scores does not dramatically alter the analysis. The numeracy 2007 and 2008 scores are slightly higher in Table 2 than in Table 1, reflecting that those individuals who dropped out of the sample in the third wave were a somewhat weaker group.

Table 2: Mean weighted scores and gain scores in Literacy and Numeracy for all 3 waves

	Literacy	Numeracy
2007 (grade 3)	20.15	29.38
2008 (grade 4)	29.59	35.50
2009 (grade 5)	37.73	47.04
Gain 2007 - 2008	9.43	6.12
Gain 2008 - 2009	8.14	11.54
2-year gain	17.57	17.66

Using box plots, Figures 2 and 3 depict the provincial breakdown of literacy and numeracy scores respectively. The thick bars extend from the 25th percentile to the 75th percentile of scores, with the median indicated somewhere between. Both figures depict a trend of moderate improvement across the three years. The Western Cape was the best performing province and also achieved a significant gain over the three years. All provinces recorded average gains over the three years of between 13 and 23 percentage points in both literacy and numeracy. It may be of concern that the Eastern Cape achieved the lowest gains in both literacy (14.39 percentage points) and numeracy (13.12 percentage points) and that off a low baseline level of performance.

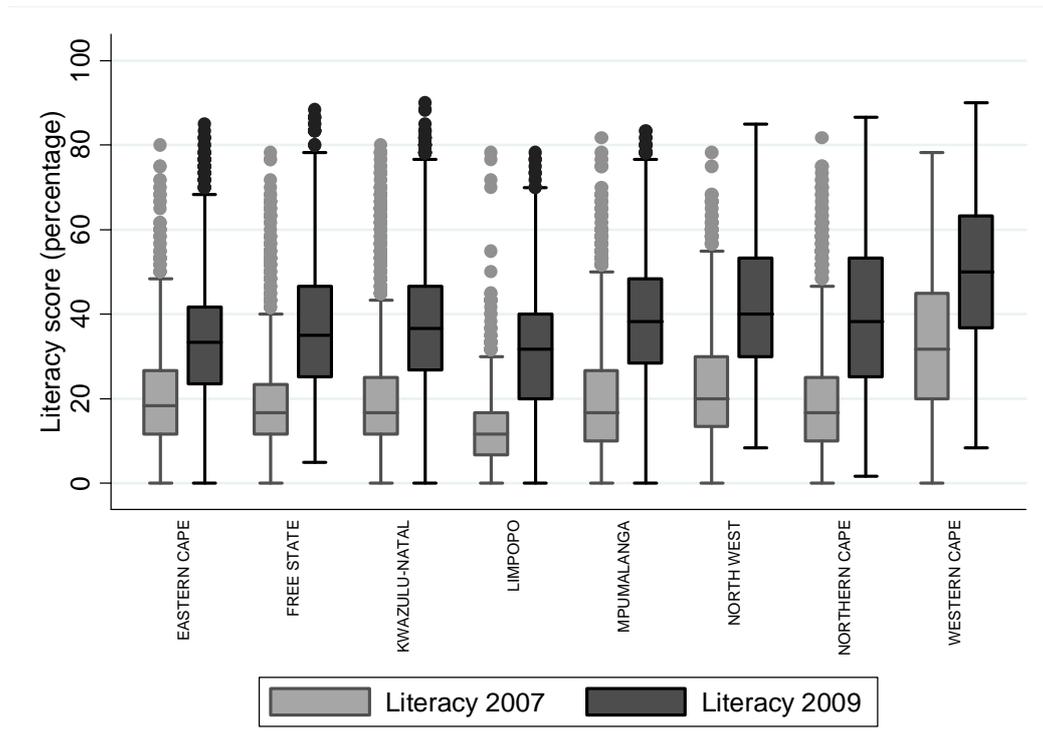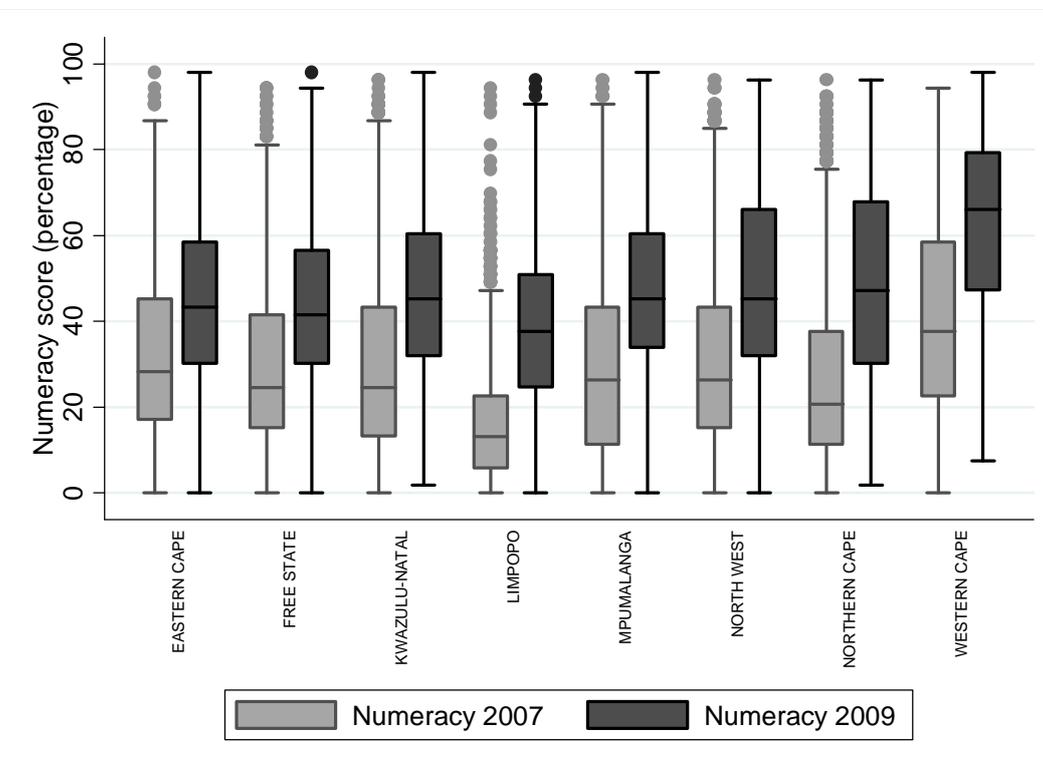**Figure 2: Literacy scores (weighted) by province**

**Figure 3: Numeracy scores by province**

3. LITERACY AND NUMERACY ACHIEVEMENT BY SES

Several studies have demonstrated that educational achievement amongst South African children is strongly associated with SES (for example, Taylor and Yu, 2009). Large surveys of educational achievement such as PIRLS, TIMSS, SACMEQ and the NSES typically do not contain information about household income or expenditure, as students cannot be expected to provide reliable income or expenditure information. It is therefore increasingly common to construct household asset-based measures of SES. Filmer and Pritchett (2001) set forth a strong case that asset-based classifications of households correspond closely to classifications by expenditure, and that asset-based indices are in fact better at predicting educational attainment than are expenditure data. One reason for this is that the presence of household assets is a more stable indicator than income or expenditure and therefore a better proxy for SES, which is fairly unresponsive to short-term household income shocks.

The student questionnaire in the NSES asked students about the presence of a number of household items at their homes. Students were asked about the presence of a fridge, tap water, a toilet, electricity, a car, a computer, a newspaper and a washing machine. Principal Component Analysis (PCA) was applied to these in order to derive appropriate weights for each variable in an SES index. The index was standardised to have a minimum value of zero and a standard deviation of 1. The mean of SES within each school was also derived in order to capture the overall SES of each school.

Figures 4 and 5 show kernel density curves of grade 5 literacy and numeracy scores, respectively, by quintile of school mean SES. The distributions show the proportion of the sample (along the vertical axis) that attained specific literacy scores (along the horizontal axis). For both literacy and numeracy, the distributions for the bottom four quintiles are remarkably similar, while that for the richest 20% of schools lies considerably to the right, indicating superior performance. This pattern is consistent with other research that has found similar levels of performance within the bottom four quintiles of South African schools and substantially higher achievement within the top quintile (Van der Berg, 2008, Taylor and Yu, 2009).

Figures 4 & 5: Kernel density curves of literacy 2009 (weighted) and numeracy 2009 by quintile of school mean SES
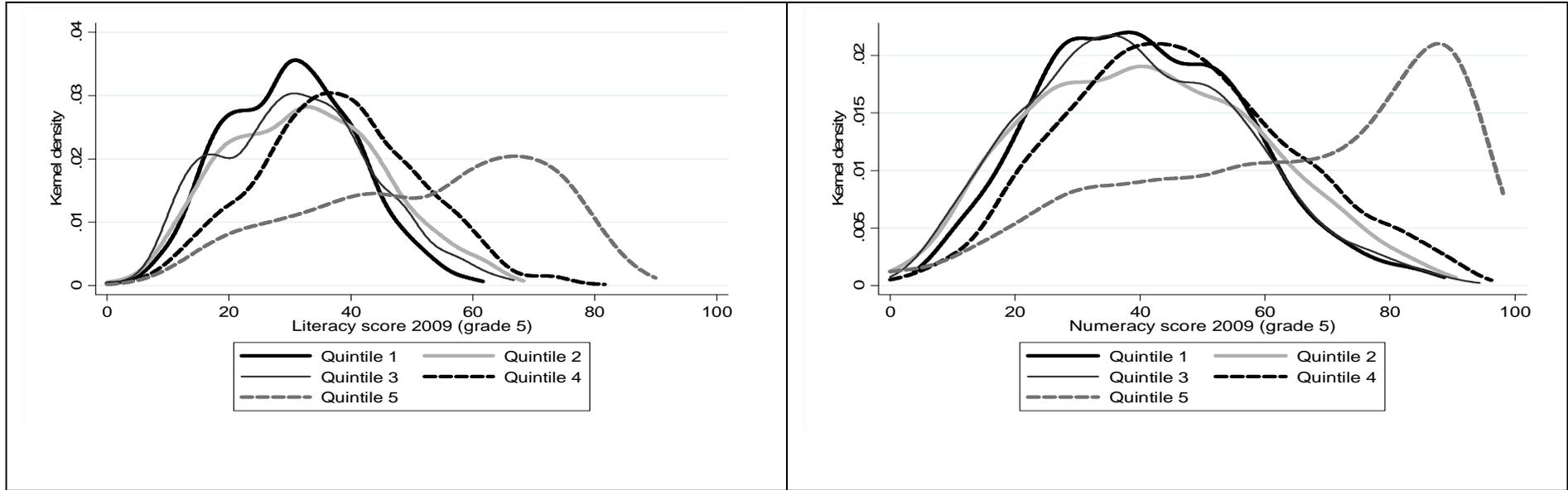
Figures 6 & 7: Lowess smoothing lines of literacy (weighted) and numeracy over the three years against SES
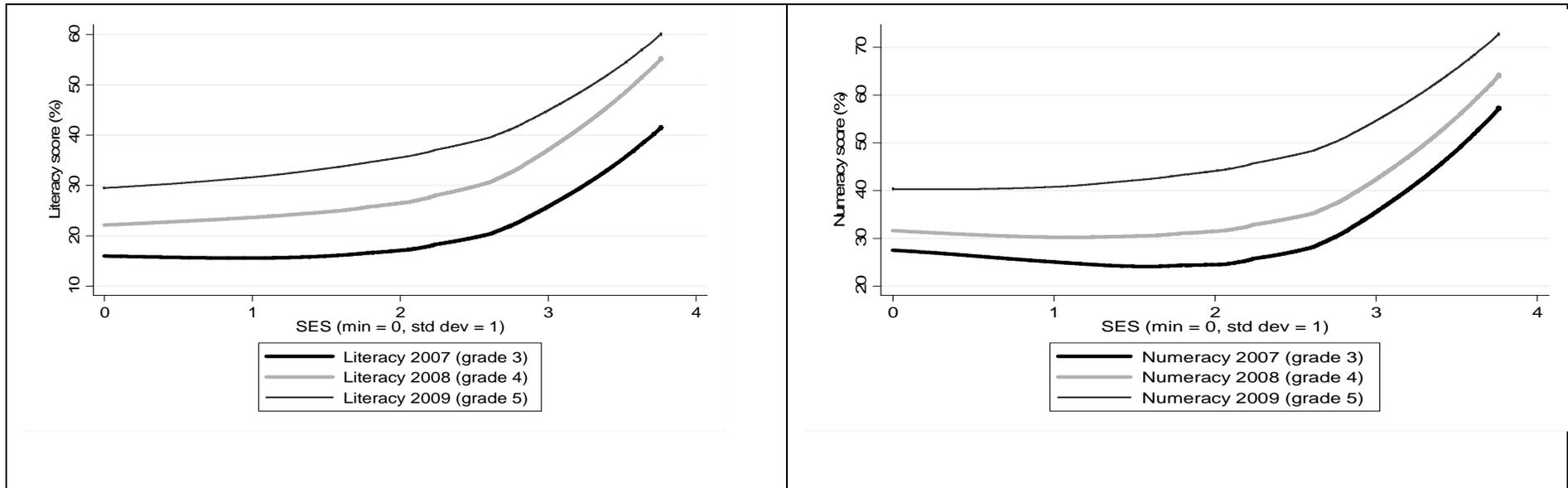
Figures 6 and 7 depict lowess-type socio-economic gradients across the three years of the survey for literacy and numeracy, respectively. A socio-economic gradient is the graphical representation of the regression relationship between SES and an outcome of interest, such as health or education. Lowess regressions do not require a linear or quadratic model specification but carry out locally weighted regressions at each data point and smooth the result through the weighting system. This means that the shape of a lowess curve is determined by the data rather than by the imposition of a model specification. It is evident that across the lower to middle range of SES the relationship is rather flat, while at higher levels of SES the relationship becomes stronger, as indicated by the steepness of the curves. This basic shape is consistent with similar estimations based on other data in previous studies, such as that done by Van der Berg (2007:857), which was reproduced in Figure 1 above. It is perhaps disconcerting that this pattern is evident as early as the third grade and stays constant through to grade 5, and later as the Van der Berg (2007) figure shows, as this reveals that the harmful impact of low SES is established early on in primary school and that no evidence can be found to suggest that primary schooling is able to reverse this. An implication of this for policy is that interventions should be made as early as possible in the educational process, including at the pre-school level and during the phase of Early Childhood Development (ECD).

Another preliminary way to analyse the influence of SES is to run an OLS regression predicting achievement based only on student SES and the mean SES in each school. Including both student and school SES allows one to assess the relative importance of these two factors. Table 3 reports the regression statistics for such a regression predicting literacy achievement in grade 4 (2008). Note that the inclusion of the squared and cubed versions of mean school SES was motivated by the sharp increase in the association of SES with achievement at higher levels of SES that was evident in Figures 6 and 7, and was justified by this third order specification providing a better model fit than either a linear or quadratic specification.

Due to the third order specification of school mean SES the coefficients reported in Table 3 are easier to interpret through graphing the results. Figure 8 depicts the predicted literacy score in 2008 according to the regression in Table 3. Movements along the horizontal axis represent changes in school mean SES, while the vertical width of the band of predicted values is due to variation in student SES at given levels of school mean SES. It is evident that variation in student SES at given levels of school mean SES was associated with fairly small changes in the predicted literacy achievement, whereas a movement to the top end of the school SES spectrum was associated with a very substantial increase in the predicted literacy score. It can therefore be said that the combined SES of a school has a more decisive impact on student achievement

than the student's own SES, although the latter may well determine what type of school students are able to attend.

Table 3: The effect of SES on literacy scores: student level and school level combined

Dependent variable: Literacy score 2008 (unweighted)	
Mean School SES	13.44***
	(1.46)
Mean school SES squared	-10.81***
	(0.75)
Mean school SES cubed	2.47***
	(0.11)
Student SES	1.51***
	(0.17)
Constant	16.03***
	(0.81)
R-squared	0.38
N	11813

* p<0.05, ** p<0.01, *** p<0.001
(Standard errors in parenthesis)

Figure 8: SES gradient for literacy: School SES combined with student SES (based on Table 3)
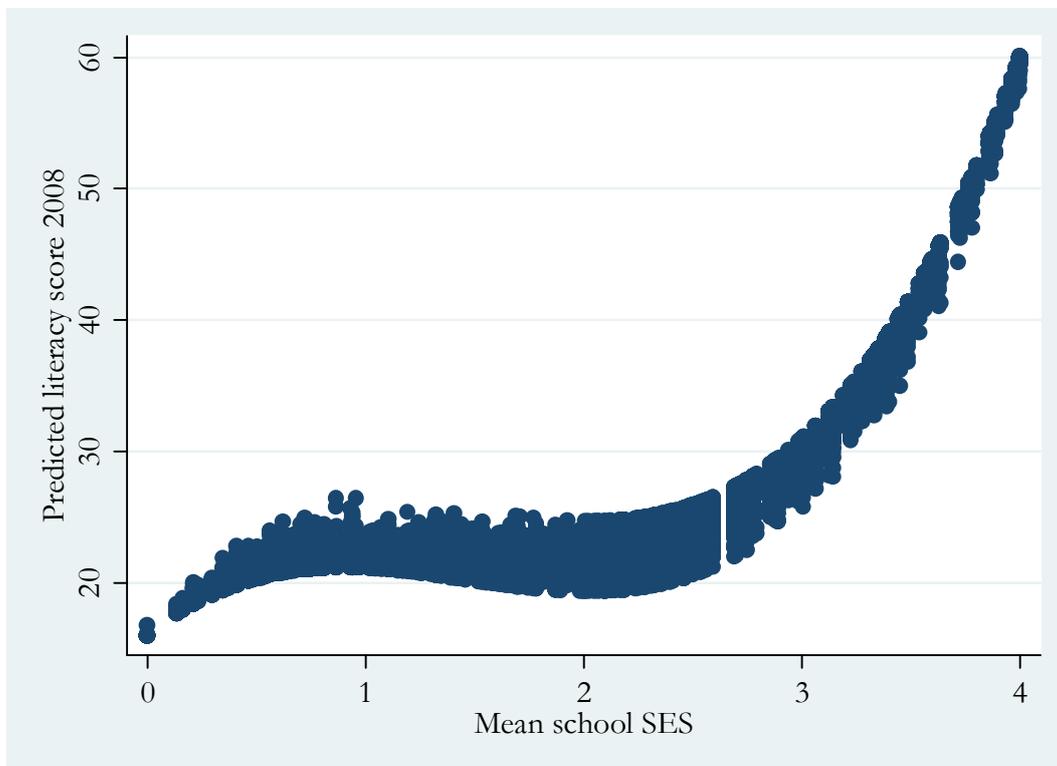
4. THE ONGOING DYNAMIC OF THE HISTORICAL FORMER EDUCATION DEPARTMENTS

As considered earlier, the influence of SES on achievement in South Africa is intertwined with the historical divisions in the governance of schools on the basis of race. Table 4 reports the mean literacy achievement (calculated as the average over three years) by former education department. The table confirms that historically black schools are achieving at lower levels than historically white and Indian schools, with historically coloured schools somewhere in between. Note that only four historically Indian schools were surveyed in the NSES, making this group too small to warrant meaningful analysis.

Table 4: Mean literacy scores (3-year average) by former education department

| Former department	Mean literacy over 3 years	Observations
Black (DET & homelands)	25.19	6776
Coloured (HOR)	39.12	880
Indian (HOD)	43.86	108
White (HOA)	58.78	619
Total	29.16	8383

It is revealing to compare the distributions of achievement for each year for historically black schools with those for historically white schools. Figures 9 and 10 depict these distributions for literacy and numeracy, respectively. The three solid lines are for historically black schools and the three broken lines for historically white schools. For both groups of schools, the distribution of achievement improved with each year (shifting to the right). It is alarming, however, that the distribution for grade 5 students in historically black schools was still a considerably weaker distribution than that of grade 3 students in historically white schools. One can therefore conclude that by the fifth grade the educational backlog experienced in historically black schools is already equivalent to well over two years worth of learning.

Figure 9: Kernel Density curves of Literacy 2007, 2008 and 2009 by ex-department
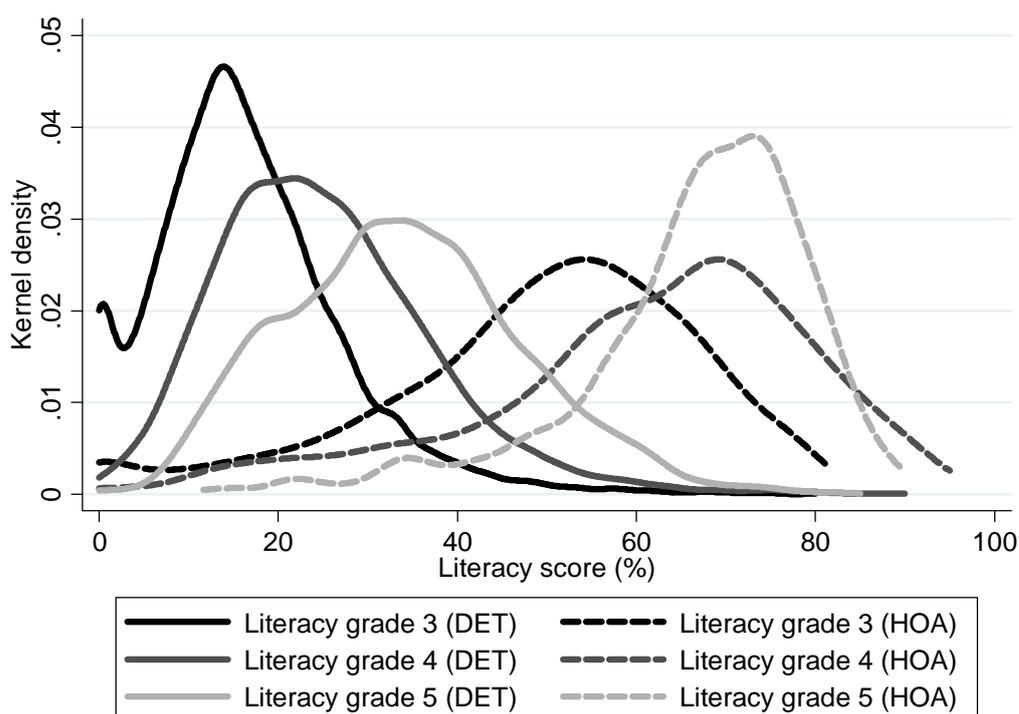
The picture for numeracy is similar. Figure 10 differs from the figure for literacy in that the distributions for historically black schools are more spread and the distributions for historically white schools are more concentrated at the top end, evidently with little room for improvement with scores in 2007 already concentrated at the high end of the spectrum. This merely reflects that the numeracy test was generally experienced as easier than the literacy test. The difference between the grade 5 distribution for historically black schools and the grade 3 distribution for historically white schools is even greater for numeracy than for literacy. This deficit despite more years of schooling may at least partly explain why the South African earnings function literature has found that white labour market participants enjoy higher returns to the same amount of education than black labour market participants (e.g. Burger and Jafta, 2006, Burger and Van der Berg, 2011). These studies suggest as the most probable explanation for this result that each additional year of education within the schools that black people typically attend does not produce the same increase in productivity than is achieved during in each additional year within the schools typically attended by white people. This is indeed what is observed in Figures 9 and 10.

Figure 10: Kernel Density curves of Numeracy 2007, 2008 and 2009 by ex-department
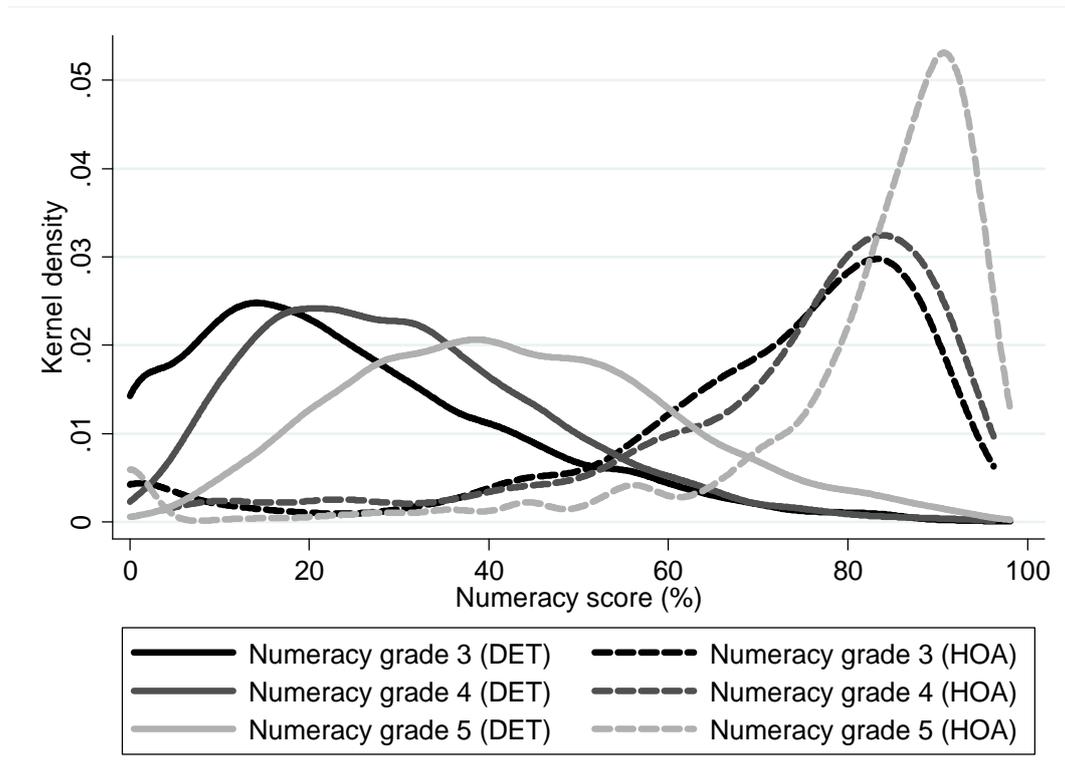
Tables 5 and 6 provide an example of how low functionality in the historically black section of the school system can act as a constraint to learning, even for those students who may enjoy other advantageous circumstances. Table 5 reveals an interesting pattern when looking at average achievement by family structure and home language. Family structure is known to have strong racial and socio-economic dimensions and to be associated with educational outcomes (Anderson, Case and Lam, 2001). For Afrikaans and English speaking students there are noticeable achievement gaps between those with no parents, a single parent and both parents. A race issue may be driving this pattern to some extent as single parent households are more common in coloured communities than in white communities. In contrast, amongst African language students average achievement in literacy and numeracy is similar for those of different family structure. At least two explanations for this might hold. It could be that the quality of parental support offered by African language parents is insufficient to substantially affect achievement. Alternatively, this pattern could reflect that most African language students are in schools with such a low level of functionality that parent support is unable to bring about a significant improvement in achievement. This latter possibility motivated the production of Table 6, which is the same as Table 5 but excludes students in historically black schools.

Table 5: Literacy and numeracy achievement by family structure and home language

Number of parents present	Literacy 2008		Numeracy 2008	
	African language	Afrikaans/English	African language	Afrikaans/English
0	23.72	35.81	30.76	44.34
1	24.21	39.98	31.63	50.37
2	24.61	47.23	30.70	60.13
Total	24.14	42.81	31.01	54.08
Number of students	9740	2048	9740	2048

Table 6: Literacy and numeracy achievement by family structure and home language excluding historically black schools

Number of Parents present	Literacy 2008		Numeracy 2008	
	African language	Afrikaans/English	African language	Afrikaans/English
0	35.88	38.60	47.25	47.50
1	43.66	42.56	56.69	53.51
2	44.13	49.05	57.71	62.49
Total	41.05	45.23	53.68	57.08
Number of students	630	1787	630	1787

Table 6 shows that African language students in historically white, coloured and Indian schools performed at a level much closer to that achieved by Afrikaans and English students. Moreover, achievement now differs with family structure for African language students. This supports the hypothesis that school functionality and parental support interact to influence achievement, and that low functionality in the historically black part of the system may be prohibiting parent support from being effective. However, this does not rule out the possibility that the quality of parental support is also driving this pattern as African language parents who value education enough to send their children to the better-performing historically white, coloured and Indian schools are themselves probably educated and therefore able to provide effective educational support.

During the years since the historically different parts of the school system were brought under a single administration, there has been some migration of black students into historically white, coloured and Indian schools, although not in the opposite direction (Soudien, 2004). Figure 11 compares the achievement of African language students in historically black schools with African language students in historically white schools. It is clear that those in historically white schools are performing at a much higher level on average. Although it is mainly an elite black middle class that attends historically white schools, Figure 11 is surely also indicative of a

different level of school effectiveness that is present in these two systems. To analyse this further requires multivariate analysis that also controls for individual SES of students in the different parts of the school system. Simple OLS regressions were therefore estimated, predicting the literacy and numeracy achievement of students whose home language was not English or Afrikaans, conditional upon student SES, mean school SES and former department. The results are reported in Table 7 and the predicted values for those in historically white and historically black schools are plotted in Figure 12.

Figure 9: Kernel density curves of numeracy achievement for African language students by historical education department
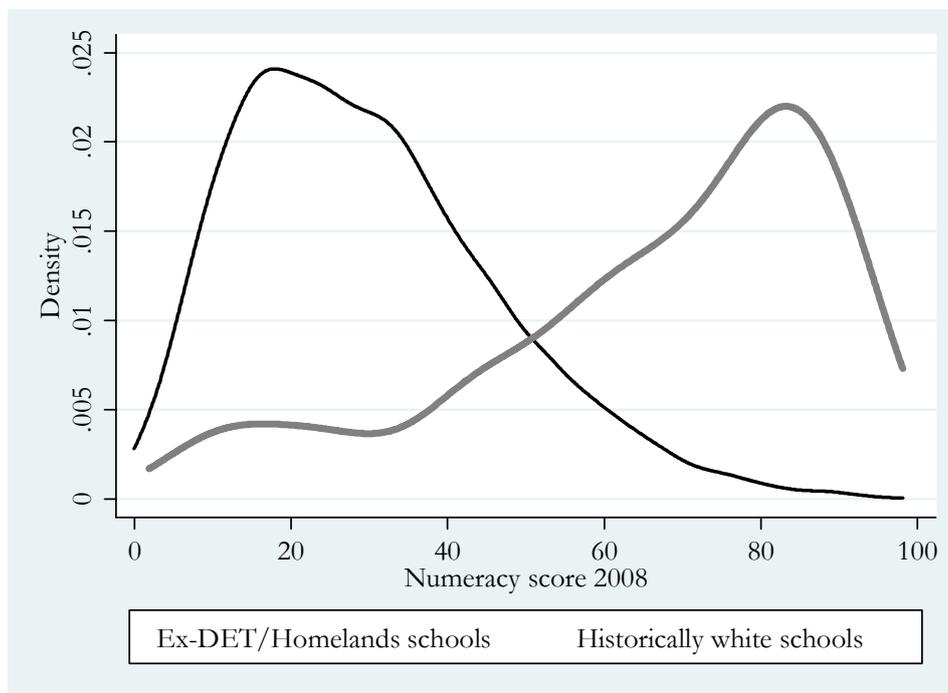
Table 7: OLS regressions predicting literacy and numeracy achievement for African language students by historical education department

Explanatory variables	For Literacy 2008	For Numeracy 2008
Student SES	0.74*** (1.56)	0.81*** (0.24)
Mean School SES	-8.41*** (2.45)	-13.40** (4.49)
Mean School SES squared	3.03*** (0.75)	3.95** (1.32)
HOR (C)	0.14 (1.41)	1.88 (2.30)
HOD (I)	6.38 (5.16)	13.50* (5.53)
HOA (W)	14.54** (4.64)	23.56** (7.56)
Constant	25.42*** (1.75)	37.39*** (3.48)
R-squared	0.2100	0.1414
Observations	9740	9740

~ p<0.10 ; * p<0.05 ; ** p<0.01 ; *** p<0.001

Note: Standard errors in parentheses

Figure 12: Predicted literacy and numeracy achievement for African language students by historical education department
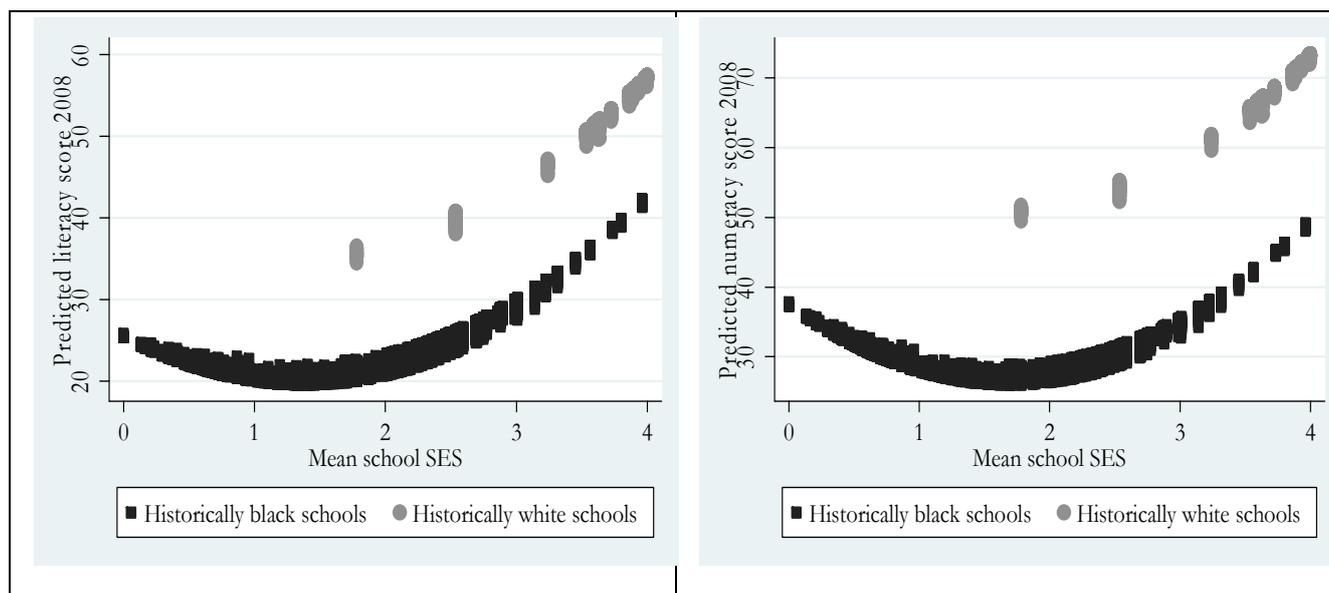
The table and figures demonstrate that even when controlling for student and school SES, African language students in historically white schools enjoy a considerable performance advantage over those in historically black schools. This difference is statistically significant and large, especially so in the case of numeracy. It is clear from this analysis that although achievement is strongly connected with student SES, much of this connection has to do with the effectiveness of schools in which students are located. Taking this analysis together with the finding shown earlier that school mean SES has a more important impact on achievement than individual SES, one might think of the effect of SES on achievement as a two-step process in which the first step is the decisive one: individual home SES may be a major determinant of the quality of schooling to which students gain access. Thereafter, home SES is limited in its ability to influence educational achievement.

Although it is clear that the historically disadvantaged and poorer parts of the school system are operating at a low level of efficiency, and that this is not completely attributable to SES, it is less clear what teaching and management practices underlie this low performance. The next section describes several school and teacher characteristics captured in the NSES that can be considered indicators of quality.

5. DESCRIPTIVE ANALYSIS OF INDICATORS OF EFFECTIVE SCHOOL MANAGEMENT AND TEACHING

The NSES boasts a rich collection of information regarding management and organisational practices within schools as well as teacher behaviour and practices. This derives from the sheer number of questions included in the principal and teacher instruments, the innovation of including short tests for teachers and an extensive review of student workbooks, which yielded several interesting indicators of curriculum coverage and the amount and type of work being done by children throughout the year. In order to avoid any bias caused by some teachers purposefully selecting the workbooks of more diligent students and other teachers selecting workbooks at random, teachers were asked to present the "best" student's workbook for inspection. Reviews of student workbooks were undertaken on this basis in 2008 and in 2009.

Student workbooks were examined to identify the number of mathematics topics (as specified in the curriculum) that had been covered up until that point in the year. Fieldworkers were looking for the 85 topics that are specified in the Revised National Curriculum Statement for grades R-9. Schools should have covered most of the curriculum by the time of the survey, although it is unlikely that exercises corresponding to all 85 topics would be identifiable in the workbooks of even the very best schools. This variable, therefore, represents a rough indicator of curriculum coverage. Table 8 reports the percentage of students located in schools where evidence was found of more than 25 maths topics being covered. This is broken down by former education department. Within the historically white part of the sample, 75% of students were in schools where evidence was found of more than 25 topics being covered, compared with just 26% of students in the historically black system.

Table 8: Percentage of students in schools where more than 25 maths topics were covered (2008)

Ex-department	Percentage > 25 topics	Number of students
DET (B)	26%	6306
HOR (C)	25%	849
HOD (I)	38%	86
HOA (W)	75%	591
Total	29%	7832

Table 9 shows the mean number of literacy exercises identified in student workbooks by former department. This demonstrates that considerably more exercises were undertaken by students within the historically advantaged parts of the system over the course of the year. Tables 8 and

9 offer some perspective on the large student achievement deficits being carried within historically black schools, as referred to earlier. If curriculum is not being covered and students are not frequently engaged in exercises it is hardly surprising that learning deficits will accumulate. It should be cautioned that there may be an element of bidirectional causality underlying the observed low curriculum coverage within historically black schools: If teachers take on students with prior learning deficits they may justifiably adopt a slower pace of curriculum coverage. However, the observed level of curriculum coverage is so low within historically black schools that it is surely safe to say that this is an aspect of school quality in need of attention.

Table 9: Mean number of literacy exercises found in the "best" learner's book (2009)

ex-department	Mean number of exercises	Number of students
DET (B)	33.43	6478
HOR (C)	62.40	837
HOD (I)	72.44	102
HOA (W)	75.21	580
Total	39.58	7997

Figure 13 provides an indication of the amount of extended writing, as measured by the number of exercises involving written paragraphs observed in student workbooks, that is undertaken by grade 5 students in South African schools. In 85 classes it would appear that no extended writing of at least a paragraph long had taken place. In only 19 classes could it be observed that students had written a paragraph at least ten times in the year.

Figure 13: Frequency of exercises consisting of paragraph length writing (grade 5)
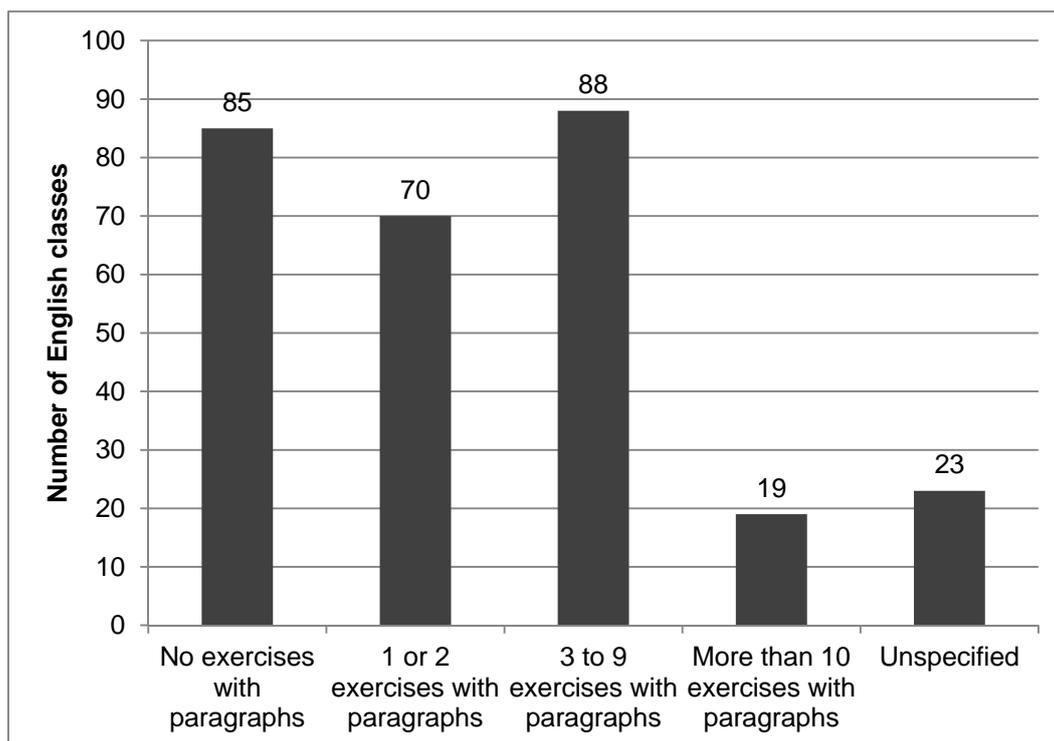
An interesting indicator of mathematics teaching quality is the frequency of complex exercises found in student workbooks. Essentially, a complex exercise was defined as an exercise consisting of more than one step. Table 10 shows the numbers of students and teachers in the various frequency categories. Nearly 22% of students in the sample were in classes where no evidence of any complex mathematics exercises could be found. Only 12% of students were in classes where more than 18 complex mathematics exercises had been completed during the year up to that point.

Table 10: The frequency of complex mathematics exercises in student workbooks (2008)

Number of complex exercises	Number of students	Percentage of students	Number of teachers
0	2586	21.89	69
1 to 4	3497	29.50	74
5 to 18	3016	25.54	73
more than 18	1429	12.09	41
unspecified	1285	10.88	23
Total	11813	100	280

Teacher knowledge has rarely been measured in large-scale sample surveys of student achievement in South Africa. The NSES administered a comprehension test with 7 questions to English teachers and a 5-mark test for mathematics teachers. The shortness of these tests means that they provide limited measures of teacher knowledge, but this feature does at least allow for the analysis to be taken one step further than before. Figure 14 shows a histogram of scores on the English teacher test. The histogram is skewed to the right indicating that most of the scores were concentrated at the higher end. Although there were few extremely low scores, there was still a lot of variation in teacher knowledge and only 16% of teachers scored 100%.

Figure 14: Histogram of English teacher test scores (2008)
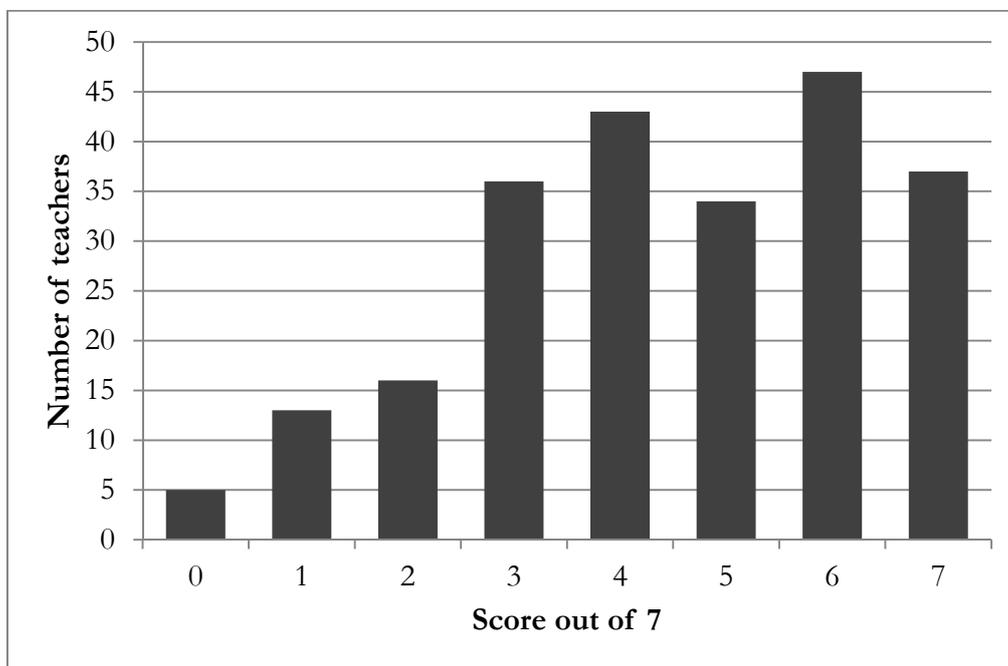
Figure 15 shows the number of teachers achieving each score out of five on the mathematics test. Most of the scores are in the middle range with only 29 teachers scoring 100%. The significance of this is better realised by looking at the distribution of mathematics teacher knowledge at the level of students, i.e. the numbers of students taught by teachers with each test score. Table 11 presents this breakdown. The table reveals that more than half of the students in this survey were taught by teachers who scored 40% or less on the simple mathematics test. Just over 12% of students were taught by teachers who scored 100%. It is not surprising that the achievement of South African students is so low given that teacher knowledge appears to be deficient in many of our schools. The far right column of Table 11 shows the mean numeracy

achievement in 2008 (grade 4) for students in each category of teacher test score. For teachers who scored anything less than 100% the mean achievement of students was very similar. However, those students taught by teachers who scored 100% performed noticeably better than the rest. This suggests that more effective teachers have sound knowledge, or at least knowledge that is sound enough to achieve 100% on this short test. In contrast, any score less than 100% is an indicator of lower teacher quality and is linked to low student achievement. However, this assertion needs to be tested using multivariate analysis as teacher knowledge may well be correlated with other aspects of school quality and with school mean SES, and these factors could be driving the pattern in Table 11. An example of the questions in the mathematics teacher test is also provided below.

10 days 75 hours can be written as .... days .... hours

Figure 15: Histogram of Mathematics teacher test scores (2008)
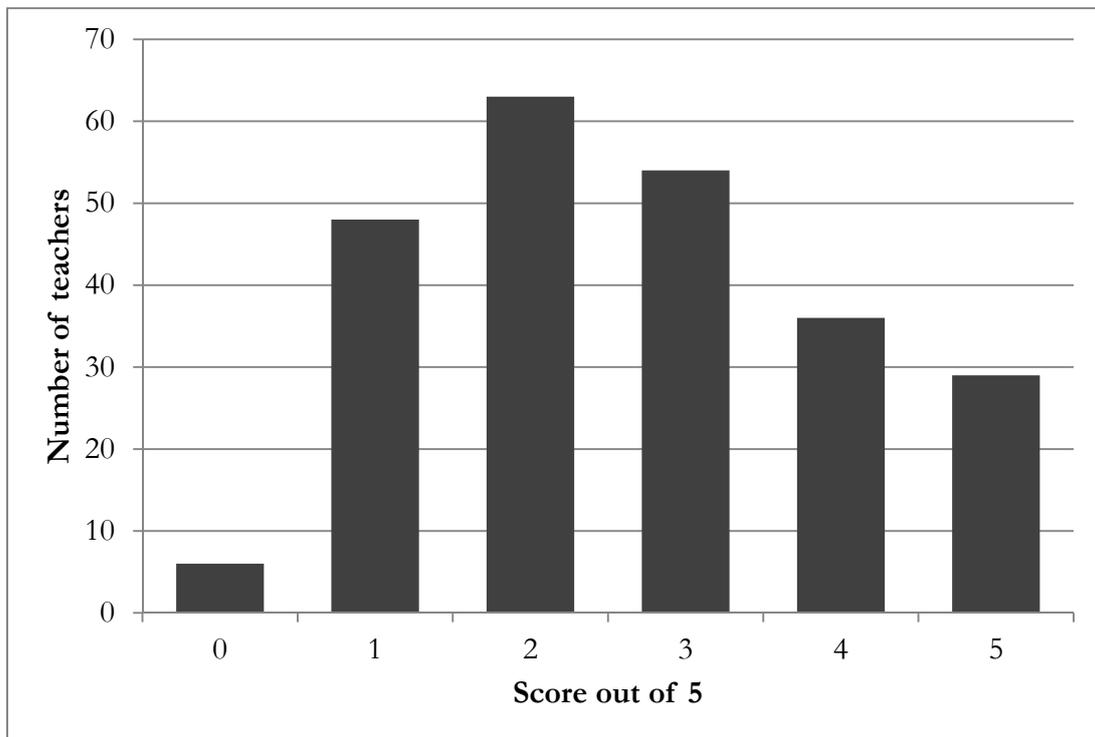
Table 11: The number and performance of students by teacher knowledge

Teacher score	Number of students	%	Cumulative %	Mean Numeracy 2008
0	210	2.12	2.12	37.27
1	2130	21.52	23.64	33.04
2	2774	28.02	51.66	33.50
3	2168	21.9	73.56	34.14
4	1408	14.22	87.79	34.77
5	1209	12.21	100	46.92
Total	9899	100	100	35.44

Another teacher characteristic captured in the NSES was the self-reported number of hours spent on actual teaching per week. This variable itself was not strongly correlated with student outcomes, although an interesting interaction between the time spent on teaching and teacher knowledge was noted. As Table 12 demonstrates, students taught by teachers who scored less than 100% in the mathematics test and who reportedly taught for less than 18 hours per week had lower numeracy achievement in grade 4 on average than students with any other combination of these two teacher characteristics. Students taught by teachers with either better knowledge or more time spent teaching but not both of these characteristics performed somewhat better than the poorest performing group. However, students whose teachers scored 100% and reportedly spent more than 18 hours teaching performed substantially better on average than the other students. Table 13 demonstrates that not only did this category of students perform at a higher level, but also showed the greatest improvement from one year to the next. This is an exciting finding as it suggests that it is only when teacher knowledge is combined with time on task that substantial student learning can be expected to occur.

Table 12: Means and frequencies of Numeracy achievement 2008 by teacher knowledge and time spent teaching

	Teacher score <100%	Teacher score 100%	Total
Less than 18 hours spent teaching	30.06 (3274)	34.84 (446)	30.64 (3720)
More than 18 hours spent teaching	36.13 (5416)	53.98 (763)	38.33 (6179)
Total	33.84 (8690)	46.92 (1209)	35.44 (9899)

Table 13: Means and frequencies of Numeracy gain score by teacher knowledge and time spent teaching

	Teacher score <100%	Teacher score 100%	Total
Less than 18 hours spent teaching	5.67 (3274)	5.01 (446)	5.59 (3720)
More than 18 hours spent teaching	7.07 (5416)	10.64 (763)	7.51 (6179)
Total	6.54 (8690)	8.56 (1209)	6.79 (9899)

Again, a word of caution regarding the interpretation of Tables 12 and 13 is necessary. It is likely that most of the more affluent of schools are located in the cell for "teacher score 100%" and "more than 18 hours spent teaching". Therefore, a multivariate analysis conditional upon SES is needed to test whether these teacher characteristics directly affect learning or whether they are better understood as indicators of the type of advantageous characteristics and practices that are present in the better-functioning part of the school system.

Another variable which is (at least anecdotally) known to be an issue in many of South Africa's schools but which has rarely been effectively linked to student achievement is teacher absenteeism. The NSES captured the number of teachers absent on the day of the survey. Taking this value as a proportion of the total number of teachers at each school it is possible to derive the percentage of teachers absent on the day of the visit. This is by no means an accurate assessment of teacher absenteeism over the full year, but it is at least one hard indicator of teacher absenteeism. Table 14 presents the interaction between the proportion of teachers absent on the day of the visit and the state of teacher attendance registers. The state of attendance registers says something about the organisational efficiency within schools and also reflects of how seriously teacher attendance is taken by the school management. The table demonstrates that teacher absenteeism was approximately twice as high in schools where the teacher attendance register was not up-to-date.

Table 14: Teacher absenteeism by state of teacher attendance register

	Percentage absent	Number of schools
Register not up-to-date	20.50	51
Register up-to-date	10.19	191

The documentary review also assessed the quality of inventories for textbooks and other learning support materials in schools. This provides another indicator of the organisational efficiency within schools. Table 15 shows how student performance was associated with the presence and completeness of LTSM inventories. It is evident that students in schools where inventories were both available and up-to-date performed better and achieved the highest gains from grade 3 to grade 4. This may mean that good management of school resources positively *affects* learning or that efficient management of learning materials is a *sign* of a good school. A combination of both of these possibilities may also underlie the association evident in Table 15. For this variable, as for many others discussed in this section, a better assessment of its impact is achievable through multivariate analysis. This is the focus of the next section.

Table 15: Student performance by state of school LTSM inventories

	Mean numeracy 2008	Mean numeracy gain	Mean literacy 2008	Mean literacy gain	number of schools
No inventory available	33.09	5.04	25.26	7.22	126
Inventory outdated	33.40	6.73	24.70	6.89	60
Inventory up-to-date	41.40	7.25	31.60	7.69	69

6. MULTIVARIATE REGRESSION ANALYSIS

The descriptive analysis presented above is useful to highlight broad trends in the data. However, if the quality of LTSM inventories is associated with student achievement, for example, it may be that this is an important factor for achievement or it may be that another factor, such as SES, is correlated with the quality of inventories and that this is the important determinant of achievement. The need for a multivariate analysis to estimate the relative impacts of such factors in combination is clear. Education production functions model cognitive skills upon individual characteristics, such as home background, measures of school quality, which can include resources, organisational practices, pedagogical methods and teacher characteristics, amongst others. This holds potential for addressing an important policy question: After accounting for the influence of SES, what school and teacher characteristics are associated with student achievement? Or slightly differently, what distinguishes better and worse-performing schools within poor communities?

The results of education production functions should be interpreted with caution, as there are numerous potential sources of bias, as Glewwe (2002) discusses. He therefore recommends an approach that sets out to gather as much evidence as possible and then make an overall judgement based on a number of estimations of production functions. Several models are presented in this section, using different techniques, and a number of other models not shown here were also estimated to provide an indication of the consistency of results. The first two models presented are cross-section models for literacy 2008 and numeracy 2008. Note that these models were based on the first two years of data only. Thereafter, two more advanced strategies are followed in an attempt to model learning gains. These represent an attempt to deal with one of the major potential sources of bias in education production functions, namely the omitted variable bias that may be present when innate ability is not controlled for. These latter two models are based on the most recent version of the NSES data, including all three waves, and should therefore be regarded as preliminary given that further rounds of analysis and fine-tuning will no doubt follow in time. A description of all explanatory variables that were used in this analysis is presented in Appendix B.

Table 16 shows the results from an OLS regression model predicting literacy achievement in 2008 (grade 4).[10] The set of student characteristics that was associated with literacy and numeracy achievement was remarkably consistent across various model specifications. One exception was the number of books at home, which was sensitive to model specification and was only a significant predictor of achievement in some models. However, student SES, age, household size, frequency of reading on one's own at home, home language and exposure to English were consistently important student characteristics to include in the models.

It is interesting to note which aspects of home background were not associated with achievement under the models in this section. Family structure, as measured by the number of parents present at home, was not significantly associated with achievement in any of the models that were estimated. This does not necessarily mean that it does not matter, as the effect of family structure may be contained within the effect of student SES. Another variable that one might have expected to influence achievement was the availability of help with homework from an adult at home. However, this variable also was not significantly associated with achievement in the multivariate analysis. Similarly, there was no significant effect of reading with an adult at home. Yet, children who reported reading on their own at home performed better than those

---

[10] In all the OLS regressions in this chapter the method of survey regression was used to account for complex sample design. The stratum variable was province (of which there were 8 due to the non-participation of Gauteng in the NSES), the Primary Sampling Unit (PSU) was the school (of which there were 266) and a person weight for each student, which differed only by province, was specified.

who did not, conditional upon all the other characteristics. As Table 16 shows, this positive effect was greater for those who read at least four times a week than for those who read between 1 and 3 times a week.

Table 16: OLS Regression model for literacy 2008

Explanatory variables		
Student characteristics		
Student SES	0.39*	(0.18)
Male	-2.48***	(0.26)
Young	-0.40	(0.46)
Old	-2.84***	(0.33)
Household size: large	-1.89***	(0.37)
Read 1 to 3 times a week	1.37**	(0.44)
Read more than 3 times	2.39***	(0.62)
Books at home: 1 to 10	0.60	(0.39)
Books at home > 10	1.17*	(0.48)
Home language English	8.42***	(1.52)
Speak English 1-3 times	1.75***	(0.38)
Speak English 4+	1.86**	(0.68)
English on TV 1-3 times	0.85*	(0.39)
English on TV 4+	3.35***	(0.44)
School characteristics		
Mean School SES	-9.13***	(1.77)
Mean School SES squared	3.35***	(0.45)
Pupil-teacher ratio	-0.18**	(0.07)
Teacher absenteeism zero	1.93*	(0.81)
LTSM Inventory good	1.66*	(0.80)
Problems with students index	-0.96*	(0.43)
Curriculum planned using year schedule	1.46~	(0.81)
Teacher characteristics		
Full year learning programme	1.55~	(0.87)
Constant	29.69***	(3.45)
R-squared statistic	0.4591	
N	10 860	

~ p<0.10, * p<0.05, ** p<0.01, *** p<0.001
(Standard errors in parentheses)

Notes:

The regression included dummy variables for the provinces but the coefficients on these are not reported in the table. Also, dummy variables controlling for non-response were included for the following characteristics: household size, frequency of reading at home on one's own, frequency of speaking English at home, frequency of hearing English on TV, teacher absenteeism and whether the teacher has a learning programme for the full year. Table A.4 in Appendix C reports the complete model statistics with the coefficients on these additional variables.

Male students performed worse than females and this effect was consistent in all the models, including the value-added models. Students who were older than the norm (10 years old) also performed worse in all the models. Students whose home language was English had a performance advantage over other students, which is understandable given that the tests were in English. A consistent pattern that emerged was that greater exposure to English through speaking and hearing English on the television was associated with higher achievement even after controlling for home language. Apart from the obvious language proficiency effect, the possibility exists that this result may be picking up a further effect of SES: those who have televisions at home and frequently speak English outside school may also be more affluent.

Student SES was itself significantly associated with literacy achievement in 2008, although a more substantial impact was attributable to the mean of school SES (and the square thereof). The positive coefficient on the squared term indicates that the effect of school SES on achievement is more pronounced at higher levels of school SES. As suggested earlier, the most important influence of student SES may therefore be through determining selection into schools of differential combined SES (and associated quality).

The pupil-teacher ratio was statistically significant in both the model for literacy and numeracy (Table 18), with small to moderate effect sizes. However, a more important factor than the ratio of students to teachers in a school is the actual number of students in a class at any point in time. Unfortunately the NSES did not capture class size in this way.

The main focus of this paper is on the identification of indicators of effective school management and teacher practice. In both the literacy and numeracy models a positive effect was obtained for schools in which no teachers were absent on the day of the survey. Although this once-off measurement may not accurately reflect teacher absenteeism over a longer period, it avoids the response bias and subjectivity often present when school principals are asked about the severity of teacher absenteeism. Moreover, the fact that this variable is significantly associated with achievement in these models indicates that it is probably capturing teacher absenteeism, and perhaps even school organisation and professional work ethic more generally, although with a certain degree of measurement error.

The coefficient on the dummy variable for schools having an inventory for Learning and Teaching Support Materials (LTSM) that is present and up-to-date was positive and significant in the model for literacy. This is a good indicator of how well resources are managed and used by schools. An index derived from a number of questions about problems with student

behaviour was negatively associated with literacy achievement. However, this variable was not consistently associated with achievement in most of the other models not reported, or in the model for numeracy achievement shown in Table 18. A positive effect was obtained for schools in which curriculum planning was reportedly done using a year schedule. School principals were read a list of ways in which curriculum planning might be undertaken and asked to identify which of those were practised in their schools. The options were not mutually exclusive. The result here may suggest that this variable is an indicator of the level and type of planning and organisation that distinguishes more effective schools from less effective ones.

Only one variable from the teacher instrument was associated with literacy achievement in Table 16. A positive and significant coefficient was obtained for the dummy variable indicating that a full-year learning programme was seen by the fieldworker. It would be unwise to conclude from this result that student achievement will improve by the size of the coefficient if teachers planned the learning programme for the full year. Rather, this variable should be regarded as a proxy for teacher organisation and preparation in general. It is noteworthy that many other teacher characteristics did not warrant inclusion in the model for literacy achievement. In particular, teacher knowledge as measured by the short comprehension test was not significantly associated with student achievement after controlling for all the other variables in the model. This may be simply a consequence of the short test not adequately differentiating between teachers with varying degrees of subject knowledge.

An alternative way of presenting the results of the models shown in Table 17 is to consider what effects various improvements in school and teacher characteristics could have on the national average of literacy achievement. The mean literacy score in 2008 for the sample of students that was included in the OLS literacy model was 26.57%. Table 17 shows the predicted changes to this sample mean associated with changing the entire sample of schools to have the positive value of the school and teacher characteristics that emerged as important in the OLS literacy model. The largest effect on the sample mean was predicted for changing schools in which some teachers were absent (on the day of the survey) to having no teachers absent. Of course, it is unrealistic always to maintain zero absenteeism. This calculation should rather be interpreted as the predicted change in the national average associated with improving teacher attendance in general (and improving whatever organisational characteristics for which this variable is acting as a proxy) in the "some teachers absent" group of schools to the level at which it is in the "zero teacher absenteeism" group. The combined effect of improving characteristics under the model parameters would be to raise the mean sample average by 3.29 percentage points. The combined effect of these characteristics is the more relevant figure as none of the

individual characteristics should be conceived of as inputs that will yield the exact outcomes predicted by the models, but rather as indicators of the underlying concepts of school and teacher effectiveness.

Table 17: *Estimated effects of change in characteristics on the literacy national average (Original sample mean = 26.57%)*

	Predicted new mean	Gain
Teacher absenteeism zero	27.84	1.27
LTSM Inventory good	27.36	0.79
Curriculum planned using year schedule	27.18	0.61
Full year learning programme	27.18	0.61
Combined effect of improved characteristics	29.85	3.29

The estimated effects of student characteristics in the model for numeracy achievement in 2008 (Table 18) were very similar to those in the literacy model, although the number of books at home was not associated with numeracy achievement. Otherwise, conditional upon all other factors in the models, those who reported having more than two siblings did worse than those in smaller households; those who read frequently at home on their own did better; and those who were exposed to English more often also did better. Interestingly, the coefficient on student SES was not significantly different from zero in the numeracy model. As it has been shown that the unconditional association between SES and achievement is very strong, this result would suggest that the effect of SES is largely bound up with the mean SES of schools to which children gain access and that little further effect of individual SES exists.

Table 18: OLS Regression model for numeracy 2008

Explanatory variables		
Student characteristics		
Student SES	0.26	(0.27)
Male	-1.13**	(0.35)
Young	-0.07	(0.72)
Old	-3.99***	(0.53)
Household size: large	-2.37***	(0.54)
Read 1 to 3 times a week	3.49***	(0.67)
Read more than 3 times	4.97***	(1.07)
Home language English	9.87***	(2.01)
Speak English 1-3 times	2.43***	(0.65)
Speak English 4+	2.01~	(1.05)
English on TV 1-3 times	0.66	(0.66)
English on TV 4+	4.50***	(0.69)
School characteristics		
Mean School SES	-16.89***	(3.38)
Mean School SES squared	4.88***	(0.78)
Pupil-teacher ratio	-0.38***	(0.11)
Media and Communication facilities index	2.45*	(1.02)
Assessment record keeping good	0.25	(1.88)
Assessment record keeping poor	-2.79	(2.16)
Assessment record keeping very poor	-4.87*	(2.41)
No timetable available	-4.87*	(2.43)
Teacher absenteeism zero	2.74*	(1.38)
Teacher characteristics		
Maths teacher test score: 100%	2.99~	(1.77)
Maths topics covered: 25 plus	4.69**	(1.54)
Constant	50.05***	(5.08)
R-squared statistic	0.4223	
N	11383	

~ p<0.10, * p<0.05, ** p<0.01, *** p<0.001
(Standard errors in parentheses)
Notes:

The regressions included dummy variables for the provinces but the coefficients on these are not reported in the table. Also, dummy variables controlling for non-response were included for the following characteristics: Household size, frequency of reading at home on one's own, frequency of speaking English at home, frequency of hearing English on TV, teacher absenteeism, teacher mathematics test result and the number of mathematics topics covered. Table A.5 in Appendix C reports the complete model statistics with the coefficients on these additional variables.

A noteworthy school-level variable that did not emerge as significant in the numeracy models was the availability of sufficient grade 4 mathematics textbooks. However, a composite index for the availability and quality of various media and communication facilities, such as projectors and copying facilities, was significantly associated with numeracy achievement. The evidence for this effect was, however, mixed and was sensitive to model specification.

A general observation pertaining to Table 18 as well as the rest of the models in this section is that school resource variables were less consistently and importantly related to achievement than were variables that can be thought of as indicators of effective school management. For example, in 15 schools no timetable was available for the fieldworker to observe, and students in these schools indeed performed worse controlling for other factors in the numeracy model. A school timetable is supposed to be a crucial and ever-present feature in the daily running of a school. It is therefore reasonable to regard not having an easily accessible timetable as a flag for a dysfunctional school. The dummy variable for zero teacher absenteeism on the day of the survey also came through strongly in the numeracy model. Although the LTSM inventory dummy did not warrant inclusion in this model, another indicator of school organisation did emerge as significant: the quality of assessment records. Students in schools where the quality of assessment records was very poor did worse than those in schools where these were both present and up-to-date.

In contrast to the literacy model, a significant effect of teacher knowledge was obtained in the model for numeracy. A student achievement advantage of slightly less than three percentage points was associated with a score on the teacher mathematics test of 100%. No significant differences in achievement were associated with variations in teacher knowledge below 100%. Sound teacher knowledge therefore seems necessary before any noticeable impact on student achievement accrues. There was also a reasonably large, positive and significant effect associated with having covered more than 25 curriculum topics, as identified in student workbooks. This variable was fairly consistently associated with numeracy achievement across other model specifications estimated but not presented here. The number of topics can therefore be considered a good indicator of curriculum coverage and, more fundamentally, of the amount of work that is being undertaken in classes. This aspect of classroom practice clearly has an important impact on student achievement.

The predicted effects of improvements in the school and teacher characteristics on the national average were greater for numeracy than for literacy. As Table 19 shows, the national average could be expected to improve most substantially in response to raising teacher knowledge and

curriculum coverage across the system. The combined effect of a universal attainment of positive values on the indicators of school and teacher quality in the model would be to raise the national average from 34.21% to 42.29%.

To summarise, the numeracy models provide evidence that good assessment practices, teacher commitment and planning, teacher knowledge and curriculum coverage vary substantially across South African schools and are strongly linked to educational achievement.

Table 19: Estimated effects of change in characteristics on the numeracy national average (Original sample mean = 34.21%)

|	Predicted new mean	Gain
Assessment record keeping	35.08	0.87
No timetable available	34.45	0.24
Teacher absenteeism zero	36.01	1.80
Maths teacher test score: 100%	36.38	2.17
Maths topics covered: 25 plus	37.20	3.00
Combined effect of improved characteristics	42.29	8.08

The remainder of this section reports on two modelling strategies that used gain scores as the outcome of interest and the most recent version of the data available (all three waves). Using gain scores has the advantage over single cross-sections of achievement in that innate ability and other factors that may have influenced cognitive development prior to the first test (in this case grade 3) can to some extent be controlled for, and the amount of learning during a given time interval can hopefully be explained by school practices during that time. However, gain scores can also suffer certain weaknesses, especially when the pre- and post-test were not far apart in time. Consider the meaningfulness of the data if the pre- and post-test were one day apart. It is highly unlikely that all students would achieve exactly the same score on both occasions, as there is an element of randomness involved every time one takes a test. The gain scores in this case would entirely represent random noise and would contain no signal of improvement. This signal-to-noise ratio is thus a weakness inherent to gain scores, and more so the closer together the tests are.

A related statistical phenomenon flowing from the random element when taking a test is called regression to the mean. This phenomenon occurs in numerous types of data whenever there is a tendency for especially large or small measurements (far from the sample mean) to be

followed by repeat measurements of the same units that are closer to the sample mean (Barnett, Van der Pols and Dobson, 2005: 15). In this case there is a tendency for especially low or high student scores in 2007 to be followed by scores of the same individuals that are closer to the mean in 2008. This happens because with each individual test score there is a degree of random error around a hypothetical true mean, which would accurately reflect the individual's true ability. If an individual were tested a sufficient number of times this true mean would become evident. A large proportion of the NSES tests were multiple choice questions. This means that the measurement error at the bottom end of test scores would be considerable. It is therefore likely that a very low score represents an underestimation of that individual's hypothetical true mean, and that a repeat test would yield a higher score even if no learning has taken place in the interim. This sort of regression to the mean would result in overestimation of gain scores at the low end of test scores and underestimation of gain scores amongst the high achieving students.

Figure 16 shows that to some extent regression to the mean was present in the literacy gain scores in the NSES. The figure shows local polynomial smoothing lines of literacy scores in each year against the pooled 1-year gains (i.e. 8383 gains from 2007 to 2008 and 8383 gains from 2008 to 2009). The figure demonstrates that larger gains were associated with low initial performance in 2007. On the other hand high scores in 2009 also were associated with high gains. The gains were highest at the mid-range of the average literacy achievement over the three years, indicating that the high gains associated with low baseline performance and with high post-test performance may be a reflection of regression to the mean. A peculiarity of the data is that a large number of students scored zero in the 2007 literacy test, which may be exacerbating regression to the mean. Various restrictions were applied to the forthcoming models to attempt to iron out the bias introduced by regression to the mean: Large negative gains (below -10 percentage points) were excluded as these are surely reflecting noise rather than a signal of learning; schools in which more than 20% of students scored zero in 2007 were excluded; and only historically black school were included in the models. This last restriction was motivated both by the relevance of exploring what management and teaching characteristics distinguish quality within the part of the school system that is in most urgent need of improvement, and by the statistical concern that historically white schools performed at initially high levels with little room for improvement resulting in low gains being associated with generally positive school characteristics thus creating perverse results.

Figure16: Local polynomial smoothing lines of literacy scores against pooled 1-year gains
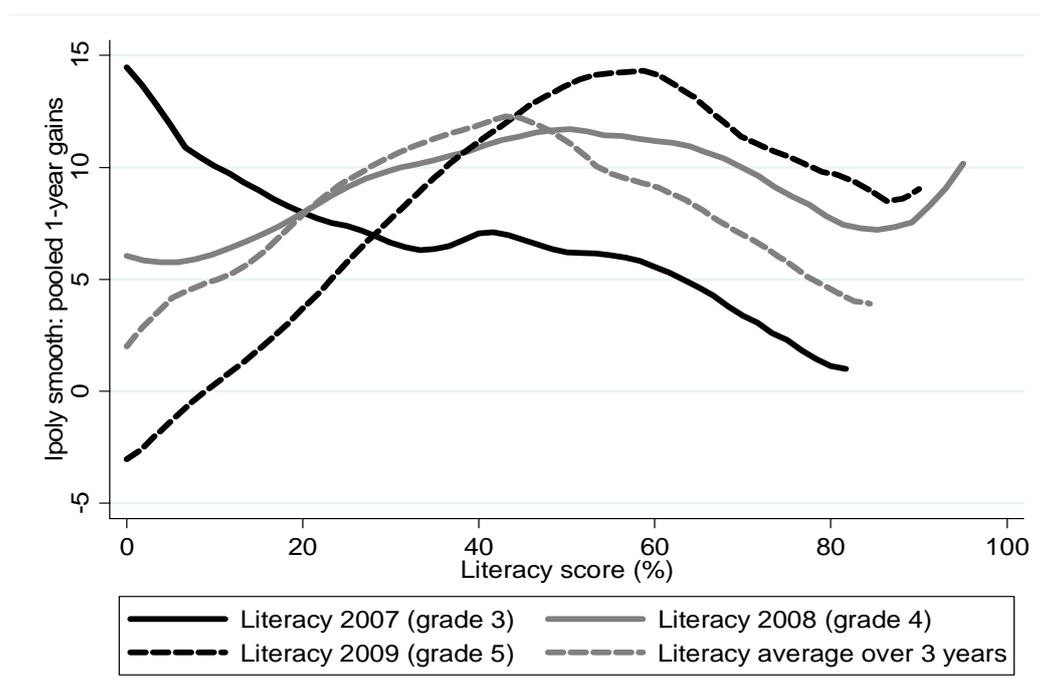
The first modelling strategy employed was a two-step school fixed effects model. The first step explained the pooled 1-year literacy gains on the basis of student characteristics as well as school fixed effects (through entering each school as a separate dummy variable into the regression). Due to the pooling of the gain scores each school was entered twice into the regression, once with reference to the gain from 2007 to 2008 and once with reference to the gain from 2008 to 2009. The results are reported in Table 20 below. Student SES and its square were not significantly associated with gains after the school fixed effects were included in the model, again suggesting that once allocation into a school has taken place there is little further effect of individual SES. The age, gender and exposure to English dummies warranted inclusion in the Step 1 model.

Table 20: Pooled literacy gains step 1 model: Student characteristics and school fixed effects

Explanatory variables	Full SA sample	
Student SES	0.56	(0.38)
Student SES squared	-0.12	(0.10)
Male	-0.95***	(0.15)
Young	-0.18	(0.50)
Old	-1.82***	(0.20)
Speak English 1-3 times	0.46**	(0.17)
Speak English 4+	0.59*	(0.27)
English on TV 4+	1.05***	(0.18)
Constant	11.12***	
R-squared	0.2519	
N	15886	

~ p<0.10, * p<0.05, ** p<0.01, *** p<0.001
Standard errors in parentheses

Note: Included in the model but not reported in the table: The school fixed effects coefficients and dummy variables for missing information regarding student gender, age, frequency of speaking English at home, frequency of hearing English on TV. Ordinary Least Squares (OLS) survey regression was used to account for the sample design.

The next step was to use the coefficients obtained on each school fixed effect dummy in the step 1 model as the outcome variable for a second school level model. The intuition here is that after controlling for student characteristics, including SES, certain schools performed better than others and that this needs to be explained by particular school and teacher characteristics. Given that the SES of students is controlled for in the step 1 model, the size of the school fixed effects coefficients can be regarded as a measure of the efficiency with which schools produce student achievement. The results of the step 2 model are reported in Column [A] of Table 21. Note that there were 195 historically black schools, in which less than 20% of the students achieved zero in 2007, and each of these schools had two coefficients to be explained due to the pooling of the gain scores in the step 1 model, thus providing 390 degrees of freedom. A time dummy was included in the model to account for overall gains being higher during one time interval than the other, but this dummy was not statistically significant.

Interestingly, the mean SES of schools was not significantly associated with the school fixed effect coefficients (the outcome variable of the regression). A small effect was obtained on an index capturing the presence and functionality of the following school facilities: running water, electricity, storerooms, toilets, administrative offices, box libraries and science kits. However, several indicators of school organisation and teacher practice did predict this outcome measure of school efficiency. A particularly revealing variable was a dummy taking a value of one if the

principal was absent on the day of the survey in either 2008 or 2009. This was negatively associated with the school fixed effect coefficient (and therefore with literacy gains) and the coefficient was fairly large and statistically significant. It is implausible that the principal's absenteeism on the day of the survey directly lowered student achievement. However, those with experience as fieldworkers will tell you that schools where the principal is either away for whatever reason or not well prepared for the survey are typically poorly run schools in general. This result exemplifies why many of these indicators of effective management should be interpreted as exactly that: indicators of good management rather than factors directly impacting on student achievement. Similarly, a dummy variable for teacher punctuality being good had a positive effect. This variable was derived from several questions in the principal questionnaires of 2007, 2008 and 2009. If in all three years the principal maintained that teacher punctuality was not a serious problem in the school this dummy variable took a value of one.

Three indicators of the extent to which work was undertaken during the course of the year proved to be significantly associated with the school fixed effects coefficients. Schools in which more than two English mark records could be observed had higher gains under this model. A large and statistically significant negative impact of having undertaken no paragraph length writing, according to student workbooks, was also obtained. Similarly, a positive effect was found for schools in which more than 27 exercises were counted in the "best" student's English workbook. This adds to the evidence presented earlier to suggest that curriculum coverage and simply ensuring that student's put skills into practice through exercises are important aspects of effective teaching.

Table 21: School level models: Pooled gain step 2 and 2-year literacy gains

Explanatory variables	[A] Pooled gains step 2		[B] 2-year literacy gains	
Mean School SES	0.39	(0.35)	1.37*	(0.63)
Facilities index (2008)	0.14~	(0.08)	0.27~	(0.15)
Monitoring through class visits			2.16*	(0.90)
No timetable available (2008)			-2.72	(1.93)
Principal absent	-1.67**	(0.65)	-4.03***	(1.13)
Teacher punctuality good	0.94~	(0.53)	3.03***	(0.91)
More than 2 English mark records	1.44*	(0.64)	3.76***	(1.13)
Paragraph writing: none	-1.72**	(0.57)	-4.12***	(1.01)
Literacy exercises: more than 27	1.34*	(0.55)	2.35*	(0.96)
Years teaching: 4 to 9			1.03	(1.87)
Years teaching: 10 to 19			2.64	(1.61)
Years teaching: 20 plus			3.83*	(1.67)
Time dummy (1st year)	0.40	(0.51)		
Constant	-5.33***		6.10**	
R-squared	0.1214		0.3976	
N	390		195	

~ p<0.10, * p<0.05, ** p<0.01, *** p<0.001
Standard errors in parentheses

Note: Both of these models were restricted to schools that were both historically black schools and in which less than 20% of the students had achieved a score of zero for literacy in 2007. Included in the models but not reported in the table: dummy variables for missing information regarding the frequency of paragraph writing in learner workbooks, whether internal monitoring occurs through class visits, the presence of a timetable, teacher experience, and the number of literacy exercises observed in learner workbooks.

Column [B] of Table 21 reports the results of a second strategy for modelling literacy gain scores. Here the school mean gain scores over two years (from grade 3 to grade 5) is the outcome variable. Again the sample was restricted to historically black schools. All those factors which came through in the pooled gains model also were found to predict the school mean 2-year gain. This is reassuring in terms of the reliability of the overall results. Several other variables that did not warrant inclusion in the pooled gains model were significantly associated with the school mean 2-year gain. Gains were higher in schools where internal monitoring occurred through classroom visits. The dummy for not being able to produce a school timetable did not come through significantly in the final model specification but only narrowly so, and probably because a relatively small number of schools fitted this category. The effect size is nonetheless substantial. Therefore this variable was retained in the model drawing attention to the organisation of time within schools. Interestingly, a positive effect was obtained for teachers with more than 20 years of experience. This result may not have immediately clear

implications for policy or regarding what constitutes effective teaching but it is nevertheless noteworthy.

The results obtained in Table 21 are significant as education production function analyses often struggle to identify management and teacher factors that are significantly associated with student achievement, especially within the historically black part of the system.

7. CONCLUSION

This paper has highlighted several indicators of effective school management and teacher practice that are associated with student achievement, even within the large historically disadvantaged and currently underperforming section of the school system. This constitutes an advance on earlier analyses which speculated about the importance of management efficiency but were limited in their ability to identify specific elements thereof.

Teacher knowledge was not consistently associated with achievement, although the evidence was stronger in the case of numeracy than for literacy. Weak evidence was found that school resources such as pupil-teacher ratios and school facilities are associated with student achievement. As other studies have argued, more important than the mere presence of resources is how well they are managed. The results pertaining to variables that can be considered indicators of management effectiveness were clearer. An organised learning environment signified by curriculum planning for the full year, a functional timetable, good-quality inventories for LTSM, low teacher absenteeism and up-to-date assessment records were all strongly linked to better student achievement, even after accounting for differences in previous student performance and SES.

The evidence suggests that effective schools offer thorough coverage of the curriculum. So there is scope for policies aimed at teachers' professional development, firstly to ensure that they are technically able to teach all of the required elements, and secondly to provide them with the necessary time management skills so that they can deliver within the set academic timeframe. Policy should also ensure that LTSM such as textbooks and workbooks are explicitly designed to facilitate the extensive coverage of curriculum and exercises, making this easier for both teachers and students to implement.

The indicators of good management identified in this research should not be interpreted as more than exactly that: indicators that point to the characteristics typically exhibited by good managers, rather than levers to be manipulated by policy to achieve improved student outcomes. Command and control measures aimed at forcing teachers to follow best practices may well empty such practices of their value through introducing the perverse incentive to window-dress those practices at the expense of focusing on the central task of teaching. A better and indeed more ambitious route for policy would be to explore ways to attract, train and support better principals, and to replace those at the head of dysfunctional schools.

LIST OF REFERENCES

Anderson, K.G., Case, A. and D. Lam. 2001. "Causes and Consequences of Schooling Outcomes in South Africa: Evidence from Survey Data." *Social Dynamics.* 27(1): 1-23.

Barnett, A.G., Van der Pols, J.C. and A.J. Dobson. 2005. "Regression to the mean: what it is and how to deal with it." *International Journal of Epidemiology.* 34(1): 215-220.

Burger, R. and R. Jafta. 2006. "Returns to race: Labour market discrimination in post-apartheid South Africa." Stellenbosch Economic Working Papers No. 04/06.

Crouch, L. and T. Magoboane. 1998. "When the residuals matter more than the coefficients: An educational perspective." *Studies in Economics and Econometrics.* 22(2): 1-14.

Filmer, D. And Pritchett, L. 2001. "Estimating Wealth Effects without Expenditure Data – or Tears: An Application to Educational Enrollments in States of India." World Bank Working Paper.

Fleisch, B. 2008. *Primary Education in Crisis: Why South African schoolchildren underachieve in reading and mathematics.* Juta.

Glewwe, P. 2002. "Schools and skills in developing countries." *Journal of Economic Literature*. XL: 436-482.

Gustafsson, M. 2007. "Using the Hierarchical Linear Model to Understand School Production in South Africa." *South African Journal of Economics.* 75(1): 84-98.

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

JET Education Services. *The National School Effectiveness Study: Data for Waves 1 to 3*. November 2010.

Soudien, C. 2004. " 'Constituting the class': an analysis of the process of 'integration' in South African schools." In *Changing Class: Education and Social Change in Post-Apartheid South Africa*. Edited by L. Chisholm. Cape Town: HSRC Press.

Spaull, N. 2011. "A preliminary analysis of SACMEQ III South Africa." Stellenbosch Economic Working Papers No. 11/2011.

Taylor, N., Muller, J. and P. Vinjevold. 2003. *Getting Schools Working: Research and Systematic School Reform in South Africa.* Cape Town: Pearson Education South Africa.

Taylor, S.G.S and D. Yu. 2009. "The importance of socio-economic status in determining educational achievement in South Africa." Stellenbosch Economic Working Papers No. 01/09.

Van der Berg, S. 2006. "The targeting of public spending on school education, 1995 and 2000." Perspectives in Education. 24(2): 49-63.

--. 2007. "Apartheid's enduring legacy: Inequalities in education." *Journal of African Economies.* 16(5): 849-880.

––. 2008. "How effective are poor schools? Poverty and educational outcomes in South Africa." *Studies in Educational Evaluation.* 34: 145-154.

Van der Berg, S. and C. Burger. 2011. "Modelling cognitive skills, ability and school quality to explain labour market earnings differentials." Stellenbosch Economic Working Papers No. 08/2011.

Van der Berg, S. and R. Burger. 2002. "Education and socio-economic differentials: A study of school performance in the Western Cape." *South African Journal of Economics.* 71(3): 496-522.

APPENDICES

Appendix A: Box plots of literacy and numeracy scores in the NSES by home language

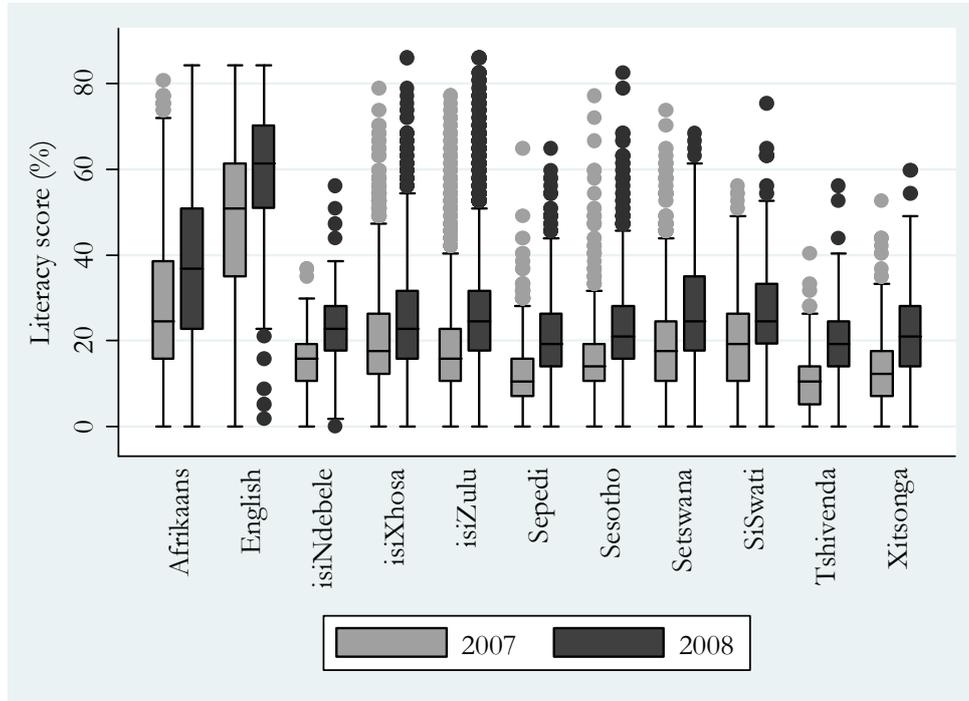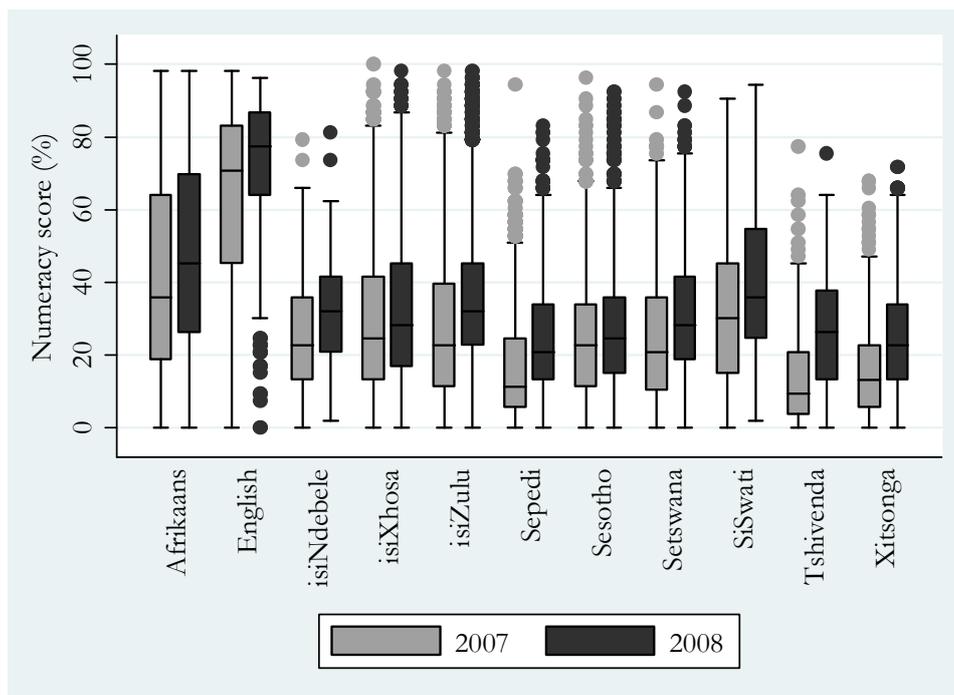*Figure A.1: Literacy scores by home language*

Figure A.2: Numeracy scores by home language

Appendix B: Description of variables used in multivariate analysis of NSES data

Table A.1: Student level variables (Student questionnaires of 2007, 2008 and 2009)

Variable name	Description
Student SES	Z-score index of socio-economic status: Min = 0, std dev = 1
Male	Dummy variable: gender is male; reference category is female
Young	Dummy variable: Younger than 10 years
Age 10	Dummy variable: Expected age at grade 4: 10 years
Old	Dummy variable: Older than 10 years
Household size: small	Dummy variable: 2 siblings or fewer
Household size: large	Dummy variable: more than 2 siblings
Read never	Dummy variable: Student never reads at home on his/her own
Read 1 to 3 times a week	Dummy variable: Student reads at home on his/her own 1 to 3 times a week
Read more than 3 times	Dummy variable: Student reads at home on his/her own more than 3 times a week
Books at home: Zero	Dummy variable: No books at student's home
Books at home: 1 to 10	Dummy variable: 1 to 10 books at student's home
Books at home > 10	Dummy variable: More than 10 books at student's home
Home language English	Dummy variable: Student's home language is English; reference category is any other language
Speak English 0	Dummy variable: Student never speaks English at home
Speak English 1-3 times	Dummy variable: Student speaks English at home 1 to 3 times a week
Speak English 4+	Dummy variable: Student speaks English at home more than 3 times a week
English on TV 0	Dummy variable: Student never hears English on TV
English on TV 1-3 times	Dummy variable: Student hears English on TV 1 to 3 times a week
English on TV 4+	Dummy variable: Student hears English on TV more than 3 times a week

Table A.2: School level variables (Principal questionnaires of 2007, 2008 and 2009)

Variable name	Description
Mean School SES	Mean of student SES within each school: Min = 0, std dev = 1
Pupil-teacher ratio	Number of enrolled students divided by the number of teachers at each school
Media and Communication facilities index	Z-scored index for the presence and functionality of the following: phone, fax, internet/email, copying facility, computer for administration, computer for staff, computers for students, TV/video and overhead projector
Teacher absenteeism zero	Dummy variable: No teachers absent on the day of the survey; Reference category: some teachers absent
LTSM Inventory good	Dummy variable: LTSM inventory complete and up to date
LTSM Inventory average	Dummy variable: LTSM inventory present but incomplete and not up to date

LTSM Inventory poor	Dummy variable: LTSM inventory not able to be seen
Problems with students index	Z-scored index combining several evaluations of the extent of problems with student discipline and work ethic in the school (mean = 0, std. dev = 1)
Curriculum planned using year schedule	Dummy variable: Principal reported that curriculum planning occurs using a year schedule. Reference category: Curriculum planning does not involve a year schedule.
No timetable available	Dummy variable: No school timetable could be shown to fieldworker. Reference category: Fieldworker saw a timetable.
Assessment record keeping very good	Dummy variable derived from a summative index combining a number of questions regarding the presence and completeness of assessment records
Assessment record keeping good	Dummy variable derived from a summative index combining a number of questions regarding the presence and completeness of assessment records
Assessment record keeping poor	Dummy variable derived from a summative index combining a number of questions regarding the presence and completeness of assessment records
Assessment record keeping very poor	Dummy variable derived from a summative index combining a number of questions regarding the presence and completeness of assessment records
Facilities index	Summative index capturing the presence and functionality of the following school facilities: running water, electricity, storerooms, toilets, administrative offices, box libraries and science kits.
Monitoring through class visits	Dummy variable: Internal curriculum monitoring takes place through class visits.
Principal absent	Dummy variable: The principal was absent in either 2008 or 2009 on the day of the survey.
Teacher punctuality good	Dummy variable: in all three years the principal maintained that teacher punctuality was not a serious problem in the school
More than 2 English mark records	Dummy variable: More than 2 mark records observed during the Principal instrument document review.

Table A.3: Teacher level variables (Teacher questionnaire of 2008 and 2009)

Variable name	Description
Full year learning programme	Dummy variable: Fieldworker was shown a learning programme for the full year
English teacher test score: 1 or 2[11]	Dummy variable: Teacher scored 1 or 2 on the comprehension test
English teacher test score: 3	Dummy variable: Teacher scored 3 on the comprehension test
English teacher test score: 4 or 5	Dummy variable: Teacher scored 4 or 5 on the comprehension test
English teacher test score: 6 or 7	Dummy variable: Teacher scored 6 or 7 on the comprehension test
Maths teacher test score: 1 out of 5	Dummy variable: Teacher scored 1 out of 5 on the maths test
Maths teacher test score: 2 out of 5	Dummy variable: Teacher scored 2 out of 5 on the maths test

---

[11] Scores of zero were relatively few and were associated with strong student performance, strangely. It is possible that many these were instances where teachers refused to take the test or for some reason other than extremely poor knowledge recorded a score of zero. Therefore, these cases were grouped together with cases of non-response.

Maths teacher test score: 3 out of 5	Dummy variable: Teacher scored 3 out of 5 on the maths test
Maths teacher test score: 4 out of 5	Dummy variable: Teacher scored 4 out of 5 on the maths test
Maths teacher test score: 100%	Dummy variable: Teacher scored 5 out of 5 on the maths test
Time spent teaching: less than 10 hours	Dummy variable: Maths teacher reported spending less than 10 hours per week on actual teaching
Time spent teaching: 10 to 18 hours	Dummy variable: Maths teacher reported spending between 10 and 18 hours per week on actual teaching
Time spent teaching: 19 to 26 hours	Dummy variable: Maths teacher reported spending between 19 and 26 hours per week on actual teaching
Time spent teaching: more than 26 hours	Dummy variable: Maths teacher reported spending more than 26 hours per week on actual teaching
Number of complex maths exercises: fewer than 18	Dummy variable: Fewer than 18 complex maths exercises found in student workbooks
Number of complex maths exercises: 18 plus	Dummy variable: More than 18 complex maths exercises found in student workbooks
Maths topics covered: < 10	Dummy variable: Fewer than 10 maths topics covered according to fieldworker review of student workbooks
Maths topics covered: 10 to 25	Dummy variable: Between 10 and 25 maths topics covered according to fieldworker review of student workbooks
Maths topics covered: 25 plus	Dummy variable: More than 25 maths topics covered according to fieldworker review of student workbooks
Years teaching: 0 to 3	Dummy variable: teacher experience: 0 to 3 years
Years teaching: 4 to 9	Dummy variable: teacher experience: 4 to 9 years
Years teaching: 10 to 19	Dummy variable: teacher experience: 10 to 19 years
Years teaching: 20 plus	Dummy variable: teacher experience: more than 20 years
Paragraph writing: none	Dummy variable: No evidence could be found in student workbooks of written exercises comprising paragraphs.
Literacy exercises: more than 27	Dummy variable: More than 27 exercises were counted in the "best" student's English workbook.

Appendix C: Full multivariate models not shown in main text

Table A.4: OLS Regression model for literacy 2008

Explanatory variables		
Student characteristics		
Student SES	0.39*	(0.18)
Male	-2.48***	(0.26)
Young	-0.40	(0.46)
Old	-2.84***	(0.33)
Age unspecified	-7.84***	(1.46)
Household size: large	-1.89***	(0.37)
Household size: unspecified	-1.15	(1.13)
Read 1 to 3 times a week	1.37**	(0.44)
Read more than 3 times	2.39***	(0.62)
Reading unspecified	-5.31*	(2.07)
Books at home: 1 to 10	0.60	(0.39)
Books at home > 10	1.17*	(0.48)
Home language English	8.42***	(1.52)
Speak English 1-3 times	1.75***	(0.38)
Speak English 4+	1.86**	(0.68)
Speak English unspecified	-5.74***	(1.56)
English on TV 1-3 times	0.85*	(0.39)
English on TV 4+	3.35***	(0.44)
English on TV unspecified	-2.70~	(1.59)
Eastern Cape	0.52	(2.28)
Northern Cape	-3.17	(2.43)
Free State	-3.71*	(1.77)
KwaZulu-Natal	1.10	(2.24)
North West Province	-1.44	(2.61)
Mpumalanga	-3.43~	(2.03)
Limpopo	-4.69*	(2.05)
School characteristics		
Mean School SES	-9.13***	(1.77)
Mean School SES squared	3.35***	(0.45)
Pupil-teacher ratio	-0.18**	(0.07)
Teacher absenteeism zero	1.93*	(0.81)
Teacher absenteeism unsp.	0.28	(2.02)
LTSM Inventory good	1.66*	(0.80)
Problems with students index	-0.96*	(0.43)
Curriculum planned using year schedule	1.46~	(0.81)
Teacher characteristics		
Full year learning programme	1.55~	(0.87)
Learning programme unsp.	1.60	(1.09)
Constant	29.69***	(3.45)
R-squared statistic	0.4591	
N	10860	

~ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

(Standard errors in parentheses)

Table A.5: OLS Regression models for numeracy 2008

Explanatory variables		
Student characteristics		
Student SES	0.26	(0.27)
Male	-1.13**	(0.35)
Young	-0.07	(0.72)
Old	-3.99***	(0.53)
Age unspecified	-9.70***	(2.75)
Household size: large	-2.37***	(0.54)
Household size: unsp.	-0.87	(1.61)
Read 1 to 3 times a week	3.49***	(0.67)
Read more than 3 times	4.97***	(1.07)
Reading unspecified	-6.17*	(2.62)
Home language English	9.87***	(2.01)
Speak English 1-3 times	2.43***	(0.65)
Speak English 4+	2.01~	(1.05)
Speak English unspecified	-8.98***	(1.87)
English on TV 1-3 times	0.66	(0.66)
English on TV 4+	4.50***	(0.69)
English on TV unspecified	-11.75***	(2.16)
Eastern Cape	2.70	(3.11)
Northern Cape	-1.24	(3.65)
Free State	-2.83	(2.66)
KwaZulu-Natal	5.90*	(2.96)
North West Province	-1.14	(3.31)
Mpumalanga	-2.32	(3.12)
Limpopo	-4.75	(2.97)
School characteristics		
Mean School SES	-16.89***	(3.38)
Mean School SES squared	4.88***	(0.78)
Pupil-teacher ratio	-0.38***	(0.11)
Media and Communication facilities index	2.45*	(1.02)
Assessment record keeping good	0.25	(1.88)
Assessment record keeping poor	-2.79	(2.16)
Assessment record keeping very poor	-4.87*	(2.41)
No timetable available	-4.87*	(2.43)
Teacher absenteeism zero	2.74*	(1.38)
Teacher absenteeism unsp.	6.51**	(2.38)
Curriculum planned using year schedule		
Teacher characteristics		
Maths teacher test score: 100%	2.99~	(1.77)
Maths teacher test unsp.	-2.59	(2.07)
Maths topics covered: 25 plus	4.69**	(1.54)
Maths topics covered: unsp.	6.99*	(3.36)
Constant	50.05***	(5.08)
R-squared statistic	0.4223	
N	11383	

~ $p<0.10$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$

(Standard errors in parentheses)