Equity & Efficiency in South African Primary Schools

A Preliminary Analysis of SACMEQ III South Africa

By Nicholas Spaull

Supervisor: Professor Servaas Van der Berg

University of Stellenbosch

Department of Economics

March 2012

Thesis presented in partial fulfilment of the requirements for the degree of Masters of Commerce in the Faculty of Economic and Management Sciences at Stellenbosch University

Stellenbosch University http://scholar.sun.ac.za

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: 10 January 2012

Copyright © 2011 Stellenbosch University

All rights reserved

Stellenbosch University http://scholar.sun.ac.za

ABSTRACT

Equity & Efficiency in South African Primary Schools:

A Preliminary Analysis of SACMEQ III South Africa

By Nicholas Spaull

The many and varied links between student socioeconomic status and educational outcomes have been well documented in the South African economics of education literature. The strong legacy of apartheid and the consequent correlation between education and wealth have meant that, generally speaking, poorer learners perform worse academically. The links between affluence and educational quality in South Africa can partially explain this outcome since the poor receive a far inferior quality of education when compared to their wealthier counterparts. This disadvantages them in the labour-market and entrenches their poverty. This thesis uses the recent Southern and Eastern African Consortium for Monitoring Educational Quality (SACMEQ III) dataset for South Africa to answer three important questions: (1) Is South African primary education efficient? (2) Is South African primary education equitable? and (3) What are the main factors that have a significant effect on student mathematics and reading performance in Grade 6. The thesis shows that a high proportion of the country's learners are functionally illiterate and functionally innumerate. The research confirms previous findings that socio-economic status, and particularly school socioeconomic status, is important when understanding student success or failure. Other factors which significantly affect student performance are homework frequency, grade repetition, and the availability of reading textbooks. In contrast, teacher-subject knowledge was found to have only a modest impact on Grade 6 performance. Policy interventions associated with the findings are also highlighted. The study concludes that South Africa is still a tale of two school sub-systems: one which is wealthy, functional and able to educate students, while the other is poor, dysfunctional, and unable to equip students with the necessary numeracy and literacy skills they should be acquiring in primary school. Finally, the thesis suggests that there are some options available to policy-makers which are expected to have a positive effect on learner performance.

Stellenbosch University http://scholar.sun.ac.za

UITTREKSEL

Ekonomiese Regverdigheid en Doeltreffendheid in Suid-Afrikaanse Primêre Skole:

'n Voorlopige analise van SACMEQ III vir Suid-Afrika

Deur Nicholas Spaull

Die vele en diverse verbande tussen studente se sosio-ekonomiese status en onderwysuitkomste is goed gedokumenteer in die Suid-Afrikaanse literatuur oor die ekonomie van onderwys. Die sterk nalatenskap van apartheid en die gevolglike korrelasie tussen onderwys en rykdom beteken dat armer leerlinge in die algemeen akademies swakker vaar. Die verband tussen welvaart en onderwysgehalte in Suid-Afrika kan hierdie uitkoms gedeeltelik verklaar, omdat arm mense 'n veel swakker gehalte van onderwys ontvang as rykes. Dit plaas hulle in 'n swakker posisie in die arbeidsmark en bevestig daarmee hulle armoede. Die tesis gebruik die onlangse SACMEQ III datastel vir Suid-Afrika (SACMEQ is die akroniem vir die *Southern and Eastern African Consortium for Monitoring Educational Quality*) om drie belangrike vrae te beantwoord: (1) Is Suid-Afrikaanse primêre skole doeltreffend? (2) Is Suid-Afrikaanse primêre onderwys regverdig verdeel? (3) Wat is die belangrikste faktore wat studente se wiskunde en leesvermoë in Graad 6 beduidend beïnvloed? Die tesis toon dat 'n groot proporsie van die land se leerlinge funksioneel ongeletterd en ongesyferd is. Die navorsing bevestig vorige bevindinge dat sosio-ekonomiese status, en veral die sosio-ekonomiese status van die skoolgemeenskap, 'n belangrike bepaler van studente se sukses is. Ander faktore wat studente se prestasie beduidend beïnvloed is hoe gereeld hulle huiswerk doen, of hulle die graad herhaal, en die beskikbaarheid van handboeke. In teenstelling daarmee is bevind dat onderwysers se vakkennis net 'n beskeie impak op Graad 6 prestasie het. Daar is ook klem op beleidsingrypings wat uit die bevindinge spruit. Die studie kom tot die gevolgtrekking dat Suid-Afrikaanse onderwys steeds die storie van twee sub-stelsels is: een wat ryk is, goed funksioneer en in staat is om studente 'n goeie opvoeding te bied, terwyl skole in die ander deel van die stelsel arm is, wanfunksioneel, en die vermoë ontbreek om studente toe te rus met die syfer- en leesvaardighede wat skole hulle behoort te bied. Ten slotte identifiseer die tesis opsies vir beleidmakers wat leerlinge se prestasie sou kon verbeter.

Stellenbosch University http://scholar.sun.ac.za

Acknowledgements

In many ways the completion of this thesis is the first of hopefully many milestones in a career of research, teaching and policy-analysis. Before I came to Stellenbosch I knew very little about education, or socioeconomic policy, and yet as I look to the future I cannot imagine myself doing something outside of the broad field of education. This is largely thanks to Prof van der Berg and his team of researcher-colleagues.

I owe an enormous debt of gratitude to Servaas van der Berg - I could not have asked for a better supervisor or mentor. He has taught me, both in the classroom and out, almost everything I know about education, social policy and research more generally. He is approachable, discerning, generous, good-natured, wise, patient and kind. I also want to thank all the researchers and lecturers in the Economics Department who made me feel welcome, and helped the newbie learn all the things which I now find myself teaching others. To the research group (and particularly my office-mates Marisa, Hendrik and Mia), thank you for all the laughs, presentations, arguments, brown-bag lunches and coffee-breaks. This wonderful group of people is further testament to the collegial and collaborative culture Prof van der Berg has created.

To my family and friends, thank you for being there for me, for giving me perspective, for loving me, for travelling with me, and for putting up with me – I love you all. Thank you Dad for proof-reading my final draft, and for all the years of linguistic punctiliousness.

Lastly, but most importantly, I am grateful to God for the experiences and opportunities I have had in my last two years at Stellenbosch. It is a rare privilege to be able to earn a living from work which you love doing, and even more so when that work is meaningful and contributes in some small way to the betterment of society.

Stellenbosch University http://scholar.sun.ac.za

Table of Contents

List of Figures 8

List of Tables 10

Chapter 1 Introduction 11

1.1 SACMEQ 15

Chapter 2 Literature Review 21

2.1 Theoretical links between Economics and Education 21

2.2 Understanding Learner Performance in South Africa 25

2.3 The Level of Learner Performance in South Africa 34

Chapter 3 Preliminary Analysis 40

3.1 **Learners** 41

3.1.1 Learner Performance 41

3.1.2 Learner Competency Levels 52

3.1.3 Selected Learner Characteristics 61

3.2 **Schools** 74

3.2.1 School Demographics 74

3.2.2 School Resources 74

3.3 **Teachers** 80

3.3.1 Teacher Absenteeism 80

3.3.2 Teacher Content Knowledge 81

3.3.3 Teacher Experience and Teacher Training 85

3.4 **Placing South Africa in Regional Context** 86

Chapter 4 Modelling Learner Performance 95

4.1 Sample Variations 97

4.2 Model Fit 100

4.3 Results from Multivariate Analysis 101

4.3.1 Individual Characteristics 102

4.3.2 Home-background Characteristics 104

4.3.3 School Characteristics 111

Stellenbosch University http://scholar.sun.ac.za

Chapter 5 Conclusion and Recommendations 125

5.1 Conclusion 125

5.2 Policy Recommendations 127

5.3 Future Research 129

Reference List 132

Appendices 139

Appendix A 139

Appendix B 142

Appendix C 149

Appendix D 151

Appendix E 163

Appendix F 180

Appendix G 182

Stellenbosch University http://scholar.sun.ac.za

List of Figures

Figure 1: Thesis Structure 14
Figure 2: Kernel Density of Student Reading Score by School SES Quartile 43
Figure 3: Kernel Density of Student Reading Score by School SES Quintile 43
Figure 4: Kernel Density of Student Reading Score by School SES Tercile 43
Figure 5: Kernel Density of Student Reading Score by School SES 43
Figure 6: Kernel Density of Student Mathematics Score by School SES Quartile 44
Figure 7: Kernel Density of Student Mathematics Score by School SES Quintile 44
Figure 8: Kernel Density of Student Mathematics Score by School SES Tercile 44
Figure 9: Kernel Density of Student Mathematics Score by School SES 44
Figure 10: Kernel Density of Student Mathematics Score by School SES Decile 45
Figure 11: Kernel Density of Student Reading Score by School SES Decile 45
Figure 12: Kernel Density of Socioeconomic Status across Provinces 49
Figure 13: Distribution of School Location by Province 49
Figure 14: Distribution of Socioeconomic Status across Provinces 49
Figure 15: Distribution of School Location by Socio-Economic Quintile 49
Figure 16: Kernel Density of Student Maths Performance by Province 50
Figure 17: Kernel Density of Student Reading Performance by Province 50
Figure 18: Kernel Density of Student Reading Performance by Socioeconomic Quintile 50
Figure 19: Kernel Density of Student Maths Performance by Socioeconomic Quintile 50
Figure 20: South African National and Provincial Student Performance 51
Figure 21: Maths Competency Levels by School Socioeconomic Status 56
Figure 22: Reading Competency Levels by School Socioeconomic Status 56
Figure 23: Reading Competency Levels by School Socioeconomic Status (CDF) 56
Figure 24: Maths Competency Levels by School Socioeconomic Status (CDF) 56
Figure 25: Maths Competency Levels by Province 57
Figure 26: Reading Competency Levels by Province 57
Figure 27: Reading Competency Levels by School Location 57
Figure 28: Maths Competency Levels by School Location 57
Figure 29: Lowess Regressions - Reading and Maths Performance 62
Figure 30: Age Distribution by School Socioeconomic Quintiles 64
Figure 31: Preschool Attainment by Learner Socioeconomic Status Quintile 65
Figure 32: Reading Textbook Availability by Province 67
Figure 33: Mathematics Textbook Availability by Province 67

Stellenbosch University http://scholar.sun.ac.za

Figure 34: Homework Frequency by School Socioeconomic Status 69
Figure 35: HIV/AIDS Knowledge Level by Province 71
Figure 36: HIV/AIDS Knowledge Level by School Socioeconomic Quintile 72
Figure 37: HIV/AIDS Knowledge Level by School Location 72
Figure 38: Variation in Student Knowledge about HIV/AIDS among SACMEQ countries and South African provinces [Source: (Moloi & Chetty, 2011, p. 5)] 73
Figure 39: Box plot of teacher absenteeism by SES quintile 81
Figure 40: Average Teacher Performance for School SES Quintiles 1 & 4 83
Figure 41: Average Teacher Performance for School SES Quintiles 1 & 5 83
Figure 42: Average Teacher Performance by School Socioeconomic Quintile 84
Figure 43: Kernel Density of Teacher and Learner Mathematics Knowledge 85
Figure 44: Kernel Density of Teacher Experience 86
Figure 45: SACMEQ Student Mathematics Performance 88
Figure 46: SACMEQ Student Reading Performance 88
Figure 47: SACMEQ Learner Reading Performance - Rural/Urban 89
Figure 48: SACMEQ Learner Mathematics Performance by Socioeconomic Status 89
Figure 49: Mathematics Teacher Content Knowledge 90
Figure 50: Reading Teacher Content Knowledge 90
Figure 51: Student Mathematics and Reading Performance in Regional Context 93
Figure 52: Textbook Distribution in Regional Context 94
Figure 53: Impact of Learner Socioeconomic Status on Reading Performance 105
Figure 54: Impact of Learner Socioeconomic Status on Mathematics Performance 106
Figure 55: School Socioeconomic Status and Reading Performance 113
Figure 56: Impact of School Socioeconomic Status on Mathematics Performance 114
Figure 57: The impact of individual SES for a given level of school SES 115
Figure 58: Distribution of Socioeconomic Status across SACMEQ countries 178

Stellenbosch University http://scholar.sun.ac.za

List of Tables

Table 1: Design Effects in SACMEQ III South Africa 20
Table 2: Levels of Reading Competency 53
Table 3: Levels of Mathematics Competency 54
Table 4: Percentage of Learners who are Non-numerate and Non-readers per Province, School Location, and School Quintile 59
Table 5: Grade Repetition by School Socioeconomic Status 69
Table 6: Diagnostic Overview (Provinces) 77
Table 7: Diagnostic Overview (School Location) 78
Table 8: Diagnostic Overview (School SES) 79
Table 9: Reading regression output: individual characteristics 104
Table 10: Mathematics regression output: individual characteristics 104
Table 11: Reading regression output: Home-background Characteristics 110
Table 12: Mathematics regression output: Home-background Characteristics 111
Table 13: Reading regression output: School Characteristics 122
Table 14: Mathematics regression output: School Characteristics 123
Table 15: Cross-tabulation - Province and School Location 139
Table 16: Cross-tabulation - Province and School Socioeconomic Quintile 140
Table 17: Cross-tabulation School Location and School Socioeconomic Status 141
Table 18: Reading Competency Levels by Province 142
Table 19: Reading Competency Levels by School Location 143
Table 20: Reading Competency Levels by School SES 143
Table 21: Maths Competency Levels by Province 144
Table 22: Maths Competency Levels by School Location 145
Table 23: Maths Competency Levels by School SES 146
Table 24: HIV/AIDS Knowledge Level by Province 147
Table 25: HIV/AIDS Knowledge Level by School Location 147
Table 26: HIV/AIDS Knowledge Level by Quintiles of School Socioeconomic Status 148
Table 27: Teacher absenteeism by province 149
Table 28: Teacher absenteeism by school location 149
Table 29: Teacher absenteeism by SES quintile 150
Table 30: Frequency of Missed Meals by School Socioeconomic Quintile 150
Table 32: Mean SACMEQ mathematics scores for participating countries for sub-groups 179
Table 31: Mean SACMEQ reading scores for participating countries for sub-groups 179

Stellenbosch University http://scholar.sun.ac.za

Chapter 1 **Introduction**

The strong legacy of apartheid and the consequent correlation between education and wealth have meant that, generally speaking, poorer learners perform worse academically. The links between affluence and educational quality in South Africa can partially explain this outcome since the poor receive a far inferior quality of education when compared to their wealthier counterparts. This is problematic for two reasons: *Firstly*, the received wisdom in economics dictates that individuals' labour-market prospects are directly correlated with their stock of human capital, which is itself correlated with the quality and duration of schooling. Offering an inferior quality of education to the poor disadvantages them in the labour-market and entrenches their poverty. The inter-generational effects of this inadequate education mean that children of impoverished parents are likely to be poor themselves. *Secondly*, given the racial dimension of poverty, and that the poor are more likely to be Black, one can go further and say that on average, Black learners receive an inferior quality of education to their White peers. In addition to the racial element of poverty, school location - both within provinces (urban/rural) and between provinces - are also important determinants of educational achievement. It is therefore necessary to improve the quality of education provided to the poor if these cycles of poverty are to be broken.

The aim of this thesis is to add to the existing body of literature which addresses the determinants of primary school performance in South Africa. The reason for the emphasis on *primary* school performance to the neglect of *secondary* school performance is based on the tenet that education is a progressive and cumulative process of acquiring knowledge, skills and attitudes. Many of the pervasive pedagogical problems in secondary schools are rooted in educational deficits that were acquired in primary school. Learners who have not mastered the basic skills of numeracy and literacy, or acquired the necessary orientations to meaning that secondary schooling requires, will not be able to assimilate new information or acquire new skills. This 'self-perpetuating handicap', to use the words of Taylor *et al.* (2003), is something which cannot be easily remediated in later school years. The focus should thus be on ensuring that the foundational skills of reading, writing, and problem-solving are acquired early on in children's learning careers, when they are most cognitively malleable. From an economic perspective, it is also most cost effective and efficient for citizens to acquire the basic educational skills in the early years of schooling, since the opportunity cost of the learners' time is here at its lowest.

The two underlying research questions which guide the analysis in this thesis are (1) Is primary education in South Africa efficient (broadly defined)? and (2) Is primary education in South Africa

Stellenbosch University http://scholar.sun.ac.za

equitable? In keeping with the latest research consensus in both economics and education, the thesis uses cognitive skills acquired rather than years of education attained as its outcome measure of both educational equity across sub-groups, and educational efficiency of the system as a whole. The data used in this thesis are from the third round of the Southern and Eastern African Consortium for Monitoring Educational Quality (SACMEQ) survey which was conducted in 2007. It includes two reliable measures of cognitive skills (Grade 6 numeracy and literacy), as well as extensive schooling, home-background and learner-level data.

Theoretical framework

In the field of economics, much time and energy has been exhausted debating the merits of the human capital approach in contrast to the screening/signalling hypothesis. Before it can be shown that the screening/signalling hypothesis is largely irrelevant for primary education, it is helpful to consider the central tenets of both theories. Economists adhering to the human capital approach believe that education is a means to increase productivity and therefore wages. In contrast, those more inclined towards the screening/signalling approach believe that individuals' productivity is largely innate and that education is merely used as a signal to employers that an individual is of high ability, productivity, or motivation. Since high ability individuals can complete education at a lower cost than can low ability individuals, education serves as a mechanism to decrease the asymmetry of information between employers and employees regarding the unobservable traits of ability or productivity. This 'Sheepskin Effect' was first posited by Spence in his paper "Job Market Signalling" (1973).

It is the position of this thesis that human capital theory is more suited to the analysis of primary school education than is the signal-theory approach. This is primarily for two reasons: (1) the ability to extract meaning from text and perform basic arithmetic functions are the foundational skills of productivity. With the exception of basic subsistence farming, it is extremely difficult to think of a scenario where a literate and numerate employee is not more productive than his illiterate and innumerate equivalent. (2) Individuals are not born with the ability to read, write and compute, but rather have to acquire these skills from someone who has already mastered them. Thus it can be said, unequivocally, that the skills of literacy and numeracy, which should be acquired in primary school, increase the productivity of individuals, and therefore increase their wages. Consequently, the impact of additional primary education is interpreted throughout the thesis using a human capital line of thinking, rather than one of the competing alternatives.

Stellenbosch University http://scholar.sun.ac.za

Thesis outline

The remainder of **Chapter 1** is devoted to a discussion of the SACMEQ III survey. Given its prominence in this thesis, an understanding of the sampling processes and design effects of a multi-stage sample are imperative, especially for interpretation purposes. In addition, the section explains how the numeracy and literacy tests were created, and provides the definitions of numeracy and literacy that were used in the SACMEQ surveys.

After using the extant theoretical literature to show the importance of human capital, and specifically cognitive skills, **Chapter 2** also reviews what we already know about the determinants of learner performance in South Africa. Specifically, the review is thematically organised around 9 topics: (1) the pupil-teacher ratio, (2) parental education, (3) grade repetition, (4) household resources, (5) managerial efficiency, (6) knowledge infrastructure, (7) ex-department[1] of education, (8) a theory of schooling, and (9) leadership of curriculum and instruction. By surveying both the literature in Economics, and in Education, the chapter shows that these disciplines are largely in agreement about some of the problems in the South African schooling system. The chapter closes with a discussion of the level of learner performance in South Africa, and does so by surveying the key findings from South African educational evaluations conducted in the past decade.

The focus of the thesis then shifts to the SACMEQ III data, with a preliminary analysis of learner performance in **Chapter 3**. After presenting a host of descriptive statistics on the characteristics of learners, teachers and schools, the performance of South African learners is placed in regional context.

The thesis culminates in **Chapter 4,** which moves beyond a bivariate analysis by modelling the generative mechanisms of learner performance in a multivariate context. The chapter presents the results from a variety of different sub-samples, and discusses the relevance of the findings.

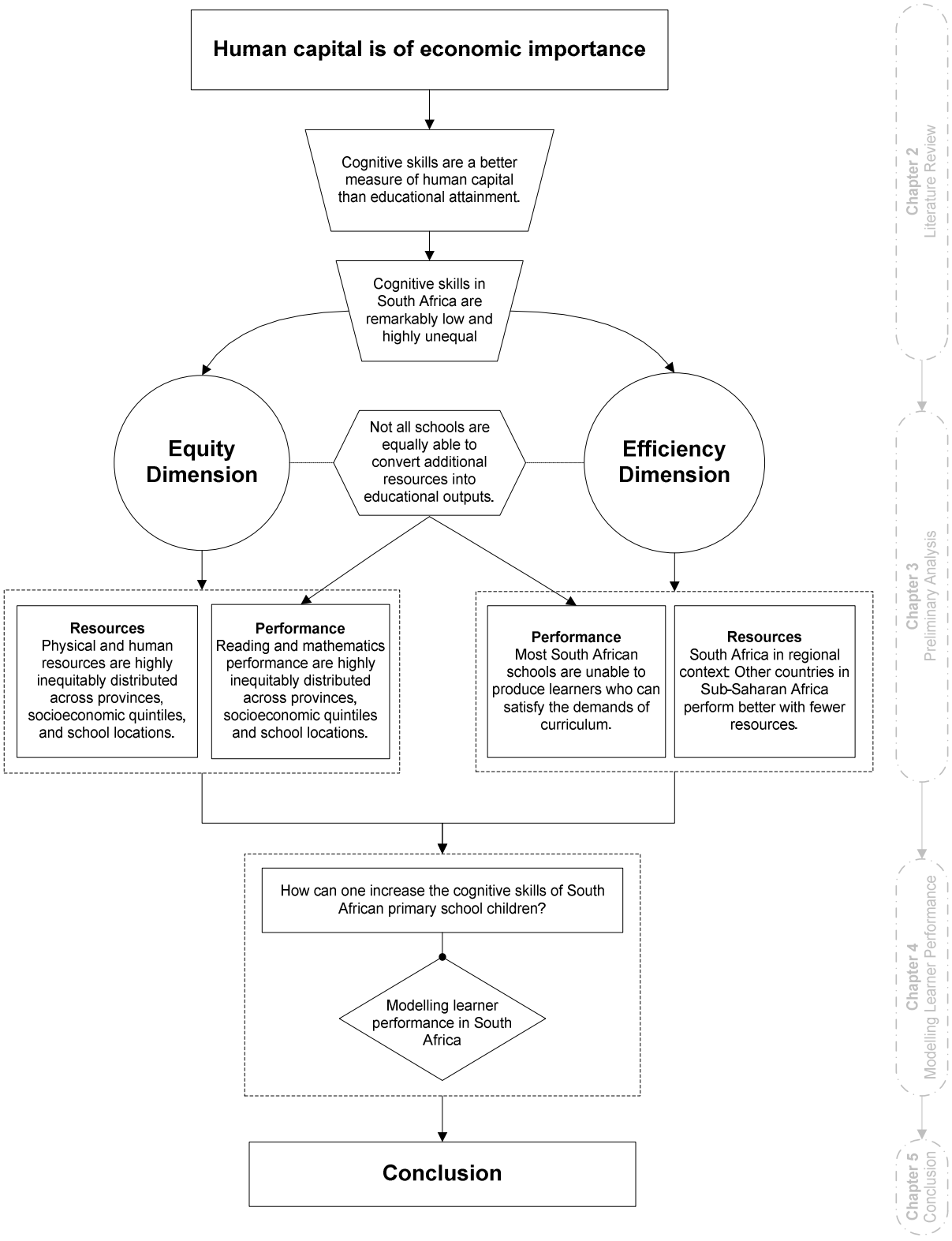
The final chapter concludes the discussion by summarising the main findings, explaining the policy implications, and suggesting possible avenues of future research (**Chapter 5**).

The above thesis outline is presented in schematic form in Figure 1 below.

[1] This refers to the system of classification that was used under apartheid, thus if a school was under the administration of the House of Assemblies (HOA) under apartheid it is referred to as an ex-HOA school.

Stellenbosch University http://scholar.sun.ac.za

Figure 1: Thesis Structure



Stellenbosch University http://scholar.sun.ac.za

Since the aim of this thesis is to expand the existing body of knowledge on learner performance in South Africa, an appropriate dataset is required. The latest SACMEQ survey - SACMEQ III - is one such dataset since it is timely, nationally representative, reliable, and comprehensive.

1.1 SACMEQ

The Southern and Eastern African Consortium for Monitoring Educational Quality (SACMEQ) is a consortium of education ministries, policy-makers and researchers who, in conjunction with UNESCO's International Institute for Educational Planning (IIEP), aims to improve the research capacity and technical skills of educational planners (Moloi & Strauss, 2005: 12). By generating information from school surveys SACMEQ enables decision-makers to monitor general conditions of schooling and the quality of basic education. SACMEQ III, was conducted in 2007, with South Africa being one of 14 participating countries.

The SACMEQ III survey instruments aimed to collect data on five themes, namely (1) learner characteristics and learning environments, (2) teachers' characteristics and viewpoints, (3) principals' characteristics and viewpoints, (4) equity in the allocation of human and material resources, and (5) achievement of teachers and learners (Lee *et al.*, 2005, p. 214). SACMEQ III South Africa surveyed 9083 Grade 6 learners and 1488 teachers from 392 schools which were randomly selected from across the country. Learners completed tests on both numeracy and literacy, and in addition, gave extensive demographic and home-background information. The teachers included 498 reading teachers, 498 maths teachers, and 492 health teachers. Each completed the Health test, with mathematics and reading teachers also completing subject-specific tests for their respective disciplines.

The SACMEQ III dataset is a particularly useful addition to existing primary school data in South Africa. The high quality and depth of the SACMEQ education data is rare in developing countries, of which South Africa is no exception.

Creation of the SACMEQ Literacy and Numeracy tests

Ross *et al* (2005, Ch2) provide a comprehensive overview of the creation of the SACMEQ II numeracy and literacy tests, which were the same as those that were used in the SACMEQ III study. For the purposes of SACMEQ, "reading literacy" was defined as: " the ability to understand and use those written language forms required by society and/or valued by the individual", while "mathematics literacy" was defined as "the capacity to understand and apply mathematical procedures and make related judgements as an individual and as a member of the wider society" (Ross, *et al.*, 2005, p. 74 & 78).

Stellenbosch University http://scholar.sun.ac.za

It is also important to understand that the SACMEQ tests were curriculum-specific and comprised of those items that were commonly agreed upon by all SACMEQ National Research Coordinators. Two excerpts from the Kenyan SACMEQ II report may serve to elucidate this process:

- On the creation of the *reading* test, "there was an initial detailed curriculum analysis undertaken across all countries in order to define – after exhaustive discussion of the most important skills contained within the reading curricula at Grade 6 level – the reading skills that were considered by all countries to be the most important" (Ross, *et al*., 2005, p. 74).
- On the creation of the *mathematics* test, the SACMEQ team decided to use the International Association for the Evaluation of Educational Achievement (IEA) mathematics domains and modify this "in order to bring it into alignment with what was actually being taught in SACMEQ classrooms in Southern and Eastern Africa" (Ross, *et al*., 2005, p. 78).

All questions in the survey were multiple-choice, with 55 reading questions and 49 mathematics questions. Using Modern Item Response Theory, a continuous and normally distributed measure of literacy and mathematical achievement was generated (Lee *et al.*, p. 217). These scores were then standardised by SACMEQ to have a mean of 500 and a standard deviation of 100 in the first survey that any of these tests was introduced, and results from subsequent surveys were converted to the same metric[2]. The SACMEQ tests and questionnaires in South Africa were only available in two languages: English and Afrikaans. Consequently, it is almost certain that learners who do not speak English or Afrikaans as a first language (the majority of South African learners) would be at a disadvantage relative to their first language peers. It is important to bear this in mind when interpreting the results from this survey. However, South Africa is not unique in this regard; for example, in Botswana, the language of communication and instruction at Grade 6 level is English even though Setswana is the most widely spoken language in the country. There is a similar situation in Mozambique where only 19.4% of Grade 6 learners spoke the language of instruction (Portuguese) at home (Passos, 2009, p. 314).

The SACMEQ III study represents an important milestone in South African educational evaluation since this is the first nation-wide education survey in South Africa where teachers were tested in

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

Stellenbosch University http://scholar.sun.ac.za

addition to learners[3]. This provides a valuable opportunity to better understand the impact of teacher-knowledge on learner performance (see Chapter 4 below).

Design effects and multi-stage sampling

The SACMEQ III survey used complex two-stage cluster sampling including weighting adjustments to compensate for variations in the probability of selection. The sample was stratified both by province (explicit strata) and school size (implicit strata). The 'province' stratification was accomplished by separating the sampling frame into provincial lists before undertaking the sample, while the 'school size' stratification used the number of Grade 6 learners in each school. The sampling method of probability proportional to size (PPS) was used to select schools[4] within strata and simple random sampling was used to select learners within schools (SACMEQ, 2010: 4). Since learners are clustered in schools, and schools are situated in provinces (which are the 9 strata in SACMEQ South Africa), the sampling errors are therefore larger than those that would have been obtained by simple random sampling. By using STATA's[5] built in *svy* command, the multi-level nature of the data is taken into account and the standard errors are calculated in light of the survey design.

However, it is still useful to approximate the effective sample size of the SACMEQ III South Africa survey which is possible by calculating the design effects of the survey. As Ross *et al.* (2005, p. 66) explain:

> "The design effect (Kish, 1965) provides an indicator of the increase in sampling variance that occurs for a complex sample in comparison with a simple random sample of the same size. The effective sample size (Ross, 1987) for a complex sample represents the size of a simple random sample that would have the same sampling accuracy as the complex sample."

The design effect for a survey that is sampling learners is directly related to the intra-class correlation coefficient (rho) for that country. The rho value for a country "expresses the variance in performance between schools as a proportion of overall variance" (Van der Berg, 2008, p. 3).The rho value for a country combines a measure of both the homogeneity of learners within schools and the heterogeneity between schools.

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

[4] Unfortunately the SACMEQ survey did not stratify based on ex-department, which is known to be an important determinant of school performance. This is perhaps an area for improvement for SACMEQ IV.

[5] This thesis used the statistical package STATA version 11.0

Stellenbosch University http://scholar.sun.ac.za

Sampling theory dictates that the more heterogeneous primary sampling units (PSUs) are, the more PSUs will be required to obtain a given level of accuracy. In educational terminology, the greater the difference in performance between schools, and the lower the difference in performance within schools (and thus the higher the intra-class correlation coefficient), the more schools will need to be sampled compared to a country where there is a lower intra-class correlation. In the SACMEQ studies, learner samples were drawn with the intention that estimates of population percentages would have a standard error of 2.5% and therefore that the population values would lie within 5% of the sample estimates (Moloi & Chetty, 2011).

With reference to the intra-class correlation coefficient in South Africa, given the exceptionally high degree of inequality in the country, it would be reasonable to expect that this would filter through to the education system, creating a highly unequal schooling system. Regrettably, the SACMEQ team had little indication of the rho value for South Africa when it participated for the first time in the SACMEQ project in 2000 (SACMEQ II). As a result, the intra-class correlation coefficient was grossly underestimated:

> "Unfortunately, the high values for South Africa and Uganda were not known beforehand, and the sample designs for these countries were based on "guesstimates" that the value of the intraclass correlation for each country was around 0.4. As a result, the number of schools in the sample designs for these two countries was too small – which resulted in a shortfall in the effective sample sizes for these countries" (Ross, *et al.*, 2005, p. 66).

It has since been found that the intraclass correlation in South Africa is approximately 0.64 (Van der Berg & Louw, 2006), far larger than the 0.4 speculated by the SACMEQ team. The aim of the SACMEQ initiative in both SACMEQ I and II was to have an *effective* sample size of 400 learners in every country in order to ensure a 2.5% standard error for population percentages. The number of schools sampled in each country would thus depend on the size of the intra-class correlation coefficient in that country. Due to the underestimation of rho in South Africa, the effective sample size in South Africa was 185 for reading and 230 for mathematics, clearly below the desired threshold of 400 (Ross, *et al.*, 2005, p. 66). One should therefore exercise caution when interpreting the SACMEQ II South Africa data.

Following the discovery of the large intra-class correlation coefficient in South Africa, the sample size for the SACMEQ III project in South Africa increased substantially, as compared to that of SACMEQ II. While SACMEQ II surveyed 3416 learners from 169 schools in South Africa, SACMEQ III surveyed 9071 learners from 392 schools – more than twice as many learners and schools than in SACMEQ II.

Stellenbosch University http://scholar.sun.ac.za

The SACMEQ III sample in South Africa was by far the largest of all SACMEQ countries, with more than double the average number of schools surveyed in other SACMEQ countries (185).

Table 1 below replicates Table 2.9(a) from Ross *et al.* (2005, p. 68) using SACMEQ III South Africa data. Ross *et al.* (2005) calculated the design effects and effective sample sizes for Botswana for a variety of pupil level, teacher level, and school-head level variables for SACMEQ II. From Table 1 below it is evident that the effective sample sizes for the teacher and principal questions are significantly smaller than those of the learner level questions. This is largely because there is only one principal and (usually) two teachers surveyed per school, compared to approximately 20 learners per school. Given that there will be no variation *within* schools for principal characteristics (since all learners have the same principal), the rho value for these variables will be higher, as will the design effects, and consequently, the effective sample size is smaller.

To further illustrate the concept of design effects, variation, and effective sample size, it is instructive to compare the effective sample size for the "Male" dummy variable with the effective sample size of the learner reading score. Given that the gender differentials between schools are unlikely to be large, it is unsurprising that the effective sample size for this variable is large (6724). In contrast, the reading score differentials between schools are likely to be large, leading to a smaller effective sample size for this variable (655).

Appendix F reports the design effects and effective sample sizes for the same variables as Table 1, but reports them for each South African province.

Prior to the analysis of the SACMEQ III data in Chapter 3 (Preliminary Analysis) and Chapter 4 (Modelling Learner Performance), it is useful to first survey the extant literature, and appreciate the context within which this research is found.

Table 1: Design Effects in SACMEQ III South Africa

Variable	Mean	Linearized SE	DEFF	DEFT	MEFF	MEFT	Sample size	
							Actual	Effective
Learner level								
Reading Score	494.86	4.55	13.85	3.72	14.14	3.76	9071	655
Maths Score	494.84	3.81	13.63	3.69	13.46	3.67	9071	666
Male	0.49	0.01	1.35	1.16	1.35	1.16	9071	6724
≥ 1yr Preschool	0.69	0.01	5.63	2.37	5.60	2.37	9071	1612
Sometimes spoke English at home	0.61	0.01	7.16	2.68	7.28	2.70	9071	1266
Maths-teacher level								
Male	0.39	0.03	25.34	5.03	24.58	4.96	9071 (498)	358
Teacher experience	15.31	0.43	27.13	5.21	26.60	5.16	9071 (498)	335
School-head level								
Teacher experience	25.25	0.40	30.47	5.52	31.34	5.60	9071 (491)	298
School library present	0.59	0.03	25.52	5.05	25.07	5.01	9071 (491)	355.7

Own calculations using SACMEQ III South Africa - based on Table 2.9(a) from Ross *et al.* (2005, p. 68)

Stellenbosch University http://scholar.sun.ac.za

Chapter 2 **Literature Review**

The aim of this literature review is threefold: firstly, to contextualise the importance of education within the framework of theoretical economics, secondly to survey the South African literature on the generative mechanisms behind learner performance, and lastly, to summarise the main studies which have assessed the level of learner performance in South Africa.

2.1 Theoretical links between Economics and Education

The interchange between the fields of Economics and Education has a long and distinguished history. If one agrees that the science of Economics was born with the publication of Adam Smith's The Wealth of Nations, then the notion of labour quality (i.e. human capital) has been present since its inception. As early as 1776, Smith had already identified that the quality of labour should be seen in the same way as that of traditional capital:

> "The improved dexterity of a workman may be considered in the same light as a machine or instrument of trade which facilitates and abridges labour, and which, though it costs a certain expense, repays that expense with a profit" (1776, p. 166).

Slightly over a decade later, Alfred Marshall (1890, p. 115) also illustrated that his understanding of capital included what we now term human capital: "Capital consists in a great part of knowledge and organisation…Knowledge is our most powerful engine of production; it enables us to subdue Nature and force her to satisfy our wants." Fisher (1906, p. 176) added some specificity to the notion by including health[6]:

> "A large part of our subjective income is due to our condition of health or disease … [A] healthy body is absolutely essential for receiving and enjoying the income from external wealth…Economists, by fixing attention exclusively on physical phenomena leave out the most essential element of all, the vigour of human life. The true 'wealth of nations' is the health of its individuals" (Fisher, 1906 as cited in Tobin, 2005).

However, it was not until the middle of the 20th century that three economists (Mincer, Schultz and Becker) focused specifically on education as a separate and important element of human capital. Jacob Mincer, the father of neoclassical labour economics published the first article in this sub-discipline: "Investment in Human Capital and Personal Income Distribution" (Mincer, 1958). However, his most enduring work was the book he wrote "Schooling, Experience and Earnings" in 1974. Soon after the publication of Mincer's first article, Schultz published "Investment in Human

[6] It is perhaps interesting to note that this was written after Fisher had overcome a serious case of tuberculosis (Tobin, 2005).

Stellenbosch University http://scholar.sun.ac.za

Capital" (Schultz, 1961), and Gary Becker published his influential book "Human Capital" (Becker, 1962).

For the purpose of this thesis, the most relevant discovery of these three economists was that labour was not a homogenous input in production, but rather that there exist large variations in the quality of labour which arise mainly from differences in health and education. This notion of a flexible quality of labour which could improve or deteriorate was termed *human capital*. Schultz explains this progression in understanding by highlighting the, now nonsensical, notion that labour is a uniform input in production:

> "The failure to treat human resources explicitly as a form of capital, as a produced means of production, as the product of investment, has fostered the retention of the classical notion of labour as a capacity to do manual work requiring little knowledge and skill, a capacity with which, according to this notion, labourers are endowed about equally. This notion of labour was wrong in the classical period and it is patently wrong now. Counting individuals who can and want to work and treating such a count as a measure of the quantity of an economic factor is no more meaningful than it would be to count the number of all manner of machines to determine their economic importance" (Schultz, 1961, p. 3).

The advances in labour economics brought about by these and other economists had implications for both a macroeconomic view of growth, and a microeconomic view of individual decision making. The improved quality of labour arising from investments in human capital, it was argued, was the main explanation behind the rises in real earnings per worker in America. On a microeconomic level, people's decisions to spend money on education, health and migration, i.e. to 'invest in themselves', was now understood from microeconomic principles – largely thanks to the theoretical foundations set out in Becker's book.

One important branch of the emerging human capital literature was the Mincerian approach to explaining wage differentials. Mincer believed that education, training, and experience explained a large part of how income was distributed in society, and consequently modelled income as a function of years of education and experience (Mincer, 1970 & 1974). These Mincerian earnings functions have become a standard tool to understand the impact of education on income. The coefficient on the *years of education* variable is typically interpreted as the private return to an additional year of schooling. Many studies have been conducted which aim to show how differing levels of education are associated with differing incomes. Important contributions include Psacharopoulos (1994), Card (1999), Psacharopoulos and Patrinos (2004), and Heckman *et al.* (2006).

Keswell and Poswell (2004, p. 851) provide a useful summary of eleven returns-to-schooling studies in South Africa up to 2004, and conclude that there are extremely low returns to education at levels

Stellenbosch University http://scholar.sun.ac.za

below matric, with a "sharp increase in the gradient of the return structure at about 12 years of schooling" (Keswell & Poswell, 2004, p. 849). This pattern of returns is robust to a variety of datasets and estimation techniques, leading the authors to conclude that South Africa's returns profile is convex, in contrast to the traditional concave return structure reported by Psacharopoulos and others. Gustafsson and Mabogoane (2010, p. 4) estimate that in South Africa, each additional year of education is associated with 22% higher earnings, with the twelfth year (matriculation) associated with an increase in earnings of 125%. These figures are calculated without taking into account any additional factors such as race, experience, location, school quality, etc.

In addition to the clear economic benefits to individuals and society, there are also numerous *noneconomic benefits* associated with increased educational attainment. While these benefits are beyond the purview of this study, it is worthwhile to mention some of the main noneconomic benefits. They include: lower fertility (Glewwe, 2002), improved child health (Currie, 2009), reduced societal violence (Salmi, 2006), promotion of a national - as opposed to a regional or ethnic - identity (Glewwe, 2002), improved human rights (Salmi, 2006), and lastly, increased social cohesion (Heyneman, 2003).

Importantly, rates of return analyses often do not take into account the *type* of education received or the *quality* of that education. For example, when looking at the returns to tertiary education, the effect is often averaged across students studying drama, engineering and medicine, for example, with little regard for the fact that these areas of specialisation have very different remuneration profiles. Similarly, and mainly due to a lack of data, there is little cognizance of the quality of the education that students receive. It is not unreasonable to believe that the average student graduating from Harvard University is likely to receive higher lifetime earnings than the average student graduating from a less prestigious institution of higher learning. While there are no doubt other factors influencing these remuneration profiles - including the selection effects from innate ability, social networks and home background - at least some of these additional earnings are likely to be an educational quality dividend. What is true for higher education is equally true for primary and secondary education, namely, that the return to an additional year of education is closely related to the quality of that education.

It is useful at this point to place the present thesis within the context of the current understanding of education's role in the economy. To this end, it is important to realise that the notion of labour quality as an essential economic input has progressed in two broad, but distinct thrusts. Firstly, economists realised that treating labour as a homogenous input in production was an unjustifiable simplification of reality, and consequently introduced the notion of variable human capital. In an

Stellenbosch University http://scholar.sun.ac.za

ironic turn of events, these same economists used educational *attainment* as the proxy measure of human capital and assumed that the quality of education received was homogenous across schools within a country. Since increased educational *attainment* raised both individual incomes and national incomes by increasing the productivity of labour, it was thought that the variation in human capital was captured by years of education completed. More recently, however, economists such as Hanushek and Woessmann (2008) have shown that cognitive skills acquired, rather than years of education completed, is the more appropriate measure of human capital. Using cognitive skills as a measure of human capital is both theoretically more plausible, and econometrically more illuminating. For example, "Models that include direct measures of cognitive skills can account for about three times the variation in economic growth than models that include only years of schooling" (Hanushek & Woessmann, 2008, p. 2, see also Hanushek & Kimko, 2000). Hanushek and Woessmann (2008) explain this progression below:

> "It is both conventional and convenient in policy discussions to concentrate on such things as years of school attainment or enrolment rates at schools. These things are readily observed and measured. They appear in administrative data, and they are published on a consistent basis in virtually all countries of the world. And they are very misleading in the policy debates. Cognitive skills are related, among other things, to both quantity and quality of schooling. But schooling that does not improve cognitive skills, measured here by comparable international tests of mathematics, science, and reading, has limited impact on aggregate economic outcomes and on economic development…We provide strong evidence that ignoring differences in cognitive skills significantly distorts the picture about the relationship between education and economic outcomes (Hanushek & Woessmann, 2008, p. 1)."

Thus, additional years of education do not necessarily increase human capital. By contrast, the cognitive skills of the population are a far more direct, intuitive, and theoretically legitimate measure of human capital. Improving these skills is likely to be associated with economic gains, as indeed Hanushek and Woessmann found.

The parallel between the two abovementioned thrusts in economic thinking is worth noting: (1) not all labour is equally productive, and (2) not all years of education are of equal quality. Thus, what began as an assertion that labour was heterogeneous and that years of education was the source of that heterogeneity, has progressed into an understanding that those years of education are themselves also heterogeneous, and these quality differentials are important when trying to understand the role of education in the economy.

The preceding discussion of human capital has demonstrated that education is economically important, and secondly, that education is best measured by cognitive skills acquired rather than school years attained. In recent times, this has changed the policy-discourse in development

Stellenbosch University http://scholar.sun.ac.za

economics, essentially shifting the focus from *access* to *quality.* It is within this context that the education production function literature is situated. Starting from the premise that cognitive skills, or educational competencies, are the true drivers behind the economic benefits of education, it is logical to ask, "What are the main determinants of cognitive skills?" This is the central question which education production functions aim to answer. By modelling educational outcomes (such as numeracy and literacy) using a variety of demographic, school-level and home-background characteristics, this type of analysis aims to identify which 'inputs' are the most important drivers of learner performance (outputs).

2.2 Understanding Learner Performance in South Africa

One of the many problems associated with the education production function literature is the inconsistency of findings between different authors, different datasets and different regression specifications. Given that these studies have direct policy relevance, it is important to avoid overstating research findings and confusing policy-makers with mixed signals. Explaining some of the methodological problems when using education production functions, Glewwe (2002, p. 475) warns:

> "Biased parameter estimates can arise due to omitted variable bias, endogenous program placement, sample selection bias, and measurement error in the explanatory variables…The problem of omitted variable bias is likely to be severe, which explains why different studies have produced very different results. Even worse, it is very difficult to overcome the problem because schools differ in so many ways, many of which are difficult to observe under even the best circumstances. Finally it is likely that measurement error problems lead to substantial biases, and there is no simple solution to this problem. Thus, all estimates of production functions for cognitive skills using conventional econometric methods should be regarded as suggestive, not definitive" (Glewwe, 2002, p. 475).

In light of the above, this review of the South African education production function literature extracts common themes rather than summarises each study, or discusses once-off findings. Before proceeding with this thematic analysis, it is instructive to compare two of the South African studies on cognitive achievement which highlight some of these complexities and contradictions that are an unfortunate feature of the education production function literature. Two of the earliest, and most interesting, studies which analysed South African cognitive performance are those by Case and Deaton (1999), and Moll (1998). A cursory reading of both papers would suggest, incorrectly, that the studies were based on two different datasets: (1) the South African Living Standards Survey (SALSS) (Case and Deaton), and (2) the Project for Statistics on Living Standards and Development (PSLSD) (Moll). Confusingly, these are the same datasets – something which is not easy to ascertain and the author found out through personal communication that these were in fact the same survey.

Stellenbosch University http://scholar.sun.ac.za

The SALDRU[7] website only has reference to the PSLSD survey, and thus, presumably, this is the correct name of this dataset. Why this is all the more interesting, is that both papers use the same dataset to answer similar questions, but find different results.

Part of Moll's (1998) study, which looked at cognitive skills and wages, addressed the impact of pupil-*classroom* ratios[8] on cognitive achievement. He finds that "Quality of schooling, as measured by the pupil-*classroom* ratio, had no impact, possibly because the measure is too coarse" (Moll, 1998, p. 278). Yet Case and Deaton (1999) find the exact opposite: "We find strong and significant effects of pupil-*teacher* ratios on enrolment, on educational achievement, and on test scores for numeracy" (Case & Deaton, 1999, p. 1047). Furthermore, rather than the ratio being "too coarse" a measure of quality, Case and Deaton find that "the unusually large variation in pupil-teacher ratios provides an excellent opportunity to examine their effects on outcomes" (Case & Deaton, 1999, p. 1048). This presents a somewhat perplexing situation, especially given that both authors use the same dataset. Perhaps these different conclusions are due to the differences between pupil-*classroom* ratios and pupil-*teacher* ratios. However, this is unlikely. Both sets of authors seem to use their ratio as a measure of school resources. Although Moll (1998) interprets his ratio as a school *quality* measure, given that the PSLSD was conducted in 1993, the correlation between school *resources* and school *quality* is likely to be sufficiently high that these two concepts are used interchangeably – something that, indeed, Moll seems to do. Case and Deaton make this more explicit in their paper where they explain that the variable of interest is really school *resources*:

> "Because we are not controlling for other school-based inputs, and because in South Africa, other inputs follow the supply of teachers, our purpose is not to assess the specific role of class size among other competing uses of resources, but to measure the effects of resources in general. Except when stated to the contrary, all subsequent references to the effects of pupil-teacher ratios should be understood in this sense" (Case & Deaton, 1999, p. 1050).

Clearly there is no obvious explanation for these different conclusions. Notwithstanding the above, Case and Deaton's study has proven to be far more influential than Moll's, even though their finding – that school resources do matter – goes against the general trend of education production functions internationally. If one looks specifically at the pupil-teacher ratio in its own right, and not as a proxy for resources in general, the evidence is clear: outside of extreme values, the pupil-teacher ratio does not seem to be significantly related to learner performance.

[7] South African Labour and Development Research Unit

[8] Moll does not provide a reason for why he chose the pupil-classroom ratio rather than the more commonly used pupil-teacher ratio, even though he acknowledges that both are available in the data (Moll, 1998, p. 268).

Stellenbosch University http://scholar.sun.ac.za

Crouch and Mabogoane (1998, 2001) in their analysis of the 1997 matric data, find only a small impact of the **pupil-teacher ratio** to the extent that they conclude that policy-makers should focus on teacher quality rather than pupil-teacher ratios. Unlike Case and Deaton (1999), however, they control for a number of important variables such as the physical condition of the school, the poverty of the school environment as well as the quantity of personnel resources. This comprehensive approach was posssible by merging data from the School Register of Needs (1996), the Education Management Information System (EMIS) (1997), as well as the socioeconomic database of the Department of Education (DOE). In a similar manner to Crouch and Mabogoane, Van der Berg and Burger (2003) use Census (1996) data in conjunction with national matric results (1997), including a more detailed dataset for the Western Cape. They find that the learner-teacher ratio in the Western Cape did not effect matric performance after controlling for a variety of socioeconomic and school-level factors. Bhorat and Oosthuizen (2006) also adopt this approach but use the matric pass rates (2000) for the whole country, the School Register of Needs (2000), and the Census (2001). This innovative approach enabled them to combine schooling outcomes with relatively accurate community-level information from each enumerator area (approximately 500 individuals). They too find that the learner-teacher ratio is not a significant predictor of matric pass rates.

The production function studies at the primary school level show mixed evidence on the insignificance of the pupil-teacher ratio. However, when this variable is significant it is always small, to the extent that it is hardly ever economically significant due to the high cost of additional teaching personnel. Van der Berg and Louw (2006), and van der Berg (2008) analysed the SACMEQ II (2000) data and found no discernable impact of the learner-teacher ratio. However, Taylor and Yu (2009), in their analysis of the Progress in International Reading Literacy Study (PIRLS 2006) dataset find a very small, but robustly significant, effect of class size on reading performance. In more recent work on the same dataset, Shepherd (2011) showed that there was a negative and moderately large impact on reading performance when class sizes were larger than 30 learners. Interestingly, this was only significant for African language schools, something which Case and Deaton (1999, p. 1080) also found. In his analysis of the National School Effectiveness Study (NSES), Taylor (2011, p. 43) found that the effect of the pupil-teacher ratio on numeracy and literacy, while statistically significant, was very small.

An intuitively appealing and consistent finding in the South African studies is the impact of **parental education** on learner educational attainment and cognitive achievement. Case & Deaton (1999) and Moll (1998) both find that educational attainment and cognitive skills are both positively and significantly related to parental education, although Moll looks at mother's education and Case and

Stellenbosch University http://scholar.sun.ac.za

Deaton look at the household head's education. Lam (1999) finds a similar effect of mother's and father's education on children's educational attainment. In addition to those studies that look at educational attainment, parental education has also been shown to influence educational achievement. Taylor & Yu (2009) show that if either parent had at least a matric education, this positively affected learner reading performance.

The negative impact of **grade repetition** is also a regular finding in the South African literature. Van der Berg & Louw (2006), Van der Berg (2008), and Gustafsson (2007) all find a strong negative impact of grade repetition on learner reading and mathematics achievement in SACMEQ II. Importantly, this may be unrelated to cognitive skills acquired as Lam, Ardington & Leibbrandt (2010) show. They develop a stochastic model of grade repetition which agrees with the empirical situation in South Africa. They find that "grade progression in African schools is poorly linked to actual ability and learning" (Lam, *et al.*, 2010, p. 1). Therefore, this grade repetition variable could be signalling the negative impact of poor assessment practices, as has been found elsewhere in the literature (Taylor, 2011).

The impact of **resources** on educational outcomes has been a topic of some interest in the local and international literature. Starting from the Coleman Report in 1966, researchers have been particularly interested in investigating the relative impacts of school resources, home socioeconomic status, and peer socioeconomic status. Coleman *et al.* (1966) found that, contrary to their initial expectations, the impact of a student's characteristics and family background on student performance in the United States was far larger than the impact of additional school resources. This finding has been supported, broadly speaking, by numerous subsequent studies (see Hanushek 1997, 2003 for an overview of these studies). Crouch & Mabogoane (1998, p. 2) succinctly summarise Hanushek's view (at least with respect to developed countries) when they state that "schools are so far inside the 'efficient frontier' that in practice inputs do not make much difference." Moving away from cognitive skills and looking at educational attainment, a learner's socioeconomic status is an important predictor of years of schooling attained (Filmer & Pritchett, 1999). There is evidence that this is also the situation in South Africa. Case and Deaton (1999) find that household resources increase educational attainment for Blacks, and Moll (1998) similarly finds that household durables are correlated with schooling attainment.

Looking at cognitive achievement, numerous South African studies have found that household resources, socioeconomic status, and the average socioeconomic status in a school are all important determinants of academic success (Van der Berg & Burger, 2003; Taylor & Yu, 2009), although sometimes these effects are limited to sub-samples of the population (Shepherd, 2011; Van der

Stellenbosch University http://scholar.sun.ac.za

berg, 2008). In contrast to the above, Bhorat and Oosthuizen (2006) find that household vulnerability (as measured by access to electricity, water, and a telephone, as well as the type of housing) was a poor indicator of academic performance in Grade 12. Also using matriculation data, Crouch & Mabogoane (1998) show that absolute performance (matric pass rates) in South Africa may not be the best measure of school success. Rather, they calculate measures of both absolute and relative measures of school performance, where the latter accounts for the socioeconomic disadvantage of the school. As they note, "It appears to be quite possible to 'normalise' for previous and current disadvantage, and to select schools that are over-achievers relative to their resources and clientele" (1998, p. 10). Given that only a small proportion of schools have adequate material and human resources, methods which account for economic disadvantage are likely to yield important insights into the determinants of schooling success in the previously disadvantaged part of the schooling system.

While school resources and home resources are both important determinants of cognitive skills, there is broad consensus in the literature that schools are not equally able to convert additional resources into improved outcomes. Thus, it is "not only the presence of school resources but how these are used which contribute to learning differentials" (Taylor, 2007). This concept of mediating resources is often termed **"managerial efficiency"** or **"school management"**. Since school management is difficult to measure directly, but is theoretically thought to be important, its importance is usually inferred from the large unexplained portion of student performance. In the same way that the residual in economic growth accounting is often attributed to the intangible concept of technological innovation, part of the residual in education production functions is plausibly attributed to managerial factors. This is most evident in Crouch and Mabogoane's (1998) paper "When the residuals matter more than the coefficients" which suggests that the large unexplained portions of matric performance were possibly due to differences in "managerial factors".

This finding is reiterated by Van der Berg & Burger (2003, p. 16) who conclude that there is need for "targetted managerial interventions to reduce inefficiencies in the educational system." Elsewhere, variables which point to the efficacy of school management are found to be significant and in some places, large. Gustafsson (2005), for example, finds that the allocation of time within schools is an important determinant of Grade 6 performance. Similarly, Van der Berg and Louw (2006) found that learners whose principals believed that monitoring student's progress was their most important activity, did significantly better than students in schools where principals expressed other priorities. In the same study, they find that students in schools where principals reported that teacher

Stellenbosch University http://scholar.sun.ac.za

absenteeism was a problem did worse than schools where this was not a problem (Van der Berg & Louw, 2006, p. 13). Perhaps the most convincing study on the impact of school management, is of Taylor's (2011) using the NSES panel dataset. The self-reported aim of this paper was to identify indicators of effective school management. He concludes his study by stating that:

> "An organised learning environment signified by curriculum planning for the full year, a functional time table, good quality inventories of 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 achievement and socioeconomic status" (Taylor S. , 2011, p. 43).

Another important finding in the South African education production function literature is the significance of certain types of **knowledge infrastructure**. These range from computer facilities and media centres (Crouch & Mabogoane, 2001; Bhorat & Oosthuizen, 2006), libraries (Case & Deaton, 1999), general school equipment such as a fax machine, photocopier, computer etc. (Van der Berg, 2008), and general media and communication facilities (Taylor, 2011). The presence of Learner and Teacher Support Materials (LTSM) has also been found to be significantly related to performance at a primary school level. Van der Berg (2008) found that learners who had their own reading textbook or shared with not more than one other performed better than those who shared with more than one other learner or had no book. Taylor (2011) found that well stocked school LTSM inventories contributed to learner performance, although he interprets this as an indicator of school management.

Linked to both school management and resources in general, the type of **educational department** a school belonged to under apartheid still comes through as a significant predictor of academic performance, illustrating the strong inertia of the apartheid era policies. Crouch and Mabogoane (2001) and Bhorat and Oosthuizen (2006) find that former department was a significant predictor of matric pass rates. Crouch and Mobogoane (1998, 2001) interpret the dummy variable for historical education department as a measure of managerial efficiency. Taylor (2011) also uses dummy variables for previous educational department in his analysis of the NSES data, and finds that African students in historically White schools perform significantly better than African students in historically Black schools, even after accounting for socioeconomic status, and school socioeconomic status. Utilizing a more indirect approach Shepherd (2011) uses language of the school as a proxy for former department, finding that the size and significance of the coefficients differ substantially by former department.

One of the earliest, and most important, analyses of the South African education system was undertaken by Andrew Donaldson in 1992. In his paper "*Content, quality and flexibility: The*

Stellenbosch University http://scholar.sun.ac.za

economics of education system change" he raises a number of important insights into South African education, only some of which are relevant to this thesis. As early as 1992, Donaldson had already identified the weak signal-value of internally conducted examinations, and called for externally validated examinations at the end of primary school (and Grade 9 & 10) – something which has only recently (2011) been implemented in the form of the Annual National Assessments. Furthermore, his prescient analysis of teacher incentives and accountability, as well as school administration and productivity, should continue to serve as theoretical blueprints for educational progress in South Africa. Following from the above, it is appropriate to include one excerpt where he discusses the concept of educational progress in somewhat abstract terms:

> "Progress cannot, however, be brought about in one great leap forward. The challenge is to organise education in ways which promote *ongoing system change, or educational technical progress,* obtaining progressively better value for money - improved content, better quality and greater flexibility - as a process driven by the internal dynamics of the education system and the external pressures to which it is subject" (Donalson, 1992, p. 2) (emphasis in original).

A more recent and comprehensive treatment of the South African schooling system is found in a book by Fiske & Ladd (2004) which focuses on the issue of equity in the education system in South Africa. In the book they discuss the political, sociological, financial, logistical and pedagogical problems which help to explain why educational equity has been so elusive in South Africa. In their closing arguments they summarise their views by highlighting three insights from the South African experience: (1) *history matters* – the incredible inertia of apartheid, and the complex ways in which it affected South African society, meant that real change was, and is, extremely difficult; (2) *resources matter* – exclusionary foci on either redistribution or improved efficiency are unlikely to paint a realistic picture of what can be achieved in South Africa – both are needed. However, they are quick to assert that resources are not the only thing that matter; (3) *implementation matters* – while there have been many well intentioned policies, few have resulted in systemic change, mainly because "serious reform requires both the managerial capacity to implement programs successfully and close attention to the design of effective implementation strategies" (Fiske & Ladd, 2004, p. 247) both of which have been lacking in the South African context.

The literature review above has focused entirely on educational research from an economic perspective, either using the framework of economics or the analytical tools of the discipline. However, given the prominence of education in the current analysis, it is important to incorporate the research findings, at least summarily, of some prominent South African educationists. Two of the most influential, and comprehensive books on the state of education in South Africa are those by Taylor, Muller & Vinjevold (2003), and another by Fleisch (2008). Taylor *et al.* set forth both an

Stellenbosch University http://scholar.sun.ac.za

analysis of the current research on education in South Africa, and go further to propose a **theory of schooling**, which has been an important contribution to the South African education literature. Their comprehensive treatment contextualises the South African schooling environment by discussing recent and past education policy, as well as elucidating issues around curriculum, pedagogy, and management. In their conclusion they highlight four "well-supported suppositions" which they show to be critical determinants of schooling performance in South Africa, namely:

> "(1) Language and early exposure to reading and writing…(2) Mechanisms which ensure coverage of the curriculum – like clear pacing signals, sequence signals, good exemplars and textbooks…(3) Explicit, regular and systematic evaluation and assessment…(4) An administrative and management climate that values and monitors high attainment for all" (Taylor, Muller, & Vinjevold, 2003, p. 135).

Fleisch (2008) presents a similar treatment of schooling in South Africa, and, furthermore, limits his treatment to the primary school sector only. While the book covers important topics such as health, expenditure, language, pedagogy, and a suggested research agenda for the future, the two most relevant chapters for this thesis are those on the bimodality of performance in the primary schooling system, and the impact of poverty on performance. On the bimodality of performance, Fleisch (2008: 30) provides some context to those learners in the weaker performing mode where the majority of learners are situated: "It is these South African children who struggle to read for meaning and to perform simple numerical operations – whose learning remains context-bound and non-generalizable." Secondly, in his survey of a number of papers linking poverty and education, Fleisch highlights numerous detrimental characteristics of poverty, including social isolation, emotional stress, low self-esteem, informal fostering, lack of education champions, text-deprived home environments, inadequate electricity, water and sanitation facilities, malnutrition, low-expectations, and psychological disempowerment (Fleisch, 2008). Like Taylor *et al.* (2003), Fleisch (2008) also concludes that poorer children, to their disadvantage, are not exposed to the literacy and linguistic codes that are generally used in the schooling environment.

A more recent discussion and summary of the classroom-based literature in South Africa is found in Hoadley (2010). This comprehensive review surveyed the relevant literature in order to understand the existing knowledge base around teaching and learning in South African primary schools. In her conclusion she highlights the seven most common classroom factors that are associated with student learning gains, namely:

> "(1) appropriate assessment and providing feedback to learners, (2) a focus on reading and writing text, (3) teacher's proficiency in the language of instruction, (4) the amount and type of reading and written work, (5) teachers adjusting pace to pupil ability, (6) greater curriculum coverage, including teacher knowledge and planning for and

Stellenbosch University http://scholar.sun.ac.za

> coverage of curriculum standards, and (7) greater opportunity to learn (content coverage by cognitive demand, content exposure as well as curriculum coherence and pacing" (Hoadley, 2010, p. 22) (numbers added to original).

Given the importance of school management and consequently, the impact of the principal on learner performance, Hoadley & Ward (2009) study the impact on learner performance of the **leadership of curriculum and instruction** in secondary schools in South Africa. They find that "the literature suggests that what principals might do that is of most importance and effect is to create containers within which effective teaching and learning can occur" (2009, p. 12). This nuanced understanding of school management as a necessary but not sufficient condition for learner performance, is likely to remain prominent in future studies of South African school leadership. Their study found that the 8 most important leadership determinants for learner performance were:

> "(1) School curriculum is covered, (2) school has a well-worked-out plan to improve student results, (3) the school day is structured for maximum student learning, (4) Positive relations between staff members at the school, (5) collaboration between teachers at the school, (6) effective management of LTSM in the school, (7) parental valuing of and support for education, and (8) school governing body's willingness to help in the school" (Hoadley & Ward, 2009, p. 53) (numbers added to original).

Interestingly many of the studies which Taylor *et al*. (2003), Fleisch (2008), and Hoadley (2010) cite, although qualitative in nature, suggest many of the same problems, and solutions, as those found in the quantitative studies surveyed earlier in the review. This is encouraging, and adds weight to the consensus that there are now a number of known factors that are consistently associated with poor performance, and at least some areas for action. This is helpful to both policy makers and theorists, since it enables them to create policies, and models of understanding that are relevant to the South African context.

In addition to understanding why some learners or schools perform better than others, it is also important to observe the system as a whole and determine the overall level of performance of the South African schooling system – this is discussed in the following section.

Stellenbosch University http://scholar.sun.ac.za

2.3 The Level of Learner Performance in South Africa

In South Africa there have been numerous initiatives to monitor the quality of education in the country. By measuring what learners know, these tests enable researchers and policy makers to assess the level of achievement of different groups of learners. Unfortunately the picture that emerges time and again is both dire and consistent: However one measures learner performance, and at whichever grade one chooses to test, the vast majority of South African primary school learners are significantly below where they should be in terms of the curriculum, and more generally, have not reached a host of normal literacy and numeracy milestones.

The discussion below focuses on the major educational evaluations conducted at the primary school level in South Africa in the last decade (2001-2011). There is unique merit in each of the evaluations since they each contribute to our understanding of primary school performance in South Africa. Since the aim of this section is to highlight the extent of underperformance in South Africa, and not to stress the severe inequalities that plague all areas of South African social policy, there has been less emphasis placed on explaining the large inequalities between learner performance.

It is commonly accepted that when looking at learner performance in South Africa there is a minority of learners (roughly 20%) who attend functional schools and perform acceptably on local and international tests while the majority of learners (roughly 80%) perform extremely poorly on these tests. Thus, there is a bimodal distribution of achievement in the country. It is for this reason that the national average on any of these tests shrouds severe inequalities. In reality, the median reading or maths score is *significantly* lower than the mean – that is to say that the better performing 20% of learners raise the extremely low average of the bottom 80%. Consequently, national averages reported here are overestimations of the average South African learner since the distribution is necessarily skewed to the right. It is ironic to note that because of this, the 'average South African learner' does not exist in any meaningful sense. Nevertheless, the statistic is widely used in local and international discussions of learner performance and is thus retained in the discussion. While often there are numerous studies which use a particular dataset, only the most authoritative study is cited below, since most studies are in broad agreement as to the extent of underperformance in a given dataset.

The **Trends in International Mathematics and Science Study** (**TIMSS 2003; Gr 8**) is a cross-national study which tests the mathematics and science knowledge of Grade 4 and Grade 8 learners in a variety of countries. South Africa participated in the Grade 8 study in 1995, 1998 and 2003, and although there was a 2007 TIMSS study, South Africa did not participate. Although the TIMSS South Africa study does not directly address primary school performance – which is the focus of this thesis

Stellenbosch University http://scholar.sun.ac.za

- the cumulative nature of education means that Grades 1-7 necessarily influence performance in Grade 8. Thus the TIMSS Grade 8 scores can be seen as a reflection of not only Grade 8 learning, but also learning in Grades 1-7.

In the most recent TIMSS study in South Africa (2003), 8952 Grade 8 learners from 255 schools across the country were assessed. Reddy's[9] 2006 country report "*Mathematics and Science Achievement at South African Schools in TIMSS 2003*" provides a detailed and comprehensive analysis of the country's performance in this study (Reddy, 2006). Of the 50 participating countries, including 5 other African countries (Botswana, Egypt, Ghana, Morocco, and Tunisia), South Africa had the lowest mean scores for mathematics and science. This echoed the 1999 TIMSS study where South Africa was also the worst performing country of the 38 that were tested.

South Africa's performance in the TIMSS 2003 study is particularly low. Only ten percent of South African learners achieved above the "low international benchmark" of 400, with a national mean of 264 (SE[10] 5.5) compared to the international average of 467 (SE 0.5). The mean of 264 shows how far below the low international benchmark the majority of South African learners are. Standardised measures of performance such as these are difficult to interpret since they are re-scaled to have a mean of 500 and a standard deviation of 100. To place South Africa's performance in perspective, if one looks at the proportion of correct answers on all mathematics and science items, this was just below 20 percent.

While it may be argued that TIMSS 2003 was administered during a period of curriculum change, and thus provides a distorted picture of learner performance, it has been shown that this is not the case. Reddy (2006, p. xiv), for example, explains that "the analysis of performance on topics which teachers said had been covered indicated that performance was still very poor, with learners achieving only 20% correct on those items." While this low level of performance in 2003 is worrying, it is even more concerning that there was no discernable improvement in mathematics and science between TIMSS 1999 and TIMSS 2003.

As in most analysis of school performance in South Africa, Reddy (2006) finds that there is considerable heterogeneity *within* South Africa and that mean mathematics and science scores actually shroud the large inequalities between different sub-groups. This is particularly acute when

[9] To avoid repetitive referencing, the figures and statistics for each sub-section are taken from the first reference in that sub-section, unless otherwise stated. For example, all statistics from the discussion of the TIMSS 2003 study are taken from Reddy (2006) which is the first reference in the discussion on the TIMSS study.

[10] Standard Error

Stellenbosch University http://scholar.sun.ac.za

looking at different provinces, at schools when split by ex-racial department, and by the proportion of economically disadvantaged learners in a school.

The **Western Cape Learner Assessment Study (2003; Gr6)** in 2003 tested every primary school in the Western Cape at the Grade 6 level. Of the 34 596 learners tested, a dismally small proportion were performing at the appropriate Grade 6 *literacy* level (35%) , and an even smaller proportion were at the appropriate Grade 6 *numeracy* level (15.6%) (Taylor, *et al.*, 2008, p. 43). Taylor *et al* go on to disaggregate these figures by ex-department and make the important point that four out of five Grade 6 children were at the appropriate reading level in former white school, compared to four children in a hundred in former Department of Education and Training schools.

The **Systemic Evaluations (2001 and 2007; Gr 3)** in South Africa tested a random sample of approximately 54 000 Grade 3 learners in more than 2000 primary schools in 2001 and 2007 (DoE, 2008a). The results of the 2007 Systemic Evaluation showed an average score of 36% for literacy (30% in 2001) and 35% for numeracy (30% in 2001). The Department of Education concluded in 2008 that there was an "urgent need to improve performance in these critical foundation skills" (DoE, 2008a, p. 12) - a statement which mirrored the earlier call for an "urgent intervention to address the situation" which appeared in the 2003 Systemic Evaluation report (DoE, 2003, p. 66).

The **Progress in International Reading Literacy Study** (**PIRLS 2006; Gr 4/5)** is an international initiative aimed at testing the reading literacy of Grade 4 learners in participating countries. Unlike almost all other countries that participated, where only Grade 4 was tested, in South Africa Grade 5 was tested in addition to Grade 4 so that one could compare Grades 4 and 5, and out of a concern that Grade 4 is a transition phase. The South African PIRLS 2006 study tested the reading proficiency of 16 073 Grade 4 learners in 429 schools, and 14 657 Grade 5 learners in 397 schools (Howie, *et al.*, 2008).

Of the 45 countries that participated in the PIRLS 2006 study, South Africa achieved the lowest score. Morocco was the only other African country that took part in the study, although there were other middle-income countries including Iran, Trinidad and Tobago, Indonesia, and Macedonia. Using the standardised PIRLS performance measure, with a mean of 500 and a standard deviation of 100, South African learners performed significantly below learners from all other countries. The mean reading score for Grade 4 learners was 253 (SE 4.6) and for Grade 5 learners was 302 (SE 5.6), both far below the international average of 500 and well below the Low International Benchmark of 400.

Stellenbosch University http://scholar.sun.ac.za

In South Africa, only 13% of Grade 4 and 22% of Grade 5 learners reached the Low International Benchmark of 400. This is in stark contrast to the majority of other participating countries. In half of the participating countries, 94% of learners reached this Low International Benchmark. Trong (2010, p. 2) elucidates the practical value of this benchmark: "learners who were not able to demonstrate even the basic reading skills of the Low International Benchmark by the fourth grade were considered at serious risk of not learning how to read." Using this framework, 87% of Grade 4 and 78% of Grade 5 learners in South Africa are deemed to be at serious risk of not learning to read.

The **Southern and East African Consortium for Monitoring Educational Quality (SACMEQ 2001/2007; Grade 6)** study is a cross-national initiative consisting of 14 countries in southern and eastern Africa. Given the centrality of the SACMEQ dataset to this thesis, and that the organisational structure has already been discussed in chapter one, only the highlights of South Africa's underperformance are given here.

SACMEQ tests the numeracy and literacy skills of Grade 6 learners in each of the participating countries. South Africa participated in the second (2000) and third (2007) rounds of SACMEQ. Of the 14 countries that participated in SACMEQ II (2000), South Africa had the 9th highest mathematics score and the 8th highest reading score (Van der Berg, 2007) - behind lower-income countries such as Botswana, Swaziland and Kenya. In the more recent SACMEQ study (2007) of the 15 countries that participated, South Africa came 10th for reading and 8th for Mathematics. More concerning than South Africa's relative position in regional context is the national prevalence of functional illiteracy and functional innumeracy. Of the Grade 6 learners that were tested, 27% were deemed to be functionally illiterate, while 40% were classified as functionally innumerate (Spaull, 2011). These figures differ substantially across the nine provinces. While half (49%) of all Grade 6 learners in Limpopo are functionally illiterate, the comparable figure in the Western Cape is one in twenty (5%). Similar differences can be seen based on the socioeconomic status of the school and the school's location – urban or rural.

The **National School Effectiveness Study (NSES 2007-2009; Grades 3-5)** is the only panel dataset on educational achievement in South Africa[11] where 266 schools were tested in numeracy and literacy in 2007 (Grade 3), 2008 (Grade 4) and 2009 (Grade 5) (Taylor , 2011). The mean scores for literacy in Grade 3 [Grade 4] were 19.4% [27%], and on the numeracy tests were 28.4% [34.6%] – all well below the levels that learners at these grades should be achieving.

[11] Gauteng did not participate in the NSES study since other testing was being administered in that province at the same time.

Stellenbosch University http://scholar.sun.ac.za

The **Annual National Assessments (ANA 2011; Gr 3/6)** are the latest in a series of initiatives by the Department of Basic Education which are aimed at measuring learner performance and identifying areas for improvement. The ANA's test both literacy and numeracy. The Grade 1-3 ANA's were available in each of the eleven official languages, while the Grade 4-6 ANA's were only available in English or Afrikaans – this mirrors the medium of instruction at each of these grades. The 'universal ANA' of 2011 was administered to all schools in Grades 2 to 7 in South Africa. In order to test the reliability of the universal ANA, 1667 schools were selected for 'verification ANA' where stricter test administration procedures were applied for the Grade 3 and Grade 6 ANA in those schools, and all verification ANA scripts were also remarked by independent assessors (DBE, 2011b).

Focusing on verification ANA, 19470 Grade 3 learners and 19367 Grade 6 learners were tested across South Africa. The average percentage scores (literacy/numeracy) for the two grades were: Grade 3 (35% / 28%) and Grade 6 (28% / 30%) – clearly the majority of South African learners are performing at unacceptably low levels, even when measured using local, curriculum specific standards of literacy and numeracy.

In addition to the above major national evaluations, Fleisch (2008, p. 22) provides a summary of some smaller school improvement project evaluations. These include the *Quality Learning Project* (2001), the *District Development Support Programme* (2001), the *Family Literacy Project* (2000), the *Early Reading Workshop*, and various projects evaluated by Eric Schollar. Each of these small-scale evaluations adds some nuance and detail to the overall picture of low and unequal performance in South Africa.

South Africa is in the fortunate position of having a number of independent and rigorous learner-performance evaluations[12] at various grades of primary schooling, testing both literacy and numeracy. While there are doubtless some methodological concerns surrounding some of the above evaluations, the overall picture that emerges is clear: South Africa's primary school performance is extremely poor by local and international standards, and is also highly unequal. Taylor *et al.* (2003, p. 41) commenting on some of the above results conclude that:

> "Studies conducted in South Africa from 1998 to 2002 suggest that learners' scores are far below what is expected at all levels of the schooling system, both in relation to other countries (including other developing countries) and in relation to the expectations of the South African curriculum."

[12] The latest PIRLS and TIMSS studies in South Africa, which were conducted in late 2011, should provide further valuable data on the performance of the South African schooling system.

Stellenbosch University http://scholar.sun.ac.za

Little has changed since 2003, and thus Fleisch (2008, p. 2) states in similarly emphatic terms: "without ambiguity or the possibility of misinterpretation, the pieces together reveal the predicament of South African primary education." Put simply, there is unequivocal consensus among researchers that South African primary schools are failing the large majority of learners in that they do not impart the necessary numeracy and literacy skills needed to function as literate and numerate members of the broader society. This situation has real economic consequences, as Schultz (1989, p. 222) explains:

> "My reasons for attributing great economic importance to schooling in low income countries are as follows. (1) The acquired abilities to *read* efficiently and to *write* with competence are essential in achieving modern economic growth and they are in general necessary prerequisites to investing in additional specialized human capital...(2) These valuable abilities are in large measure acquired during the early years of schooling. (3) The real costs of learning to read and write is at its lowest during the early years of schooling...this cost increases as the value of time of the maturing student rises. (4)The abilities to read and write are critical components of the quality of the human capital of any population."

Stellenbosch University http://scholar.sun.ac.za

Chapter 3 **Preliminary Analysis**

The high levels of inequality prevalent in South Africa are not limited to income and wealth, they are also found in the provision of social services such as education. In a sense, the educational inequalities found in South Africa are simply a by-product or reflection of the country's inherent socioeconomic inequalities. This is largely a legacy of the racial discrimination practiced under apartheid, although these inequalities were apparent even before the apartheid period, i.e. during colonial rule (Van der Berg, 2010, p. 3).

These inequalities manifest themselves in numerous, often inter-locking ways. It is useful to think about five different types of inequalities that are found throughout the country: racial, social, spatial, economic and linguistic inequalities. All of these types of inequalities filter through to the education system. On average, the most advantaged South African citizens are wealthy, White, English or Afrikaans speaking individuals who live in big cities in either Gauteng or the Western Cape. By contrast, the most disadvantaged citizens are poor, non-English speaking Black individuals who live in rural areas in the Eastern Cape, Mpumalanga KwaZulu-Natal or Limpopo provinces. Since the SACMEQ III data did not include questions on home language or race, the descriptive analyses below focus on the spatial and economic differences found between the various provinces. However, census data show that race and language are highly correlated, at least on the macro level, with income as well as regional location (rural-urban). This is one of the consequences of the apartheid-era policies of racial and spatial segregation, whereby non-white citizens were not permitted to live in more affluent urban areas and were restricted to townships and rural areas. All of the above factors contribute to the present situation where these five dimensions of disadvantage are usually found together.

This chapter provides background information and descriptive statistics on learners, schools, teachers and the regional context within which South Africa is situated. By illustrating and explaining the distributions of performance, as well as school and teacher characteristics, the reader is better positioned to appreciate the often nuanced complexities in interpreting multivariate analyses, which are found in the following chapter. Furthermore, the bivariate statistics presented here give a broad outline of the inherent inefficiencies and inequalities in the South African primary school system.

In SACMEQ III, as is the case with most surveys which target children, it is not possible to get an accurate representation of the monetary value of family income. Consequently, socioeconomic status (SES) was inferred from a series of possession questions. In SACMEQ III, learners were asked

Stellenbosch University http://scholar.sun.ac.za

whether or not each of 31[13] items was found in the place where they stayed during the school week. To construct the SES variable, all 31 of these items were used in a Multiple Correspondence Analysis (MCA) forming the SES index. The SES variable was transformed to be the negative of the MCA index to ensure that the largest positive value of MCA was assigned to the wealthiest learner for ease of interpretation. An MCA was used rather than Principal Component Analysis (PCA) since MCA is more appropriate for categorical variables (Booysen, *et al.*, 2008).

3.1 Learners

3.1.1 Learner Performance

Material resources enable parents to send their children to the best schools and provide the highest level of support, thus, it should come as no surprise that South Africa has an extremely unequal distribution of attainment *across* grades and performance *within* grades. It is the latter issue – performance differentials within a grade – that is of particular interest here. To be more specific, the distribution of learner performance in South Africa is bimodal, that is to say it has two relatively distinct modes: one at a very low level of performance, and the other at a substantially higher level. The mode associated with the lower performance consists largely of poorer schools, usually ex-Department of Education and Training (DET) schools, while the higher performance mode is made up of wealthier schools which are usually ex-Model C schools.

This bimodal distribution of South African learner performance has been well documented in the South African literature. The unequal distribution can be found at all levels throughout the education system, and is impervious to the dataset used. Notable examples of this bimodality have been found in the matriculation pass rates in 1999/2000 (Van der Berg, 2007), the Grade 6 SACMEQ II (2000) data (Gustafsson, 2005, p. 12; Van der Berg, 2008; Fleisch, 2008, p. 21) as well as at the Grade 5 level in the PIRLS (2006) dataset (Taylor & Yu, 2009, p. 41). Most disturbingly, it can already be seen as early as Grade 3 (Taylor, 2011, p. 18). Given that language and socioeconomic status are highly correlated in South Africa, the bimodality can also be seen quite clearly if schools are split according to language, as Shepherd (2011, p. 32) does with the PIRLS (2006) data.

While the bimodality of the schooling system is quite noticeable, it is less clear where the distinction between the two systems actually occurs. The data suggests that the upper distribution is relatively small (between 15 and 30 percent) while the lower distribution is where the bulk of schools are

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

Stellenbosch University http://scholar.sun.ac.za

located (between 70 and 85 percent). This bimodality in performance is largely along socioeconomic lines, and can be seen when splitting performance either by individual socioeconomic status, *or* school socioeconomic status (SES) (where school SES is the average SES of learners in that school). The method adopted in this thesis is to use school SES rather than individual SES since there is reason to believe that the two distributions are driven by school level factors rather than home level factors: economically disadvantaged learners who attend well-functioning schools do only slightly worse than their wealthier peers, as will be shown in Chapter 4 below. For these reasons, splitting the sample by school socioeconomic status seemed most logical. In order to ascertain where the split occurs, kernel density curves were used, with each additional graph splitting the sample into smaller sub-groups of school socioeconomic status. Figure 4 to Figure 11 show the distributions of reading and mathematics performance when the sample is split according to the wealthiest/poorest schools and thereafter according to terciles, quartiles, quintiles and deciles of school socioeconomic status. It soon becomes clear that lumping all learners into the same distribution of performance masks the very different underlying distributions.

Figure 5: Kernel Density of Student Reading Score by School SES

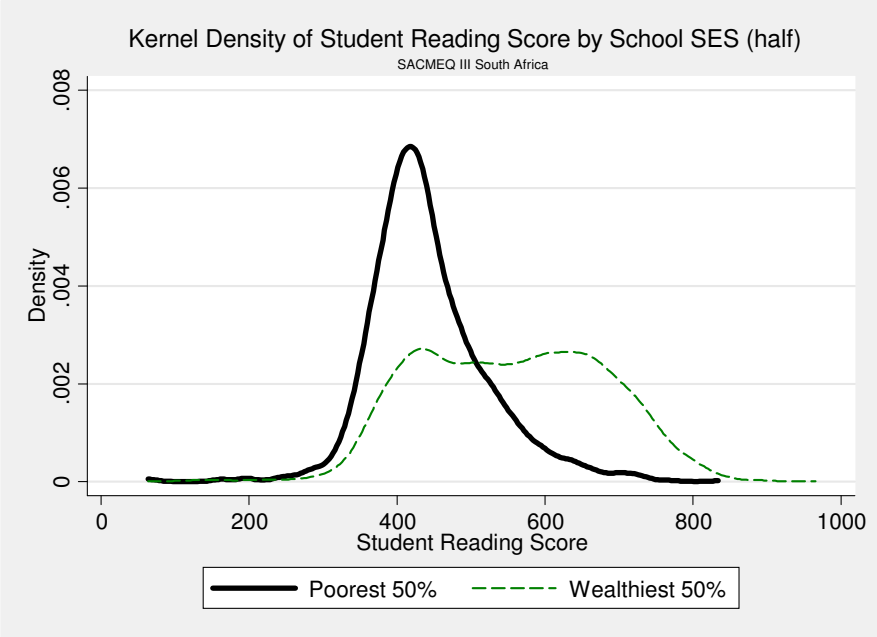
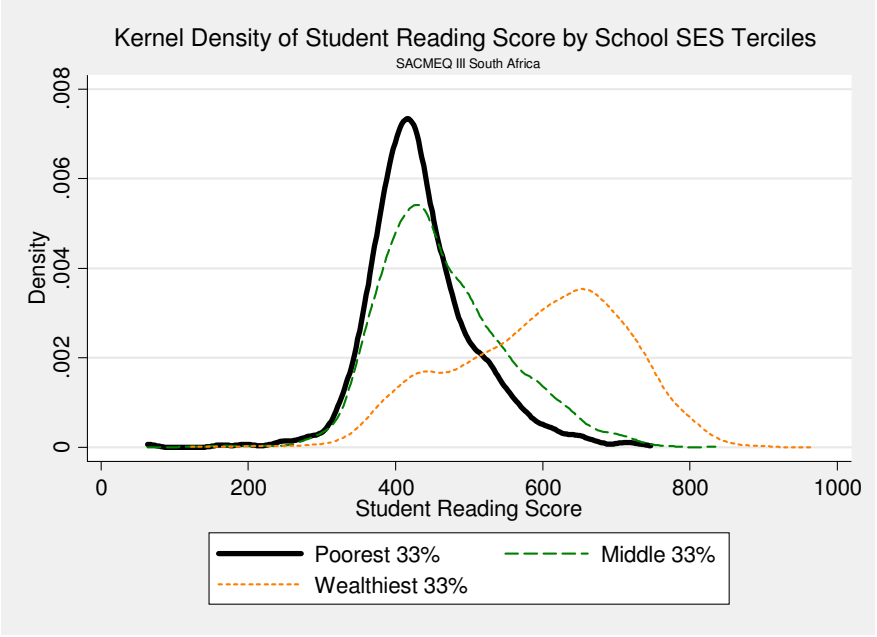
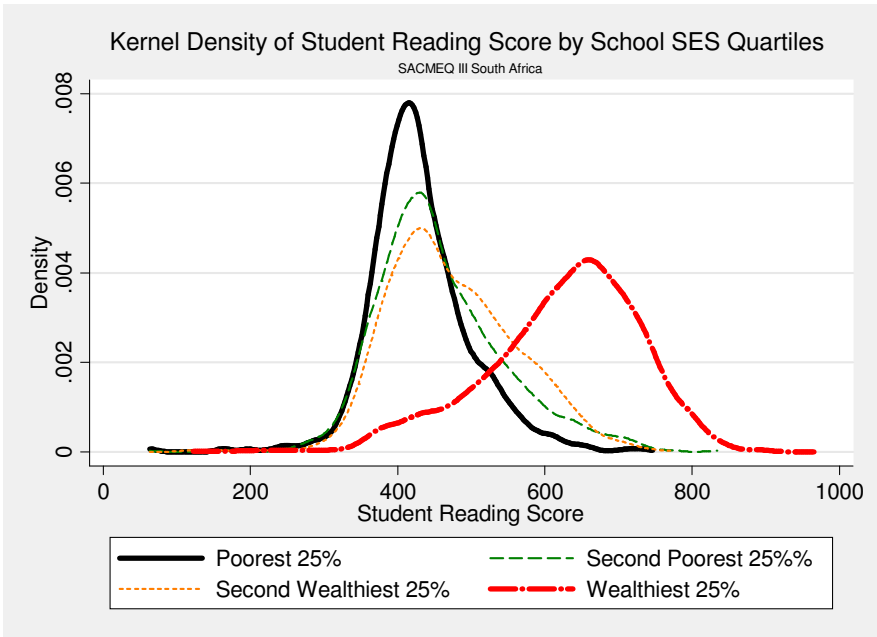


Figure 4: Kernel Density of Student Reading Score by School SES Tercile





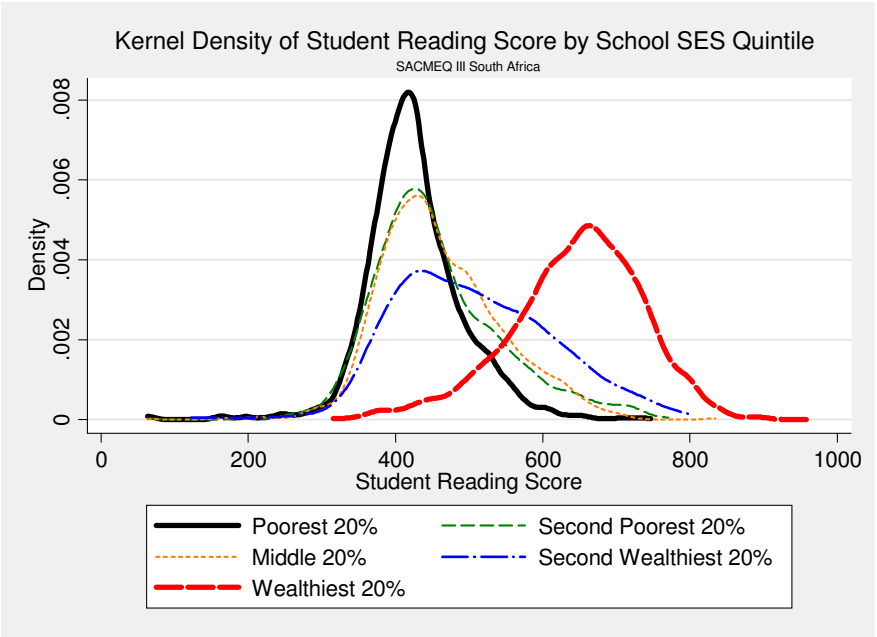


Figure 2: Kernel Density of Student Reading Score by School SES Quartile

Figure 3: Kernel Density of Student Reading Score by School SES Quintile

Figure 9: Kernel Density of Student Mathematics Score by School SES

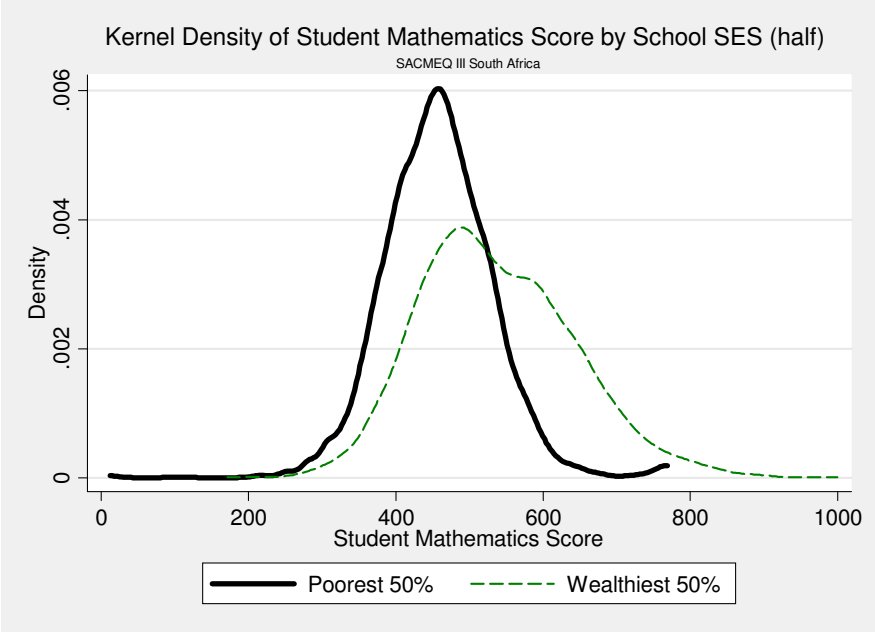
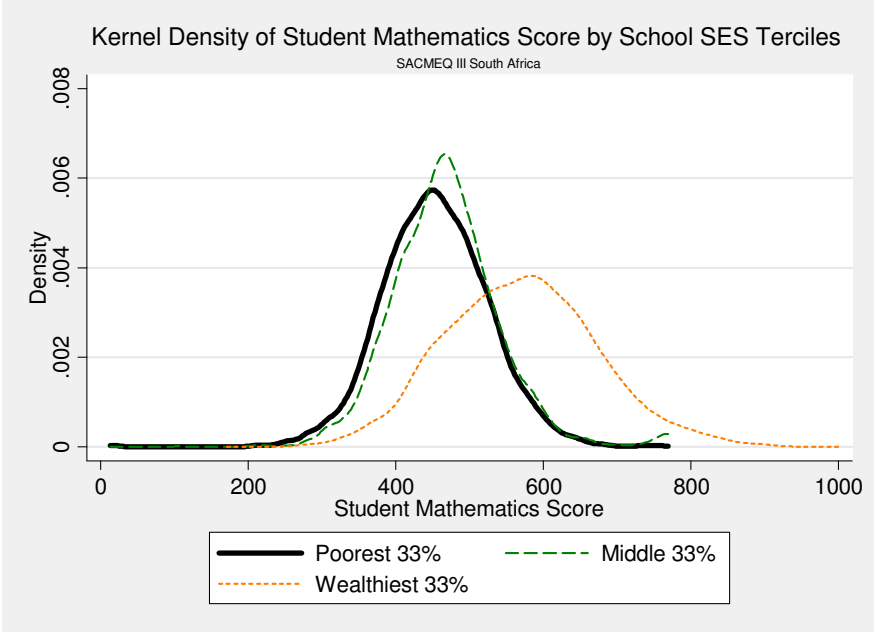


Figure 8: Kernel Density of Student Mathematics Score by School SES Tercile



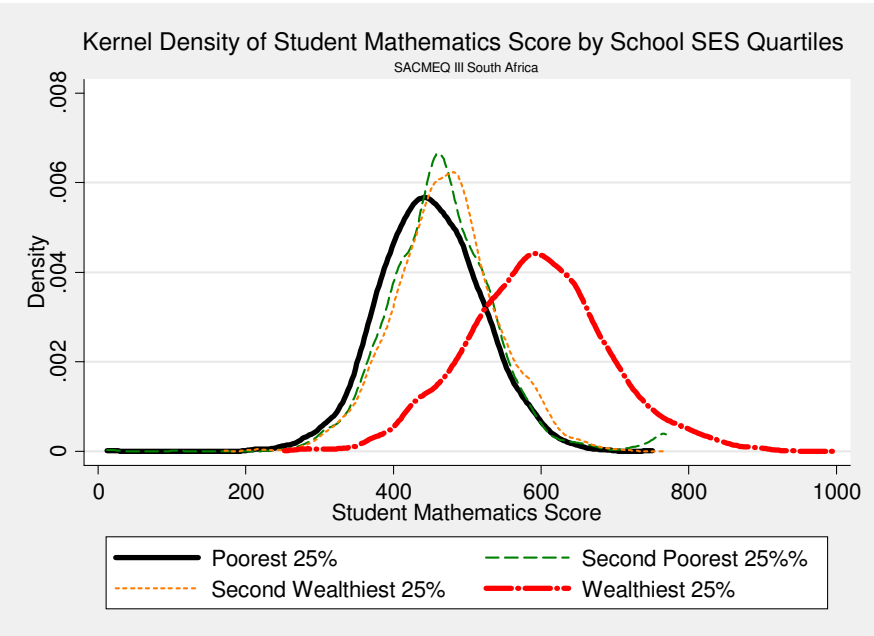


Figure 6: Kernel Density of Student Mathematics Score by School SES Quartile

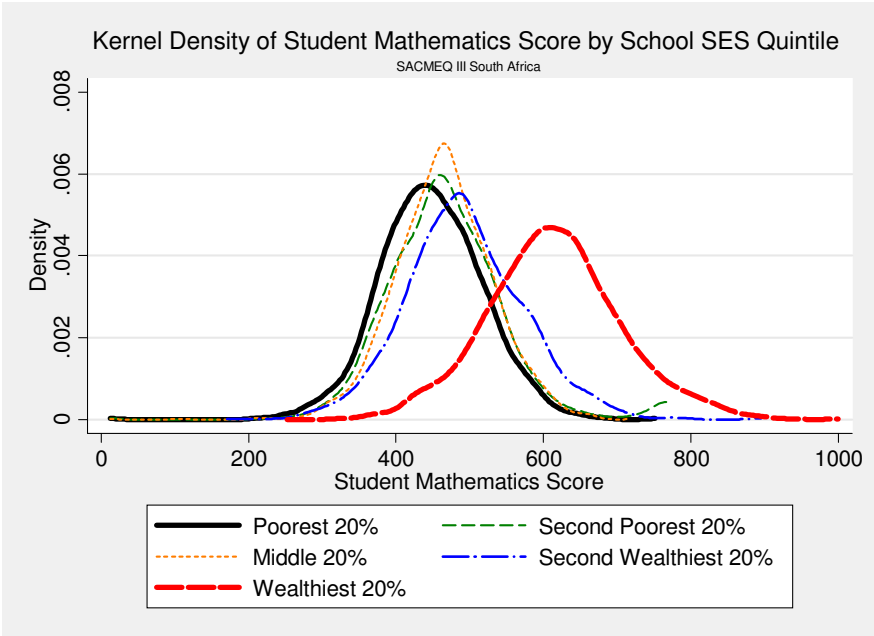


Figure 7: Kernel Density of Student Mathematics Score by School SES Quintile

Figure 11: Kernel Density of Student Reading Score by School SES Decile

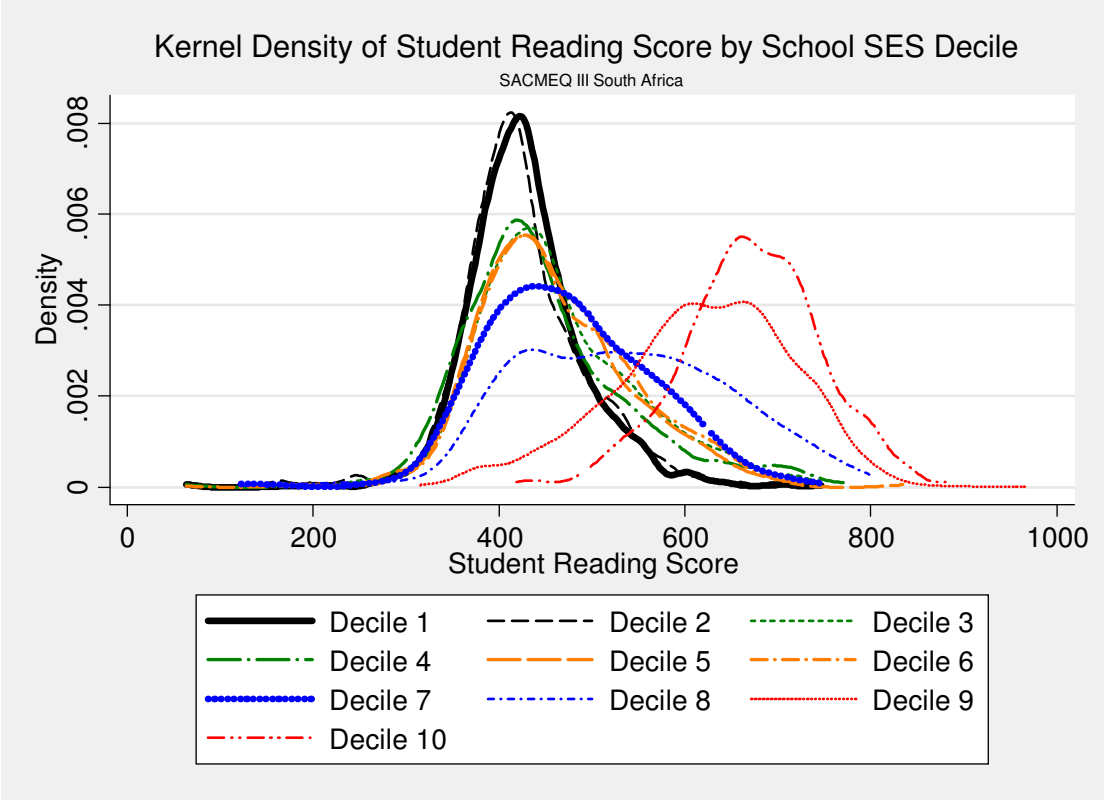
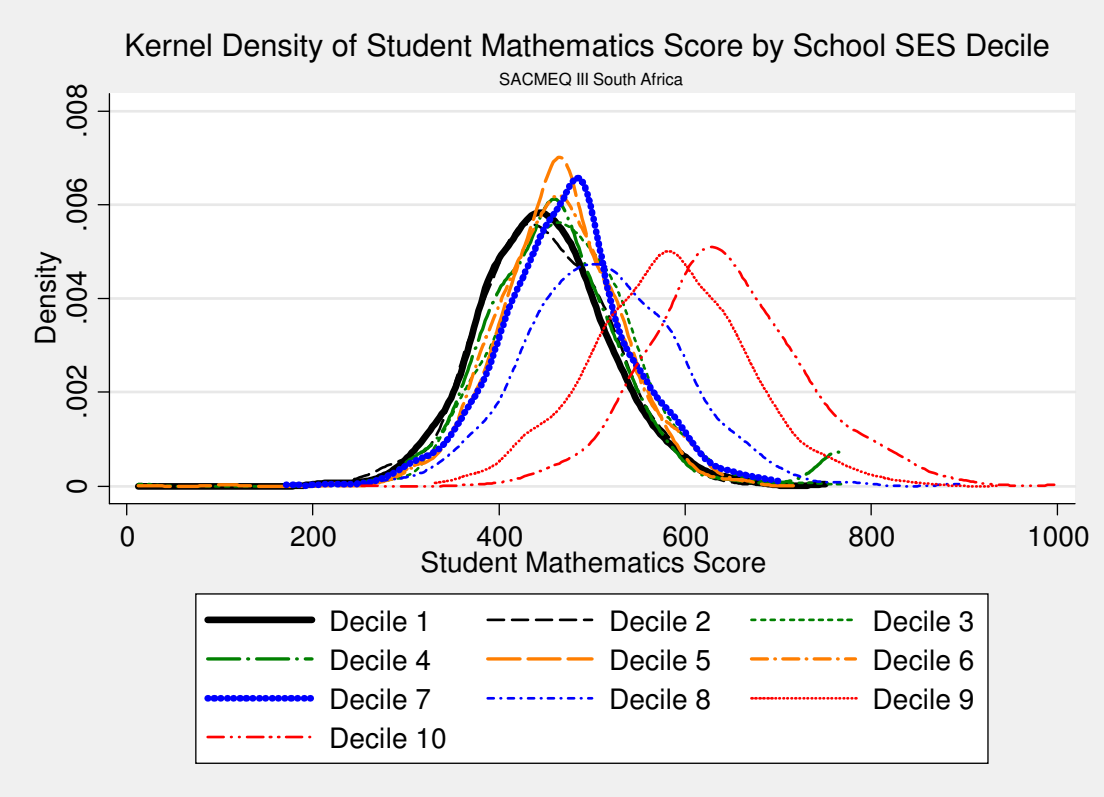


Figure 10: Kernel Density of Student Mathematics Score by School SES Decile



Stellenbosch University http://scholar.sun.ac.za

Figure 11 above clearly shows that the top two deciles of school SES have very different distributions to the bottom seven deciles, which all have very similar distributions. The eighth decile of school SES seems straddled between the two systems with the lower half having a mode of 400 while the other half has a mode of 600. This agrees with the previous kernel densities which split reading by quartiles and quintiles of school SES (Figure 2 & Figure 3). It would seem that based on these distributions, the most logical split would be between the top quartile (or quintile) of school SES, and the bottom four quintiles (or three quartiles[14]).

Van der Berg (2008) postulates that this bimodal distribution is indicative of two different data generating processes in the South African schooling system. Given the *a-priori* assumption that learner performance is characterised by a normal distribution, the normality of each of the four quartile distributions adds weight to the argument that there are indeed two data generating processes at work. A similar, although less-pronounced, situation can be found by observing the distributions for mathematics.

Comparing the average[15] reading scores of each quartile of school socioeconomic status shows that the mean of the bottom three quartiles (approximately 400) is more than one SACMEQ standard deviation (100) below the SACMEQ III average (approximately 510), while the mean of the top quartile (approximately 650) is more than one and a half standard deviations above the SACMEQ average. This epitomises the inequality in learner performance in South Africa.

Provincial differentials

The South African reality that certain provinces are wealthier than others means that the above mentioned socioeconomic differentials invariably extend to geographic differentials as well (Figure 13 to Figure 15). The wealthiest two provinces, the Western Cape and Gauteng, have a different distribution of learner mathematics and reading scores as compared to the other seven provinces.

The strong correlations between learner test scores, SES, and the wealthiest two provinces may indicate some form of causation between wealth and performance. However, it should be stressed that SES is not necessarily the main reason why the wealthier provinces perform better academically. It may be that parents with higher than average ability are more likely to work and live in the commercial hubs of Gauteng and the Western Cape and thus learners in these provinces may have higher than average ability. Or it could be that the education departments and schools in these

[14] Given that the distinction between quintiles and quartiles is really a discussion of whether the break happens at the 75th or 80th percentile, either one of these seem adequate.

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

Stellenbosch University http://scholar.sun.ac.za

two provinces are better managed and more able to create environments where learners excel. Or it could also be that wealthier parents place a higher premium on education as compared to poorer parents and thus spend more time, energy and money on improving their children's education and holding government departments accountable. More than likely, it is a combination of all of the above effects. In each of these cases, SES still plays an indirect role in improving learner performance.

Figure 13 to Figure 15 below show the distributions of socioeconomic status and school location for each province. It soon becomes clear that, crudely speaking, there are three types of province in the country: **Type 1** - *Wealthy and urbanised* (Gauteng and the Western Cape), **Type 3** - *Poor and rural* (Eastern Cape, KwaZulu-Natal and the Limpopo), with **Type 2** being some combination of Type 1 and 3 (Freestate, Northern Cape, Mpumalanga, and the North West). It is self-evident that this crude classification has many exceptions since there are pockets of wealth in all provinces, particularly in larger cities such as Durban, Port Elizabeth and Nelspruit. Conversely, there are regions in both Gauteng and the Western Cape that are either rural, poor, or both.

However, the classification is useful for those who are unfamiliar with the provincial contexts of South Africa. Figure 13 to Figure 15 show that the majority of rural schools (78.5%), and quintile one schools (89.6%) are located in only three provinces: Eastern Cape, KwaZulu-Natal and the Limpopo (Type 3). By contrast, the majority of urban schools (59.8%) and quintile five schools (58.5%) are located in only two provinces: Western Cape and Gauteng (Type 1). As one would expect, the majority of quintile one schools are also rural schools (87.7%) and the majority of quintile five schools are situated in large cities (78.5%). Table 15, Table 16, and Table 17 in Appendix A provide the exact numerical distributions for Figure 13 to Figure 15.

Given that learner performance is significantly different across the sub-groupings of these three variables (province, school location and school socioeconomic quintile), most descriptive statistics provided in this chapter are reported for each of these sub-groups.

The process of reporting the mean value of each variable by province, location or quintile, should be thought of as taking many two dimensional slices of a three dimensional picture. Only when these two dimensional slices are seen together (i.e. the intersection of the slices), is there an accurate representation of the underlying data. Seen independently of each other, or without sufficient contextual information, these mean scores can be misleading. For example, after looking at the mean learner test-scores for the various provinces, one may conclude that the Limpopo province has the worst education system in the country. However, this simplistic conclusion fails to acknowledge

Stellenbosch University http://scholar.sun.ac.za

that 89.1% of schools in Limpopo are situated in rural areas. Or if one focuses instead on the poor performance in the Eastern Cape, a similar demographic is evident: 70% of schools are located in rural areas, and the majority (62%) of schools in the province are quintile one schools, i.e. some of the poorest in the country. Indeed, almost 40% of all quintile one schools are located in the Eastern Cape.

Illustrating this data in a different manner, Figure 20 below shows the provincial averages for both reading and mathematics, with the size of each bubble being proportional to the Grade 6 enrolments in that province. Clearly KwaZulu-Natal has the largest share of Grade 6 learners, and the Northern Cape has the smallest share. If one compares these provincial averages to the SACMEQ averages for reading and mathematics (dotted lines in Figure 20), one can see that seven of the nine provinces perform worse than the SACMEQ averages. This is surprising given that South Africa's economy is by far the largest amongst the SACMEQ countries and it is also one of the richest in per capita terms. South Africa's performance in regional context is covered in more depth towards the end of this chapter.

Figure 13: Distribution of School Location by Province

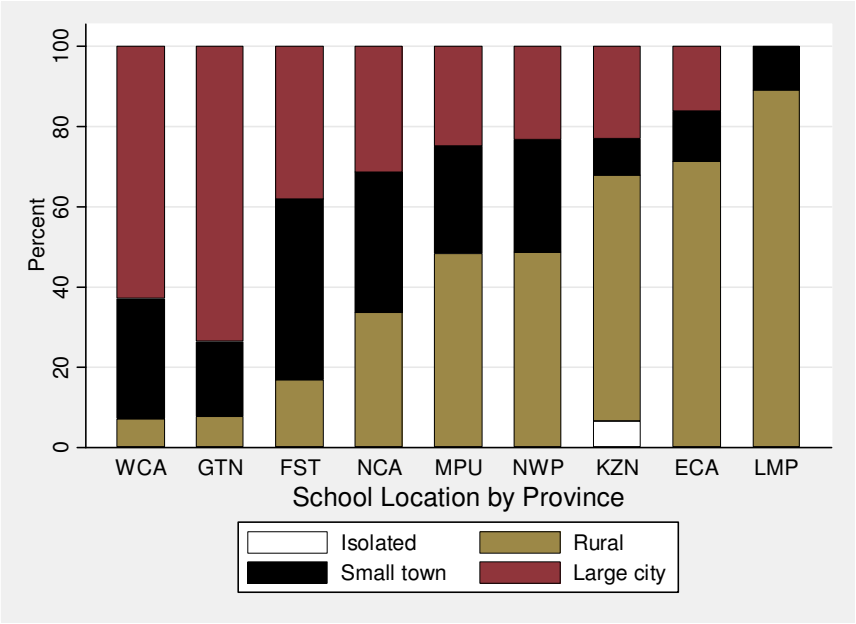
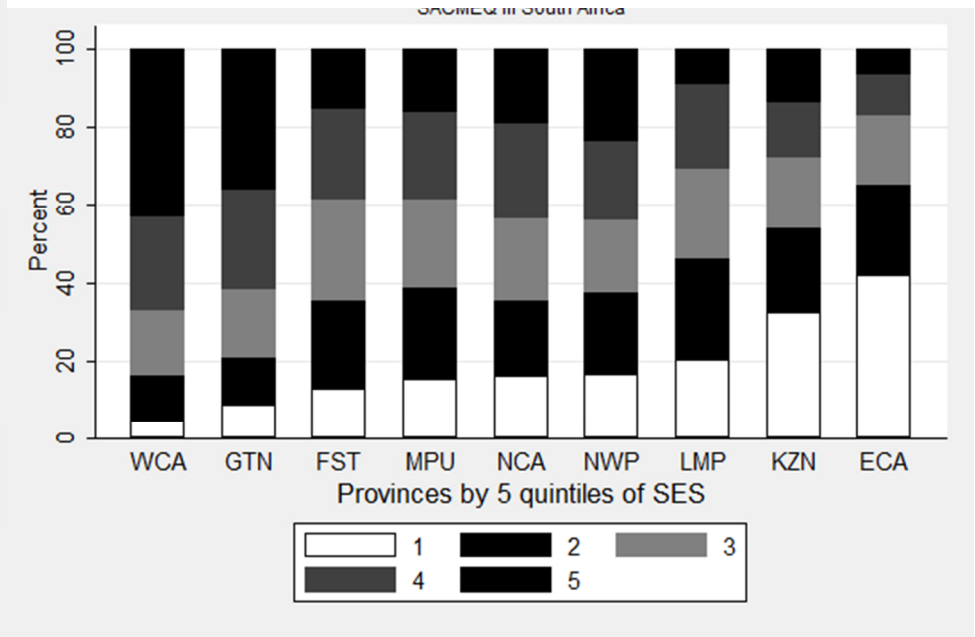


Figure 14: Distribution of Socioeconomic Status across Provinces



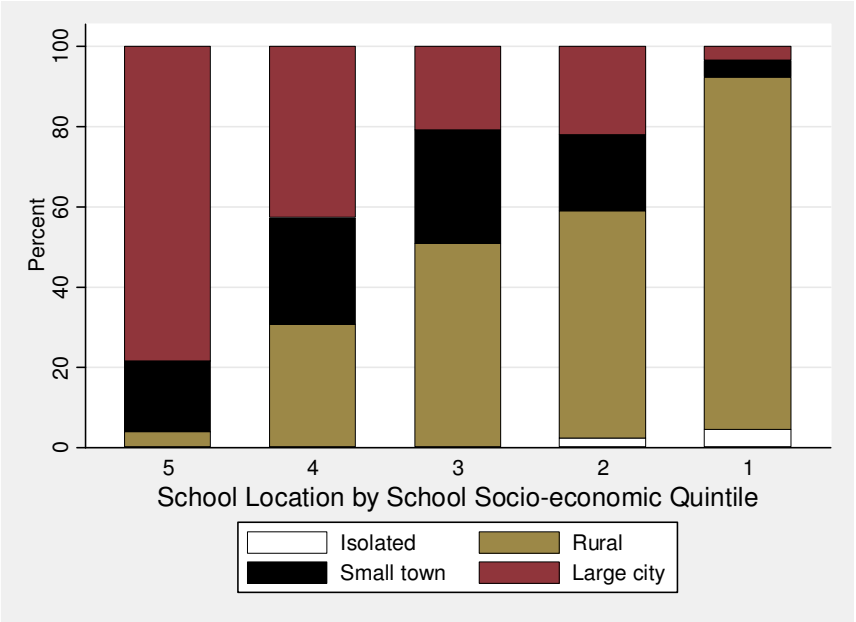


Figure 15: Distribution of School Location by Socio-Economic Quintile

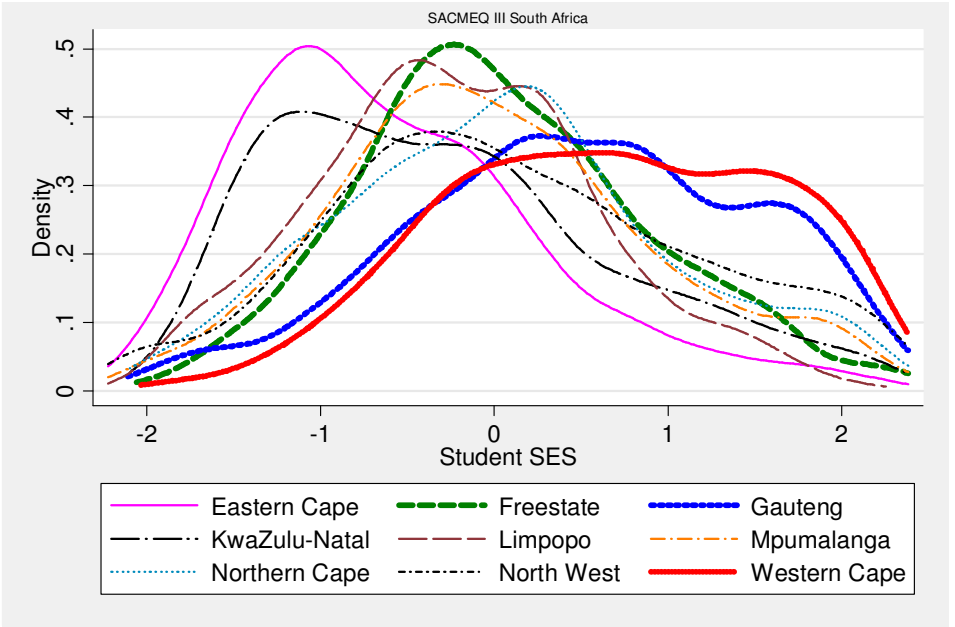


Figure 12: Kernel Density of Socioeconomic Status across Provinces

Figure 18: Kernel Density of Student Reading Performance by Socioeconomic Quintile

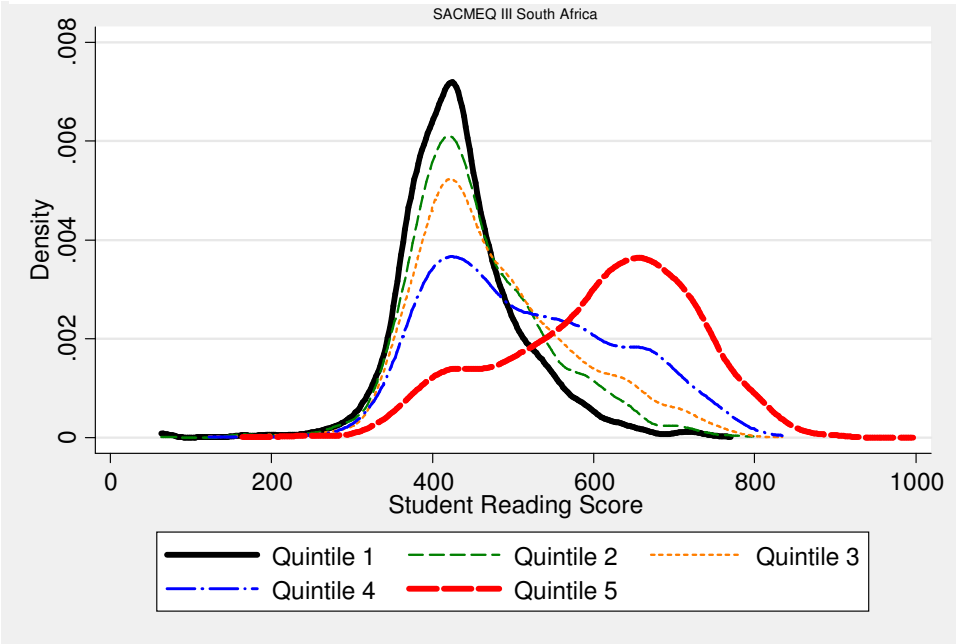
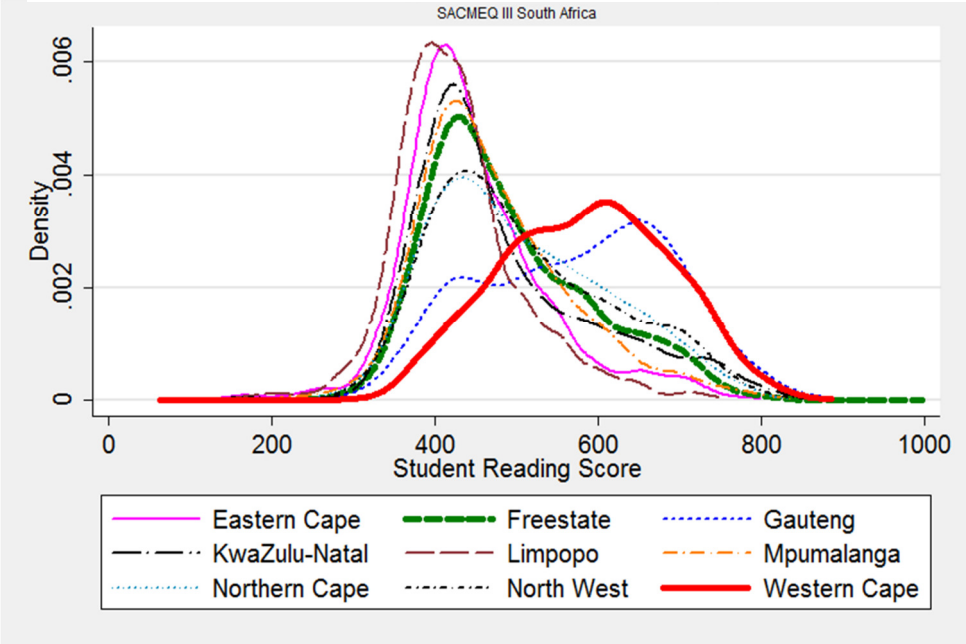


Figure 17: Kernel Density of Student Reading Performance by Province



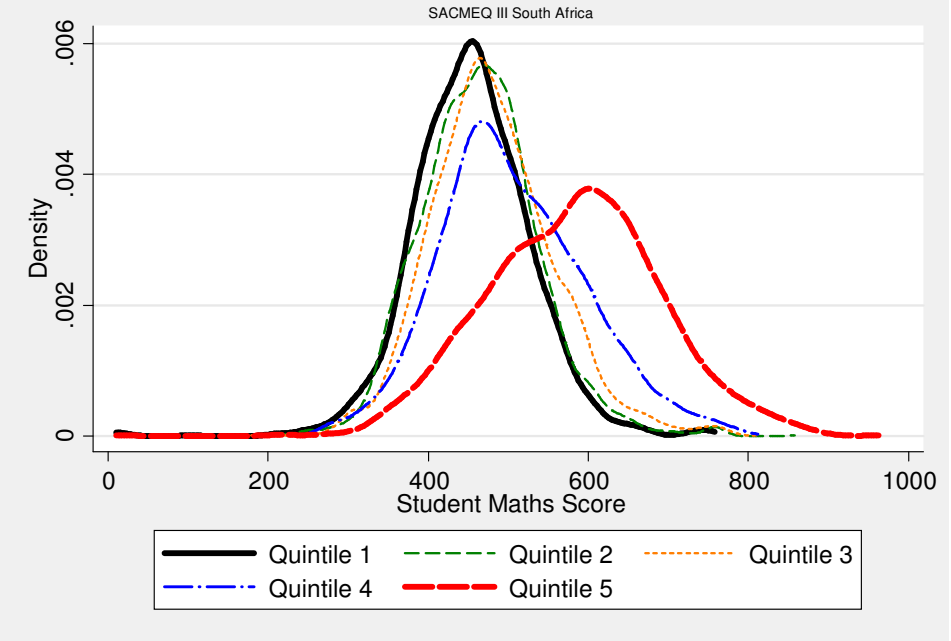


Figure 19: Kernel Density of Student Maths Performance by Socioeconomic Quintile

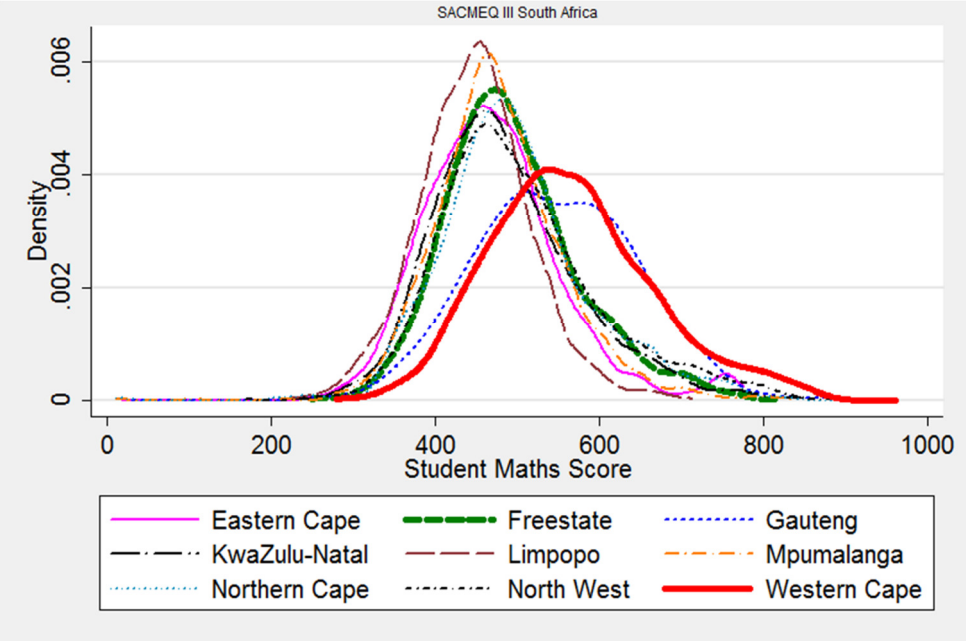
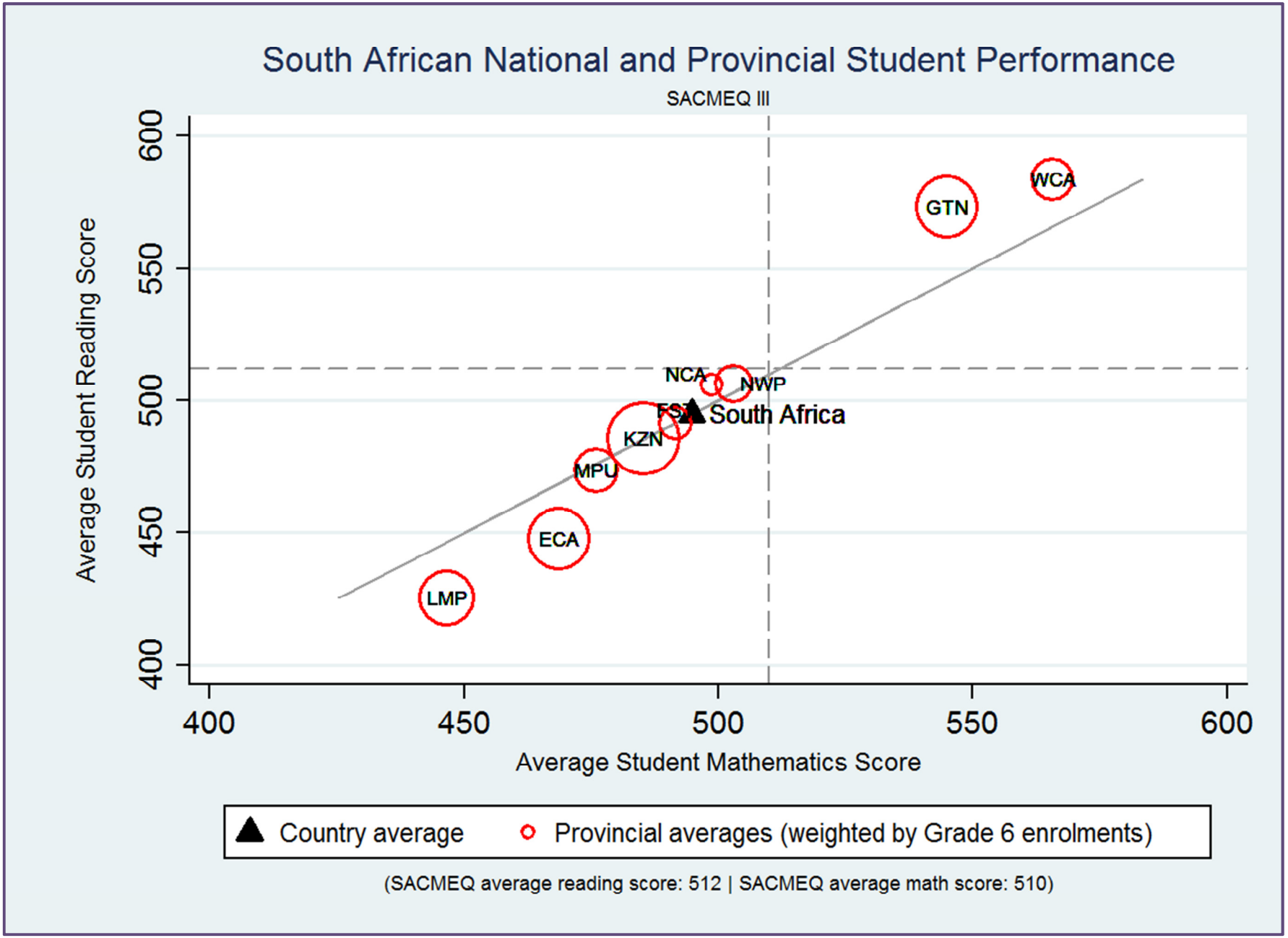


Figure 16: Kernel Density of Student Maths Performance by Province

Figure 20: South African National and Provincial Student Performance



Stellenbosch University http://scholar.sun.ac.za

3.1.2 Learner Competency Levels

In addition to the raw learner scores, SACMEQ calculated competency levels for reading and mathematics. By arranging test items in order of difficulty and then grouping these items by common themes, the SACMEQ team was able to construct eight levels of reading competency and eight levels of mathematics competency (Hungi, *et al.*, 2010). These range from Level 1 items which require only the most basic skills to answer correctly (such as *Pre numeracy* and *Pre literacy*), to Level 8 items which are more challenging and complex, and require higher order thinking and reasoning processes (such as *Critical reading* and *Abstract problem solving*).

In order to classify test items, it was necessary for the SACMEQ team to decide which skills were required to answer the various questions. Following this, the team had to group these skills into meaningful themes, which constitute the eight competency levels, and provide an overview of the skills required to successfully answer test items in that category. As one might expect, reaching consensus on exactly what thought processes were involved in correctly answering a particular test item is notoriously difficult. In conjunction with national curriculum experts, the SACMEQ team agreed on eight competency levels for mathematics, and eight for reading (Hungi, *et al.*, 2010). These were given meaningful names and descriptions in order to explain the underlying thought processes that were involved. A summary of the different skills associated with each competency level is provided in Table 2 (Reading) and Table 3 (Mathematics) below.

In order to calculate the score ranges for each competency level, Rasch analysis was used to calculate the difficulty of each item. This made it possible to match the ability of pupils with the difficulty of the test items (Ross, *et al.*, 2005, p. 94). By calculating the cumulative score based on the difficulty of the items which pupils answered correctly, pupils could be grouped based on their score, and the range within which that score fell. For example, a score of 400 in reading meant that the pupil was at the *Emergent Reading* level (Level 2).

Stellenbosch University http://scholar.sun.ac.za

Table 2: Levels of Reading Competency

Description of levels	Range on 500 point scale[16]	Skills
Level 1 *Pre-reading*	< 373	Matches words and pictures involving concrete concepts and everyday objects. Follows short simple written instructions.
Level 2 *Emergent reading*	373 → 414	Matches words and pictures involving prepositions and abstract concepts; uses cuing systems (by sounding out, using simple sentence structure, and familiar words) to interpret phrases by reading on.
Level 3 *Basic reading*	414 → 457	Interprets meaning (by matching words and phrases, completing a sentence, or matching adjacent words) in a short and simple text by reading on or reading back.
Level 4 *Reading for meaning*	457 → 509	Reads on or reads back in order to link and interpret information located in various parts of the text.
Level 5 *Interpretive reading*	509 → 563	Reads on and reads back in order to combine and interpret information from various parts of the text in association with external information (based on recalled factual knowledge) that "completes" and contextualizes meaning.
Level 6 *Inferential reading*	563 → 618	Reads on and reads back through longer texts (narrative, document or expository) in order to combine information from various parts of the text so as to infer the writer's purpose.
Level 7 *Analytical reading*	618 → 703	Locates information in longer texts (narrative, document or expository) by reading on and reading back in order to combine information from various parts of the text so as to infer the writer's personal beliefs (value systems, prejudices, and/or biases).
Level 8 *Critical reading*	703+	Locates information in a longer texts (narrative, document or expository) by reading on and reading back in order to combine information from various parts of the text so as to infer and evaluate what the writer has assumed about both the topic and the characteristics of the reader – such as age, knowledge, and personal beliefs (value systems, prejudices, and/or biases).

Source: (Hungi, *et al.*, 2010)

[16] See Ross *et al.* (2005, p. 95).

Stellenbosch University http://scholar.sun.ac.za

Table 3: Levels of Mathematics Competency

Description of levels	Range on 500 point scale[17]	Skills
Level 1 *Pre-numeracy*	< 364	Applies single step addition or subtraction operations. Recognizes simple shapes. Matches numbers and pictures. Counts in whole numbers.
Level 2 *Emergent numeracy*	364 → 462	Applies a two-step addition or subtraction operation involving carrying, checking (through very basic estimation), or conversion of pictures to numbers. Estimates the length of familiar objects. Recognizes common two-dimensional shapes.
Level 3 *Basic numeracy*	462 → 532	Translates verbal information presented in a sentence, simple graph or table using one arithmetic operation in several repeated steps. Translates graphical information into fractions. Interprets place value of whole numbers up to thousands. Interprets simple common everyday units of measurement.
Level 4 *Beginning numeracy*	532 → 587	Translates verbal or graphic information into simple arithmetic problems. Uses multiple different arithmetic operations (in the correct order) on whole numbers, fractions, and/or decimals.
Level 5 *Competent numeracy*	587 → 644	Translates verbal, graphic, or tabular information into an arithmetic form in order to solve a given problem. Solves multiple-operation problems (using the correct order of arithmetic operations) involving everyday units of measurement and/or whole and mixed numbers. Converts basic measurement units from one level of measurement to another (for example, metres to centimetres).
Level 6 *Mathematically skilled*	644 → 720	Solves multiple-operation problems (using the correct order of arithmetic operations) involving fractions, ratios, and decimals. Translates verbal and graphic representation information into symbolic, algebraic, and equation form in order to solve a given mathematical problem. Checks and estimates answers using external knowledge (not provided within the problem).
Level 7 Concrete problem solving	720 → 806	Extracts and converts (for example, with respect to measurement units) information from tables, charts, visual and symbolic presentations in order to identify, and then solves multi-step problems.
Level 8 *Abstract problem solving*	> 806	Identifies the nature of an unstated mathematical problem embedded within verbal or graphic information, and then translate this into symbolic, algebraic, or equation form in order to solve the problem.

Source: (Hungi, *et al.*, 2010)

By decomposing the test items according to the competencies required to answer them successfully, one is better able to understand learner performance. Furthermore, the differences in the abilities

[17] See (Ross, *et al.*, 2005, p. 95).

Stellenbosch University http://scholar.sun.ac.za

of various types of learners become more evident when they are presented as competencies acquired rather than simply test score achieved. As Hungi *et al.* (2010, p. 5) explain, "The eight competency levels provide a more concrete analysis of what pupils and teachers can actually do, and they also suggest instructional strategies relevant to pupils who are learning at each level of competence."

Observing Figure 21 to Figure 27 below, it soon becomes clear that learner competencies in both mathematics and reading are very different for each of the different sub-categories. Figure 25 and Figure 26 show once again that the Western Cape and Gauteng provinces far outperform the other seven, most notably in reading. Similarly, learners from large cities have acquired a significantly larger repertoire of academic skills than those learners who are situated in rural and isolated areas. This contrast within provinces and geographies is all the more evident when the sample is split by school socioeconomic status. The vast majority of quintile five learners have mastered the basic skills of reading, and to a lesser extent mathematics, and a significant proportion have mastered the higher order competencies in reading. The cumulative distribution functions for reading and mathematics competency levels (Figure 24 and Figure 23) show this stark contrast between the skills acquired by rich and poor.

Figure 22: Reading Competency Levels by School Socioeconomic Status

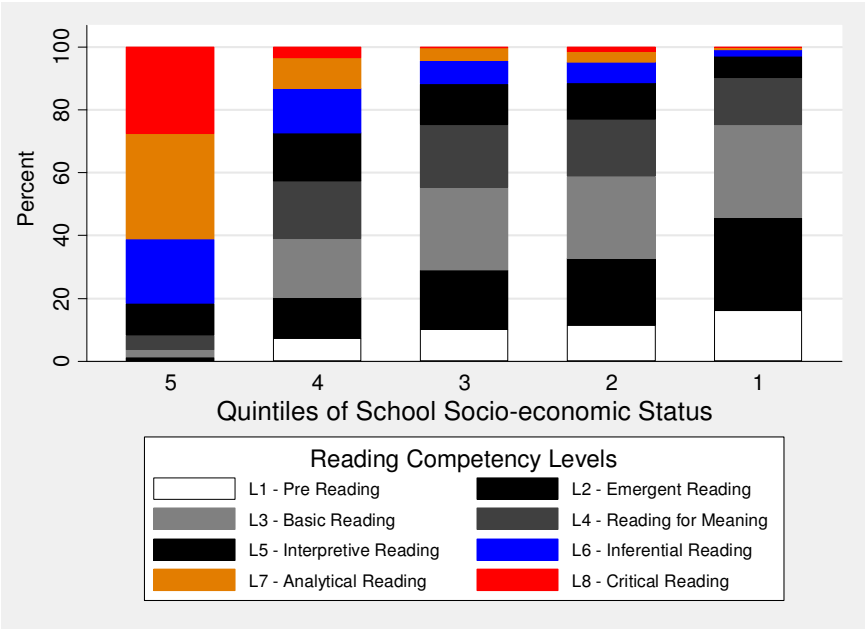
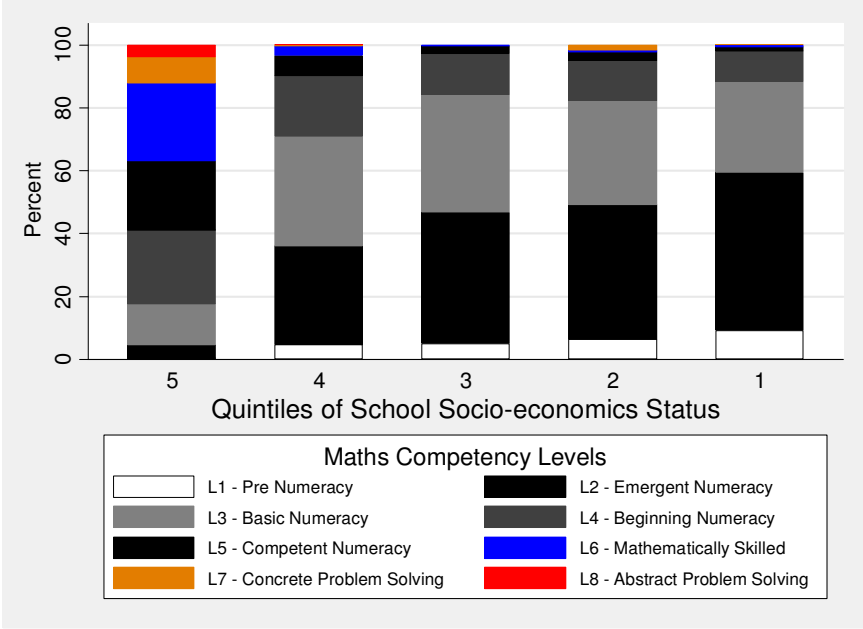


Figure 21: Maths Competency Levels by School Socioeconomic Status



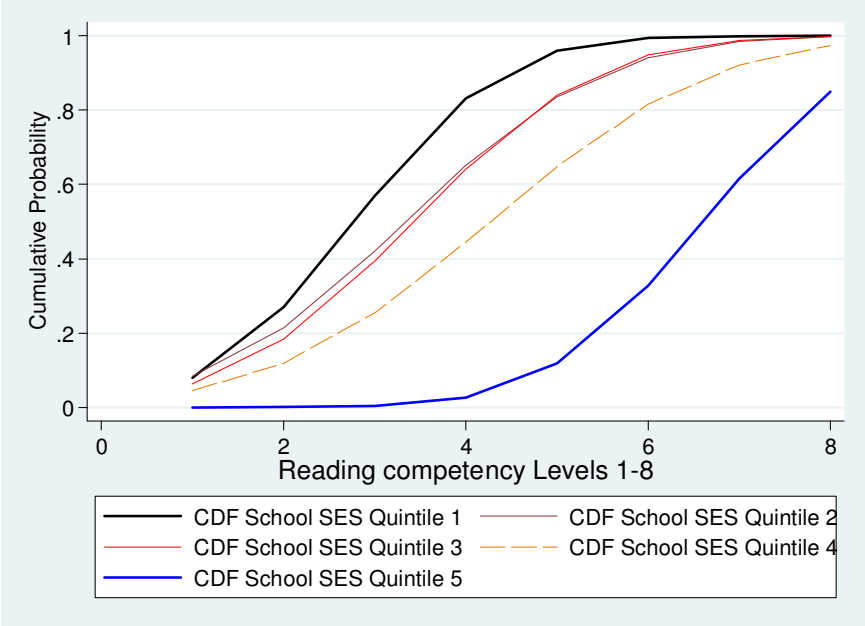


Figure 23: Reading Competency Levels by School Socioeconomic Status (CDF)

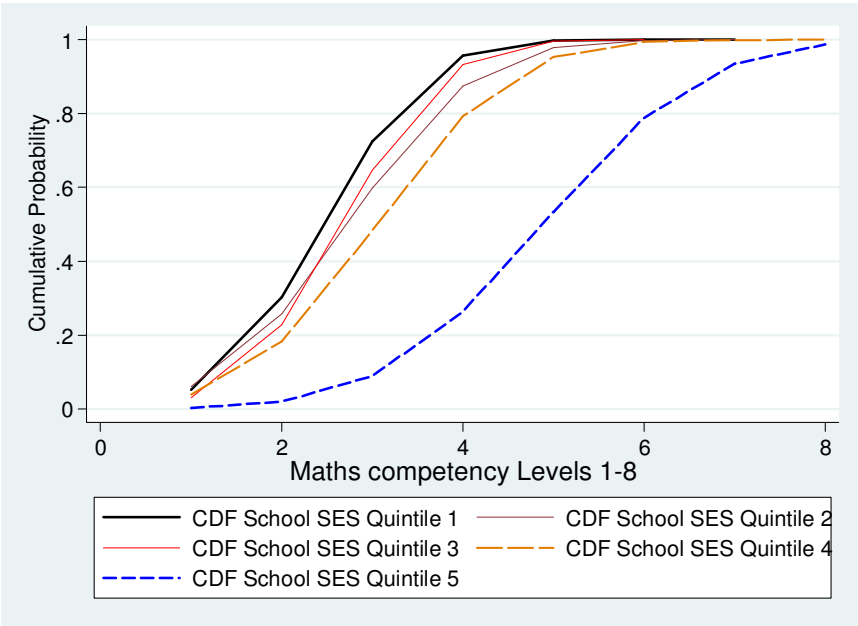


Figure 24: Maths Competency Levels by School Socioeconomic Status (CDF)

Figure 26: Reading Competency Levels by Province

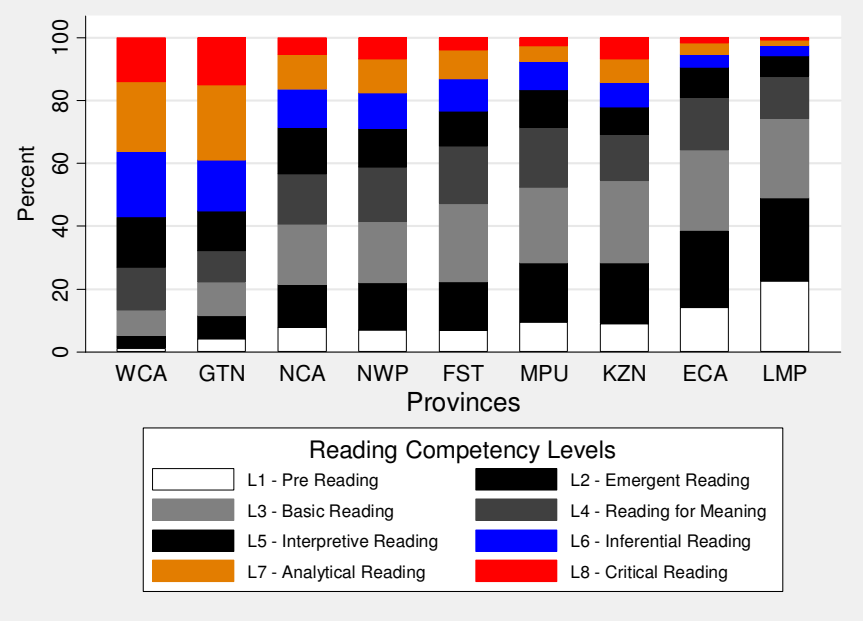
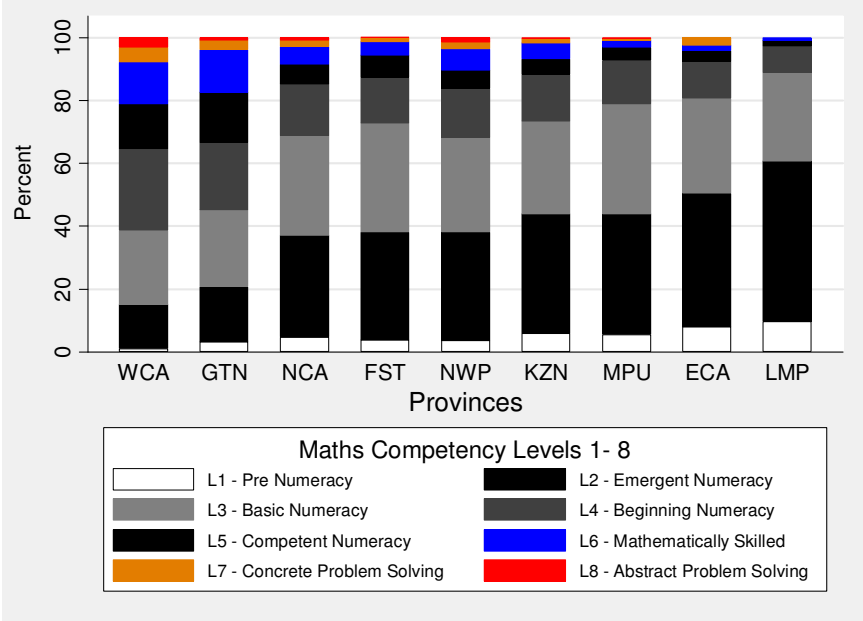


Figure 25: Maths Competency Levels by Province



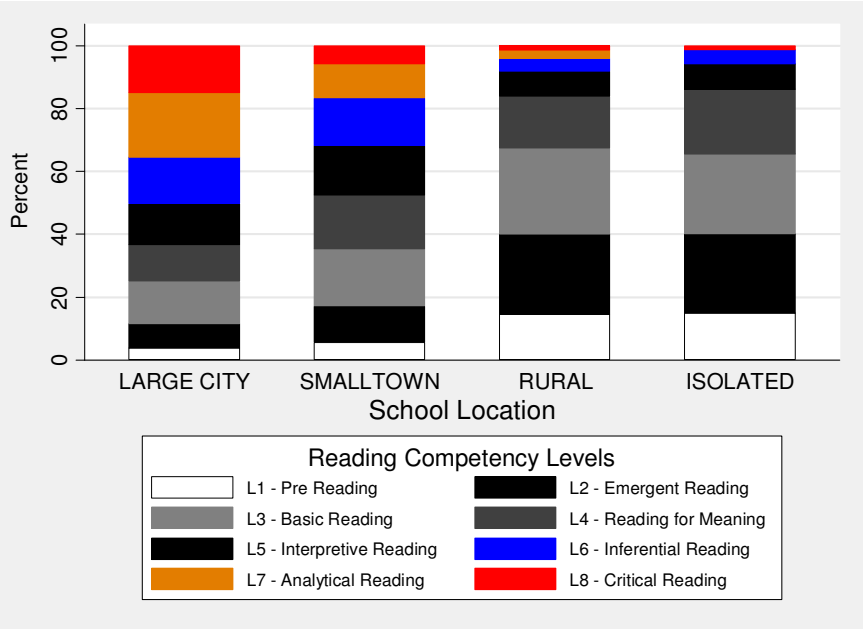


Figure 27: Reading Competency Levels by School Location

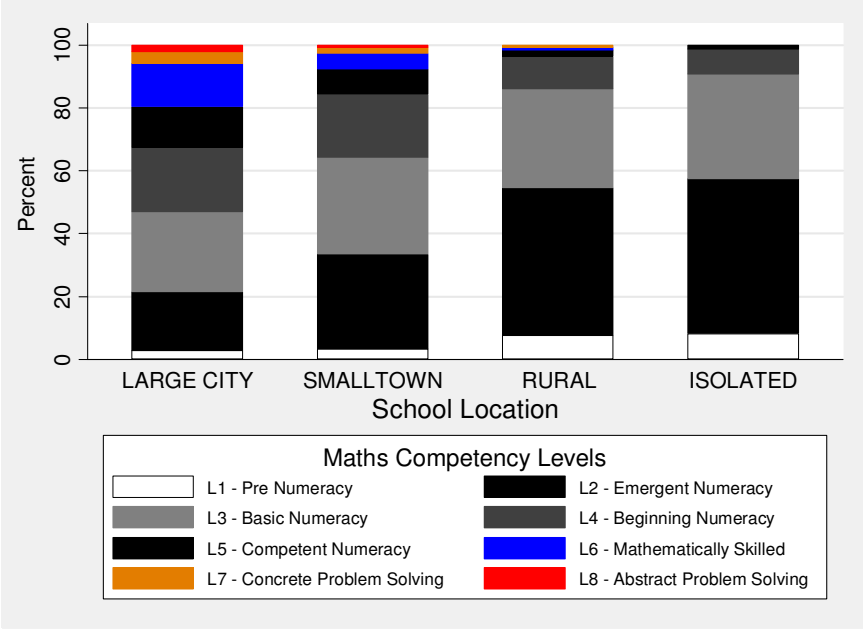


Figure 28: Maths Competency Levels by School Location

Stellenbosch University http://scholar.sun.ac.za

What is perhaps most egregious about the above distributions is the large number of learners failing to acquire even the most basic numeracy and literacy skills. If one takes the lowest two competency levels for reading (*pre literacy* and *emergent literacy*) and mathematics (*pre numeracy* and *emergent numeracy)* as a lower-bound threshold, there are still many learners falling below this elementary threshold. Shabalala (2005) places these two lower bound thresholds in perspective by explaining their practical implications:

> "The two lower levels of reading competence are concerned with 'pre-reading' and 'emergent reading'. Pupils at these two levels should be able to undertake simple decoding tasks and match words to pictures and very simple phrases. However, neither of these levels requires pupils to read even simple sentences in order to extract meaning. Therefore, pupils at these lowest two levels could be categorised as 'non-readers' in the sense that they cannot 'interpret meaning in a short and simple text'" (Shabalala, 2005, p. 222).

Similarly for numeracy:

> "Pupils at the lowest levels of numeracy competence ('pre- numeracy' and 'emergent numeracy') … are only able to count, recognise shapes and numbers, carry out simple operations, and link simple verbal and graphic forms with simple arithmetic operations. Neither of these two levels requires pupils to work with three-dimensional shapes, use multi-step arithmetic operations, or undertake conversions using division. Therefore pupils at these lower two levels could be categorised as 'non-numerate' in the sense that they have not moved beyond the mechanical skills related to basic calculation and simple shape recognition" (Shabalala, 2005, p. 225).

Continuing with Shabalala's distinctions, this thesis uses these thresholds as benchmarks for functional illiteracy and functional innumeracy. These terms are used to indicate whether an individual has acquired sufficient numeracy and literacy skills such that he or she is able to satisfactorily *use* those skills in everyday life. It is of little use if children can write down and read a memorised paragraph if they do not understand what they are reading or writing. Similarly, if children cannot relate basic arithmetic skills into real world situations, it is questionable whether they have actually acquired those skills.

It is revealing to calculate the percentage of Grade 6 South African learners who can be classified as functionally illiterate and functionally innumerate per province, school location, and school socioeconomic quintile (Table 4 below). Observing the national averages, it is disconcerting to see that 40.2% of South African Grade 6 learners are functionally innumerate, and 27.2% are functionally illiterate. As with all South African data, these averages shroud the severe inequalities between sub-groups of learners. For example, only 5.1% of learners in the Western Cape are functionally illiterate, and only 1.4% of quintile five learners are functionally illiterate. This is in stark contrast to the Eastern Cape where 38.6% of learners are functionally illiterate, and the Limpopo province where

Stellenbosch University http://scholar.sun.ac.za

the figure is 49%. Almost half (44.7%) of all quintile one learners are classified as functionally illiterate, and more than half (58.7%) of learners in quintile one schools are functionally innumerate.

For every sub-group, there are more learners classified as functionally *innumerate* than those classified as functionally *illiterate*. There is an alarmingly large percentage of functionally innumerate learners across all regions, all school locations, and all socioeconomic quintiles (except quintile 5). The specific percentages of functionally innumerate learners for selected sub-groups are: Limpopo (60.6%), Eastern Cape (50.3%), KwaZulu-Natal (44%), Mpumalanga (43.8%), Rural schools (55.2%), Quintile one schools (58.7%), Quintile two schools (48.9%) and Quintile three schools (47.4%).

Table 4: Percentage of Learners who are Non-numerate and Non-readers per Province, School Location, and School Quintile

Provinces	**% Functionally Illiterate**	**% Functionally Innumerate**
Eastern Cape	38.6	50.3
Freestate	22.3	38.1
Gauteng	11.6	20.5
KwaZulu-Natal	28.4	44
Limpopo	49	60.6
Mpumalanga	28.4	43.8
Northern Cape	21.4	37.1
North West	21.9	38.1
Western Cape	5.1	15
Total	**27.2**	**40.2**
Quintiles of School SES	**% Functionally Illiterate**	**% Functionally Innumerate**
Quintile 1	44.7	58.7
Quintile 2	34.4	48.9
Quintile 3	30.4	47.4
Quintile 4	20.1	35.4
Quintile 5	1.4	4.6
Total	**27.2**	**40.2**
School Location	**% Functionally Illiterate**	**% Functionally Innumerate**
Isolated	38.8	56.2
Rural	41.3	55.2
Small town	16.7	32.9
Large city	11.4	20.7
Total	**27.2**	**40.2**

Stellenbosch University http://scholar.sun.ac.za

From an educational perspective it is important to realise that a large number of children - particularly those from disadvantaged backgrounds - acquire learning deficits very early in their educational careers. Given that education is a cumulative process, these deficits in numeracy and literacy are likely to stay with these children for the rest of their lives. As time passes and children proceed to higher grades, both teachers and the curriculum presume that children have acquired the skills taught in previous grades. However, grade progression is often not determined solely, or even mainly, by skills acquired but rather by the need to maintain steady grade-enrolments and 'normal' pass rates. Also, teachers often do not know what the appropriate level of testing is for the grade which they teach, and thus mistakenly believe that their learners have attained this artificially low benchmark. Consequently learners (and parents) do not realise the extent of their own (or their children's) underperformance until it is too late to do anything about it – usually matric, which is often the first examination these learners write which is set at the appropriate level. The recently implemented Annual National Assessments (ANA) which tested Grades 1-6[18] using a standardised national test may go some way to rectifying this situation.

The fact that a large proportion of South African children are not acquiring even the most basic numeracy and literacy skills in six years of full-time schooling is, from an efficiency perspective, a serious failure on the part of teachers and the Department of Basic Education as a whole, and begs the question: How difficult is it to teach children to read? Whether this situation is due to an inappropriate curriculum, a lack of opportunity to learn, poor teaching methods, or a host of other factors, is important from a policy perspective, but immaterial from an accountability perspective. The fact remains that in 2007 almost one in three South African primary school children could not read for meaning. Given that the various provincial Departments of Education have been mandated by the government to provide education to South African children, the blame for this inexcusable situation is firmly and squarely located with the national Department of Basic Education, as well as the various provincial Departments themselves. While it is true that the government cannot change the socioeconomic status of learners from poor schools, the fact that it is able to do so little given the large amount of resources and time allocated to primary schooling indicates a fundamental inefficiency and inability to find methods that work. It is disappointing to see that since the dawn of democracy in South Africa, the government has lacked the political will for serious reform in the primary education sector.

[18] The Universal ANA tested all learners in Grades 2-7, examining content from the previous grade (Grades1-6). The Verification ANA applied more rigorous procedures to a sample of 1800 Grade 3 and Grade 6 schools in order to verify the results from the Universal ANA (DBE, 2011c, p. 5)

Stellenbosch University http://scholar.sun.ac.za

3.1.3 Selected Learner Characteristics

Socioeconomic Status (SES)

The positive impact of socioeconomic status on learner performance has been well established in both the local and international economics of education literature. Learners from more affluent backgrounds perform better academically than those from poorer households. The term socioeconomic status usually refers to a concept that is more nuanced than simply material resources. Willms (1997, p. 24) defines SES as "A person's access to, and control over, wealth, prestige and power. It is typically measured through factors such as income, the prestige of the person's occupation and their level of education." Thus SES usually includes notions of class, and other ways of explaining the ordinal structure of societal advantage. In South Africa, more so than most developing countries, there is a very strong correlation between material wealth (income) and social class (occupation, education and power). This is largely as a result of the policies of apartheid where spatial segregation, labour-market discrimination, and general race discrimination led to a polarization of social class – especially along racial lines. While this has changed somewhat since the end of apartheid, there is still a highly significant relationship between economic wealth and social class. This relationship is important when justifying why an economic index can proxy for a socioeconomic index in this thesis.

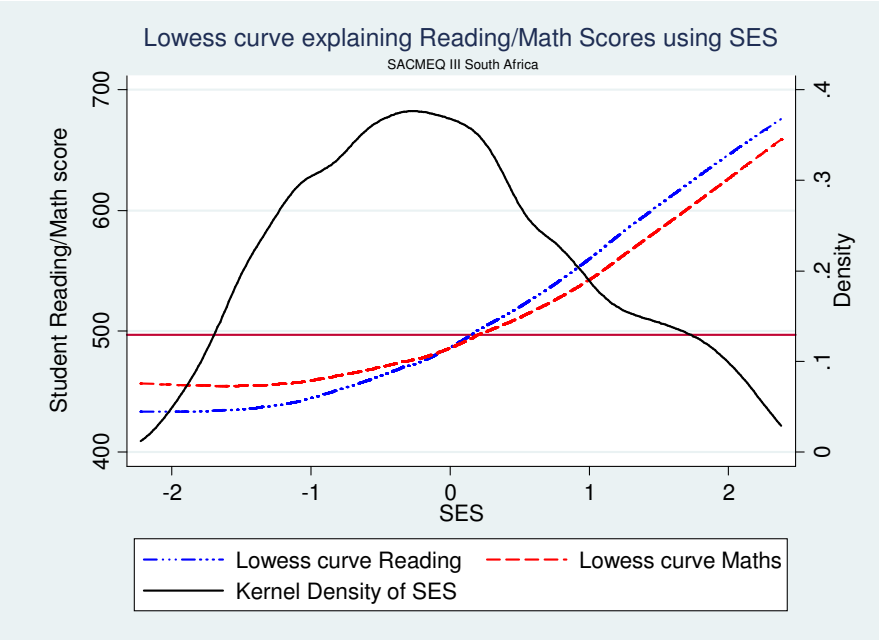
Although parental education is usually included in an SES index, in this thesis parental education is excluded. The reason for this was to determine if there was an impact of parental education on learner performance *over and above* the impact of material wealth. This would not be possible in the multivariate analysis if parental education was included in the SES index. Taylor and Yu (2009, p. 62) provide a comprehensive explanation for excluding parental education in the SES index. Given that the SES index used here only includes material possessions, it is prudent to ask why it has been termed a *socio*economic index rather than simply an *economic* index. The reason for this is found in the strong correlation between social class and material wealth in the South African context. While it would perhaps be more technically correct to interpret this variable as the impact of material wealth alone, the abovementioned relationship between material wealth and socioeconomic status suggests that the impact of this economic variable is likely to be conflated with noneconomic benefits associated with social class. If true, ascribing the positive impact of additional resources solely to economic resources would be misleading. Therefore it was deemed more appropriate to interpret the variable as socioeconomic status.

To further understand the links between socioeconomic status and learner performance, SES was used as the explanatory variable on learner mathematics and reading performance in two locally

Stellenbosch University http://scholar.sun.ac.za

weighted regressions (Lowess). These are shown in Figure 29 below. The shapes of both curves suggest that learner SES is only significantly positively related to learner performance at higher levels of SES. The kernel density of SES was superimposed on the Lowess curves to show that only a relatively small sub-sample of learners fall in the range where additional SES is beneficial. To be specific, the fourth quintile of SES begins at an SES value of 0.19 and the fifth quintile at 0.89. Thus, the majority (60%) of learners lie to the left of an SES value of 0.19 in the graph below, i.e. the flatter portion of the Lowess curves. It is interesting to note that this threshold area corresponds to a learner reading/mathematics score of 500, which is approximately the SACMEQ mean.

Figure 29: Lowess Regressions - Reading and Maths Performance



Importantly, the Lowess regressions above do not control for factors other than socioeconomic status. Thus one should interpret these results with caution since socioeconomic status in South Africa is also highly correlated with linguistic advantage, preschool exposure and school functionality, for example. In a sense, all that the Lowess regressions show is the relationship between these two variables (performance and SES) without parameterising the relationship. This has the benefit of "letting the data speak" and avoiding the possibility that the results are driven by a particular regression specification rather than the underlying data or processes.

Age of learners

There is reason to believe that learners who are younger or older than their age-appropriate peers will perform worse in tests of academic ability. Learners who enter the education system late may find it difficult to adjust to a cohort that is sometimes significantly younger than they are. Similarly,

Stellenbosch University http://scholar.sun.ac.za

young learners may not fit in as well with their older peers, or even be at the same level of biological cognitive development. Learners who repeat grades as they progress through school will be older than the majority of their non-repeating peers. This too is likely to affect their performance, since grade repetition often does not improve academic performance, and learners carry through their educational deficit to the following grade.

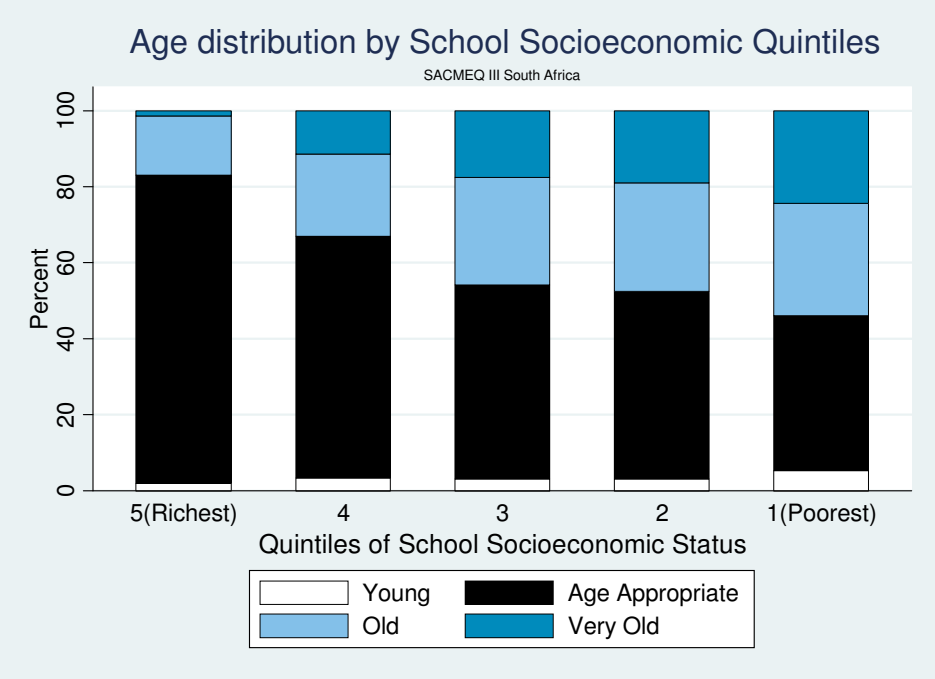
The age variables constructed for this analysis were in accordance with South African educational policy. According to the 2002 amendments of the South African Schools Act (2002, Section 5) a child may be admitted to Grade 1 if he or she is five turning six by 30 June in the year of admission, or he or she must wait until the following year to be admitted. Consequently, children can be deemed age-appropriate if they are between five years six months and seven years old at the beginning of Grade 1. Thus in Grade 6, the age range would be between 10 years six months and 12 years old. Given that the SACMEQ survey was adminsitered in September, the age range for learners writing the SACMEQ tests would be between 11 years and three months and 12 years and nine months. The age categories were calculated as young (less than 11 years 3 months), age appropriate (11 years 3 months – 12 years 8 months), old (12 years 8 months – 13 years), and very old (14 years and older).

The distribution of age-appropriate learners across the five quintiles of school socioeconomic status illustrates the stark differences between Grade 6 classes in each of the five quintiles. While more than 80% of learners in quintile five schools are at the appropriate age, only half as many (approximately 40%) are at the appropriate age in quintile one schools (Figure 30 below). It would also seem that very few learners are young for their cohort, and relative to overage, this does not seem to be a problem.

If one separates the school system into two categories, only 1.3% of quintile five learners are *very old* while almost one in five (18.6%) quintile one to four learners are.

Stellenbosch University http://scholar.sun.ac.za

Figure 30: Age Distribution by School Socioeconomic Quintiles



Nutrition

The importance of nutrition for cognitive development has been well established in the literature. Del Rosso (1999, p. 5) provides a concise summary of the impact of poor nutrition:

> "Children who lack certain nutrients in their diet (particularly iron and iodine), or who suffer from protein-energy malnutrition, hunger, parasitic infections or other diseases, do not have the same potential for learning as healthy and well-nourished children. Weak health and poor nutrition among school-age children *diminish their cognitive development* either through physiological changes or by reducing their ability to participate in learning experiences – or both … Children with diminished cognitive abilities and sensory impairments naturally perform less well and are more likely to repeat grades and drop out of school than children who are not impaired; they also enrol in school at a later age, if at all, and finish fewer years of schooling."

Following from the above, it is instructive to look at the incidence of missed meals for each of breakfast, lunch and supper. Table 30 in Appendix C provides the percentages of learners who reported normally eating each of these three meals every day, three to four times per week, one to two times per week, or not at all. From these variables, three other variables were created for the regression analysis of Chapter 4: *missed breakfast*, *missed lunch* and *missed dinner.* Each of these is a dummy variable taking on the value one if a learner reported that they normally missed that meal at least once per week.

The Diagnostic Statistics in Appendix E illustrate that missing either breakfast or lunch at least once per week is quite common, while missing dinner at least once per week is less common. Only 5.6% of

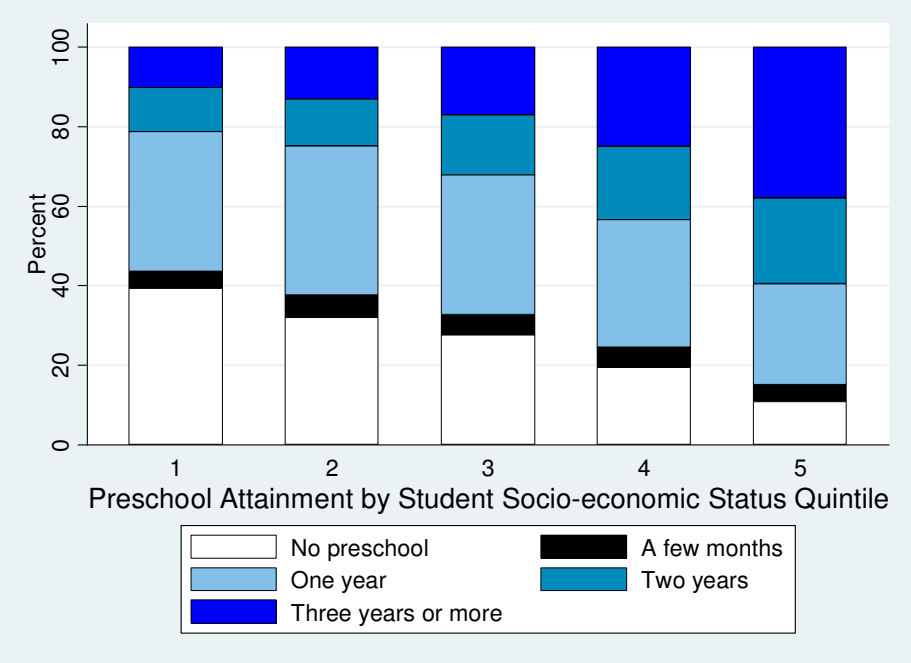
Stellenbosch University http://scholar.sun.ac.za

learners attending quintile five schools reported that they normally missed dinner at least once per week, while 14.4% of learners in quintile one to four schools did. Similarly, 11.8% of learners in quintile five schools reported normally missing lunch at least once per week, while almost twice as many learners (22.4%) in quintile one to four schools did.

Preschool Education

The importance of preschool education for future academic success has been increasingly stressed in the international literature. Early childhood exposure to pedagogical resources and enrolment in preschool education is associated with learning gains in primary schooling. Given that this is so, it is worrying to see the strong correlation between the level of preschool education and family wealth, as measured by learner socioeconomic status (Figure 31). As one would expect, poorer quintiles have less preschool education and wealthier quintiles have more preschool education. Almost 40% of learners in the poorest quintile receive no preschool education whatsoever[19].

Figure 31: Preschool Attainment by Learner Socioeconomic Status Quintile



The preschool question in the SACMEQ III study was included in the learner questionnaire. It is interesting to compare these results to those found by Howie *et al.* (2008) in the PIRLS 2005/6 study. In the PIRLS study, parents were asked to complete a questionnaire and indicate how many years of preschool education their children had received. It was found that 86% of learners had attended at

[19] It should be noted, however, that the situation has improved somewhat since 2007, with increased access to pre-school education seen across the country, and specifically on the part of poorer learners (Gustafsson, Policy note on pre-primary schooling: An empirical contribution to the 2009 Medium Term Strategic Framework, 2010).

Stellenbosch University http://scholar.sun.ac.za

least one year of preschool education (Howie, *et al.*, 2008, p. 39). By contrast, only 68.8% of learners in the SACMEQ III survey indicated that they had had at least one year of preschool education. This is quite surprising since the SACMEQ III study was conducted one and a half years after the PIRLS study. One would expect there to be more learners with preschool exposure arising from government initiatives to expand Early Childhood Development.

Textbook availability

Textbooks are a fundamental resource to both teachers and learners. Teachers can use textbooks for lesson-planning purposes, as a source of exercises and examples, and also as a measure of curriculum coverage. Learners can use textbooks to 'read-ahead' if they have sufficiently mastered the current topic, preventing gifted learners from being held back. Textbooks can, to a certain extent, also mitigate the effect of a bad teacher since they facilitate independent learning.

The problem of a lack of textbook access is now commonly accepted in the South African research literature. For example, in Hoadley's (2010, p. 11) review of the classroom-based literature research in South Africa, she finds that one of the dominant descriptive features of primary schools is "a lack of print materials in classrooms, especially textbooks." Figure 32 and Figure 33 below show the distributions of reading and mathematics textbooks across each of the nine provinces. KwaZulu Natal has the lowest textbook access for both reading and mathematics. In all provinces except Mpumalanga and the Free State, more than 20% of learners are in classrooms where there are no mathematics textbooks, or where only the teacher has a textbook.

Stellenbosch University http://scholar.sun.ac.za

Figure 32: Reading Textbook Availability by Province

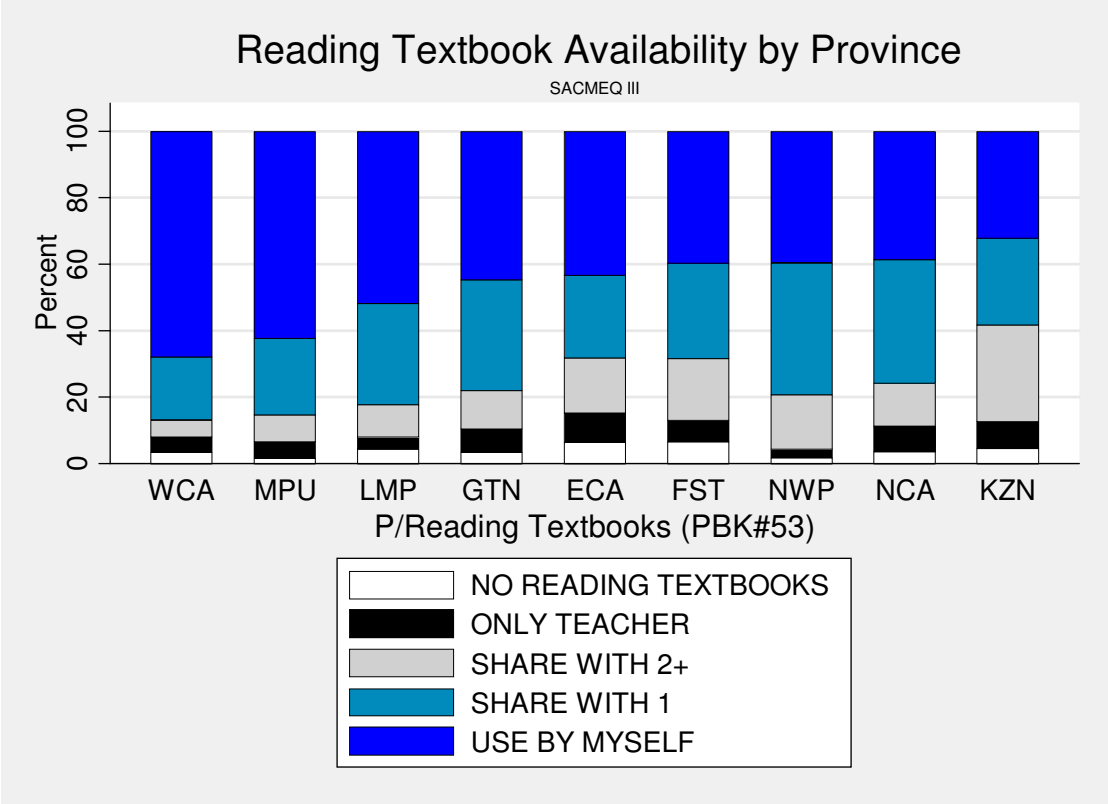
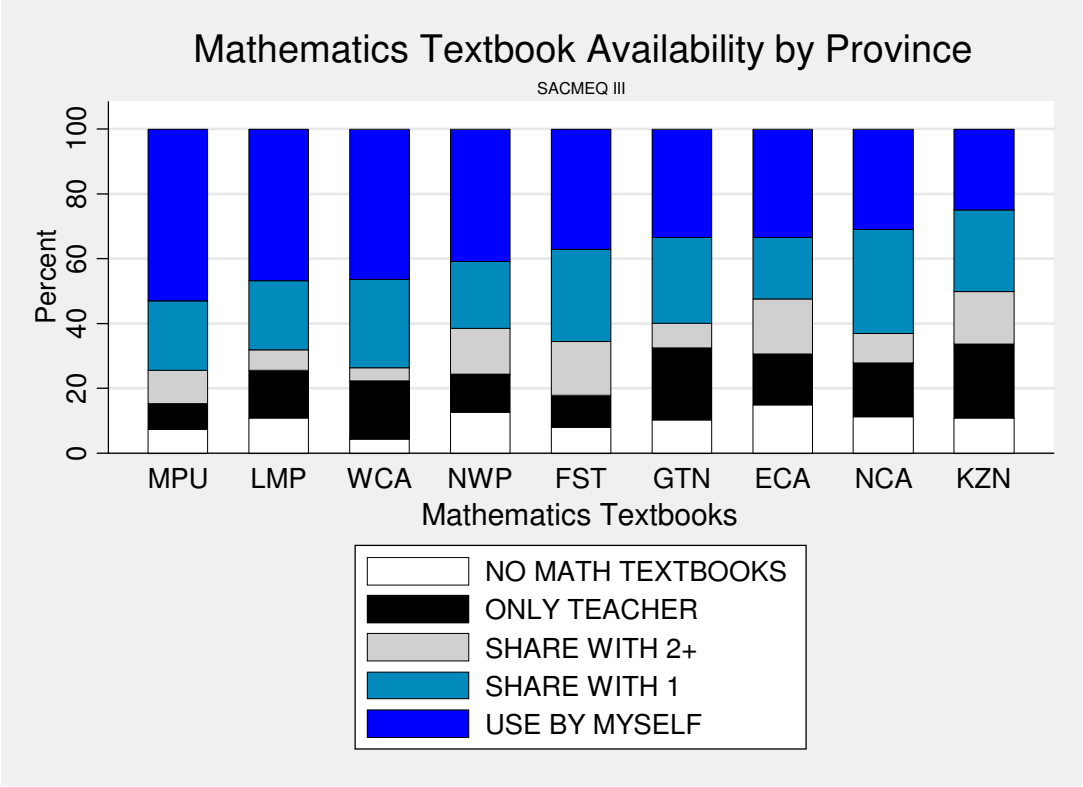


Figure 33: Mathematics Textbook Availability by Province



Stellenbosch University http://scholar.sun.ac.za

Grade Repetition

Excessive grade repetition is a common problem in many developing countries, and South Africa is no exception. Learners who do not achieve the required grades are held back and repeat that same grade in the hope that they will acquire the skills the second time around. However, in a report on learner retention commissioned by the Department of Education, a team of experts concluded that:

> "Grade repetition is generally ineffective as an intervention to address learning problems, regardless of when the repetition occurs. Learners repeating grades should have special programmes that are not a mere repetition of the material and content covered during the first year in the grade" (DoE, 2008b, p. xix).

The point is well taken that repetition is likely to produce more of the same results (failure) unless it caters to the specific reasons *why* learners did not acquire the skills they were meant to during the first year in that grade. The solution to grade repetition is therefore not to simply push learners through to higher grades when they clearly lack the capabilities necessary to acquire the skills and knowledge - and thus benefit - from that level of education, but rather some form of remedial education.

In the SACMEQ III survey, learners were asked how many times they had repeated a grade (including Grade 6, the current year) since they started school. Table 5 below shows the distribution of grade repetition for each of the five quintiles of school socioeconomic status. Clearly learners who attend top quintile schools repeat fewer grades, and are far less likely to repeat more than one grade by Grade 6. Roughly 10% of all learners in quintile one to four schools have repeated at least 2 grades, compared to only 1% of learners in quintile five schools. Some amount of grade repetition is not, in and of itself, a bad thing. Given that the figures in Table 5, below, are cumulative (i.e. for all grades up to Grade 6), it would not be worrisome to see that 20% of learners had repeated only one grade since this would amount to only 3% grade repetition for any one year in the previous six grades. However, a significant number of learners in poorer schools are repeating more than one grade, with only 65% reporting that they had never repeated a grade. The SACMEQ data thus seems to reiterate the widely held notion that grade repetition is a problem in South Africa. However, given that the SACMEQ questionnaire did not ask learners *which* grade they repeated, one cannot draw strong conclusions from the data. This is mainly because excessive grade repetition in Grade 1 has more to do with admission policies than educational factors; i.e. admitting learners who are too young and then allowing them to repeat Grade 1.

Stellenbosch University http://scholar.sun.ac.za

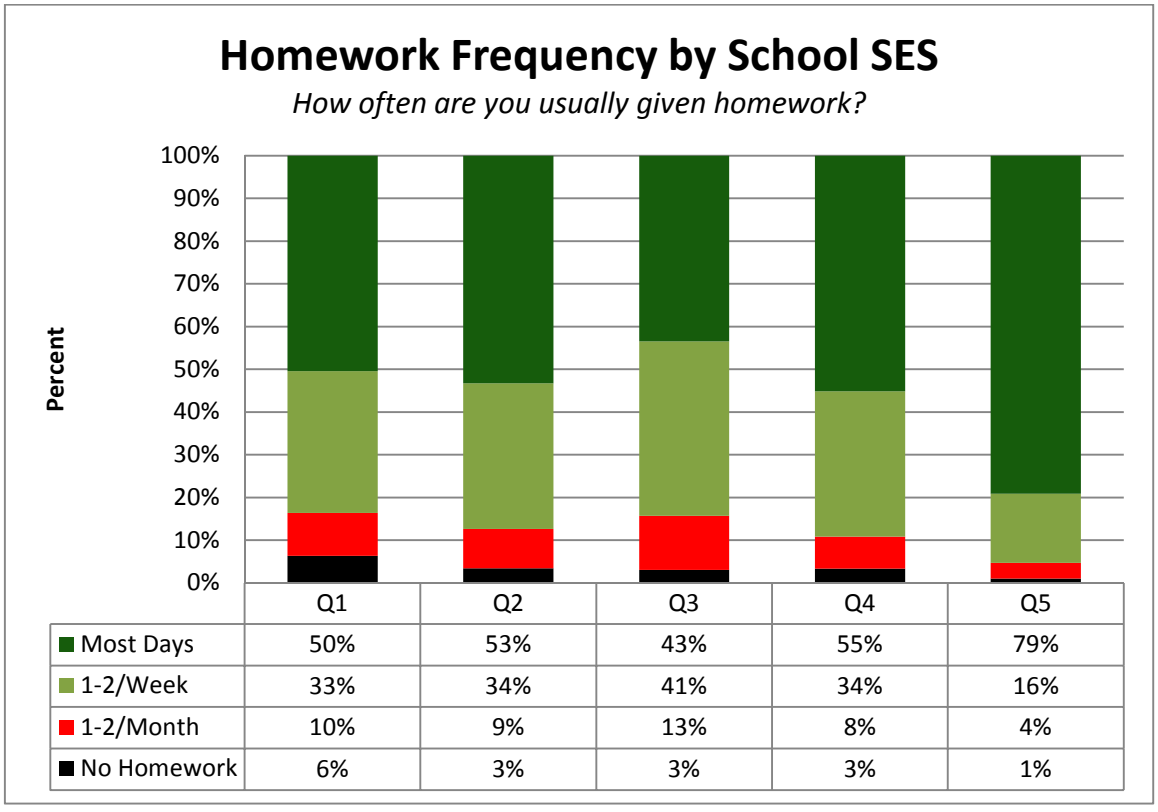
Table 5: Grade Repetition by School Socioeconomic Status

		Quintiles of School SES					
		Q1	**Q2**	**Q3**	**Q4**	**Q5**	**Total**
How many times have you repeated a grade since you started school including Grade 6?	Never	65%	69%	66%	75%	85%	72%
	Once	24%	22%	24%	18%	13%	20%
	Twice	7%	6%	7%	4%	1%	5%
	Three +	4%	3%	4%	3%	0%	3%
	Total	100%	100%	100%	100%	100%	100%

Homework Frequency

According to the received wisdom in the pedagogical literature, practice is imperative for student learning. Therefore, it seems logical that the benefits of homework are derived because learners are practising to read and practising mathematical problems and that this process improves learning. It is unfortunate then, that poorer schools are far less likely to prescribe homework than wealthier schools. More than ten percent of learners in quintile one to four schools received homework either once or twice per month, or not at all (Figure 34, below). This is compared to five percent of learners in quintile five schools. Similarly, 79% of learners in quintile five schools reported receiving homework *most days of the week* compared to only 50% (quintile 1), 53% (quintile 2), 43% (quintile 3), and 55% (quintile 4).

Figure 34: Homework Frequency by School Socioeconomic Status



Stellenbosch University http://scholar.sun.ac.za

Learner HIV/AIDS knowledge

In addition to reading and mathematics tests, SACMEQ III included an HIV/AIDS Knowledge Test (HAKT). Although the health knowledge of learners is not the focus of this thesis, given the importance of HIV/AIDS in the South African context and the availability of the HAKT test, it was decided to include a brief account of learner HIV/AIDS knowledge, at least in this descriptive chapter.

The SACMEQ research teams used the official HIV/AIDS curricula of all participating SACMEQ countries in the development of the HAKT. The test consisted of 86 true-or-false questions which covered five main knowledge domains: definitions and terminology, transmission mechanisms, avoidance behaviours, diagnosis and treatment, and myths and misconceptions (DBE, 2011b). Learners were asked to indicate if a variety of statements were true or false, with statements ranging in complexity from statements which were relatively simple such as "*A person can get HIV by swimming in dirty rivers*" to more complex statements, such as "*HIV can be found in mother's breast milk*" and "*A test for HIV is detecting HIV antibodies*".

From the 86 question HAKT, the SACMEQ researchers created two levels of performance: "minimum level of HIV/AIDS" and "desirable level of HIV/AIDS". Learners are regarded as reaching the minimum level when they answer at least 50% of the questions correctly. Similarly, learners have a desirable level of knowledge if they answer at least 75% of the questions correctly. In their report on the SACMEQ HAKT test, the Department of Basic Education elaborates further on these two thresholds:

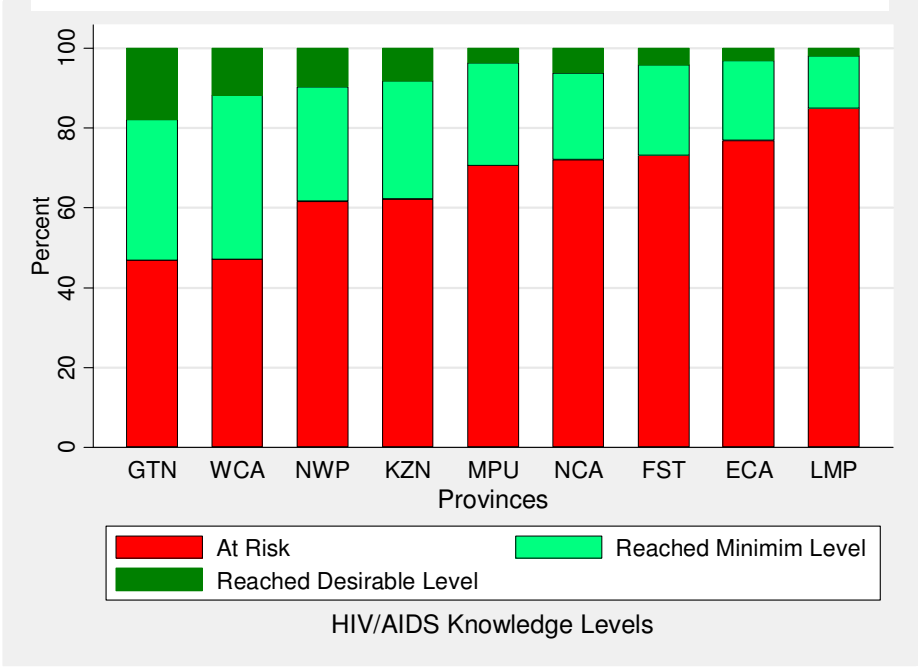
> "Performance below 50% in the test...indicated that the respondent lacked knowledge of basic but essential facts about HIV and AIDS. A 'desirable level', on the other hand, was defined as a mean score of 75%. Performance at this level and above would include sufficient knowledge to inform sound decisions on critical facts about HIV and AIDS" (DBE, 2011b, p. 5).

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

Stellenbosch University http://scholar.sun.ac.za

Figure 35, Figure 36, and Figure 37 below show that the vast majority of South African Grade 6 learners who participated in the SACMEQ survey do not know the "basic but essential facts about HIV and AIDS" (DBE, 2011b, p. 5), and as such, are at risk. This is in stark contrast to the *a priori* expectations of the South African SACMEQ National Research Coordinators who state that: "It was expected that 100% of learners in all SACMEQ countries should reach the minimal knowledge level" (Moloi & Chetty, 2011, p. 3).

Figure 35: HIV/AIDS Knowledge Level by Province



Stellenbosch University http://scholar.sun.ac.za

Figure 37: HIV/AIDS Knowledge Level by School Location

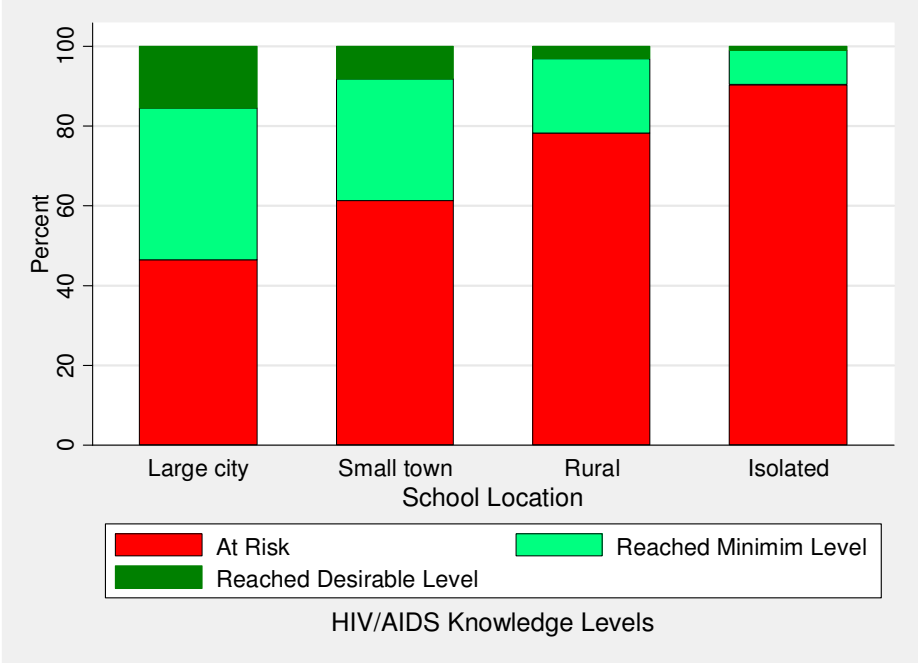
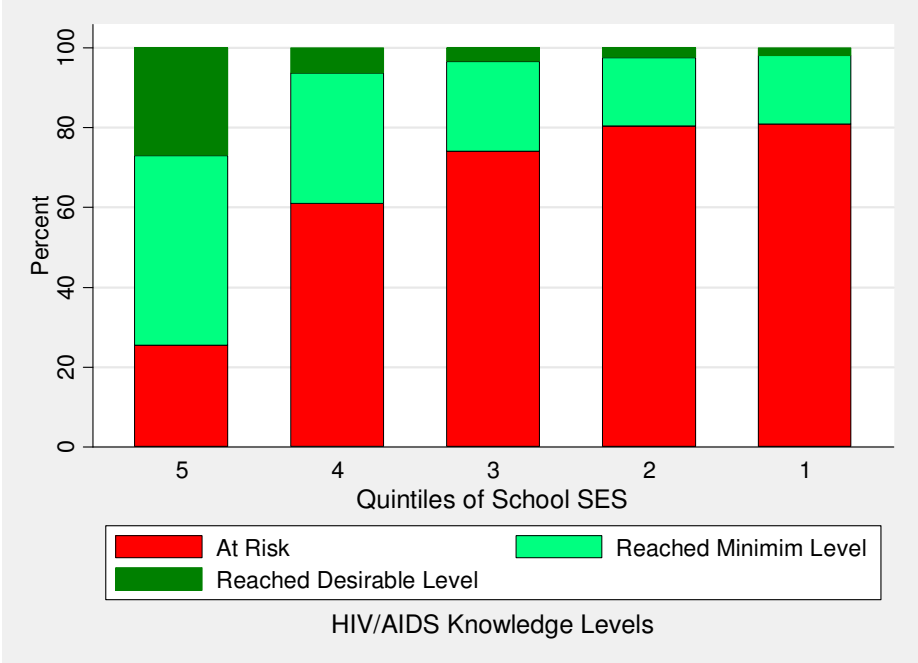


Figure 36: HIV/AIDS Knowledge Level by School Socioeconomic Quintile

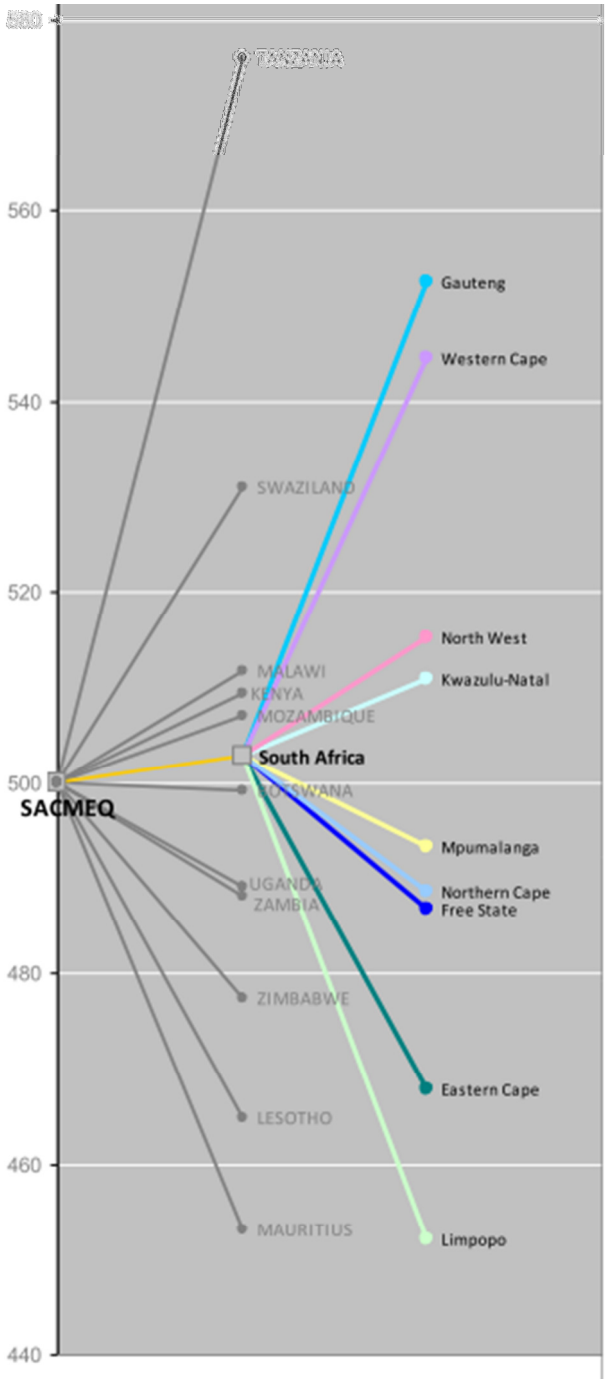


Stellenbosch University http://scholar.sun.ac.za

When seen in regional context, the health knowledge of South African learners seems to be average, if not slightly above average. However, this average of 503 on the HAKT masks significant variation at the provincial level, as can be seen in Figure 35 and Figure 38, as well as Table 24 in Appendix B. For example, 85% of learners from the Limpopo province were 'at risk', while just more than half as many (47.2%) were at risk in the Western Cape (Table 24). Seen along socioeconomic lines, 80% of the poorest 40% of learners (quintiles 1 and 2) are 'at risk', compared to only 25.5% of learners in the top quintile (Table 26). Moloi and Chetty (2011, p. 4) are correct in questioning the efficacy of the existing AIDS education initiative in South Africa:

> "It is clear from the SACMEQ III Project research results that the time has come to take stock of the impact of current HIV and AIDS prevention education programmes for young people in South Africa. The SACMEQ results showed that during 2007 two-thirds of Grade 6 learners did not have the minimal level of knowledge about HIV and AIDS that was required to preserve and promote their health."

Figure 38: Variation in Student Knowledge about HIV/AIDS among SACMEQ countries and South African provinces [Source: (Moloi & Chetty, 2011, p. 5)]



Stellenbosch University http://scholar.sun.ac.za

3.2 Schools

Before proceeding with the analysis of school level variables, it is important to first explain how the SACMEQ study was conducted since this has bearing on the way one interprets school level characteristics. Given that the SACMEQ study used the pupil as the unit of analysis, all school level averages need to be interpreted in terms of pupils. This thesis follows[20] the approach of Ross *et al.* (2005, p. 144) when interpreting school level variables, namely:

> "Pupils are the unit of analysis – even though some variables in this report refer to teachers or schools. Where a percentage for a variable that describes teachers is presented, this percentage should be interpreted as 'the stated percentage of pupils who were in schools with teachers having the particular characteristic.' Similarly, a percentage for a variable that describes schools should be interpreted as 'the stated percentage of pupils who were in schools with the particular characteristic.' Where a mean for teachers or schools of schools is presented, then the mean should be interpreted as 'the average pupil in Kenya who had a teacher with such and such characteristics or was in a school with such and such characteristics.'"

3.2.1 School Demographics

When comparing school size, the average learner in a quintile 1 school was in a school with 540 learners enrolled which employed 15 teachers, compared to the average learner in a quintile 5 school which had 763 learners enrolled and employed 27 teachers (Table 8 below). Wealthier schools also employ more non-permanent teachers as a proportion of total teachers, compared to poorer schools. Looking at school location, the average rural school is smaller (566 learners) than one in a large city (839 learners), and consequently employs fewer teachers (16 teachers) compared to the average school in a large city (26 teachers) (Table 7 below).

3.2.2 School Resources

School resources are believed to be an important aspect when trying to make the school environment conducive to learning. The presence or absence of these resources could explain some variation in learner performance. There are various measures of school resources, some of which have direct bearing on learner performance (such as library books), while others are indicative of the school environment more generally (such as sanitation quality). As one would expect, the distribution of school resources across provinces, school locations, and socioeconomic quintiles is by no means equal.

In order to measure the presence or absence of numerous school resource items, two indicators were created: the *School Buildings Index*, and the *School Equipment Index*. The *School Buildings*

[20] Where interpreting variables in this technically correct way is overly cumbersome, a more conventional (although not technically correct) method has been used. In these instances, the reader should bear in mind that the unit of analysis is always the learner and not the school.

Stellenbosch University http://scholar.sun.ac.za

Index was calculated as the count of seven[21] items from question 19 of the School Head questionnaire, while the *School Equipment Index* was calculated as the count of 18[22] items from the same questionnaire.

Table 6, Table 7, and Table 8 provide an overview of several different elements of school resources. By showing the mean values of each measure by province, school location, and school socioeconomic quintile, it soon becomes clear that school resources, not only home resources, are highly unequal. Some of these results are highlighted below:

- **Provincial differences**: In accordance with *a priori* expectations, the Western Cape and Gauteng rank as the top two provinces in terms of school resources, while the Eastern Cape and Limpopo provinces are the least resourced of the provinces. For example, the average learner in a school in either Gauteng or the Western Cape has a school library stocked with more than 2000 books and a computer room with more than 24 computers. In stark contrast, the average learner in a school in the Eastern Cape and the Limpopo province has only 92 library books, and 1.9 (Eastern Cape) and 6.9 (Limpopo) computers per school. The situation is again similar when observing sanitation quality. Looking at the number of non-flushing school toilets (latrine places, squat holes, or pit toilets) as a percentage of total toilets (non-flushing and flushing toilets combined) the abovementioned provincial differences are reiterated. 80% of learners in the Eastern Cape and Limpopo provinces were in schools with only non-flushing toilets. The figures for learners in the Western Cape and Gauteng were 7% and 0.2% respectively. Looking more specifically at the Eastern Cape, 35% of learners were in schools where the classrooms are either temporary[23] classrooms or open-air classrooms. This lack of physical school infrastructure is also confirmed in the *School Building Index*, where the Eastern Cape average is only 2.1 school building items out of a possible 7 items.
- **Urban-Rural differences:** Ignoring the *Isolated* category of school location (only 1.55% of schools were in this category and all were in KwaZulu Natal), the significant differences between school resources in large cities and rural towns becomes evident. The average learner in a rural school had access to only a few computers (3.9), a few library books (237), and low sanitation quality

[21] These seven items were: school library, school or community hall, teacher/staff room, separate office for School Head, store room, special area for guidance and counselling, and cafeteria/shop/kiosk.
[22] These 18 items were: first aid kit, clock, telephone, typewriter, duplicator, electricity (mains or generator), radio, tape recorder, TV, audio cassette player, CD player, VCR machine, DVD player, fax machine, photocopier, overhead projector, computer(s), computer room.
[23] "**Permanent** classrooms are completed classrooms that have been built using materials in compliance with approved specifications; **temporary** classrooms include, for example, temporary / thatch roof, roof only, walls only; **open-air** teaching areas are areas that have no floors, walls, or roofs and are usually located under a tree." (Student questionnaires are available from the SACMEQ website)

Stellenbosch University http://scholar.sun.ac.za

(80.6% of toilets are non-flushing). By contrast, the average learner in a school in a large city has access to many computers (26.9), a well-stocked library (2934 books), and high sanitation quality (only 4.3% of toilets are non-flushing).

- **Socioeconomic differences:** Although the differences between provinces and school locations are significant, school resource differentials are most pronounced when seen across the five socioeconomic quintiles. Comparing learners in the poorest 20% of schools (quintile 1) and the wealthiest 20% of schools (quintile 5), the inequality in school resources becomes conspicuous. The average learner in a quintile five school had access to 16.8 of the 18 items in the *School Equipment Index*, and 6.1 of the 7 items on the *School Building Index*. The corresponding figures for the average learner in a quintile one school are 6.2/18 (equipment) and 1.6/7 (buildings). The average learner in a quintile five school has access to more library books than the average learner in a quintile one school (4568 books compared to 35 books), as well as more computers than the average quintile one school (37.4 computers compared to 2 computers). Only 3.5% of learners in quintile one schools have access to flushing toilets, while 98.5% learners in quintile five schools have access to flushing toilets – the remaining toilets are either latrines, squat holes or pit toilets.

Table 6: Diagnostic Overview (Provinces)

Theme	Indicator	Provinces									Total
		ECA	FST	GTN	KZN	LMP	MPU	NCA	NWP	WCA	Total
Sample	Distribution of Learners (%)	16.29 %	4.95 %	17.22 %	23.68 %	13.43 %	8.54 %	2.05 %	6.02 %	7.83 %	100 %
Learner Performance	Avg Learner Maths Score	468.8	491.6	545.0	485.2	446.7	476.1	498.7	503.1	565.7	**494.8**
	Avg Learner Maths Competency Level (1-8)	2.8	3.0	3.9	3.0	2.4	2.8	3.2	3.2	4.1	**3.1**
	Avg Learner Reading Score	447.8	491.1	573.1	485.6	425.3	473.6	505.6	506.3	583.4	**494.9**
	Avg Learner Reading Competency Level (1-8)	3.2	4.0	5.4	3.8	2.8	3.7	4.2	4.2	5.6	**4.0**
	Avg Learner Health Score	467.8	486.8	552.4	510.9	452.0	493.1	488.5	515.2	544.5	**502.8**
	% Learners reaching min Health Knowledge	23.2	26.9	53.1	37.9	15.0	29.4	28.1	38.4	52.8	**34.8 %**
School Demographics	Total number of teachers in the school (H)	14.6	28.5	28.1	18.6	15.7	21.8	24.1	22.3	24.2	**20.7**
	Non-permanent teachers as % of total teachers in the school (permanent + non-permanent) (H)	7.1	10.1	12.8	10.6	3.4	8.1	14.0	13.3	16.6	**9.9 %**
	Total school enrolment (number of learners)	545.6	894.0	919.6	647.7	543.0	738.7	811.0	687.5	823.0	**703.3**
Resources	Avg School SES (x10) : range [-20, +20]	-5.9	0.5	4.8	-2.9	-2.2	0.0	0.6	1.1	6.5	**-0.5**
	School Building Index : range [0, 7]	2.1	3.9	4.7	3.0	1.4	2.3	4.0	3.6	5.0	3.1
	School Equipment Index : range [0, 18]	5.2	15.5	15.9	10.1	8.2	9.8	14.9	13.2	16.0	11.0
	Number of school PC's in working order (H)	1.9	18.3	24.6	7.3	5.5	6.9	18.8	13.4	39.4	**12.8**
	Number of library books in the school library (H)	92	1258	2749	1373	92	568	824	1200	2273	**1203.7**
	Latrine places, squat holes, or pit toilets as % of total toilets (latrine etc. + flushing toilets)	80.0	7.3	7.0	53.1	79.8	42.7	10.2	34.6	0.2	**43.4 %**
	Temporary class-rooms + open-air classrooms as % of total (temporary + open-air + permanent)	35.0	5.2	9.9	5.4	13.8	12.3	3.8	4.8	7.4	**12.7 %**
Miscellaneous	School days lost due to disruptions (natural disaster, strikes, unrest etc) (H)	11.3	8.4	7.2	5.5	10.4	8.0	11.3	7.8	3.6	**7.9**
	Avg* self-reported teacher absenteeism /year in days (Reading/Maths/Health teachers*) (T)	20.8	16.5	12.8	25.4	24.2	19.3	17.1	18.8	12.0	**19.7**

Table 7: Diagnostic Overview (School Location)

Theme	Indicator	School Location				Total
		Isolated	Rural	Small town	Large city	Total
Sample	Distribution of Learners (%)	1.55	48.49	18.11	31.86	**100 %**
Learner Performance	Average Learner Maths Score	450.9	456.9	504.6	549.3	**494.8**
	Average Learner Maths Competency Level (1-8)	2.5	2.6	3.2	3.9	**3.1**
	Average Learner Reading Score	441.9	440.7	514.5	568.9	**494.9**
	Average Learner Reading Competency Level (1-8)	3.1	3.0	4.4	5.3	**4.0**
	Average Learner Health Score	462.3	471.5	511.9	547.4	**502.8**
	% Learners reaching minimum Health Knowledge Level	9.7	21.8	38.7	53.5	**34.8 %**
School Demographics	Total number of teachers in the school (H)	18.3	15.8	24.4	26.2	**20.7**
	Non-permanent teachers as % of total teachers in the school (permanent + non-permanent) (H)	20.4	6.0	10.6	15.0	**9.9 %**
	Total school enrolment (number of learners) (H)	634.2	566.3	837.4	838.9	**703.3**
Resources	Average School SES (x10) : range [-20, +20]	-8.7	-4.9	1.1	5.6	**-0.5**
	School Building Index : range [0, 7]	3.3	1.8	3.4	5.0	3.1
	School Equipment Index : range [0, 18]	6.6	7.6	13.0	15.4	11.0
	Number of school PC's in working order (H)	0.5	3.9	12.9	26.9	**12.8**
	Number of library books in the school library (H)	213	237	818	2943	**1203.7**
	Latrine places, squat holes, or pit toilets as % of total toilets (latrine etc. + flushing toilets)	73.7	80.6	13.4	4.3	**43.4 %**
	Temporary class-rooms + open-air classrooms as % of total (temporary + open-air + permanent)	13.8	15.7	13.5	7.7	**12.7 %**
	School days lost due to disruptions (natural disaster, strikes, unrest etc) (H)	5.7	8.2	7.9	7.5	**7.9**
	Avg* number of days absent for teachers per year (self-reported by Reading/Maths/Health teachers*) (T)	28.0	24.1	17.5	14.1	**19.7**

Table 8: Diagnostic Overview (School SES)

Theme	Indicator	5 Quintiles of School SES					Total
		1	2	3	4	5	Total
Sample	Distribution of Learners (%)	25.64%	18.25%	17.80%	19.31%	19.01%	100%
Learner Performance	Avg Learner Maths Score	448.0	468.3	464.9	492.9	613.9	**494.8**
	Avg Learner Maths Competency Level (1-8)	2.5	2.7	2.7	3.1	4.9	**3.1**
	Avg Learner Reading Score	426.3	456.8	460.8	505.4	645.5	**494.9**
	Avg Learner Reading Competency Level (1-8)	2.8	3.3	3.4	4.3	6.6	**4.0**
	Avg Learner Health Score	462.8	470.0	485.1	510.1	597.7	**502.8**
	% Learners reaching minimum Health Knowledge Level	19.1	19.6	25.9	38.9	74.5	**34.8 %**
School Demographics	Total number of teachers in the school (H)	14.7	21.7	20.9	21.8	26.9	**20.7**
	Non-permanent teachers as % of total teachers in the school (permanent + non-permanent) (H)	7.2	7.6	5.1	10.4	19.8	**9.9**
	Total school enrolment (number of learners) (H)	539.6	763.9	746.6	764.6	763.0	**703.3**
Resources	Avg School SES (x10) : range [-20, +20]	-9.3	-3.5	-1.2	2.8	11.5	**-0.5**
	School Building Index : range [0, 7]	1.6	2.3	2.4	3.7	6.1	**3.1**
	School Equipment Index : range [0, 18]	6.2	10.1	10.7	13.0	16.8	**11.0**
	Number of school PC's in working order (H)	2.0	6.6	6.0	15.1	37.4	**12.8**
	Number of library books in the school library (H)	35	333	374	1032	4568	**1203.7**
	Latrine places, squat holes, or pit toilets as % of total toilets (latrine etc. + flushing toilets)	96.5	52.6	34.3	17.7	1.5	**43.4 %**
	Temporary class-rooms + open-air classrooms as % of total (temporary + open-air + permanent)	22.8	16.6	10.7	4.0	6.1	**12.7 %**
Miscellaneous	Number of periods/lessons taught by School Head per week	14.2	14.1	13.2	11.2	6.4	**11.9**
	School days lost due to disruptions (natural disaster, strikes, unrest etc) (H)	7.5	8.8	9.9	9.3	4.1	**7.9**
	Avg* number of days absent for teachers per year (self-reported by Reading/Maths/Health teachers*) (T)	24.8	21.0	23.9	19.8	8.1	**19.7**

Stellenbosch University http://scholar.sun.ac.za

3.3 Teachers

3.3.1 Teacher Absenteeism

Following from the premise that teachers are essential to the learning process, it is logical to suspect that higher rates of teacher absenteeism are associated with lower learner performance. This is largely due to inadequate coverage of the curriculum and shorter time-on-task. High rates of teacher absenteeism can also have a reciprocating effect on learner absenteeism whereby learners choose not to attend school because they are unsure about whether their teachers will be at school on that particular day.

In SACMEQ III, teachers were asked how many days they were absent 'during this school year', as well as the reasons[24] for that absence. This self-reported measure of absenteeism is likely to be affected by a number of issues, most prominently the tendency to under-report absenteeism. Nevertheless, the figures which teachers report may provide a useful lower-bound estimate of teacher absenteeism in the country.

Table 6 and Table 7 above, and Figure 39 below show that self-reported teacher absenteeism varies substantially across provinces, school locations and socioeconomic quintiles. The average Grade 6 learner in South Africa was taught by a teacher who was absent approximately 19.7 days in 2007. Given that there are only 5 days in a school week, this amounts to an entire month of school time. The provincial differences are large with learners in the Western Cape and Gauteng experiencing less teacher absenteeism (around 12 days), than learners in KwaZulu Natal and Limpopo, where the corresponding figures are 25.4 and 24.2 days respectively.

A similar trend is seen when comparing quintile 1-4 learners (poorest 80%) with quintile 5 learners (richest 20%) in South Africa: quintile 1-4 learners experienced almost three times as much teacher absenteeism (22.4 days) compared to quintile 5 learners (8.1 days). Figure 39 below shows the stark contrasts in teacher absenteeism by socioeconomic quintile.

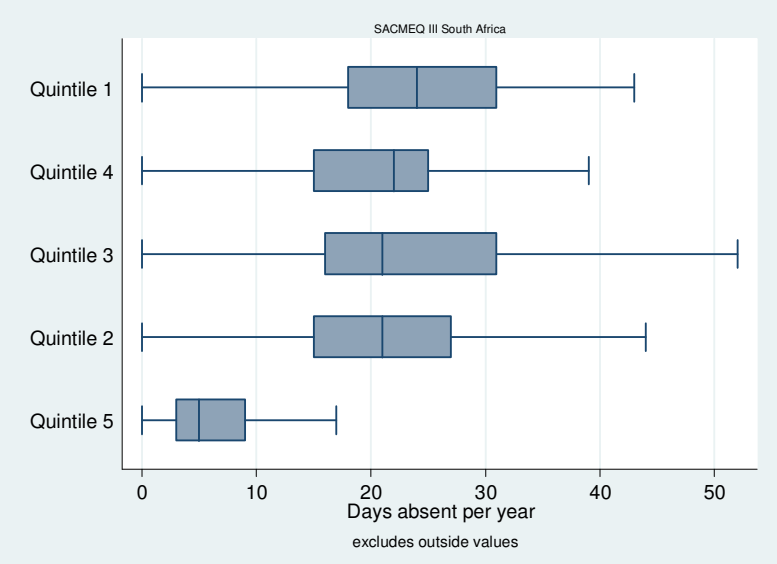
There is one potential caveat to the above findings: teachers may have a mental reference point for what is an 'acceptable' or 'normal' level of absenteeism, and thus in the self-reported answers may tend not to deviate substantially from that reference point. This is especially true given that some legitimate teacher absence is to be expected (sickness, union meetings, etc.). Following on from this, teachers in different parts of the country may have different mental reference points for what is

[24] The 12 categories were: own illness, own injury, family member's illness, family member's injury, funerals, medical appointments, bad weather/ road not accessible, official business, maternity leave, security reasons, teachers' strikes, other reasons.

Stellenbosch University http://scholar.sun.ac.za

acceptable and this may account for some of the large differences between provinces or school locations. However, even if this is the case, the fact that most South African teachers believe that being absent for an entire month in a year (19.7 days) is 'normal' or 'acceptable' is particularly disconcerting, especially when one considers that this is more than 10% of the number of school days in a year.

Figure 39: Box plot of teacher absenteeism by SES quintile



3.3.2 Teacher Content Knowledge

One of the numerous factors that are thought to impact learner performance is teacher content knowledge. Taylor (2008, p. 24) states the obvious, but important, reality: "teachers cannot teach what they do not know." While pedagogical skills, teacher motivation and classroom resources are all important inputs into the student learning process, sufficient teacher content knowledge of the subject being taught is a necessary condition for student learning. An overview of the literature on teacher content knowledge in South Africa is provided in Appendix G.

In addition to learners that were tested in numeracy and literacy in SACMEQ III, the learners' teachers were also tested in the subjects that they taught. As Ross *et al.* (2005, p. 257) note, "The major reason for measuring the achievement of teachers is the belief that their mastery of the subject matter is critical in curriculum implementation." While the majority of the teacher test was the same as the learner test, some more challenging questions were added, and some elementary questions were removed. Using Rasch scaling, the teacher scores were adjusted so that they would be comparable with the learner test scores.

The inclusion of the teacher test in the SACMEQ III study enables one to answer two interesting and important questions; 1) Are teachers in wealthier schools more knowledgeable than teachers in

Stellenbosch University http://scholar.sun.ac.za

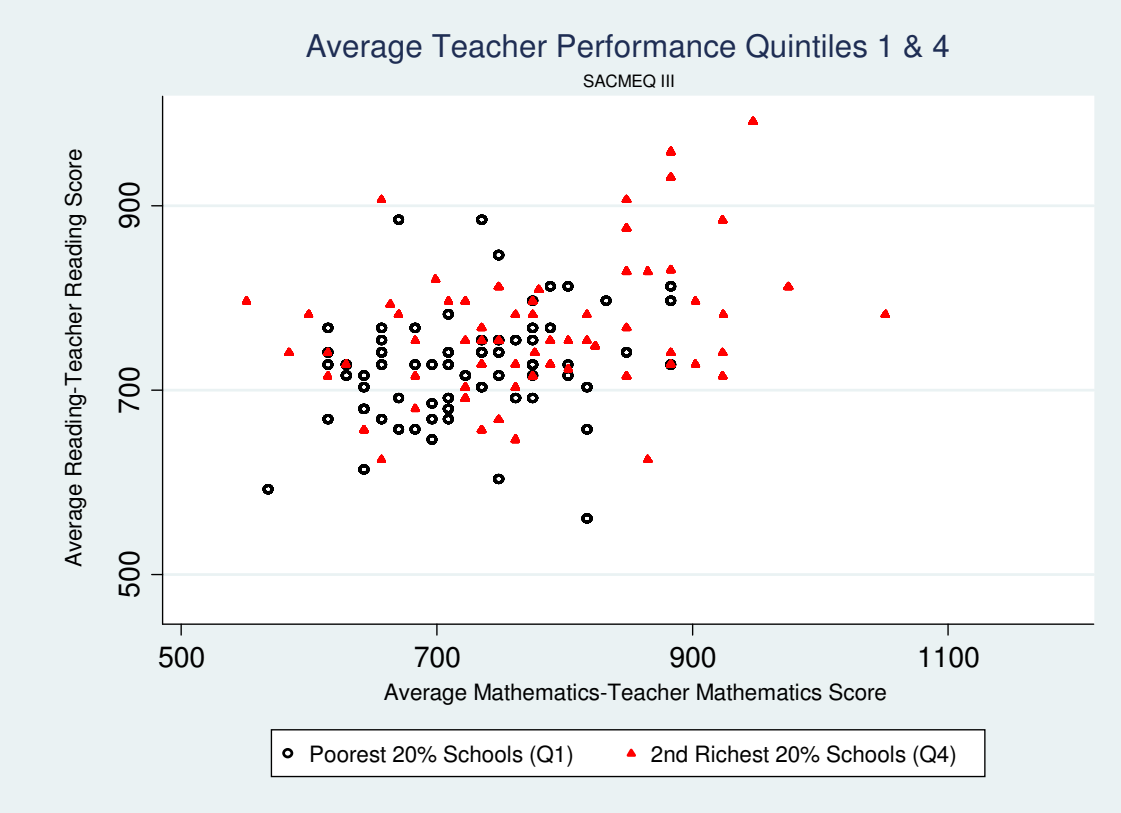
poorer schools? 2) Do more knowledgeable teachers produce more knowledgeable learners? It is only possible to answer the second question in a multivariate framework, and thus this is left to Chapter 4 *Modelling Learner Performance.* With respect to the first question, it is reasonable to expect that teachers who are more knowledgeable, better qualified, and who have more experience, are more likely to teach in wealthier urban schools than their less qualified, less knowledgeable and less experienced counterparts. This is largely because most teachers would prefer to live in an urban area rather than a rural one, and because employment conditions in wealthier schools appear to be more favourable.

Figure 40, Figure 41, and Figure 42 below illustrate that this is indeed the case in South African primary schools. Figure 42 shows that learners in the wealthiest 20% of schools are most likely to be taught by teachers who were found at the upper end of the teacher reading and mathematics distributions. An interesting pattern emerges when one compares the difference in teacher knowledge for teachers in each of the five quintiles of school SES. While there is a noticeable difference between teachers from quintiles one and *five* (Figure 41), the difference between quintiles one and *four* (Figure 40) is much less pronounced. Indeed, there is very little difference in teacher knowledge between the bottom four quintiles. This again supports the hypothesis that there are two very different schooling systems present in South Africa.

To place the above scores of teacher knowledge in perspective, it is instructive to compare teacher knowledge to learner knowledge. This is possible since the teacher knowledge test was transformed (using Rasch scaling) to a score that is directly comparable to the learner scores. It is worrisome to note that the top 5% of Grade 6 *learners* (559 learners) scored higher marks on the same mathematics test than the bottom 12.5% of Grade 6 mathematics *educators* (62 teachers[25]) in the sample (Figure 43).

[25] It is worth noting that of the 498 mathematics teachers and 498 reading teachers, 114 individuals appear to be both the reading teacher *and* the mathematics teacher. This is based on a matching algorithm using age, gender and experience, since teachers could not be matched using teacher ID's.

Figure 40: Average Teacher Performance for School SES Quintiles 1 & 4



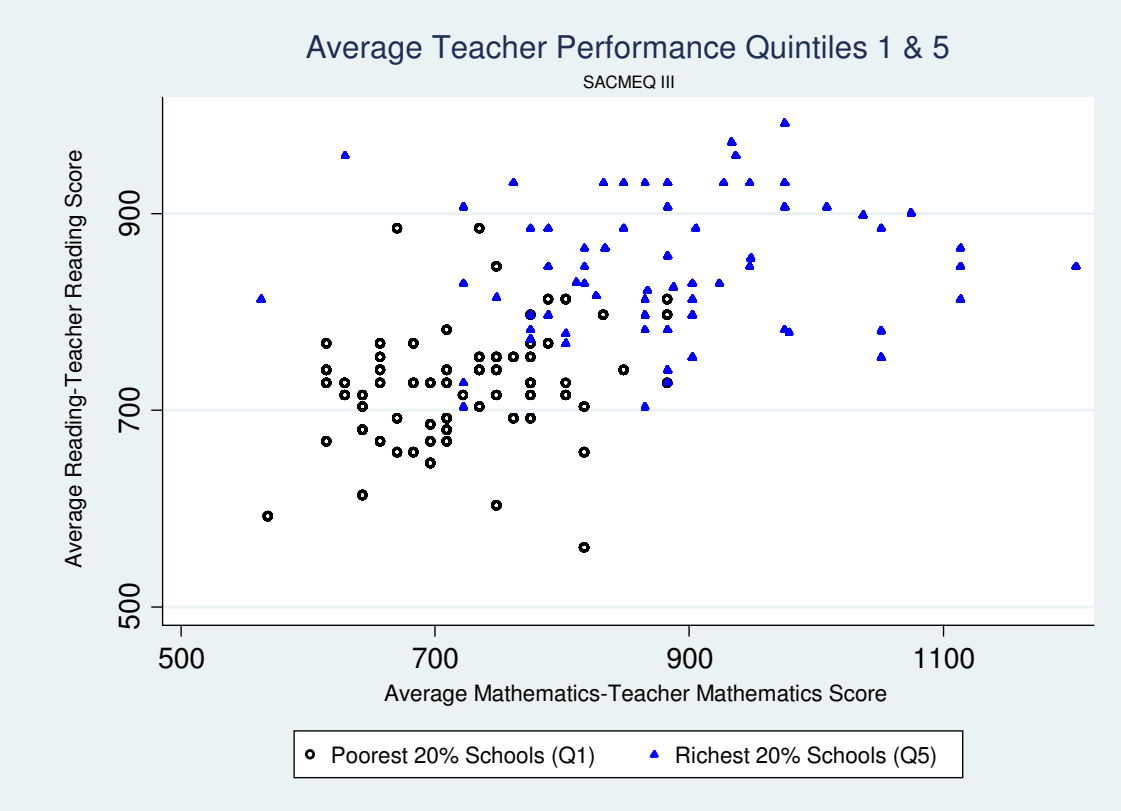


Figure 41: Average Teacher Performance for School SES Quintiles 1 & 5

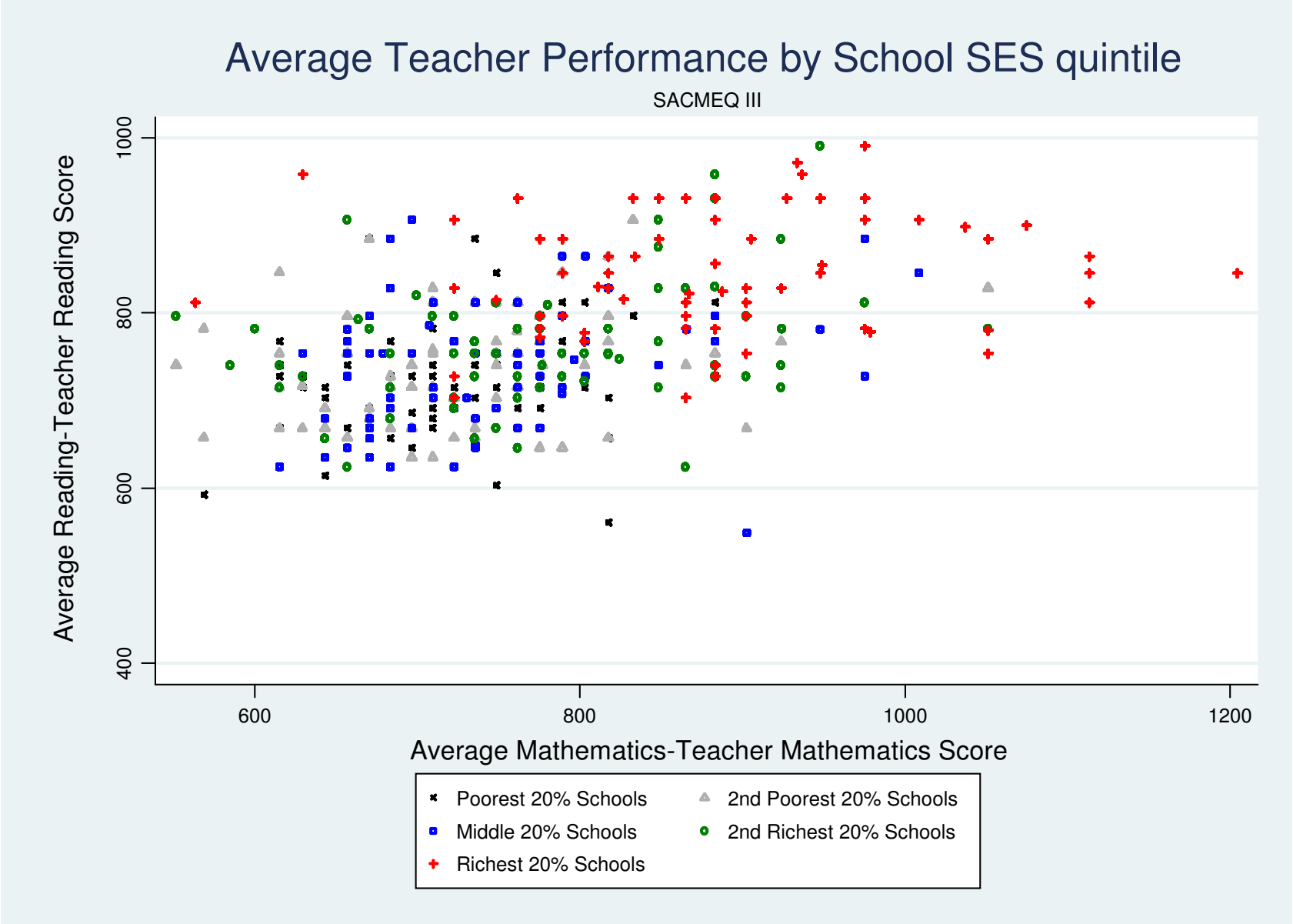
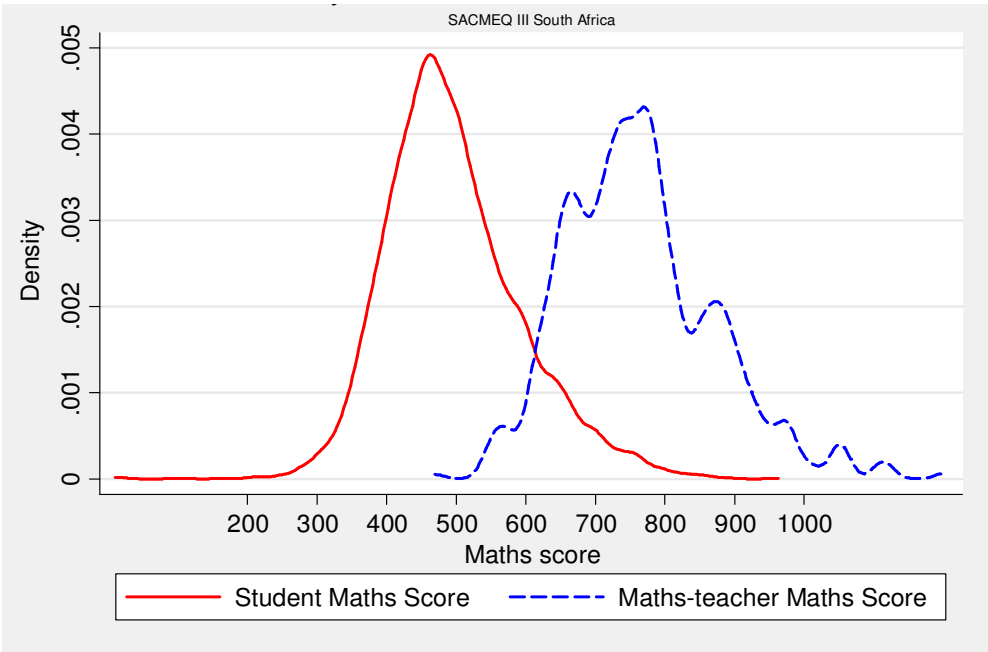


Figure 42: Average Teacher Performance by School Socioeconomic Quintile

Stellenbosch University http://scholar.sun.ac.za

Figure 43: Kernel Density of Teacher and Learner Mathematics Knowledge



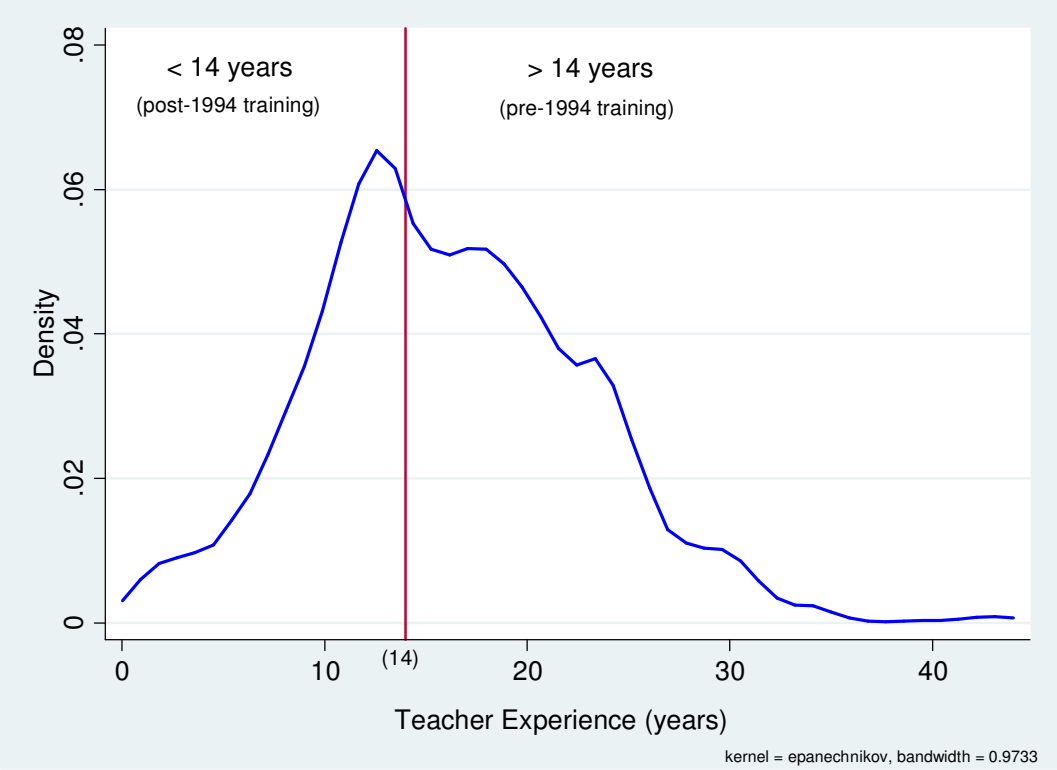
3.3.3 Teacher Experience and Teacher Training

One of the questions in the SACMEQ III teacher test asked teachers "*How many years altogether have you been teaching for?"* This enables one to calculate when teachers received their teacher training. Assuming that teachers received the bulk of their training or education *before* they started teaching, one can calculate if teachers were trained before 1994 or after 1994. This is potentially an important question since the training methods and content knowledge acquired by non-white teachers during apartheid were inadequate. Figure 44 shows the distribution of average[26] teachers experience for the SACMEQ III sample. Given that SACMEQ III was conducted in 2007, if teachers indicated that they had more than 17 years of experience, one can deduce that they received their training under apartheid. Clearly most teachers have more than 14 years of teaching experience, and thus were most probably trained under the apartheid system, something that Carnoy *et al.* (2008, p. 15) also found.

[26] Since both mathematics and reading teachers answered this question, the distribution shows the average experience across mathematics and reading teachers.

Stellenbosch University http://scholar.sun.ac.za

Figure 44: Kernel Density of Teacher Experience



3.4 Placing South Africa in Regional Context

The preceding analysis has shown that when the South African primary school population is split along socioeconomic lines, it is difficult to describe the system as anything but inequitable. Learners who are fortunate enough to attend one of the top quartile schools have usually acquired the basic numeracy and literacy skills as outlined in the curriculum. However, the majority of learners, who do not attend these functional schools, have not met the demands of the curriculum, or even the most basic benchmarks of educational progress: functional numeracy and functional literacy. Importantly, this is not due to resource constraints: the education budget makes up the largest single line item in the national budget. From this alone one could conclude that South African primary education is both inequitable and inefficient.

Comparing learner performance to the national curriculum is an internal measure of assessing the level of underperformance in the system; however, it is also possible to use external comparisons as a measure of underperformance. Placing South African learner and teacher performance in regional context is one such measure of external comparison. Using the SACMEQ data, it is possible to compare South Africa's performance to 14 other education systems on the continent. While there are a host of comparability problems associated with such a comparison, when interpreted with

Stellenbosch University http://scholar.sun.ac.za

caution the comparison can yield useful insights. It will be helpful to first elucidate three important comparability problems: resources, enrolment and repetition.

Resources **–** Comparing Gross Domestic Product (GDP) and GDP per capita between the 14 SACMEQ countries shows that there are large differences in the size of each country's economy as well as the size of each country's population. While some countries have a high GDP per capita, such as Botswana ($14,313), Mauritius ($11,106), Seychelles ($21,350) and South Africa ($9,778), others have an extremely low GDP per capita such as Malawi ($759), Mozambique ($843), and Zimbabwe ($189) (Hungi, *et al.*, 2011, p. 5). As one would expect, the per child spending on primary education is largely a function of GDP per capita, and thus per pupil spending on primary education differs significantly across the 14 SACMEQ countries. For example, the amount spent on the average Mozambican child in primary school per year is only 5.8% ($79) of what is spent on the average South African child ($1356), and 7.9% of what is spent on the average Namibian child ($999) (UNESCO, 2011, p. 337). Clearly this is likely to have an impact on learner performance.

Enrolment **–** In Sub-Saharan Africa, unlike the developed world, many countries have not reached universal primary school enrolment. Given that the learners who never enrol in school are more likely to be situated in rural areas, and come from poorer families with lower average levels of education, if one were to increase access to include these learners, the average performance of that country would most probably decline. This is an important consideration when comparing two countries with very different enrolment profiles, since it overestimates the performance of the low enrolment country.

Drop out **–** Another factor which influences cross-national comparisons is learner drop out. Given that the SACMEQ tests were administered to Grade 6 learners, any learner who dropped out of school prior to Grade 6 would not have been included in the sampling frame for the SACMEQ survey. Based on the plausible assumption that learners who drop out are, on average, the weakest-performing learners in the school, it is not unreasonable to conclude that the learners who progress to Grade 6 are better performing learners than those who drop out.

To avoid the complications that arise from comparing South Africa to countries that have very different enrolment and drop out profiles, the comparisons in this section will focus on countries that have similar enrolment and drop out patterns.

Figure 45 and Figure 46 below show the mean reading and mathematics scores for each of the fifteen education systems that participated in SACMEQ. In both subjects South Africa performs

Stellenbosch University http://scholar.sun.ac.za

worse than the SACMEQ mean and the country ranks only 10th/15 for reading, and 8th/15 for mathematics.

Figure 46: SACMEQ Student Reading Performance

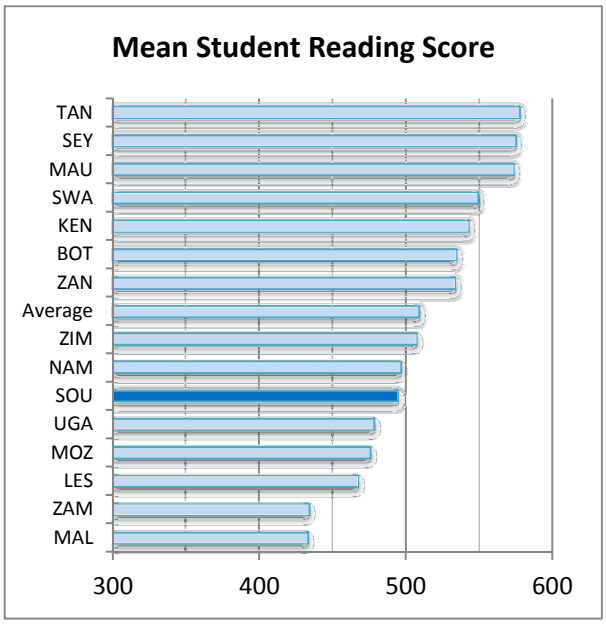
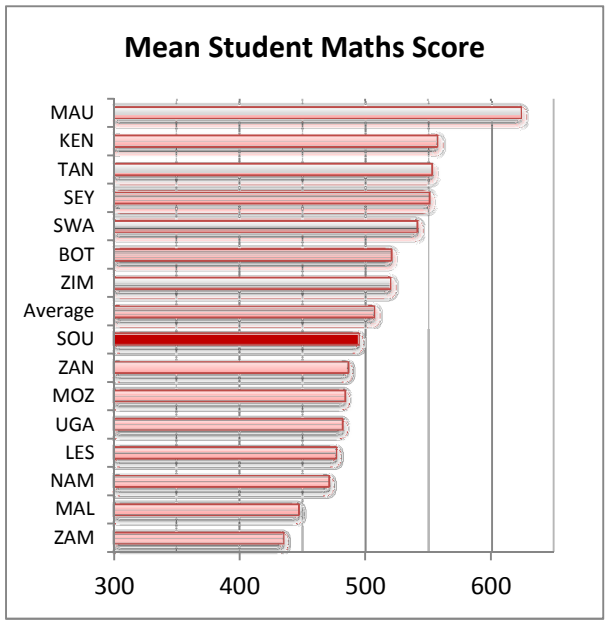


Figure 45: SACMEQ Student Mathematics Performance



While the above charts are helpful in that they easily convey the relative position of each country, average measures of performance often shroud the sometimes severe inequalities within countries, as is the case in South Africa. Figure 47 shows the mean reading scores for both rural and urban learners in each of the fifteen SACMEQ education systems ranked by mean rural score. Rural learners in South Africa perform substantially worse than rural learners in comparable countries such as Botswana, Namibia and Mauritius.

Figure 48 shows the mean mathematics scores for the wealthiest 25% of learners *in each country,* and the mean scores for the poorest 25% of learners *in each country*. Countries are ordered based on the mean score of the poorest 25% of learners. Clearly South Africa's poorer learners perform worse than poorer learners in Mauritius and Botswana, but perform marginally better than Namibia's poorer learners. The "high SES" distribution shows South Africa as an outlier. This is because the wealthier learners in South Africa perform in the upper end of the "high SES" distribution, while the poorer South African learners perform at the bottom of the "low SES" distribution.

Stellenbosch University http://scholar.sun.ac.za

Figure 47: SACMEQ Learner Reading Performance - Rural/Urban

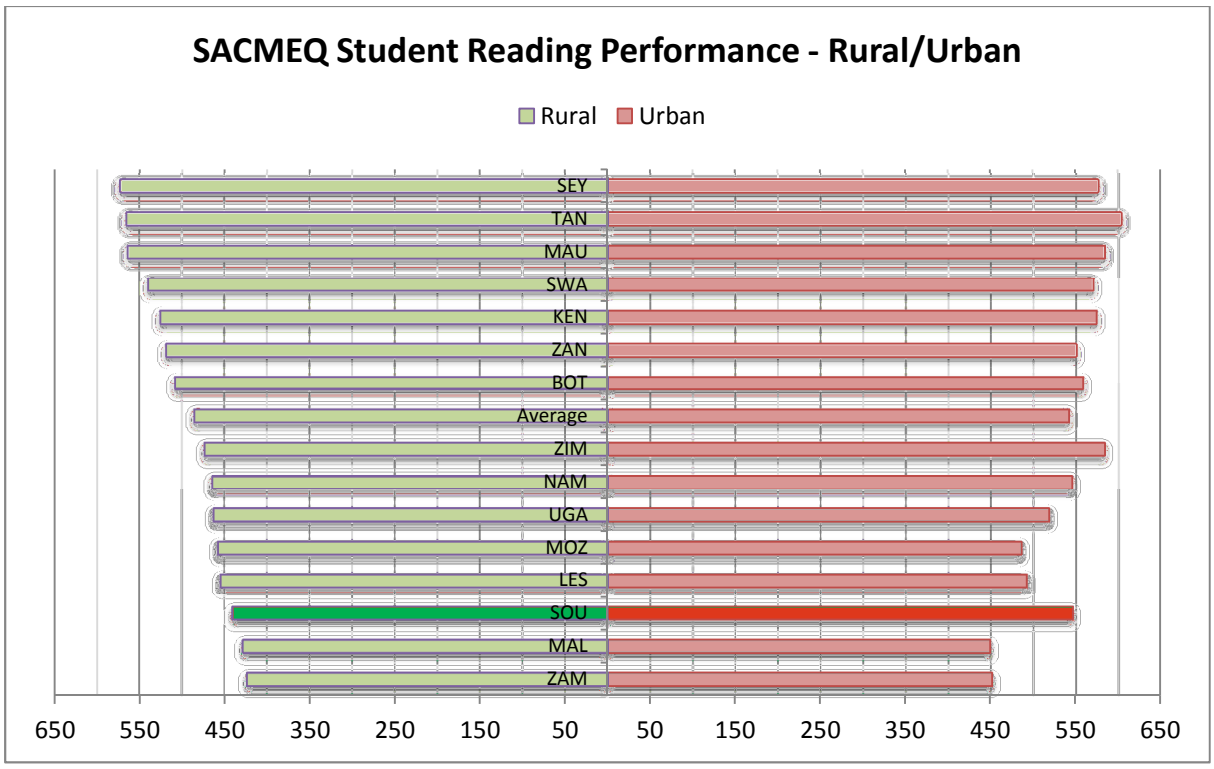
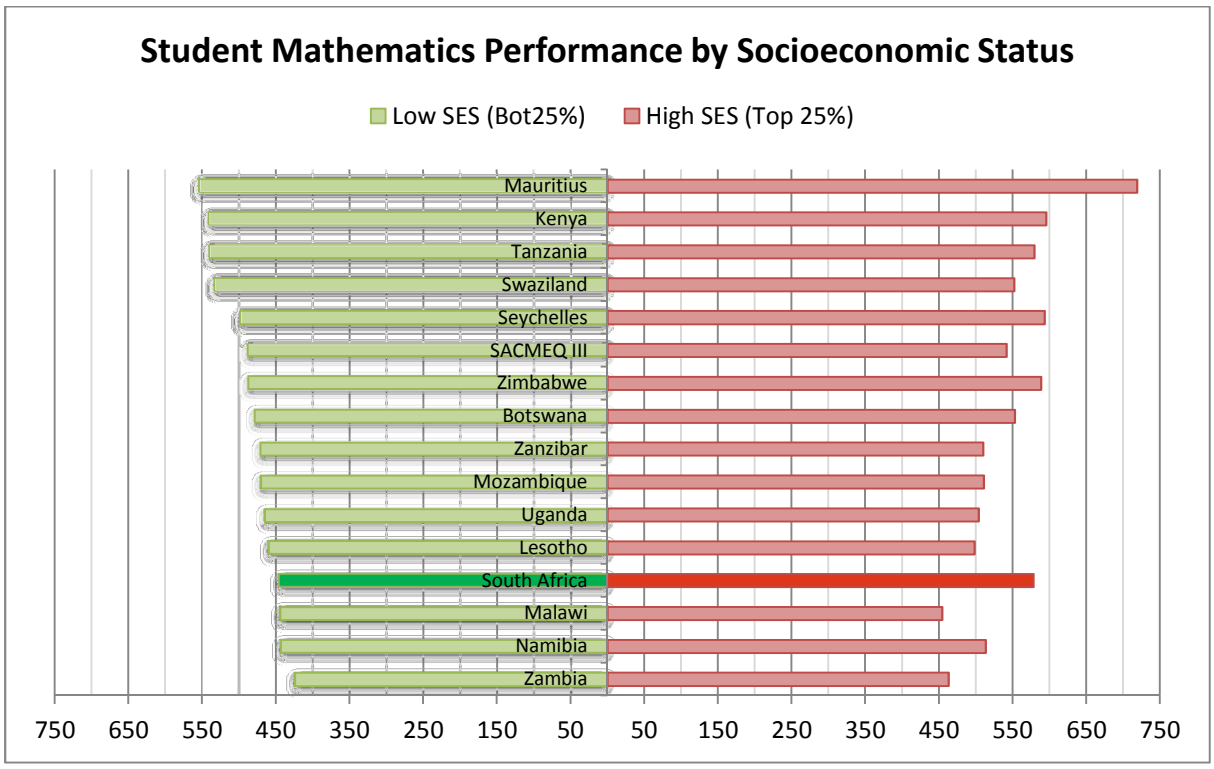


Figure 48: SACMEQ Learner Mathematics Performance by Socioeconomic Status



Stellenbosch University http://scholar.sun.ac.za

Teacher Content Knowledge

One of the interesting features of the SACMEQ data is that one can compare teacher content knowledge across each of the fifteen education systems. Since all reading teachers wrote the same reading test across all participating countries (and similarly for mathematics teachers and the mathematics test) one can easily determine the differences in average teacher content knowledge. South African teachers rank 6th out of 15 for reading content knowledge and 9th out of fifteen for mathematics content knowledge. This is surprising given that South African teachers have, on average, more tertiary education and training than most of their SACMEQ counterparts.

Figure 50: Reading Teacher Content Knowledge

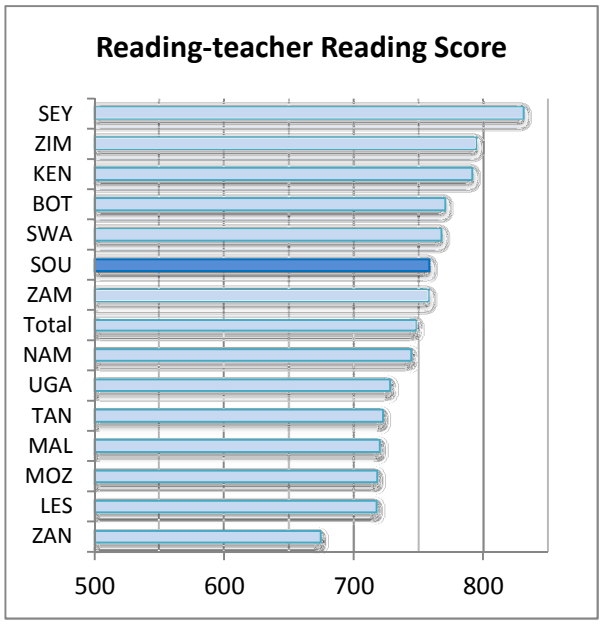
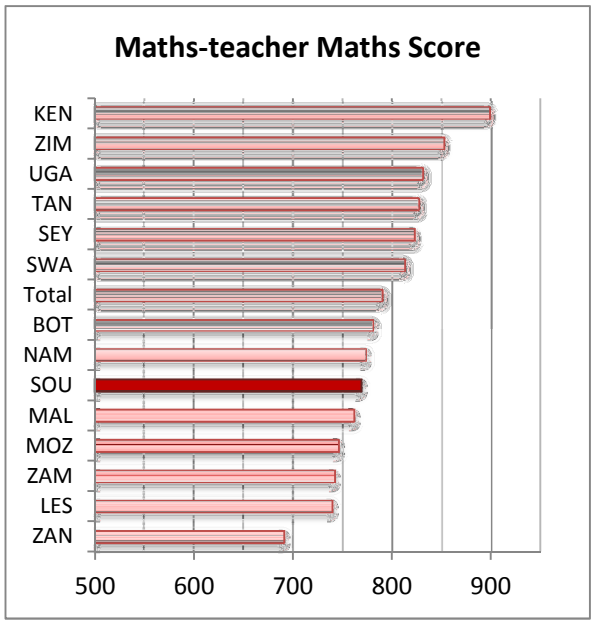


Figure 49: Mathematics Teacher Content Knowledge



Provincial Differences in Performance and Resources

As is the case with most South African data, the national averages hide severe inequalities within the country. Figure 51 below shows the national and provincial averages for reading and mathematics for South Africa and three of its neighbours; Botswana, Mozambique and Namibia. From this one can see that South Africa has both the second-best[27] performing province (Western Cape), and the worst performing province (Limpopo) of all 40 provinces of the four countries. It is also surprising to note that seven of the nine South African provinces are below the SACMEQ averages for reading and mathematics (dotted lines), with only Gauteng and the Western Cape achieving average scores higher than the average SACMEQ country. Another interesting trend that can be seen in Figure 51 is the seemingly lock-step relationship between reading and mathematics performance. While it is

[27] Gaborone is the best performing region.

Stellenbosch University http://scholar.sun.ac.za

reasonable to expect that higher mathematics performance is associated with higher reading performance (and vice versa), it is less clear why this relationship would be different for different countries. Comparing South Africa and Namibia for example, if one were to draw a line through these observations, the slope of the two lines would be very similar, but the intercept would be quite different. For some reason, Namibian learners perform noticeably better in reading than in mathematics, at least when compared to regional trends.

Shifting the focus from performance to pedagogical resources, Figure 52 shows the proportion of learners in each country that had access to their own reading and mathematics textbooks. Given that South Africa spends considerably more than Mozambique on pedagogical resources (such as textbooks), it is perplexing to see that more Mozambican learners have access to their own reading and mathematics textbooks than their South African peers. There is one important caveat when reporting sole textbook ownership – it takes no cognizance of those learners who share their textbook with only one other learner, as many learners do in South Africa. As shall be seen in the following chapter, it is only when learners have to share their textbook with more than one other learner that there are losses to learning.

Conclusion

Looking at bivariate distributions of learner performance and various input factors above could lead one to suspect that they are causally linked, especially when the links between the input factors and learner performance are conceptually appealing. For example, learners in smaller classes do better than learners in larger classes; learners who have had some preschool exposure perform better than those who have not; learners with more knowledgeable teachers do better than learners with less knowledgeable teachers; learners who attend a school equipped with a swimming pool do better than learners who attend schools without swimming pools. All of these statements are true and empirically verifiable. However, while learner performance and these four input-factors are *correlated* it is not possible to say that there is a *causal* relationship without further investigation.

This would be a classic case of the *post hoc ergo propter hoc* fallacy*:* i.e. that correlation does not imply causation. The sun rises because the rooster crows. Simply because better performing learners have preschool exposure, and are taught in smaller classes with more knowledgeable teachers does not mean that these factors are *driving* the better performance. The last example of the swimming pool makes this point abundantly clear: swimming facilities are unlikely to increase mathematics knowledge, although they are correlated with each other. However, wealthier schools in suburban areas are more likely to have swimming pools, but they are also more likely to be better managed,

Stellenbosch University http://scholar.sun.ac.za

better resourced, have more motivated teachers, etc. Unlike swimming facilities, these other factors are likely to improve mathematics knowledge. Which of the abovementioned factors, or, more accurately which combination of these factors, is the main driving force behind learner performance? It is not possible to answer this question using bivariate analysis.

Multivariate analysis, by contrast, enables one to control for possible correlates in an attempt to determine which of a myriad of variables is truly important for learner performance – this is presented in the next chapter.

Figure 51: Student Mathematics and Reading Performance in Regional Context

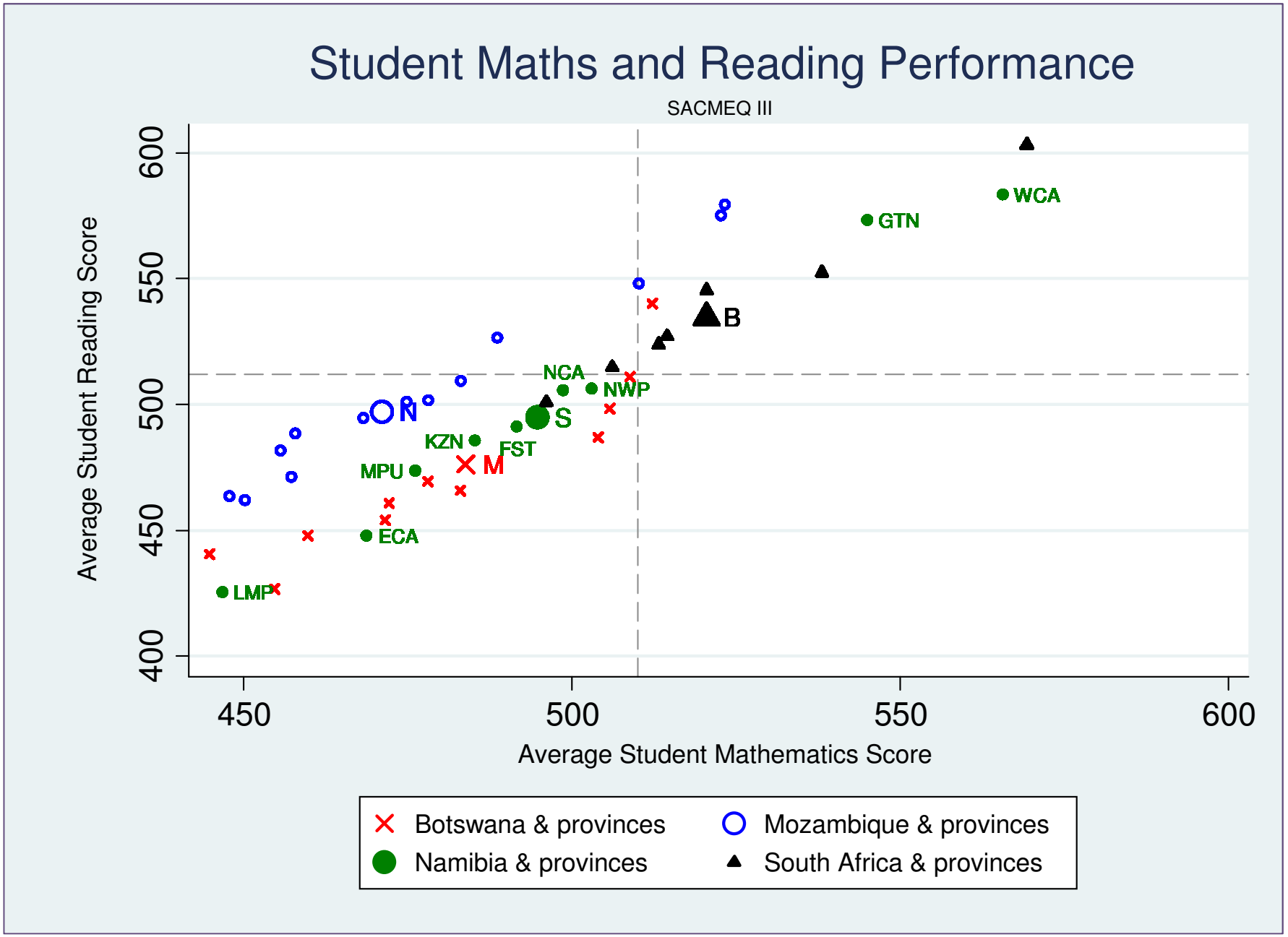
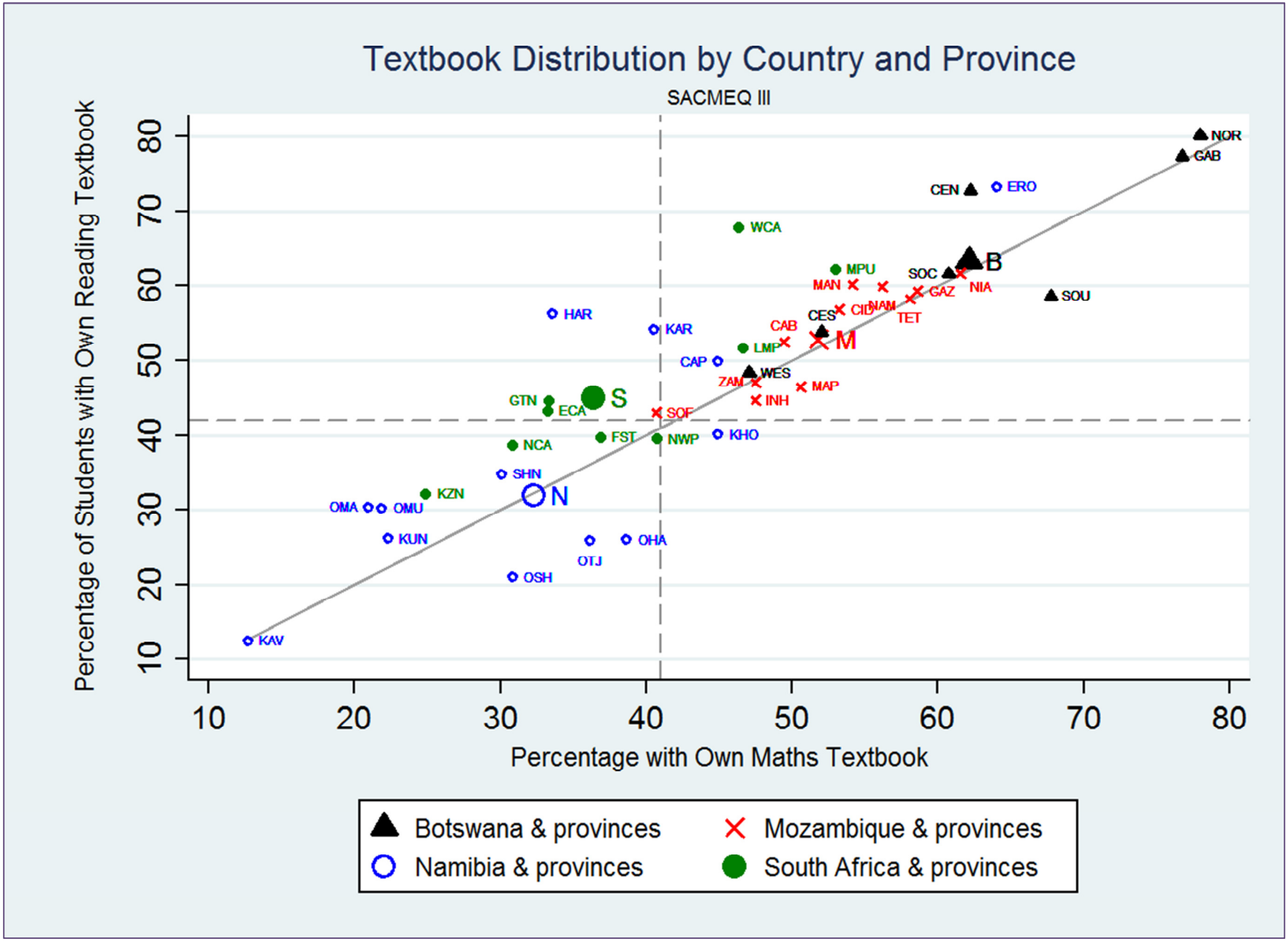


Figure 52: Textbook Distribution in Regional Context



Stellenbosch University http://scholar.sun.ac.za

Chapter 4 **Modelling Learner Performance**

From the preceding literature review, we already know that there are systemic inefficiencies in the South African primary education system. Looking internationally, South Africa performs poorly relative to comparable countries in Sub-Saharan Africa, and other middle-income countries around the world. It would seem that South Africa fails to convert its material advantage into academic performance. Leaving international comparisons aside, comparing learner performance to the expectations of the national curriculum again paints a dismal picture of general underperformance and high inequality. Similarly, the descriptive analysis showed that an unacceptably high proportion of learners, especially the poor and those in rural areas, are functionally innumerate and functionally illiterate, i.e. they have not acquired even the basic academic skills required by the curriculum.

However, the preceding bivariate analysis is purely descriptive and of limited use when trying to draw conclusions as to the generative mechanisms of learner performance. Given that many theoretically important variables are highly correlated with each other, it is often misleading to draw inference from these bivariate distributions. The unconditional correlation between two variables will often fall away when other variables are controlled for, indicating that some variables are merely proxying other important variables, and due to their correlation with those important variables, there seems to be a strong relationship. Multivariate analysis can go some way to account for these correlations and provide some indication of the true generative mechanisms.

Regression analysis allows the researcher to control for numerous factors, enabling one to make conditional recommendations: for specific values of grade-repetition, preschool exposure, learner-absenteeism, socioeconomic status and so on, a one unit increase in teacher absenteeism (a potential variable of interest) is associated with reading scores that are x percent higher/lower. In order to generalize, we measure the average impact of a given variable holding all other variables constant at their mean values. Following this methodology, a selection of variables was regressed on learner reading performance and learner mathematics performance. Variable selection was based on theory, and any available variable that could plausibly be thought to impact learner performance was included.

To account for the complex two-stage survey design of SACMEQ, STATA's built-in '*svy*' command was used in all regressions, with clustering by school and stratification by province – in accordance with the sampling structure used in the SACMEQ survey. Thus all regression specifications used STATA's *svyset with* the '*linearized*' sub-option within *svy regress.* In so doing, it accounts for both weighting and variance estimation in a complex two-stage survey design environment.

Stellenbosch University http://scholar.sun.ac.za

More specifically, I used the Ordinary Least Squares (OLS) estimation procedure with Taylor linearization to approximate the variance of the point estimators. This is also known as the delta method or the Huber/White/robust variance estimator[28] and is invoked using the *linearized* command in *svy*. The interest here is in modelling the relationship between learner performance (Y_j) and a number of explanatory variables (x_j) and thus one needs to estimate the vector of parameters β to solve the following equation:

$$G(\beta) = \sum_{j=1}^{M} S(\beta;\ Y_j, x_j) = 0 \qquad \text{(eq.1)}$$

Thus to solve the OLS estimator, we calculate:

$$G(\beta) = X'Y - X'X\beta = 0 \qquad \text{(eq.2)}$$

Where Y is the vector of learner performance in the population and X is the matrix of explanatory variables in the population. Since we do not have the full population data, we estimate using a weighted sample which yields:

$$\hat{G}(\beta) = \sum_{j=1}^{m} w_j S(\beta;\ y_j, \boldsymbol{x}_j) = 0 \qquad \text{(eq.3)}$$

Where the jth observation in the sample represents w_j elements in the population. Stata's *svy* command incorporates the complex two-stage sampling procedure. A first-order matrix Taylor-series expansion gives:

$$\hat{\beta} - \beta \approx -\left\{\frac{\partial \hat{G}(\beta)}{\partial \beta}\right\}^{-1} \hat{G}(\beta) \qquad \text{(eq.4)}$$

and a variance estimator for $\hat{\beta}$ of:

$$\hat{V}(\hat{\beta}) = \left[\left\{\frac{\partial \hat{G}(\beta)}{\partial \beta}\right\}^{-1} \hat{V}\{\hat{G}(\beta)\} \left\{\frac{\partial \hat{G}(\beta)}{\partial \beta}\right\}^{-T}\right]\Bigg|_{\beta = \hat{\beta}} = D\hat{V}\{\hat{G}(\beta)\}\Big|_{\beta = \hat{\beta}} D' \qquad \text{(eq.5)}$$

where D is $(X_s' W X_s)^{-1}$, W is the diagonal matrix of the sampling weights and X_s is the matrix of sampled explanatory variables. Thus $\hat{G}(\beta)$ can be written as

$$\hat{G}(\beta) = \sum_{j=1}^{m} w_j \boldsymbol{d}_j \qquad \text{(eq.6)}$$

where $\boldsymbol{d}_j = s_j \boldsymbol{x}_j$ and s_j is a residual for linear regression.

[28] All technical details on equations and formulae from this section are from StataCorp (2009, p. 161).

Stellenbosch University http://scholar.sun.ac.za

4.1 Sample Variations

Splitting the sample by sub-groups of school socioeconomic status

Since both the extant literature and the descriptive analysis point to bimodality in learner performance, it is of some interest to test the hypothesis that there are indeed two data-generating processes at play, and consequently to determine if the same factors are equally important for each of these two sub-sets of learners. Following the methodology of van der Berg (2008), I split the sample of 9071 learners into two groups using school socioeconomic status – a small group of wealthier schools, and a larger group of poorer schools - and then re-ran the regressions on each of these sub-samples. The reason for choosing *school* socioeconomic status rather than *individual* socioeconomic status is based on the understanding that the data-generating process is *school*-based rather than *home*-based. According to this view, learner performance is mainly determined by school level factors such as school discipline, school management, and teacher quality rather than home level factors such as individual wealth, parental education or educational resources at home. Thus, learner success or failure can be more easily predicted based on school-level variables rather than home-level variables, particularly so for poorer learners. It comes as little surprise then, that some of the largest impacts across all regression specifications are found at the school level. This is further justification for using school, rather than home, SES for the split regressions, i.e. school SES absorbs more of the explanatory power in these regressions than does home SES.

The descriptive analysis presented in Chapter 2 seemed to suggest that the appropriate place to draw the distinction between the two samples is somewhere between the 75th and 80th percentiles. To test whether or not the results are overly-sensitive to where the sample is split, I run the regressions for both splits: 1) the poorest three quartiles of school SES (75%) compared to the wealthiest quartile of school SES (25%), and 2) the poorest four quintiles (80%) compared to the wealthiest quintile (20%). This can be seen as a form of sensitivity analysis, and aims to test the robustness of the explanatory power of the variables, as well as their invariance to relatively minor sample changes.

It is important to remember that regression coefficients for dummy variables are calculated with reference to a base category *in that sample*. Therefore, one must take care when comparing coefficients between these two sample variations due to the drastically different samples. For example, in both the reading and mathematics regressions, the coefficient on 'extra tuition' is negative, large, and significant for the top quartile regression and the top quintile regression, while it is not significant for the other two, poorer samples. Since we expect the quality of education offered to the richest 20 to 25% of learners to be much higher than that offered to the poorest 75 to 80%,

Stellenbosch University http://scholar.sun.ac.za

learners in the top quintile will only attend extra tuition if they are performing particularly badly (i.e. they are weak learners). Thus this variable is most probably indicating which learners are underperforming and therefore attending extra classes. This may not be the case for poorer learners. Average learners, not only underperforming learners, may attend extra lessons due the lower quality of education provided to poorer learners in their normal school hours. Consequently, average learners attending extra lessons may moderate the signalling effect of this variable.

Following from the above discussion, it is important to document the mean value, standard deviation, minimum, maximum, and number of observations of each variable for each regression specification. This is provided in Appendix E. In addition, the "Diagnostic statistics" table (also in Appendix E) shows the mean value for each variable in each sample, as well as the percentage change in mean value between the two quartile groups, the two quintile groups, and quintile five as a percentage of quartile four and quartile one to three as a percentage of quintile one to four. The percentage changes have been visualized to facilitate easy identification across the many variables.

While limiting the sample to specific socioeconomic groups was an explicit objective of this research, there is also an implicit sample restriction in any regression. Since it is only possible to include learners who have observations on the full list of variables included in the model, including a variable with many missing values can also limit the sample. One such example is the teacher test score variable which is discussed below.

Stellenbosch University http://scholar.sun.ac.za

Although teachers were asked to complete the teacher test, they were allowed to refuse to write it. Subsequently, of the 498 reading teachers, 83 did not write the reading-teacher's test (16.7%); of the 498 maths teachers, 97 did not write the maths teacher's test (19.5%). This creates a problem if one wishes to include the teacher test score variable since doing so reduces the sample size by approximately the same percentage as the proportion of those teachers that did not write the test: roughly 15%. This is because it is only possible to include those learners in the sample whose teachers wrote the test, and thus have non-missing values for this variable. Since there is likely to be a sample selection issue at play, with weaker teachers refusing to take the test, it is possible that limiting the sample could bias the results[29]. If the missing values are not missing-at-random (MAR), which they are unlikely to be in this case, their exclusion will necessarily bias the coefficients. The question is therefore the severity of that bias, and not the presence or absence of it.

Although it may be possible to impute teacher-test scores, the fact that the selection process likely depends on the same variable as that which would be imputed (i.e. teacher knowledge) means that any imputation method would have its own complications. As such, teacher test scores were not imputed. In order to see whether limiting the sample would change coefficients in a material way, two regressions were run for reading and mathematics; one including teacher test score (and thus a smaller sample size), and one excluding teacher test score (with the full sample).[30] There are non-trivial benefits to including teacher test-scores in the analysis, but equally non-trivial are the assumptions required to include it. These benefits and costs are briefly outlined below:

The main *benefit* of including the teacher knowledge variable in the multivariate analysis is that South Africa has, to date, not conducted a nationally representative survey which has collected content knowledge on both learners and their teachers. Consequently, there has been little research on the impact of teacher knowledge on learner performance - at least not in a multivariate framework. Given the theoretical importance of teacher content knowledge, it would seem that the SACMEQ teacher test-scores present a valuable opportunity to increase our understanding of the impact of this important variable.

The main *costs* of including the teacher knowledge variable relate to the accuracy and representivity of the resulting smaller sample. The Diagnostic Statistics table in Appendix E show the mean value

[29] Another possible reason for refusing to write the test is that teachers may feel that their professional integrity is disrespected by asking them to write the tests, thus refusing to write the test is a form of protest rather than trying to hide weak subject knowledge. In this event, the sample selection effect is unclear.

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

Stellenbosch University http://scholar.sun.ac.za

for each variable in both the full sample and the (smaller) sample where the teacher test was included, as well as the percentage change in mean value between the two samples. The most noticeable, and most problematic, change occurs in the provincial distribution of learners. Clearly more teachers in the North West, Eastern Cape and Western Cape refused to write the test than did the other provinces. Consequently, by including the teacher test score variable there will be fewer learners from these provinces. This is most noticeable in the North West. Of the full sample, 10.1% consists of learners from the North West, while only 6.1% of the restricted sample (where the teacher test is included) is made up of learners from the North West. Put differently, there are some 40% fewer North West learners in the restricted regression compared to the full sample regressions. This is clearly undesirable. When observing the regression output in Appendix D one can see that, apart from a few relatively minor variables becoming significant or insignificant where previously they were not, the coefficients on most variables did not change between the two specifications. In addition, the full regression was re-run but the sample was limited to those learners where the teacher test score was available, even though the teacher test score was not included in this robustness-check regression. This was done in order to compare the exact same sample of learners with and without the teacher test score variable. Again, the size and significance of the coefficients hardly change between the two regressions.

Thus, while it is unfortunate that we do not have teacher test-scores for around 15% of learners and that there is no easy way to impute this variable, the opportunity to include a teacher subject-knowledge variable is valuable enough to warrant limiting the sample. Therefore, teacher test-scores were included in all other regressions[31].

4.2 Model Fit

The R-squared output of the two standard full-sample models, shows that the variables are better able to explain Reading-scores (0.6), than Mathematics-scores (0.5). This difference in explanatory power between Reading and Mathematics has been found elsewhere in the literature for similar data (Van der Berg, 2008: 27).

When the sample is split into two groups, (1) the top quintile of school socioeconomic status and (2) the bottom-four quintiles of school SES, it is interesting to see that the included variables are able to explain more variation in wealthy learner performance (42% for reading and 40% for mathematics) than poorer learner performance (35% for reading and 20% for mathematics). The most likely cause

[31] The full regression output for both the full sample and the limited sample can be found in Appendix D.

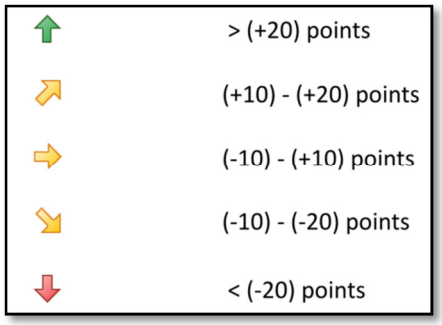
Stellenbosch University http://scholar.sun.ac.za

of this difference is that variables that are important for understanding poor learners' performance have been excluded from the model. For example, variables such as school management and teacher quality are thought to be extremely important in understanding why some poor schools perform better than others. If the variation in school management and teacher quality is greater between poor schools than between wealthy schools, as we expect to be the case, then the exclusion of these variables will affect the bottom-four-quintile regression more than the top-quintile regression. On a slightly different note, capturing these variables in a survey questionnaire is a difficult task, but necessary if one is to explain why some poor schools perform well in spite of their disadvantaged background.

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

4.3 Results from Multivariate Analysis

While the full output of all regressions used in this thesis can be found in Appendix D, selected results can be found in Table 9 to Table 14 below. The analysis is broadly split into four themes 1) Individual characteristics, 2) Household characteristics, 3) School characteristics, and 4) Provincial characteristics. These tables only report those coefficients that were significant at the 10% level or below. They also include diagrammatic indicators illustrating the size of the coefficient. A variable is deemed to have a large effect if the coefficient is larger than +20 or smaller than -20. This is roughly equivalent[32] to 0.2 standard deviations in the reading and mathematics scores. A variable is said to have a moderate effect if the coefficient is significant and ranges between +10 and +20, or -10 and -20. (I.e. between 0.1 and 0.2 standard deviations). If a coefficient is between -10 and +10, that variable is said to have a minor impact. These five levels of impact are shown graphically with arrows as indicated in the alongside legend. The full regression output of significant coefficients has been split according to these four themes with the interpretations interspersed between the tables. It is important to note that these are not separate regressions (one individual regression, one school regression, etc.) but are all included in a single regression – only the output here is reported separately. A complete list of



[32] The exact standard deviation for learner reading score is 116 and for learner mathematics score is 98.

Stellenbosch University http://scholar.sun.ac.za

regression output including significance levels, standard errors, and non-significant coefficients can be found in Appendix D.

4.3.1 Individual Characteristics

Age

The regression output[33] in Table 9 and

Table 10 below shows the impact of age, gender and preschool exposure on learner performance. Department of Education regulations were used to calculate which learners were age appropriate (11yrs 3mo - 12yrs 8mo) and thus which learners were underage (less than 11yrs 3mo), overage (12yrs 8mo – 13yrs) or severely overage (14 years and older). While the negative impact of being an underage learner ranges from minor to moderate, the impact of being an overage learner ranges from moderate to large. This effect is even more pronounced when considering severely overaged learners, where the impact is large and consistent across the regressions. It is also worth noting that a large proportion (15%) of the total Grade 6 sample was classified as *severely* overage, and an even larger proportion (25%) were classified as overage (see Diagnostic Statistics in Appendix E). Furthermore, overaged learners are far more likely to be found in quartile[34] one to three schools (28.3%) compared to quartile four schools (15.1%). The difference is starker when comparing the distribution of severely overaged learners: poorest three quartiles (19.7%) compared to the wealthiest quartile (2%).

Gender and Preschool education

The impact of gender on learner performance seems more pronounced for reading, where males do moderately worse than females, while for maths, males do moderately better than females – but only in the top quartile and top quintile regressions. Since SACMEQ III was the first of the SACMEQ surveys to ask learners about their preschool education, it is of particular interest to see the extent that such education impacts on literacy and numeracy performance. Learners who have at least one year of preschool education do better than learners with no preschool education, especially for reading where the impact is consistent across all regression specifications. The size of the impact ranges from moderate in the full sample and lower quintiles[35] sample, to large in the top quintile

[33] Since only the significant coefficients are shown in these tables, blank spaces indicate that those coefficients were not significant. The point estimates of these non-significant coefficients can be found in Appendix D.

[34] Throughout the multivariate analysis, any reference to quartiles or quintiles is referring to quartiles or quintiles *of school socioeconomic status.*

[35] Unless otherwise stated, throughout this chapter the interpretation refers to quintiles *and* quartiles, although only one is mentioned. This is to avoid using the cumbersome 'quintiles/quartiles' or repetitive use of 'quintiles and quartiles.'

Stellenbosch University http://scholar.sun.ac.za

sample. For all of the regressions, the coefficient on one year of preschool is not statistically different from two or three years of preschool. These results may speak to the issue of preschool quality, since the size of the impact in the top quintile sample is larger than that of the other regressions. It is reasonable to assume that the quality of preschool education for learners attending top quintile schools is substantially higher than that offered to other learners.

If these findings are correct, this may have important policy implications, namely that the Department of Education should focus on providing one year of quality preschool education to all children, rather than trying to offer multiple years of preschool education at the expense of quality. Preschool education is of particular interest to policy makers since it could well be a mechanism through which social mobility is improved. There is a large body of international literature indicating the importance of preschool education, and conversely, the negative effects of forgoing early education (see Gustafsson (2010) for a recent discussion of preschool education in South Africa). A lack of preschool education could place economically disadvantaged learners at an educational disadvantage[36] which further increases the number of hurdles these learners must overcome if they are to succeed at school and in later life.

Given the above, it is worrying to see the strong correlation between preschool education and wealth, as measured by socioeconomic status (see Chapter 2). As one would expect, poorer quintiles have less preschool education and higher quintiles have more preschool education. Almost 40% of learners in the poorest quintile receive no preschool education whatsoever[37]. Since the difference in performance between one year and more than one year of preschool education is small, the point of emphasis should be on those learners who received no preschool education.

Learner absenteeism

One perplexing result which does not seem to be related to a small sample problem or influential outliers, is why the *'absent for more than five days in the previous month'* variable has a *positive* impact on learners performance in the top quintile regression (in reading and mathematics), the bottom three quartiles regression (mathematics only), and in the full sample regression (mathematics only). In each regression there are between 2.1 % and 2.9% of learners who were absent for more than 5 days – i.e. for more than one week in the preceding month. It is unclear why this effect is positive (and significant) in these regressions and further investigation is required.

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

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

Stellenbosch University http://scholar.sun.ac.za

Table 9: Reading regression output: individual characteristics

	Student Reading Score (Mean: 498 / SD: 115)											
Sample:	Teacher-test only		Full Sample		Top Quartile		Bottom 3 Quartiles		Top Quintile		Bottom 4 Quintiles	
Young (<11y3m)	↘	-11.3	↘	-10.2			↘	-10.3				
Old (>11y3m-12y8m)	↘	-17.9	↘	-17.3	↓	-22.8	↘	-16.9	↓	-27.7	↘	-16.7
Very old (14y+)	↓	-22.6	↓	-21.9	↓	-58.1	↓	-21.9	↓	-94.4	↓	-21.0
Male	↘	-12.8	↘	-11.9	↘	-15.6	↘	-10.4	→	-9.8	↘	-12.7
> 5 Days absent									↑	32.9		
Preschool - months									↑	24.4		
Preschool - 1 year	↗	10.3	↗	10.1	↗	19.5	→	7.0	↑	22.0	→	9.7
Preschool - 2 years	↗	16.7	↗	15.6	↑	24.8	↗	11.2	↑	28.2	↗	14.8
Preschool - 3 years or more	↗	11.4	↗	12.8	↗	19.7	→	7.1	↑	26.3	→	7.8

Table 10: Mathematics regression output: individual characteristics

	Student Mathematics Score (Mean: 497 / SD: 99)											
Sample:	Teach. test only		Full Sample		Top Quartile		Bottom 3 Quartiles		Top Quintile		Bottom 4 Quintiles	
Young (<11y3m)	→	-9.7	→	-8.5								
Old (>11y3m-12y8m)	↘	-15.8	↘	-16.5	↓	-23.1	↘	-12.7	↓	-28.8	↘	-13.4
Very old (14y+)	↘	-15.4	↘	-16.9	↓	-36.5	↘	-14.0	↓	-42.7	↘	-14.9
Male					↗	10.8			↗	12.2		
> 5 Days absent			↗	12.8			↗	15.8	↗	15.4		
Preschool - months												
Preschool - 1 year	→	7.5	→	6.4	↗	14.8			↑	22.6	→	6.6
Preschool - 2 years									↑	21.2		
Preschool - 3 years or more	→	9.8	→	8.9	↑	21.0			↑	33.5		

4.3.2 Home-background Characteristics

Home Socioeconomic Status

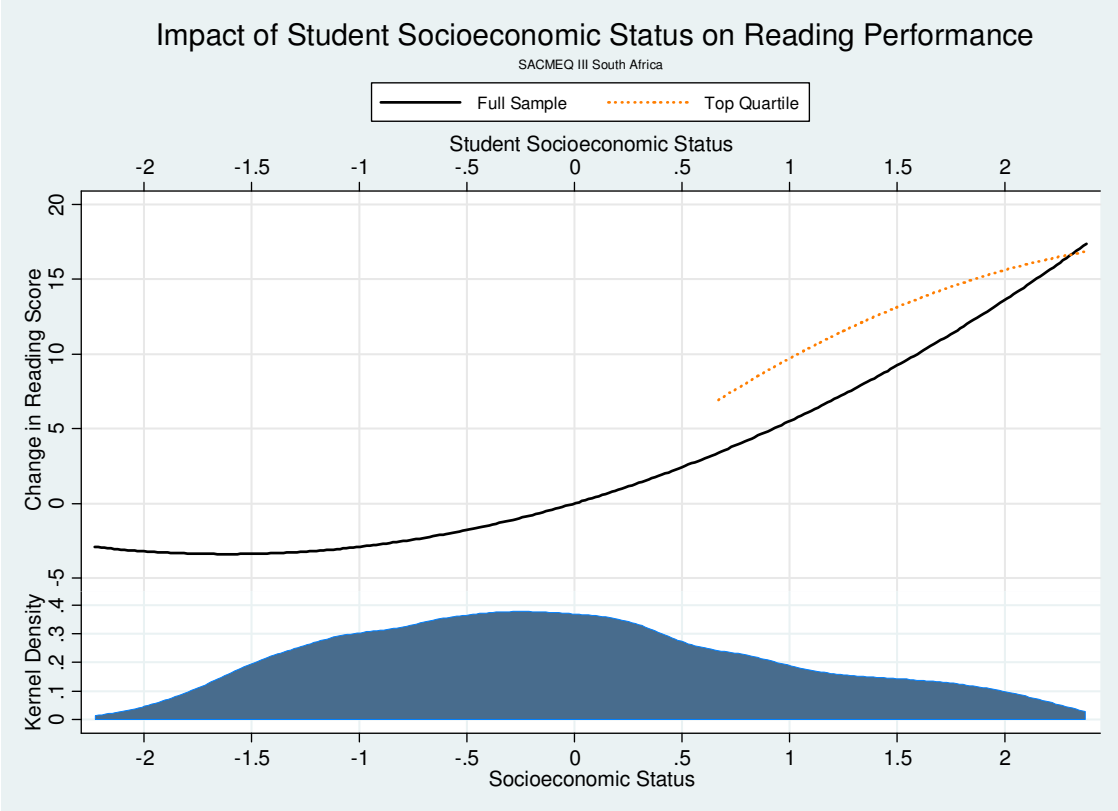
The impact of home-socioeconomic-status (SES) is minor in 11 of the 12 regressions (Table 11 and Table 12 below). This is primarily because school socioeconomic status - which is the average SES of learners in that school - absorbs most of the positive impact of belonging to a wealthier family. If one excludes the school-SES variables, the coefficients on the home-SES variables increase substantially. Nevertheless, when both sets of variables are included, the school SES variables absorb most of the explanatory power. One interpretation of this is that the average socioeconomic status of a learner's school is more important for academic success than that learner's individual

Stellenbosch University http://scholar.sun.ac.za

socioeconomic status. Put differently, a poor learner in a wealthy school will perform better than a wealthy learner in a poor school, *ceteris paribus.*

Figure 53 and Figure 54 below illustrate the impact of a change in learner socioeconomic status on reading and mathematics performance respectively. These figures were calculated using the regression coefficients on socioeconomic status and socioeconomic status squared from the regressions where these variables were significant. Only the full sample and top quartile samples had learner SES coefficients which were significant or jointly significant. These figures provide a better indication of the true impact of SES than do Lowess curves, for example, since they represent the impact of SES on reading performance *after* taking account of all other variables in the regression. More important than the *shape* of these curves, the *size* of the impact of a one standard deviation increase in SES should be noted. One full standard deviation increase in SES leads to, at most, a 10 point increase in learner reading or mathematics performance.

Figure 53: Impact of Learner Socioeconomic Status on Reading Performance



Stellenbosch University http://scholar.sun.ac.za

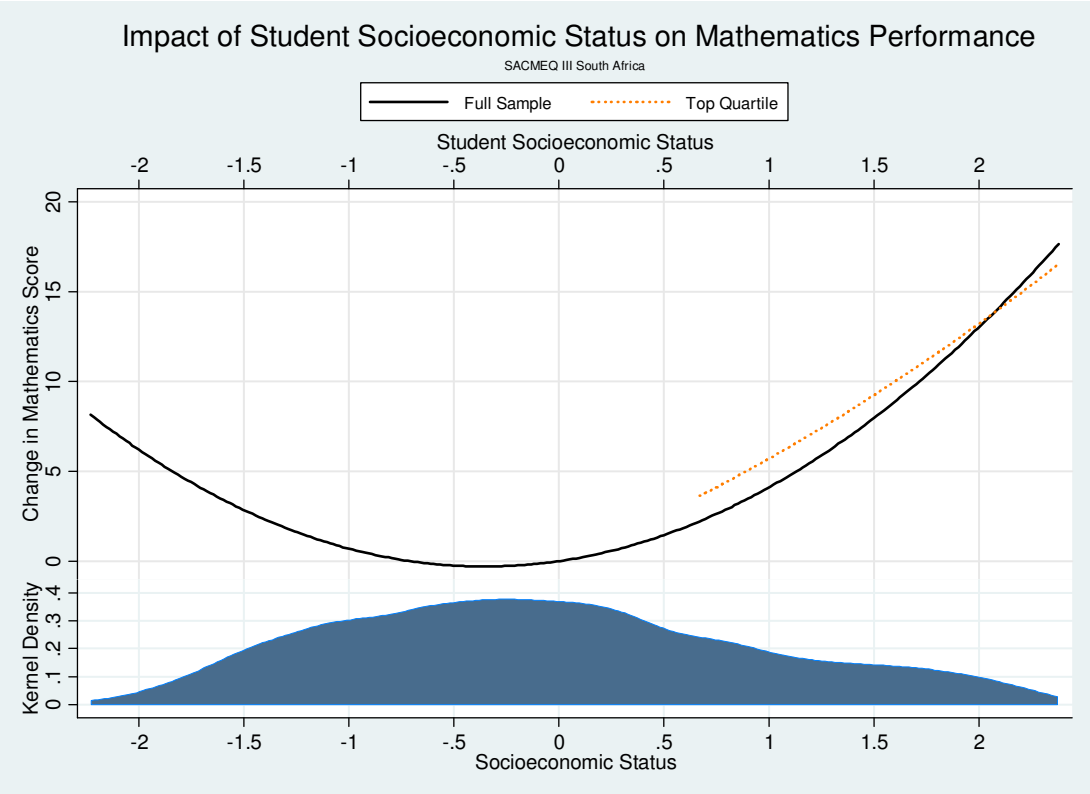


Figure 54: Impact of Learner Socioeconomic Status on Mathematics Performance

Familial background

If learners are living with their parents they perform slightly worse than those who do not live with their parents, however this effect is limited to reading performance, and even then only four of the six reading regressions. It is perhaps worth noting that 26% more quintile 5 learners live with their parents than quintile 1-4 learners (see Diagnostic Statistics in Appendix E). The impact of having three or more siblings is slightly negative and is significant in 10 of the 12[38] regressions. 60% fewer quintile five learners are in families with three or more siblings, compared to quintiles one to four.

Parental education seems to positively affect both reading and mathematics performance, as one would expect. If at least one parent has completed matric, a learner's performance is slightly higher in both reading and mathematics, with a large impact in the top quartile regression. For the majority of quintile five learners (83.4%), at least one parent had matric. Since the reference category (neither parent has matric) is relatively small, the coefficients are better interpreted as the negative impact of a lack of parental education rather than the benefit of additional parental education. If at

[38] There are twelve regressions since there are 6 for reading performance and six for mathematics performance. As can be seen in the tables, the 6 specifications are where the sample has been limited to (1) only those learners were the teacher test score variable was present, (2) the full sample, (3) the top quartile of school SES, (4) the bottom three quartiles of school SES, (5) the top quintile of school SES, and (6) the bottom four quintiles of school SES.

Stellenbosch University http://scholar.sun.ac.za

least one parent has a degree, this is also associated with small increases in reading in 3 of the 6 regressions, and slightly higher mathematics scores in all six of the mathematics regressions.

Missed Meals

Somewhat counter-intuitively, learners who reported that they normally missed breakfast at least once per week had slightly *higher* learner scores in 8 of the 12 regressions. However, in all 12 of the regressions learners who reported that they normally missed dinner at least once per week had scores which were moderately or considerably lower than those who ate dinner every day. The differential effect between breakfast and dinner suggests that missing dinner is a better indicator of food-poverty or malnutrition than missing breakfast. The Diagnostic Statistics in Appendix E seem to support this: 58% fewer quartile 4 learners reported missing *dinner* as compared to quartiles 1-3, compared to only 4% fewer quartile 4 learners who reported missing *breakfast*, as compared to quartiles 1-3. It is perhaps not unusual that children might miss breakfast once per week for reasons other than poverty.

Books at home

Learners who reported that there were more than 10 books in the place (home) where they stay during the school week achieved moderately better than those who had fewer than 10 books at home, with the effect being more pronounced in the wealthy sub-sample regressions. For the mathematics regressions, this variable is not significant for the poorer sub-sample regressions. The distribution of home reading resources across the top quintile and bottom four quintiles groups is highly unequal. While 73.1% of learners attending quintile five schools reported having more than ten books at home, only 25% of learners attending quintile one to four schools did.

Computer Use

The coefficients on the variable 'used *a computer before*' indicate that those learners who have used a computer before are more academically able than those that have not. These coefficients are positive and significant across all twelve specifications. The size of the impact ranges from moderate to large. While this relationship may be causal, i.e. computer use improves learner numeracy or literacy skills, it is also possible that this variable could simply be distinguishing between already better performing learners and weaker learners through socioeconomic status. On a technical note, it is prudent to ask whether this variable introduces multicollinearity (with SES) in the regression. Consequently the regressions were run with and without the '*Used PC*' variable. Since the coefficients on SES did not change substantially in size or significance, 'Used PC' was kept as a separate variable.

Stellenbosch University http://scholar.sun.ac.za

There is a stark contrast in computer use between learners attending the wealthiest 20% of schools compared to those attending the poorest 80% of schools. Almost all learners (97.8%) attending quintile five schools reported that they had used a computer before, compared to only 36.1% of learners in quintile one to four schools.

Frequency of English Spoken at Home

One of the largest effects found in the regressions is the impact of speaking English '*sometimes'* or '*always'* in the home environment. Learners who spoke English '*always'* in the home environment scored 0.4 standard deviations higher on average than those who did not. This effect is larger for reading than for mathematics, as we might expect.

These positive returns to speaking English at home can partially be explained by the fact that the SACMEQ III tests were only conducted in English and Afrikaans. As argued earlier, given that most South African learners do not speak either English or Afrikaans as a first language, these learners are likely to be at a disadvantage relative to their native English-speaking counterparts. Hence, one would expect those who spoke English '*sometimes'* or '*often'* in the home environment to better understand the literacy and numeracy tests, and thus to perform better.

Orphans and Orphanages

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

A particularly startling finding is the large negative effect of being in an orphanage or children's home. The regression results indicate that those learners who lived in an orphanage or children's home fared substantially worse in both reading and mathematics. The effect is large and stable across ten of the twelve regression specifications. Initially, one would be prudent in thinking that these results could be driven by only a few learners who live in the same orphanage and all attend a few underperforming schools. However, upon closer inspection of the data it becomes clear that this is not the case. The 58[39] learners (0.67% of the total) who indicated that they lived in an orphanage

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

Stellenbosch University http://scholar.sun.ac.za

or children's home each attended one of 46 different schools in 35 different districts across the nine provinces. Since these 58 learners were distributed across such a large number of schools and districts, one would expect that the orphanage dummy variable is not picking up school-level factors.

Alternatively, one could perhaps argue that orphanages or children's homes send learners to under-performing schools (perhaps due to resource constraints), in which case the '*orphanage'* dummy might simply be capturing poorly performing schools[40]. To ensure that this was not the case, the regressions were re-run including a school-level dummy variable, which took a value of 1 if there was a learner who lived in an orphanage or children's home in the school and zero otherwise. This variable was negative and significant at conventional levels, and took a value of -12.35 in the reading regression, and -10.59 in the mathematics regression. Even after including these dummy variables, the orphanage variable was still significant and remained large and negative in the reading regression (-25.1), and the mathematics regression (-28.98).

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

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

Stellenbosch University http://scholar.sun.ac.za

Table 11: Reading regression output: Home-background Characteristics

	Student Reading Score (Mean: 498 / SD: 115)											
Sample:	*Teacher-test only*		*Full Sample*		*Top Quartile*		*Bottom 3 Quartiles*		*Top Quintile*		*Bottom 4 Quintiles*	
SES	→	3.6	→	4.2	↗	11.6					→	3.7
SES squared	→	1.4	→	1.3	→	-1.9						
Lived with parents	→	-5.9	→	-5.2			→	-4.9			→	-4.7
3 or more siblings	→	-9.6	↘	-10.4			→	-8.3			→	-9.9
Missed breakfast	→	6.7	→	8.2			→	9.4			→	7.5
Missed lunch	→	-8.0	→	-8.7			→	-8.1			→	-6.9
Missed dinner	↘	-14.2	↘	-16.4	↓	-28.0	↘	-12.5	↘	-17.4	↘	-13.4
More than 10 books at home	↗	10.1	↗	10.3	↗	14.3	→	6.0	↗	19.3	→	7.4
Used PC before	↗	19.5	↗	17.1	↑	32.9	↑	20.9	↑	34.5	↗	19.8
Urban	↘	-11.3	↘	-10.7			↘	-16.3			↘	-12.2
Mother or father has matric	↗	10.9	↗	10.2	↑	21.8	→	8.6	↗	14.8	↗	10.7
Mother or father has degree	→	7.4	→	9.3			↗	11.0				
Speak Eng. at home sometimes	↑	20.9	↑	21.8	↗	19.4	↑	21.9	↗	14.0	↑	22.0
Speak Eng. at home always	↑	38.8	↑	40.6	↑	47.7	↗	12.4	↑	41.7	↑	31.0
Orphan (double-orphan)					↓	-23.1			↓	-31.1		
Orphanage or children's home	↓	-34.5	↓	-36.7			↓	-30.2	↓	-23.2	↓	-32.1

Stellenbosch University http://scholar.sun.ac.za

Table 12: Mathematics regression output: Home-background Characteristics

	Student Mathematics Score (Mean: 497 / SD: 99)					
Sample:	*Teach. test only*	*Full Sample*	*Top Quartile*	*Bottom 3 Quartiles*	*Top Quintile*	*Bottom 4 Quintiles*
SES	1.8	1.7	4.8	-0.8	-2.1	1.6
SES squared	2.5	2.4	0.9			
Lived with parents						
3 or more siblings	-6.7	-6.2	-6.9	-4.2	-7.7	-5.4
Missed breakfast	6.3	3.9		8.5	-9.5	8.8
Missed lunch						
Missed dinner	-15.8	-15.2	-34.0	-12.6	-26.1	-14.3
More than 10 books at home	5.7	6.1	16.1		23.1	
Used PC before	14.5	14.2	25.1	15.6	22.1	17.1
Urban	-9.1			-14.3		
Mother or father has matric	6.1	7.0	15.8	4.3	13.7	5.7
Mother or father has degree	10.1	9.4	7.7	11.4	9.3	7.8
Speak Eng. at home sometimes	17.6	17.5		19.3	12.4	18.3
Speak Eng. at home always	17.7	19.2	11.9			18.1
Orphan (double-orphan)			-26.9		-36.1	
Orphanage or children's home	-35.9	-33.2		-37.6	-20.1	-34.7

4.3.3 School Characteristics

The impacts of school characteristics on learner performance are likely to be large. To be sure, the main reason why the full sample of learners has been split along *school* socioeconomic quintiles, rather than *individual* socioeconomic quintiles for the regression analysis, is that there is reason to believe that the data-generating processes are *school-based* rather than *home-based,* as was discussed in Chapter 3.

School Socioeconomic Status

Of all the variables included in the model, school socioeconomic status has the largest impact on learner performance. However, this impact is concentrated among learners from the wealthiest 20-25% of schools. For these learners, attending a school with a higher average socioeconomic status drastically increases their reading and mathematics performance. Observing the slope of the full sample regression (solid line) in Figure 55 below illustrates this point well. At school socioeconomic levels below 0, attending a school with a higher average socioeconomic status has no noticeable impact on either reading or mathematics performance. To place this in perspective, we can calculate the impact of a one unit increase in school SES on learner reading performance. If there were 100

Stellenbosch University http://scholar.sun.ac.za

schools ranked from poorest (School 1) to richest (School 100), the impact would differ substantially across this distribution. To illustrate, if we moved a child from School 25 to School 79 (a one unit increase in school SES) her reading score would only increase by *6.85* points *ceteris paribus* - an extremely small impact given that the mean score is 500. By contrast, moving a learner from School 70 to School 97 (also a one unit increase in school SES) would increase her reading score by *37.35* points, almost 0.4 of a standard deviation in reading. It may be helpful to show the derivations behind the above example. To calculate the impact of a one unit increase in school SES, we first solve the partial derivative of:

$$\hat{G}(\beta) = \sum_{j=1}^{m} w_j S(\beta;\, y_j, \boldsymbol{x}_j) = 0 \qquad \text{(eq.7)}$$

with respect to k where k is the column element in the $\boldsymbol{x}_j$ matrix representing school socioeconomic status. This yields:

$$\frac{\partial \hat{G}(\beta)}{\partial k} = \{\, \beta_{SSES} + 2\, \beta_{SSES-sq}(k_i)\} \qquad \text{(eq.8)}$$

where k_i is the i^{th} observation of school socioeconomic status. We take k_i to be -0.574 which is the mean school SES for the 25th percentile of the school SES distribution. From Table 13 we see β_{SSES} equals 31.3 and $\beta_{SSES-sq}$ equals 21.3 such that equation 8 solves to 6.85 [31.3 - 2(21.3)(0.574)] when school socioeconomic status is -0.574[41]. Thus 6.85 is the marginal impact of a one unit increase in school SES at the 25th percentile of school SES. Similarly if we take k_i to be 0.142 which is the mean school SES at the 70th percentile of the school SES distribution, equation 8 solves to 37.35 [31.3 + 2(21.3)(0.142)]. Thus 37.35 is the marginal impact of a one unit increase in school SES at the 70th percentile of school SES.[42]

[41] Linking with the preceding example, moving a child from School 25 to School 79 is equivalent to increasing school SES from -0.574 (mean at p_{25}) to 0.59 (mean SSES at p_{79}) i.e. approximately a one unit increase in school SES.

[42] Linking with the preceding example, moving a child from School 70 to School 97 is equivalent to increasing school SES from 0.142 (mean at p_{70}) to 1.38 (mean at p_{97}) i.e. approximately a one unit increase in school SES.

Figure 55: School Socioeconomic Status and Reading Performance

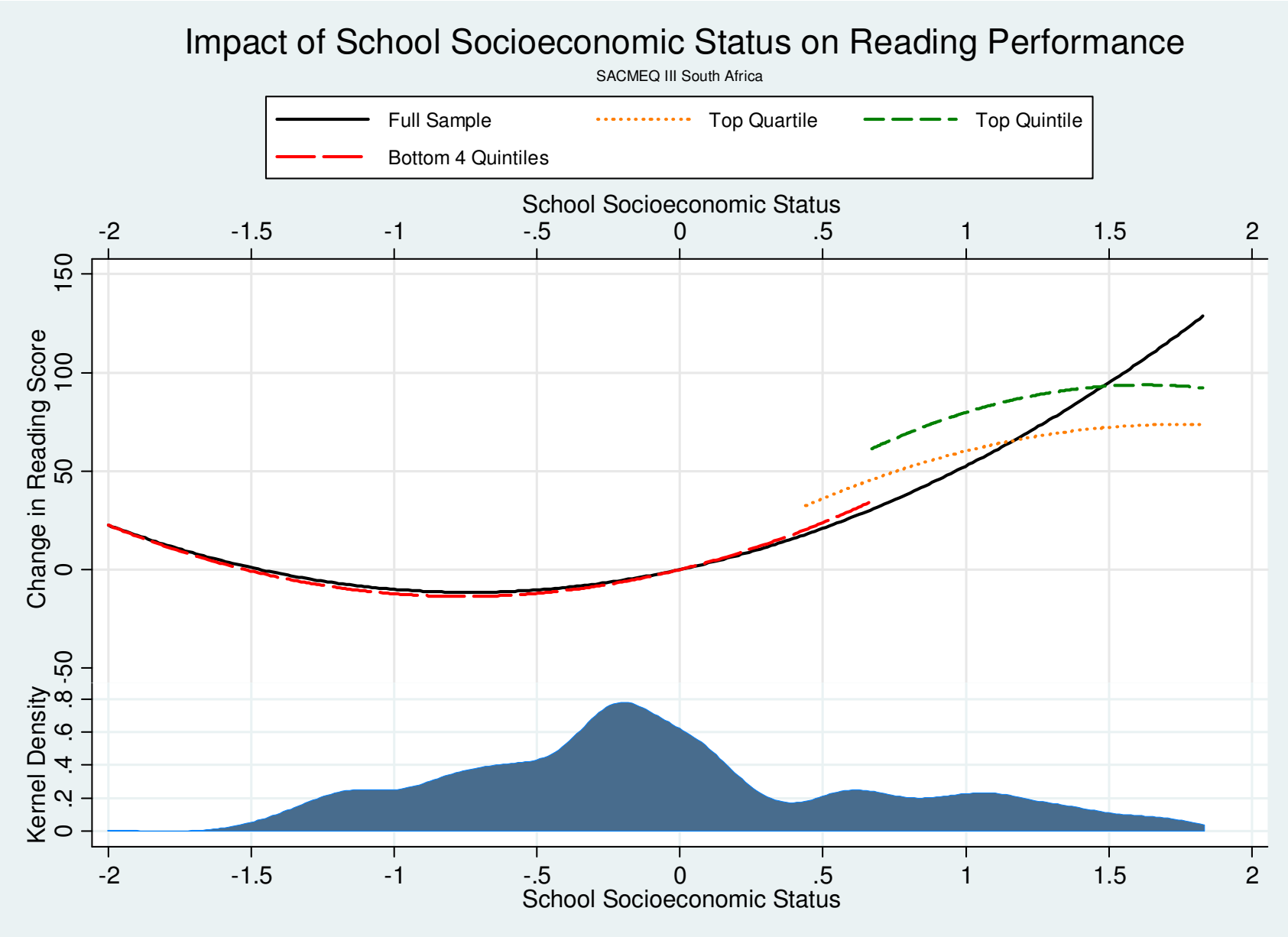
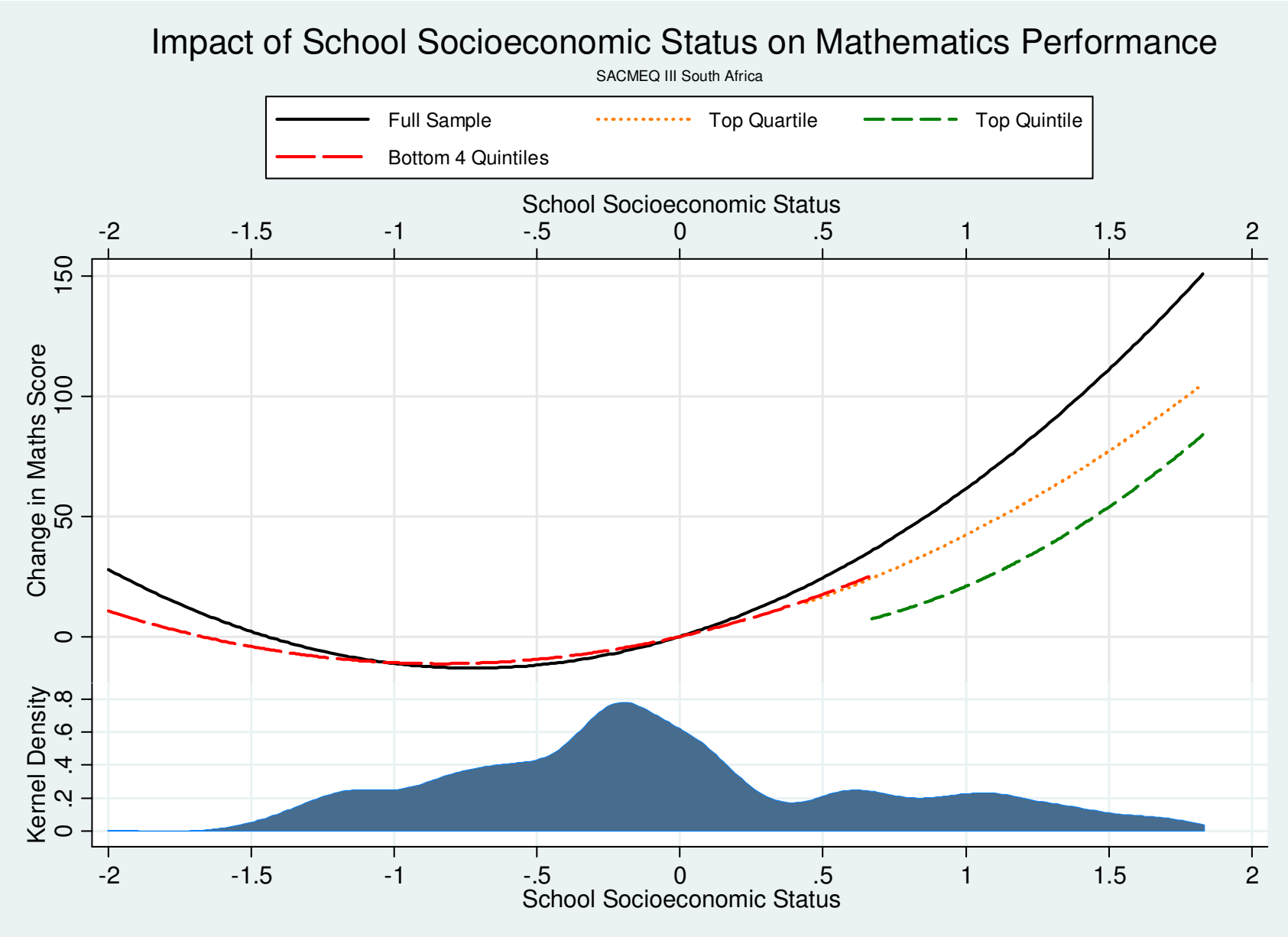


Figure 56: Impact of School Socioeconomic Status on Mathematics Performance

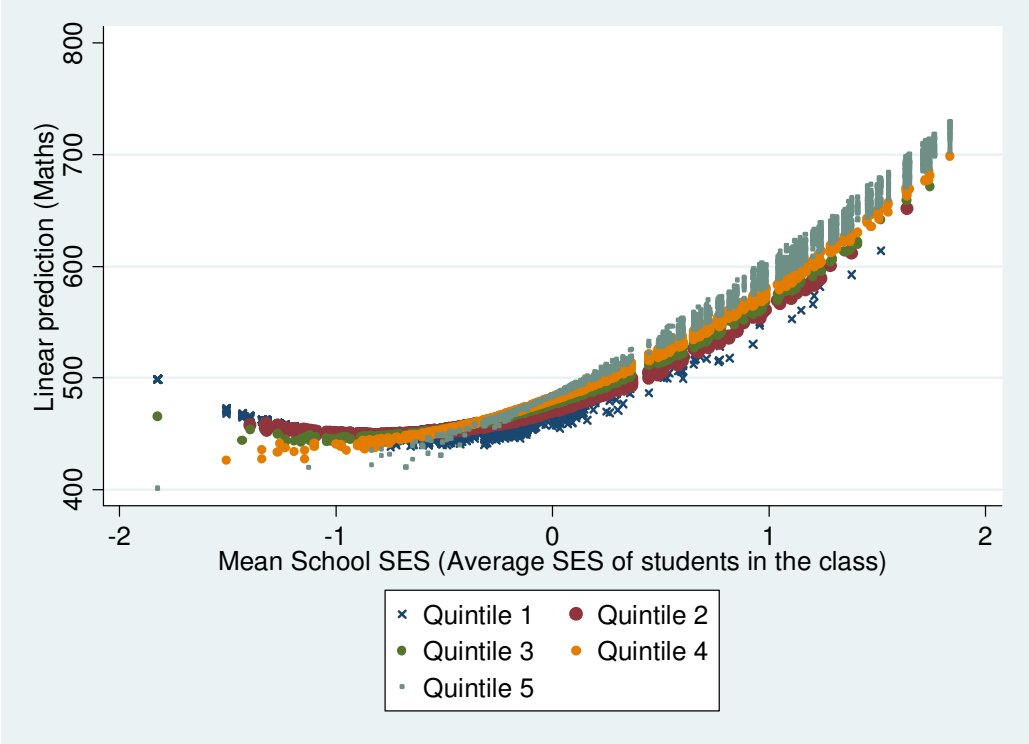


Stellenbosch University http://scholar.sun.ac.za

It is important to remember that the school SES variable included in the regressions is most probably also capturing elements of school *quality*, not simply school *wealth*. Affluent schools are more likely to exhibit those characteristics that we know are important for learner success. These include better school management, greater parental and governing-body involvement, sufficient school discipline, little teacher absenteeism, high teacher quality and motivation, and generally a more functional school environment, all of which aid learner learning and thus performance.

While the preceding graphs and analysis have shown that *school* SES is more important than *learner* SES in determining learner performance, it is unclear how these two variables are related. Since both were included in a quadratic form, interacting all four variables would make the interpretation of this combined effect cumbersome. Instead of including interaction effects, I predicted the learner mathematics score conditional on learner SES, school SES, and their quadratics. By plotting these predicted values on school SES (Figure 57 below) it becomes possible to see the variation in learner maths performance due to *individual* SES for each level of *school* SES. One important observation is that poor learners (the crosses in Figure 57) do not significantly underperform relative to their richer counterparts *for a given level of school SES.* Those poorer learners who attended wealthy schools experienced gains to school SES similar to wealthy learners. In Figure 57, this can be seen by the fact that the crosses (poorest learners) rise in a similar fashion to wealthier learners as school SES increases.

Figure 57: The impact of individual SES for a given level of school SES



Stellenbosch University http://scholar.sun.ac.za

Homework Frequency

Observing the full sample regressions in Table 13 and Table 14 below, it is clear that learners who received homework either once or twice a week or most days of the week, performed significantly better than learners who received homework less frequently. This effect is large across most of the regressions and is especially noticeable when learners received homework '*most days of the week*.' There was no discernable effect for learners who only received homework once or twice a month for either reading or mathematics in the full sample regressions. Learners in the lowest two quintiles of SES received the least homework overall. Between 12% and 15% of learners from the bottom four quintiles received homework only once or twice a month or not at all. This is compared to only 6% of top-quintile learners. 79.2% of quintile five learners reported receiving homework most days of the week compared to only 50.7% of learners from quintile one to four schools (Diagnostic Statistics, Appendix E).

Interestingly, the impact of homework frequency is larger for mathematics than for reading. In all six regression specifications for mathematics performance, the impact of receiving homework 'most days of the week' is large and significant – sometimes as large as 0.5 standard deviations of the mathematics distribution (top quintile regression).

Clearly learners who receive homework frequently are more literate and more numerate than those with less prescribed homework. The fact that the positive impact of homework is highly significant and stable across most regression specifications lends credibility to the notion that homework is important for learner performance. Given that this is a relatively easy and almost cost-free policy option, teachers should be encouraged to prescribe regular homework to learners.

Grade Repetition

The large negative impact of grade repetition is significant in all twelve regression specifications with the negative effect increasing as the number of grade-repetitions increases. Looking at the full sample regressions for both reading (and mathematics) the impact of repeating a grade once is -19.6 (-13.1), twice is -23.5 (-14.8), and repeating a grade three times or more is -38.7 (-24.0). Importantly, grade repetition is far more common in quintiles one to four where 9.7% of learners have repeated at least two grades, compared to only 1.5% of learners in quintile five schools. Thus, the negative impacts associated with grade repetition are experienced more frequently in quintile one to four schools.

Stellenbosch University http://scholar.sun.ac.za

Considering that grade repetition is meant to bring learners up to the required level by holding them back a year, it is disconcerting that even after repeating a grade, these learners perform consistently worse than those that did not repeat.

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

Average Parental Education of a School

An interesting finding arising from the multivariate analysis is the impact of the average level of parental education in a school. Learners in schools which had a higher proportion of learners who had parents with degrees did marginally better than learners in schools with lower levels of parental education. Although the size of this coefficient is small[43], it is significant in 8 of the 12 specifications. Interestingly, it is not significant in any of the lower SES regressions. One intuitively appealing interpretation of this variable is that the average level of parental education of a school has an impact on school functionality over and above all of the variables included in the regression. Perhaps parents with higher levels of education are more likely to hold the school accountable or be more involved in the school. This could be because these parents have a greater preference for education, indicated by their own higher levels of education, or, alternatively, due to more psychological factors such as the power relations between teachers, principals and parents, where more educated parents may feel less intimidated to confront school authority figures. It is important to remember that this impact is additional to the positive impact of own-parent education, since dummy variables for own-parent education have already been included. If this is the case, it would support Shepherd's (2011, p. 26) analysis of the PIRLS 2005/6 dataset, where she finds that parental involvement is only significant in the English/Afrikaans schools, which she uses as a proxy for former school department.

[43] The variable "Average Parent Degree" is also standardised to have a mean of zero and a standard deviation of one. Thus the coefficient represents the increase in learner performance for a one standard deviation increase in average parental degree education. The fact that this variable is standardised makes it difficult to interpret in an intuitive way, but the point remains that the higher the proportion of parents in a school who have a degree, the better is learner performance in that school, even after accounting for all the other variables in the regression.

Stellenbosch University http://scholar.sun.ac.za

Impact of textbook availability:

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

Similar to the trends seen in grade repetition and homework frequency, richer learners are far more likely to have access to reading textbooks than their poorer counterparts. Amongst the poorest 20% of learners, 36.8% either do not have a reading-textbook or must share with two or more learners. The figure for the richest 20% of learners is only 15.3%. By contrast, 71.3% of learners in the wealthiest quintile of schools have their own reading textbooks, compared to only 28.8% of learners in the poorest four school quintiles. Given that the reading-performance gains to reading textbooks are only evident when learners either have their own textbook or share with not more than one other, policy should focus on ensuring that no learner need share with more than one learner. Given the well-defined, and relatively low cost of this policy option, it would seem that providing reading textbooks where they are in short supply – particularly in poor schools – is the low hanging fruit of the South African primary education system.

Extra Tuition

The coefficient on 'extra tuition' is uniformly negative across all regression specifications except the bottom-four quintile and bottom-three quartile regressions. While initially this may seem counter-intuitive – attending extra lessons makes learners perform worse – this is only so if one interprets this coefficient in a causal way. If, as is likely to be the case, the 'extra-tuition' variable is instead correlated with the weakest performing learners, this coefficient is simply picking up the unexplained impact of this underperformance. As was explained in Section 4.1 (Sample Variations) above, the coefficients on extra tuition are not significant for the poorer sub-samples. This is most likely because learners who attend extra lessons in this part of the school system are not necessarily only underperforming learners, but could also be average learners who, given the low quality of education in their schools, may seek additional tuition to supplement the poor quality education they receive at school. This highlights the importance of interpretation in multivariate analysis.

Stellenbosch University http://scholar.sun.ac.za

Impact of teacher content knowledge on learner test scores

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

The full-sample-regression results indicate that teacher knowledge is statistically significant, with reading teacher knowledge and mathematics-teacher knowledge having a similarly sized impact on learner performance. However, it must be recognised that while this relationship exists, it is exceedingly small. The coefficients on the two teacher test score variables are 5.5 (reading-teacher), and 5.6 (mathematics-teacher). Given that both of these variables have been standardised to have a mean of zero and a standard deviation of one, increasing teacher content knowledge by one standard deviation (approximately 100[44] points) in each of the teacher test-score distributions, would only raise learner reading-scores by 5.5 points, and learner maths-scores by 5.6 points. When seen in light of the size of some of the other coefficients in the learner regressions, clearly teacher knowledge is not a significant determinant of learner test performance. This is in stark contrast to the initial assumptions of most researchers who would expect teacher content knowledge to have a large impact on learner performance. However, it is important to remember that these are Grade 6 (primary school) teachers, and the effect may be very different at higher levels of schooling.

To stress the small size of the teacher test-score coefficients, it is revealing to consider how learner performance would change if all teachers performed satisfactorily in the teacher tests. Since the teacher tests contained many of the same questions as the learner tests[45], one would expect all teachers to score almost full marks on the teacher test. This was certainly not the case.

Comparing the lowest performing decile of teachers with the best performing decile of teachers shows that there are large discrepancies between teacher knowledge: Reading from 641 to 931, and Maths from 612 to 991. Applying the coefficients from the standard regressions to these differences shows how little teacher knowledge impacts on learner performance. Thus the learner reading gain from raising the weakest performing 10% of teachers (with a score of 641) to be equivalent to the best-performing 10% of teachers (with a score of 931), is only slightly over 20 points. Similarly small results are found for mathematics teacher knowledge. These figures are comparable with the impact

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

Stellenbosch University http://scholar.sun.ac.za

size of far less dramatic changes such as increasing learner homework frequency to most days of the week.

This does not mean, however, that teachers do not matter, only that teacher knowledge is not as strongly correlated with teacher quality as one might have expected – at least not in this dataset. Factors such as teacher motivation or the ability of teachers to convey their subject-knowledge may better capture what makes a 'good' teacher, however this is only speculation. Thus, it would seem that the ability to teach learners well at the Grade 6 level is not very dependent on subject knowledge. The low impact of teacher content knowledge may also be the result of 'binding-constraints' in the school, which could include excessively large class sizes or poor school discipline. Thus, whether or not a teacher has high content knowledge may not matter if there are 60 learners in the class, or if the classroom environment is chaotic.

Interestingly, there is no significant impact for any level of teacher training or teacher education in either the bottom three quartiles, or the bottom four quintiles regressions. However, in the wealthy school regressions, the teacher education and teacher training coefficients are somewhat counter-intuitive. In the reading regressions, reading teachers with some further study did *worse* than teachers who reported that they had no formal schooling. In the mathematics regressions, teachers who reported that they had received three years of training did worse than those who had no teacher training. Given that these results are entirely counterintuitive, more investigation is needed regarding the cause of significance of these coefficients.

Provincial School Location

The provincial dummy variables included in the regressions show some important trends in province-level schooling characteristics. Interestingly, even after accounting for class size, socioeconomic status, school socioeconomic status, and the myriad of other factors in the regressions, there is still an unexplained portion of learner performance that can be distinguished based on province-level distinctions. In both reading and mathematics the only province to perform worse than the North West province (reference category) is Limpopo, with Mpumalanga having an insignificant coefficient. The Western Cape has the largest coefficient of approximately 30 points for reading and mathematics. Placing this in perspective, if we observed two learners with identical characteristics, with the only difference being that one was in the Western Cape and one was in Limpopo, the predicted difference in reading score would be approximately 60 points (the coefficient on Limpopo in the reading regression is -30.7). Given that education in South Africa is

Stellenbosch University http://scholar.sun.ac.za

managed at the provincial level, the fact that these coefficients remain after taking into account a variety of important factors seems to suggest that provincial differences matter a lot.

As has been indicated previously, the fact that 40% of the North West sample is excluded when the teacher test-score variable is included is cause for some concern. However, as a robustness check, all regressions were rerun using the Northern Cape as a reference category since the Northern Cape's representation in both the teacher-test-only regression and the full-sample regression is almost identical (there is a 1% difference between the two samples). While the provincial coefficients and the constant do change (as one would expect when changing a reference category), the point estimates and standard errors do not change when the reference category is changed.

Stellenbosch University http://scholar.sun.ac.za

Table 13: Reading regression output: School Characteristics

	Student Reading Score (Mean: 498 / SD: 115)					
Sample:	*Teach. test only*	*Full Sample*	*Top Quartile*	*Bottom 3 Quartiles*	*Top Quintile*	*Bottom 4 Quintiles*
School SES	33.7	31.3	84.3	20.4	115.2	35.9
School SES squared	19.6	21.3	-24.1	8.8	-35.4	23.6
Homework - 1 or 2 times a month					40.3	
Homework - 1 or 2 times a week	21.5	20.7		23.7		21.8
Homework - Most days	22.5	23.0		21.7		19.8
Repeated a grade once	-19.1	-19.6	-29.2	-16.2	-22.8	-17.8
Repeated a grade twice	-22.7	-23.5		-21.7		-22.5
Repeated a grade three or more	-40.4	-38.7	-52.1	-35.1	-74.6	-37.2
Sanitation (Std)		-6.6		-9.1		-6.9
Building Index (Std)		8.6	14.1		18.4	
Avg. Parent degree (school)	7.6	8.6	7.5		6.3	
R/M Textbook -Teacher only					34.3	
R/M Textbook -Share 2+			-32.7			
R/M Textbook -Share with 1	16.7	17.1	23.7	21.7	23.4	19.5
R/M Textbook - own textbook	14.9	11.7		17.6	17.6	14.8
Extra tuition (English/Maths)	-16.3	-15.9	-40.9		-43.5	
Reading-teacher Reading score	5.6				6.2	5.3
Teacher Male					15.7	
Teacher A-level / further study			-27.6		-31.5	

	Student Reading Score (Mean: 498 / SD: 115)					
Sample:	*Teach. test only*	*Full Sample*	*Top Quartile*	*Bottom 3 Quartiles*	*Top Quintile*	*Bottom 4 Quintiles*
Eastern Cape		21.5	29.7		25.6	
Free State						
Gauteng	27.9	30.3	22.1	15.1	16.5	22.2
KwaZulu Natal	17.5	20.8			20.1	15.1
Limpopo	-30.7	-27.9	-48.2	-26.0		-27.3
Mpumalanga						
Northern Cape	22.3	24.1		19.4		20.5
Western Cape	29.8	38.0		45.4		43.0
Constant	434.8	431.7	397.9	433.7	362.2	425.0
N	**7974**	**8822**	**8807**	**8238**	**8882**	**8163**
F-stat	**81.62857**	**83.2893**	**407.9038**	**17.23041**	**307.134**	**24.30282**
Prob > F	**0**	**0**	**0**	**0**	**0**	**0**
R-squared	**0.59966**	**0.60186**	**0.51788**	**0.26963**	**0.4169**	**0.34484**

Stellenbosch University http://scholar.sun.ac.za

Table 14: Mathematics regression output: School Characteristics

	Student Mathematics Score (Mean: 497 / SD: 99)					
Sample:	*Teach. test only*	*Full Sample*	*Top Quartile*	*Bottom 3 Quartiles*	*Top Quintile*	*Bottom 4 Quintiles*
School SES	35.6	36.4	23.9	7.4	-9.3	27.1
School SES squared	24.4	25.2	18.4	1.8	30.2	16.2
Homework - 1 or 2 times a month	12.1			14.1	47.0	
Homework - 1 or 2 times a week	26.1	19.9		29.3	36.1	26.1
Homework - Most days	27.6	22.5	25.4	27.6	50.2	24.8
Repeated a grade once	-12.5	-13.1	-22.0	-8.9	-21.4	-9.5
Repeated a grade twice	-17.6	-14.8		-16.4		-17.2
Repeated a grade three or more	-26.3	-24.0		-23.3		-25.3
Building Index (Std)			17.1		34.1	
Avg. Parent matric (school)				9.8		7.2
Avg. Parent degree (school)	6.1	6.8	19.1		16.3	
Extra tuition (English/Maths)	-14.1	-14.8	-24.3		-23.4	
Reading-teacher Reading score	5.5					6.3
Teacher A-level / further study			32.0		38.9	
Teacher Degree			33.2		29.5	
Teacher training: 2yrs			42.8			
Teacher training: 3yrs			-18.4		-21.9	
Teacher training: >3yrs						
Eastern Cape	23.1	24.5		23.5		23.7
Free State		13.2				
Gauteng	14.0	14.3	20.1			
KwaZulu Natal	13.7	14.0	31.1		31.4	12.8
Limpopo	-20.1	-19.9		-19.8		-19.6
Mpumalanga						
Northern Cape	18.7	19.7		15.2		14.3
Western Cape	29.5	33.2	42.2	33.0	37.9	34.4
Constant	435.1	443.0	379.9	441.3	359.0	435.3
N	7740.0	8803.0	8644.0	8167.0	8729.0	8082.0
F-stat	39.4	41.2	.	9.7	264.9	13.3
Prob > F	0.0	0.0	.	0.0	0.0	0.0
R-squared	0.5	0.5	0.4	0.2	0.4	0.2

Stellenbosch University http://scholar.sun.ac.za

Insignificant variables

While the above interpretation and analysis have focussed on only those variables that came through significant in at least one regression, there were also a number of variables that were included in the models but were uniformly insignificant (as can be seen in Appendix D) and consequently were not included in the tables above. Three variables were not significant in any of the twelve regression specifications (six reading and six mathematics). These were "*no class library*", "*Class size greater than 40*", and "*equipment index*". These results contradict those of Shepherd (2011) who found a significantly negative impact for class sizes greater than 30 using the PIRLS 2006 data, and van der Berg (2008) who found that school equipment[46] was a significant correlate of learner performance using the SACMEQ II data. To explore this dissimilarity, it would be instructive to use one common model, with as similar variables as is possible, and apply it to all three datasets (SACMEQ II, III and PIRLS 2006). This would eliminate the possibility that these differences are being driven by underlying specification differences, rather than differences in the generative mechanisms of learner performance.

The proportion of parents in a school who had a matric did not significantly influence reading performance, while the gender dummy *"teacher male",* and the sanitation index were both not significant in the mathematics regressions. Textbooks had no discernable effects in any of the mathematics regressions.

In addition to the above, many of the teacher training and teacher education variables did not come through significant

[46] Van der Berg (2008, p. 9) uses an index which was created by measuring if each of the following items were present in a school: first aid kit, fax machine, typewriter, duplicator, radio, tape recorder, overhead projector, TV, VCR, photocopier, and computer (11 items). The equipment index used in this study includes all of these items and 7 others, namely a clock, telephone, DVD/VCD player, cassette player, electricity (mains or generator), projector for computer images, and computer room.

Stellenbosch University http://scholar.sun.ac.za

Chapter 5 **Conclusion and Recommendations**

5.1 Conclusion

It is without question that South Africa's schooling system is dysfunctional in that it fails to fulfil its mandated role in society: educating South African youth. The consensus in the extant literature is unequivocal and additional studies seem to play the same tune in a different key, adding nuance and detail to our understanding, but never departing from the general theme of underperformance and inequality. In keeping with the trend, this thesis has used a newly available dataset - SACMEQ III - to better understand primary education in the country. Essentially, South Africa is still a tale of two school sub-systems: one which is wealthy, functional and able to educate students, while the other is poor, dysfunctional, and unable to equip students with the necessary numeracy and literacy skills they should be acquiring in primary school.

After a preliminary analysis of the data in Chapter 3, it became clear that a large proportion of Grade 6 learners were functionally illiterate and functionally innumerate. Reporting the prevalence of functional innumeracy and functional illiteracy (defined as the lowest two SACMEQ competency levels) was one of the important contributions of the thesis since this measure is easier to interpret than SACMEQ's standardised scores. Comparing these measures across a variety of subgroups illustrated that South Africa's average performance shrouds a host of inequalities, most notably those between provinces, school socioeconomic quintiles, and geographic location. While only 1 in 20 learners in the Western Cape are functionally illiterate, the comparable figure in Limpopo is 10 in 20. Similarly, 59% of learners in the poorest socioeconomic quintile are functionally innumerate, while only 5% of learners in the wealthiest socioeconomic quintile are thus classified. Given that educational performance is largely driven by socioeconomic status - which is only loosely correlated with ability or motivation - one can conclude that South African primary education is inequitable. It does not offer equal educational opportunities to all children.

In addition to the descriptive statistics, the large and significant impact of school socioeconomic status shows that educational outcomes are largely a function of wealth: if a family is able to send their children to affluent, functional schools, those children's educational prospects are auspicious. Unfortunately only 20-25% of schools fall into this category. Consequently, the majority of learners in the country must attend schools that are not vehicles for social mobility, but rather propagating mechanisms which entrench poverty and social disadvantage.

Stellenbosch University http://scholar.sun.ac.za

South Africa's performance in regional context provides an indication of the inefficiency of the primary school system. South Africa's public expenditure on primary education is magnitudes higher than that of Kenya, Tanzania or Zimbabwe, and yet these countries' learners significantly outperform South African learners in both reading and mathematics. Sample selection issues arising from differences in enrolments and drop outs in these countries may explain some of their high performance. However, Botswana and Mauritius have similar enrolment and drop-out rates to South Africa - and also similar public expenditure on primary education - and yet South Africa performs well below both of these countries. Clearly South Africa is less able to convert material inputs into educational outputs. In short, South African primary education is inefficient - it does not fulfil its mandate of educating children to the standards of the curriculum, and it produces worse educational outcomes with more resources than other comparable African countries.

After highlighting the low absolute and relative performance of South African learners, the thesis indicated that some pedagogical inputs seem to be more important than others. Homework frequency and reading-textbook availability impact learner performance more than teacher content knowledge. This was one of the more surprising findings of the research, and given the centrality of teachers to primary education, more research is required before strong conclusions are drawn.

In addition to the above determinants, many learners enter Grade 6 with a variety of pre-existing disadvantages. Learners who were under or over-aged, had no preschool exposure, or were repeating a grade, did markedly worse than their peers, as did those who had less-educated parents, less household wealth, fewer books at home, and less exposure to English at home. While the state has no control over some of these variables, there are also policy options available to educational planners that, at least according to these results, will have a significant impact on student learning. These are elucidated in the *Policy Recommendations* section below.

One of the explicit goals of the post-apartheid government has been to try and level the playing field such that life opportunities are not determined by parental wealth, race, gender, or any other factor over which learners have no control. Given the close links between education and labour market outcomes, this meritocratic ideal is not possible without first addressing South Africa's unequal education, something which the country has, to date, failed to accomplish. While almost all South African children have access to education, the quality of that schooling is very different for different sub-groups. It is unfortunate but true that the current educational system lacks the ability to educate most of the youth in South Africa. It is not effective. It is not efficient. It is not fair. Until such a time as the education system can provide a quality education to all learners, not only the wealthy, the current patterns of poverty and privilege will remain unabated.

Stellenbosch University http://scholar.sun.ac.za

5.2 Policy Recommendations

The descriptive and multivariate analyses presented in this thesis have highlighted some areas that are significant determinants of learner performance. In addition to contributing to the broader understanding of primary school learner performance in the country, some of the findings have direct policy relevance, and should be part of the national educational policy discourse. Of the many findings presented in this thesis, eight have logical policy implications. These are outlined below in increasing order of complexity and cost.

I. Homework frequency

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

However, due to differing home-backgrounds, homework frequency may not have the same impact for poor learners compared to more affluent learners. This is because wealthier learners are more likely to be encouraged to complete their homework by their parents or caregivers. This is in stark contrast to poorer learners who often have many chores to do after school, may have no access to electricity, and little private time to work. In addition, poorer learners may have to work after school hours in order to supplement the low household income. Consequently, innovative solutions such as "after-school home-work clubs" or similar initiatives may be necessary if poorer learners are to reap the benefits of increased homework.

II. Access to reading textbooks

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

III. Teacher absenteeism

Given the central and fundamental role of teachers in the learning process, the high levels of teacher absenteeism in the country are unacceptable. The fact that the average Grade 6 teacher in South Africa is - by their own admission - absent for an entire month in the school year should provoke

Stellenbosch University http://scholar.sun.ac.za

education officials to implement corrective measures. Teachers need to be in class teaching to provide learners with sufficient opportunity to learn.

IV. Grade repetition

The large and consistent negative association of grade repetition requires policy attention. Clearly the current approach to grade repetition is not working since students are not learning the skills they should have acquired the first time around. Some form of remedial teaching seems to be required. Further analysis, particularly of appropriate panel data, is likely to yield greater insight into this important issue.

V. Preschool education

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

VI. Teacher knowledge and quality

Teachers' subject expertise has a very small positive impact on learner performance at a Grade 6 level. Given the centrality of teachers to the education industry, further analysis is needed to ascertain *why* teacher content knowledge translates so weakly to improved learner performance.

VII. School quality

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

VIII. South Africa's regional performance

Given South Africa's status as a middle-income country, education policy makers should ask how it is possible that primary schooling systems in neighbouring low-income countries are able to outperform South Africa, when South Africa has a clear resource advantage. A poignant example is how Tanzania outperforms South Africa for every sub-population (rural-urban, rich-poor, male-

Stellenbosch University http://scholar.sun.ac.za

female, overall) when South Africa's GDP per capita is more than ten times higher than that of Tanzania[47].

5.3 Future Research

Social policy research, like all research, moves forward incrementally. New findings are reviewed, tested, challenged by fellow researchers and policy makers, and if the findings or methods hold up to such scrutiny, the findings are eventually incorporated into the existing body of evidence, and so the discourse changes over time. This verification process is important, especially in education where the policy implications affect so many people in such real, tangible, and lasting ways. In addition to suggestions that the findings of this thesis be verified by other, independent researchers, possible areas of future research are also highlighted in this section.

One avenue of future research involves the verification of the above results using other South African datasets. It is reasonable to assume that the underlying generative mechanisms of learner performance do not change quickly so it should be possible to verify the above results by using an alternative source of data but using the same education production function. The most logical of these alternative data sources is the PIRLS 2006 study. Given that it was conducted only one year before SACMEQ III, tested a similar primary school grade (Grade 5), and that many of the background questions were very similar, it becomes possible to create a generic education production function that can be run on both the SACMEQ III reading data and the PIRLS 2006 reading data. If the same variables are consistently significant, and of similar magnitudes (after standardising and adjusting them), this would add credibility to the assertion that they are key generative mechanisms of learner performance in South Africa. Unfortunately, since PIRLS did not test teachers, that dataset, or any other large South African dataset for that matter, is unable to confirm or deny the significance of the teacher content knowledge variable.

One of the unique features of the SACMEQ study is that the survey was administered in each of the 14 participating countries in southern and eastern Africa. This provides the research community with comparable data on the performance and characteristics of learners and their teachers, as well as accurate data on the distribution of human and physical resources. This cross national comparability should be exploited, especially since SACMEQ represents the only independent educational evaluation in many of the participating countries. An indication of the scope and magnitude of the SACMEQ III study is that 61,396 Grade 6 learners, 8,026 teachers, and 2,779 schools were surveyed in nationally representative samples in each of the 14 participating countries. Even if this survey did

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

Stellenbosch University http://scholar.sun.ac.za

not collect information on the cognitive abilities of learners and their teachers (which it did), such a large nationally-representative and cross-nationally comparable survey would be immensely useful. A clear area of future research is to compare the generative mechanisms of learner performance across countries, and for those countries that participated in SACMEQ II, also across time. Some researchers have already begun to conduct such research, see for example Hungi & Thuku (2010).

A limitation of the existing structure of the SACMEQ survey is that it is cross-sectional rather than longitudinal. While this means that the costs of data collection are lower, and it is less administratively complex, it does limit the types of analyses that can be conducted using the data. For example it is not possible to accurately estimate the impact of grade repetition without longitudinal data. Similarly, longitudinal data allows one to model the *gain* in learner performance over the period rather than simply learner performance at one point in time. Panel data of this type allows one to control for prior learning (and in a sense prior resources) and thus isolate the causal mechanisms *in that particular year* - something that is far more useful to policy makers. Lee *et al.* (2005) provide practical guidance as to how such a longitudinal design could be incorporated into SACMEQ:

> "As an experiment, we suggest that in perhaps three countries that are known for especially high-quality data collection (e.g., Kenya, Botswana, Namibia) in any SACMEQ data-collection year there be two different data collections in the same schools - perhaps in fourth and sixth grades ... we suggest that SACMEQ consider at least a pilot design for collecting panel data on the same students at different time points. Quite simply, solid studies of school effectiveness require longitudinal data" (Lee *et al.,* 2005, p. 240).

Although the time intervals between SACMEQ surveys is not usually known beforehand, the possibility to retest the same individuals at a later grade, and thus create a panel dataset, warrants further exploration into issues such as tracking learners and sampling over time.

The fact that SACMEQ III South Africa has an accurate measure of teacher content knowledge means that this dataset, and specifically this content knowledge variable, is likely to be the subject of much future research. Modelling the content knowledge of teachers by using a variety of demographic and school level variables is likely to yield some insight into the conditional correlates of teacher content knowledge. A better understanding of such knowledge may help to explain why this variable has such a seemingly small impact on learner performance.

The extant research on primary education in South Africa has shown the value of independent assessments of learner performance. These survey datasets help both policy-makers and researchers advance their respective agendas. Consequently, while they may be expensive and administratively

Stellenbosch University http://scholar.sun.ac.za

onerous, nationally representative surveys which include cognitive assessments should be encouraged and expanded. In the same way that previous rounds of the SACMEQ study (II and III) have yielded numerous insights into the level and drivers of learner performance in South Africa, SACMEQ IV is likely to be as valuable in contributing to our understanding of these important issues.

Stellenbosch University http://scholar.sun.ac.za

Reference List

Act, S. A. (1996). *The South African Schools Act: Act no.84 of 1996.* Pretoria: National Government RSA.

Act, S. A. (2002). *No. 50 of 2002: Education Laws Ammendment Act, 2002.* Pretoria: Government Gazette.

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

Becker, G. (1962). Investment in Human Capital: A Theoretical Analysis. *The Journal of Political Economy, 70*(5), 9-49.

Bertram, C. (2006). Exploring teachers' reading competencies: A South African case study. *Open Learning, 21(1)*, 5-18.

Bhorat, H. & Oosthuizen, M. (2006). *Educational Outcomes in South Africa: A Production Function Approach.* Cape Town: Development Policy Research Unit (DPRU).

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

Bray, M. (1986). If UPE is the Answer, What is the Question? *International Journal of Educational Development, 6*(3), 147-158.

Burns, J. (2001). Inheriting the future: the role of family background and neighborhood characteristics in child schooling outcomes in South Africa. *DPRU/FES conference.* Johannesburg.

Card, D. (1999). Causal effect of education on earnings. In O. Ashenfelter, & D. Card, *The handbook of Labor Economics* (pp. 1801-1863). Amsterdam: North-Holland.

Carnoy, M., Chisholm, L., Addy, N., Arends, F., Baloyi, H. & Irving, M., (2011). *The Process of Learning in South Africa: The Quality of Mathematics Teaching in North West Province.* HSRC Press.

Carnoy, M., Chisolm, L., *et al.* (2008). *Towards Understanding Student Academic Performance in South Africa: A Pilot Study of Grade 6 Mathematics Lessons in South Africa.* Pretoria: HSRC.

Stellenbosch University http://scholar.sun.ac.za

Case, A. & Deaton, A. (1999). School Inputs and Educational Outcomes in South Africa. *Quarterly Journal of Economics, 114*(3), 1047-1084.

Coleman, J., Campbell, E., Hobson, C., McPartland, J., Mood, A., Weinfeld, F. et al. (1966). *Equality of educational opportunity.* Washington, DC: US Government Printing Office.

Crouch, L. & Mabogoane, T. (1998). When the residuals matter more than the coefficients. *Studies in Economics and Econometrics, 22*(2), 1-14.

Crouch, L. & Mabogoane, T. (2001). No magic bullets, just tracer bullets: The role of learning resources, social advantage, and education management in improving the performance of South African schools. *Social Dynamics, 27*(1), 60-78.

Currie, J. (2009). Healthy, Wealthy, and Wise: Socioeconomic Status, Poor Health in Childhood, and Human Capital Development. *Journal of Economic Literature, 47*(1), 87-122.

DBE. (2011a). *Macro Indicator Trends in Schooling: Summary report 2011.* Retrieved from Department of Basic Education: http://www.education.gov.za/LinkClick.aspx?fileticket=lkzSZYKJGno%3D&tabid=422&mid=1261

DBE. (2011b). *A SACMEQ South Africa Report on the Health Knowledge off Learners on HIV and AIDS.* Pretoria: Department of Basic Education.

DBE. (2011c). *Report on the Annual National Assessments of 2011.* Pretoria: Department of Basic Education .

Del Rosso, J. (1999). *School Feeding Programs: Improving effectiveness and increasing the benefit to education. A Guide for Program Managers.* Oxford: Partnership for Child Development.

DoE. (2003). *Systemic Evaluation: Foundation Phase (Mainstream).* Pretoria: Department of Education.

DoE. (2008a). *Grade 3 Systemic Evaluation 2007 Leaflet.* Pretoria: Department of Education.

DoE. (2008b). *Ministerial Committee on Learner Retention in the South African Schooling System report.* Pretoria: Department of Education.

Donalson, A. (1992). *Content, Quality and Flexibility: The Economics of Education System Change. Spotlight 5/92.* Johannesburg: South African Institute of Race Relations.

Stellenbosch University http://scholar.sun.ac.za

Filmer, D. & Pritchett, T. (1999). The Effect of Household Wealth on Educational Attainment: Evidence from 35 Countries. *Population and Development Review, 25*(1), 85-120.

Fisher, I. (1906). *The Nature of Capital and Income.* The Macmillan Company.

Fiske, E. & Ladd, H. (2004). *Elusive Equity: Education Reform in Post-apartheid South Africa.* Washington: Brookings Institution Press / HSRC Press.

Fleisch, B. (2008). *Primary Education in Crisis: Why South African schoolchildren underachieve in reading and mathematics.* Cape Town. : Juta & Co.

Glewwe, P. (2002). Schools and Skills in Developing Countries: Education Policies and Socioeconomic Outcomes. *Journal of Economic Literature, 40*(2), 436-482.

Gustafsson, M. (2005). *The Relationship Between Schooling Input and Output in SA - Methodologies and Policy Recommendations based on 2000 SACMEQ.* Paris: SACMEQ.

Gustafsson, M. (2010). Policy note on pre-primary schooling: An empirical contribution to the 2009 Medium Term Strategic Framework. *Stellenbosch Working Paper*.

Gustafsson, M. & Mabogoane, T. (2010). South Africa's economics of education: A stocktaking and an agenda for the way forward. *Stellenbosch Economic Working Papers: 06/10*.

Hanushek, E. (1997). Assessing the effects of school resources on student performance: an update. *Educational Evaluation and Policy Analysis, 19*(2), 141-64.

Hanushek, E. (2003). The Failure of Input-Based Schooling Policies. *The Economic Journal, 113*(February), F64-98.

Hanushek, E. & Kimko, D. (2000). Schooling, labor force quality, and the growth of nations. *American Economic Review, 90*(5), 1184-208.

Hanushek, E. & Woessmann, L. (2008). The Role of Cognitive Skills in Economic Development. *Stanford Institute for Economic Policy Research Working Paper No. 07-34*.

Heckman, J., Lochner, L. & Todd, P. (2006). Earnings functions, rates of return and treatment effects: The Mincer equation and beyond. In E. Hanushek, & F. Welch, *Handbook of the Economics of Education* (pp. 307-458). Amsterdam: North-Holland.

Heyneman, S. (2003). Education and Social Cohesion. In J. (. Guthrie, *Encycolpedia of Education* (pp. 2242-2250). New York: MacMillan Publishers.

Stellenbosch University http://scholar.sun.ac.za

Hoadley, U. (2010). *What do we know about teaching and learning in primary schools in South Africa?* Stellenbosch: Appendix B to Van der Berg, S; Meyer, H; Reeves, C; van Wyk, C; Hoadley, U; Bot, M; & Armstrong, P 2010. 'Grade 3 Improvement Project: Main report and Recommendations" for Western Cape Education Department.

Hoadley, U. & Ward, C. (2009). *Managing to Learn: Instructional Leadership on South African Secondary Schools.* Cape Town: HSRC Press.

Howie, S., Venter, E., Van Staden, S., Zimmerman, L., Long, C., & Du Toit, C. (2008). *PIRLS 2006 Summary Report: South African Children's Reading Literacy Achievement.* Pretoria: Center for Evaluation and Assessment.

Hungi, N., & Thuku, F. (2010). Variation in reading achievement across 14 southern African school systems: What factors matter? *International Review of Education*, 63-101.

Hungi, N., Makuwa, D., Ross, K., Saito, M., Dolata, S., & Van Capelle, F. (2010). *SACMEQ III Project Results: Pupil Achievement Levels in Reading and Mathematics.* Paris: Southern and Eastern Africa Consortium for Monitoring Educational Quality.

Hungi, N., Makuwa, D., Ross, K., Saito, M., Dolata, S., & Van Capelle, F. (2011). *SACMEQ III Project Results: Levels and Trends in School Resources among SACMEQ School Systems.* Paris: Southern and Eastern Africa Consortium for Monitoring Educational Quality.

Keswell, M. & Poswell, L. (2004). Returns to Education in South Africa: A Retrospective Sensitivity Analysis of the Available Evidence. *South African Journal of Economics, 72*(4), 834-860.

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

Lam, D., Ardington, C. & Leibbrandt, M. (2010). Schooling as a lottery: Racial differences in school advancement in urban South Africa. *Journal of Development Economics* .

Lee, V., Zuze, T. & Ross, K. (2005). School Effectiveness in 14 Sub-Saharan African Countries: Links with 6th Graders' Reading Achievement. *Studies in Educational Evaluation 31*, 207-246.

Marshall, A. (1890). *The Principles of Economics.* Macmillan.

Mincer, J. (1958). Investment in Human Capital and Personal Income Distribution. *The Journal of Political Economy*, 281-302.

Stellenbosch University http://scholar.sun.ac.za

Mincer, J. (1970). The distribution of labor incomes: a survey with special reference to the human capital approach. *Journal of Economic Literature, 8*(1), 1-26.

Mincer, J. (1974). *Schooling Experience and Earnings.* New York: NBER.

Moll, P. (1998). Primary schooling, cognitive skills, and wage in South Africa. *Economica, 65*(258), 263-284.

Moloi, M. & Chetty, M. (2011). *SACMEQ Policy Brief: Learner and Teacher Knowledge about HIV and AIDS in South Africa.* Pretoria: Southern and Eastern Africa Consortium for Monitoring Educational Quality.

Passos, A. (2009). *A comparative analysis of teacher competence and its effect on pupil performance in upper primary schools in Mozambique and other SACMEQ countries.* Pretoria: PhD thesis: Policy Studies. University of Pretoria.

Psacharopoulos, G. (1994). Returns to Investment in Education: A Global Update. *World Development, 22*(9), 1325-1344.

Psacharopoulos, G. & Patrinos, H. (2004). Returns to investment in education: A global update. *Education Economics, 12*(2), 111-134.

Reddy, V. (2006). *Mathematics and Science Achievement at South African Schools in TIMSS 2003.* Cape Town: HSRC Press.

Ross, K., Saito, M., Dolata, S., Ikeda, M., Zuze, L. & Murimba, S. (2005). The Conduct of the SACMEQ III Project. In E. Onsomu, J. Nzomo, & C. Obiero, *The SACMEQ II Project in Kenya: A Study of the Conditions of Schooling and the Quality of Education.* Harare: SACMEQ.

SACMEQ. (2010). *SACMEQ III Project Results: Pupil Achievement Levels in Reading and Mathematics.* Retrieved January 2011, from Southern and Eastern Africa Consortium for Monitoring Educational Quality: http://www.sacmeq.org/downloads/sacmeqIII/WD01_SACMEQ_III_Results_Pupil_Achievement.pdf

Salmi, J. (2006). Violence, democracy and education: An analytic framework. In Schweitzer, E (ed), *Promoting social cohesion through education.* Washington, DC: World Bank Institute.

Schultz, T. (1961). Investment in Human Capital. *The American Economic Review, 51*(1), 1-17.

Stellenbosch University http://scholar.sun.ac.za

Schultz, T. (1989). Investing in People: Schooling in Low Income Countries. *Economics of Education Review*, 219-223.

Shabalala, J. (2005). *The SACMEQ II Project in Swaziland: A Study of the Conditions of Schooling and the Quality of Education.* Retrieved 2009, from Southern and Eastern Africa Consortium for Monitoring Educational Quality: http://www.sacmeq.org/education-swaziland.htm

Shepherd, D. (2011). Constraints to School Effectiveness: What prevents poor schools from delivering results? *Stellenbosch Economic Working Papers 05/11*.

Smith, A. (1776). *The Wealth of Nations.* Digireads.com.

Spaull, N. (2011). Primary School Performance in Botswana, Mozambique, Namibia and South Africa: A Comparative Analysis of SACMEQ III. *SACMEQ Working Papers*, 1-74.

Spence, M. (1973). Job Market Singalling. *Quarterly Journal of Economics, 87*(3), 355-374.

StataCorp. (2009). *Stata: Release 11.* Texas: StataCorp Statistical Software.

Stols, G., Olivier, A. & Grayson, D. (2007). Description and impact of a distance mathematics course from Grade 10 to 12 teachers. *Pythagoras, 65*, 32-38.

Taylor, N. (2007). Equity, Efficiency and the Development of South African Schools. In T. Townsend, *International Handbook of School Effectiveness and Improvement.*

Taylor, N. (2008). *What's Wrong with South African Schools? .* Retrieved from JET Education Services: http://www.jet.org.za/events/conferences/What%20works%20in%20school%20development/Papers/Taylor%20Whats%20wrong%20with%20SA%20schools%20JET%20Schools%20Conf%20final.pdf/at_download/file

Taylor, N. & Moyana, J. (2005). *Khanyisa Education Support Programme: Baseline study part 1: communities, schools and classrooms.* Johannesburg: JET Education Services.

Taylor, N., Fleisch, B. & Shindler, J. (2008). *Changes in Education since 1994.* Pretoria: The Presidency: Input into the 15 year review process.

Taylor, N., Muller, J. & Vinjevold, P. (2003). *Getting Schools Working.* Cape Town: Pearson Education.

Taylor, S. (2011). Uncovering Indicators of Effective School Management in South Africa using the National School Effectiveness Study. *Stellenbosch Economic Working Papers*.

Stellenbosch University http://scholar.sun.ac.za

Taylor, S. & Yu, D. (2009). The Importance of Socioeconomic Status in Determining Educational Achievement in South Africa. *Stellenbosch Economic Working Papers*.

Tobin, J. (2005). Fisher's "The Nature of Capital and Income". *American Journal of Economics and Sociology*, 207-214.

Trong, K. (2010). *The Relative Risk-Percentage Equity Index: Measuring Equity in Reading Achievement Across PIRLS 2006 Countries.* Hamburg: International Association for the Evaluation of Education.

UNESCO. (2011). *Education For All Global Monitoring Report 2011.* Paris: UNESCO Publishing.

Van der Berg, S. (2007). Apartheid's Enduring Legacy: Inequalities in Education. *Journal of African Economies, 16*(5), 849-880.

Van der Berg, S. (2008). *How effective are poor schools? Poverty and educational outcomes in South Africa.* Centre for European, Governance and Economic Development Research (CEGE) Discussion Papers 69.

Van der Berg, S. (2010). Current Poverty and Income Distribution in the Context of South African History. *Stellenbosch Economic Working Paper*, 1-23.

Van der Berg, S. & Burger, R. (2003). Education and Socio-Economic Differentials: A Study of School Performance in the Western Cape. *Development Policy Research unit Working Paper 03/73*.

Van der Berg, S. & Louw, M. (2006). Lessons Learnt from SACMEQII: South African Student Performance in Regional Context. *Investment Choices for Education in Africa.* Johannesburg.

Van der Sandt, S. & Niewoult, H. (2003). Grade 7 Teachers' and Prospective Teachers' Content Knowledge of Geometry. *South African Journal of Education, 23(3)*, 199-205.

Webb, P., Boltt, G., Austin, P., Cloete, M., English, V. & Feza, N. (1998). *The nature and impact of accredited in-service education on under-qualified science and mathematics teachers: What factors generate best practice in the classroom.* Johannesburg: Joint Education Trust.

Willms, J. (1997). Literacy Skills and Social Class. *Options Politiques (Jul/Aug)*, 22-26.

Appendices

Appendix A

Table 15: Cross-tabulation - Province and School Location

Province	Isolated	Rural	Small town	Large city	Total
Eastern Cape	0	71.1	12.5	16.4	**100**
	0	23.9	11.2	8.4	**16.3**
Freestate	0	16.7	45.2	38.1	**100**
	0	1.7	12.4	5.9	**5**
Gauteng	0	7.7	18.6	73.7	**100**
	0	2.7	17.7	39.8	**17.2**
KwaZulu-Natal	6.5	61.2	9	23.2	**100**
	100	29.9	11.8	17.2	**23.7**
Limpopo	0	89.1	10.9	0	**100**
	0	24.7	8.1	0	**13.4**
Mpumalanga	0	48.3	26.7	25	**100**
	0	8.5	12.6	6.7	**8.5**
Northern Cape	0	33.5	35.2	31.3	**100**
	0	1.4	4	2	**2**
North West	0	48.6	28.1	23.4	**100**
	0	6	9.3	4.4	**6**
Western Cape	0	7.1	29.8	63.1	**100**
	0	1.1	12.9	15.5	**7.8**
Total	**1.5**	**48.5**	**18.1**	**31.9**	**100**
	100	**100**	**100**	**100**	**100**

Table 16: Cross-tabulation - Province and School Socioeconomic Quintile

Province	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5	Total
Eastern Cape	61.5	11.2	13	12.2	2	**100**
	39	10	11.9	10.3	1.7	**16.3**
Freestate	0	32.7	34.4	16.4	16.5	**100**
	0	8.9	9.6	4.2	4.3	**5**
Gauteng	4.6	13.1	9.5	29.6	43.1	**100**
	3.1	12.4	9.2	26.4	39	**17.2**
KwaZulu-Natal	44.8	12.9	17.6	11.8	12.9	**100**
	41.4	16.8	23.4	14.5	16.1	**23.7**
Limpopo	17.6	36.8	21.9	18.4	5.3	**100**
	9.2	27.1	16.5	12.8	3.8	**13.4**
Mpumalanga	10.2	27.9	32.4	17.3	12.2	**100**
	3.4	13.1	15.6	7.7	5.5	**8.5**
Northern Cape	8.1	25.9	27.1	21.4	17.5	**100**
	0.6	2.9	3.1	2.3	1.9	**2**
North West	13.8	20.5	12.6	26	27	**100**
	3.2	6.8	4.3	8.1	8.6	**6**
Western Cape	0	5	14.5	33.8	46.6	**100**
	0	2.1	6.4	13.7	19.2	**7.8**
Total	**25.6**	**18.3**	**17.8**	**19.3**	**19**	**100**
	100	**100**	**100**	**100**	**100**	**100**

Table 17: Cross-tabulation School Location and School Socioeconomic Status

		School Socioeconomic Status					
		Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5	Total
School Location	Isolated	4%	2%	0%	0%	0%	2%
	Rural	88%	57%	51%	31%	4%	48%
	Small Town	4%	19%	28%	27%	18%	18%
	Large City	4%	22%	21%	43%	78%	32%
	Total	100%	100%	100%	100%	100%	100%

Appendix B

Reading Competency Levels

Table 18: Reading Competency Levels by Province

Reading Competency Levels	Province									
	ECA	FST	GTN	KZN	LMP	MPU	NCA	NWP	WCA	Total
L1 - Pre Reading	23.1	3.4	7	21.3	30.4	8.1	1.6	4.2	0.9	100
	14.1	6.9	4	9	22.5	9.4	7.7	6.9	1.2	9.9
L2 - Emergent Reading	23	4.4	7.5	26.5	20.5	9.4	1.6	5.2	1.8	100
	24.5	15.4	7.6	19.4	26.5	19	13.7	15	3.9	17.3
L3 - Basic Reading	19.9	5.8	8.6	29.3	16.2	9.8	1.9	5.5	3	100
	25.8	24.7	10.5	26.1	25.4	24.1	19.4	19.4	8.1	21.1
L4 - Reading for Meaning	18.5	6.2	11.9	23.7	12	11	2.2	7.1	7.3	100
	16.7	18.5	10.2	14.7	13.1	19	16	17.4	13.7	14.7
L5 - Interpretive Reading	14.8	5.2	20.5	19.7	8.6	9.7	2.8	7	11.8	100
	9.7	11.2	12.6	8.9	6.8	12	14.6	12.3	16	10.6
L6 - Inferential Reading	6.7	5.3	29.2	19.5	4.6	7.9	2.6	7.2	17	100
	3.9	10.2	16.2	7.9	3.3	8.8	12.1	11.4	20.7	9.6
L7 - Analytical Reading	5.8	4.4	40.3	17.3	2.3	4.2	2.2	6.4	17.1	100
	3.7	9	23.8	7.5	1.7	5	11.1	10.8	22.3	10.2
L8 - Critical Reading	4.2	3	39.4	24	1.3	3.5	1.7	6.1	16.7	100
	1.7	4	15.1	6.7	0.7	2.7	5.4	6.7	14.1	6.6
Total	16.3	5	17.2	23.7	13.4	8.5	2	6	7.8	100
	100	100	100	100	100	100	100	100	100	100

Table 19: Reading Competency Levels by School Location

Reading Competency Levels	School Location				
	Isolated	Rural	Small town	Large city	Total
L1 - Pre Reading	2.3	74.9	10.7	12.1	100
	14.6	15.4	5.9	3.8	9.9
L2 - Emergent Reading	2.2	72.5	11.3	14	100
	24.2	25.9	10.8	7.6	17.3
L3 - Basic Reading	1.8	63.5	15.3	19.3	100
	24.6	27.6	17.8	12.8	21.1
L4 - Reading for Meaning	2.3	52.6	21.2	23.9	100
	22.2	16	17.2	11	14.7
L5 - Interpretive Reading	1.3	34.9	28.1	35.8	100
	8.8	7.7	16.5	11.9	10.6
L6 - Inferential Reading	0.7	17.8	29.6	51.8	100
	4.6	3.5	15.6	15.5	9.6
L7 - Analytical Reading	0	12.1	18.8	69.1	100
	0	2.5	10.6	22.1	10.2
L8 - Critical Reading	0.3	10.9	15.2	73.7	100
	1.1	1.5	5.5	15.2	6.6
Total	1.5	48.5	18.1	31.9	100
	100	100	100	100	100

Table 20: Reading Competency Levels by School SES

Reading Competency Levels	5 Quintiles of School SES					
	1	**2**	**3**	**4**	**5**	Total
L1 - Pre Reading	40.1	24.6	19.3	15.3	0.7	100
	15.6	13.4	10.8	7.9	0.4	9.9
L2 - Emergent Reading	43.1	22.1	20.1	13.6	1.1	100
	29.1	21	19.6	12.2	1	17.3

L3 - Basic Reading	37.2	22.5	21.9	16.2	2.2	100
	30.6	25.9	25.9	17.6	2.5	21.1
L4 - Reading for Meaning	26.4	20.7	24.2	23.4	5.3	100
	15.1	16.7	20	17.8	4.1	14.7
L5 - Interpretive Reading	16.5	18.8	20.2	28.1	16.4	100
	6.9	10.9	12.1	15.5	9.2	10.6
L6 - Inferential Reading	5.4	12.1	13.8	28.3	40.3	100
	2	6.3	7.4	14	20.3	9.6
L7 - Analytical Reading	1.3	7.4	6.8	19.4	65.1	100
	0.5	4.1	3.9	10.2	34.9	10.2
L8 - Critical Reading	0.8	4.5	0.9	13.7	80.1	100
	0.2	1.6	0.3	4.7	27.8	6.6
Total	25.6	18.3	17.8	19.3	19	100
	100	100	100	100	100	100

Maths competency levels

Table 21: Maths Competency Levels by Province

Maths Competency Levels	**Provinces**									
	ECA	FST	GTN	KZN	LMP	MPU	NCA	NWP	WCA	Total
L1 - Pre Numeracy	23.3	3.4	9.9	24.5	23.5	8.4	1.7	4	1.3	100
	7.9	3.8	3.1	5.7	9.6	5.4	4.6	3.6	0.9	5.5
L2 - Emergent Numeracy	19.9	4.9	8.7	26.1	19.8	9.5	1.9	6	3.2	100
	42.4	34.3	17.4	38.3	51	38.4	32.5	34.5	14.1	34.7
L3 - Basic Numeracy	17.1	5.9	14.7	24.2	13.1	10.3	2.2	6.3	6.3	100

	30.3	34.7	24.6	29.6	28.2	34.9	31.7	30.2	23.4	29
L4 - Beginning Numeracy	12.4	4.6	24	22.3	7.5	7.7	2.2	6	13.3	100
	11.7	14.4	21.4	14.5	8.6	13.9	16.5	15.3	26.2	15.4
L5 - Competent Numeracy	8.1	5.1	39.1	17	3.3	5.1	1.8	5.1	15.5	100
	3.5	7.2	16.1	5.1	1.7	4.2	6.2	6	14.1	7.1
L6 - Mathematically Skilled	5	3.6	39.5	20.2	2	3.3	2	6.8	17.5	100
	1.8	4.3	13.5	5.1	0.9	2.3	5.7	6.7	13.3	5.9
L7 - Concrete Problem Solving	20.1	3	29	16.8	0	2.4	2.2	7.4	19.1	100
	2.3	1.1	3.1	1.3	0	0.5	2	2.3	4.6	1.9
L8 - Abstract Problem Solving	0	0.7	19.4	16.5	0	4.5	2.5	13.7	42.8	100
	0	0.1	0.7	0.4	0	0.3	0.7	1.3	3.2	0.6
Total	16.3	5	17.3	23.6	13.4	8.6	2	6	7.8	100
	100	100	100	100	100	100	100	100	100	100

Table 22: Maths Competency Levels by School Location

Maths Competency Levels	Isolated		Rural	Small town	Large city	Total
L1 - Pre Numeracy		2.3	71	12.2	14.5	100
		8.2	8.1	3.7	2.5	5.5
L2 - Emergent Numeracy		2.1	66	15.2	16.7	100
		48	47.1	29.2	18.2	34.7
L3 - Basic Numeracy		1.8	51.2	19.3	27.6	100
		34.2	30.6	31.1	25.1	29
L4 - Beginning Numeracy		0.9	32.2	24.7	42.3	100
		8.5	10.2	21.1	20.4	15.4
L5 - Competent Numeracy		0.2	15.8	20.3	63.6	100
		1.1	2.3	8	14.2	7.1
L6 - Mathematically Skilled		0	7.3	14.6	78.2	100
		0	0.9	4.8	14.5	5.9

L7 - Concrete Problem Solving		0	23.1	15.9	61	100
		0	0.9	1.6	3.6	1.9
L8 - Abstract Problem Solving		0	0	12.7	87.3	100
		0	0	0.4	1.6	0.6
Total		1.5	48.6	18	31.9	100
		100	100	100	100	100

Table 23: Maths Competency Levels by School SES

Maths Competency Levels	**5 Quintiles of School SES**					
	1	2	3	4	5	Total
L1 - Pre Numeracy	44.8	22	17.1	15.7	0.4	100
	9.6	6.6	5.3	4.5	0.1	5.5
L2 - Emergent Numeracy	36.4	22.3	21.6	17.2	2.5	100
	49.1	42.3	42.1	30.9	4.5	34.7
L3 - Basic Numeracy	25.8	20.8	22.6	23.1	7.8	100
	29.1	33	36.7	34.6	11.9	29
L4 - Beginning Numeracy	16.6	15.3	15	24.3	28.8	100
	9.9	12.9	13	19.4	23.4	15.4
L5 - Competent Numeracy	6.1	6	6	17.6	64.3	100
	1.7	2.3	2.4	6.5	24.1	7.1
L6 - Mathematically Skilled	1.9	2.6	1.4	12.2	82	100
	0.4	0.8	0.5	3.7	25.5	5.9
L7 - Concrete Problem Solving	0.9	19.7	0	2.8	76.6	100
	0.1	2	0	0.3	7.5	1.9
L8 - Abstract Problem Solving	0	0	0	3.5	96.5	100
	0	0	0	0.1	3	0.6
Total	25.7	18.3	17.8	19.3	19	100
	100	100	100	100	100	100

Learner HIV/AIDS Knowledge levels

Table 24: HIV/AIDS Knowledge Level by Province

HAKT level	Province									
	ECA	FST	GTN	KZN	LMP	MPU	NCA	NWP	WCA	Total
At risk	19.2	5.6	12.4	22.6	17.5	9.2	2.3	5.7	5.7	100
	76.8	73.1	46.9	62.1	85	70.6	71.9	61.6	47.2	65.2
Minimum level	12.1	4.2	22.6	26.3	6.5	8.2	1.7	6.5	12	100
	19.9	22.7	35.1	29.7	13	25.7	21.8	28.7	41	26.8
Desirable level	6.6	2.6	38.8	24.2	3.3	4	1.6	7.3	11.6	100
	3.2	4.2	18	8.1	2	3.8	6.3	9.7	11.8	8
Total	16.3	5	17.2	23.7	13.4	8.5	2	6	7.8	100
	100	100	100	100	100	100	100	100	100	100

Table 25: HIV/AIDS Knowledge Level by School Location

HAKT level	School Location				
	Isolated	Rural	Small town	Large city	Total
At risk	2.1	58.1	17	22.7	100
	90.3	78.2	61.3	46.5	65.2
Minimum level	0.5	33.8	20.6	45.2	100
	8.6	18.7	30.4	38	26.8
Desirable level	0.2	19.2	18.7	61.9	100
	1.1	3.2	8.3	15.5	8
Total	1.5	48.5	18.1	31.9	100
	100	100	100	100	100

Table 26: HIV/AIDS Knowledge Level by Quintiles of School Socioeconomic Status

HAKT level	Quintiles of School SES					
	1	2	3	4	5	Total
At risk	31.8	22.5	20.2	18.1	7.4	100
	80.9	80.4	74.1	61.1	25.5	65.2
Minimum level	16.4	11.6	14.9	23.5	33.6	100
	17.1	17.1	22.4	32.6	47.4	26.8
Desirable level	6.3	5.9	7.9	15.4	64.5	100
	2	2.6	3.5	6.4	27.1	8
Total	25.6	18.3	17.8	19.3	19	100
	100	100	100	100	100	100

Appendix C

Teacher Absenteeism

Table 27: Teacher absenteeism by province

	Maths teacher			Reading teacher		
Province	**Days absent**	**SD**	**Percent**	**Days absent**	**SD**	**Percent**
Eastern Cape	21.6	14.3	11.0%	23.9	11.6	10.9%
Freestate	17.2	14.7	11.0%	16.6	16.2	11.1%
Gauteng	13.5	23.6	11.3%	11.6	9.4	11.1%
KwaZulu-Natal	25.3	20.1	16.8%	28.4	24.0	16.9%
Limpopo	20.3	8.9	10.7%	21.5	7.1	11.0%
Mpumalanga	20.9	15.4	10.0%	29.1	52.5	9.8%
Northern Cape	18.5	14.5	9.9%	20.5	12.8	9.9%
North West	22.3	20.0	10.2%	16.7	12.8	10.4%
Western Cape	12.5	11.1	9.0%	11.8	12.4	9.0%
Total	**20.0**	**18.0**	**100.0%**	**21.3**	**22.5**	**100.0%**

Table 28: Teacher absenteeism by school location

	Maths teacher			Reading teacher		
School location	**Days absent**	**SD**	**Percent**	**Days absent**	**SD**	**Percent**
Isolated	23.2	16.4	1%	33.8	5.8	1%
Rural	23.5	18.0	44%	24.7	16.9	45%
Small town	19.3	16.6	23%	23.6	38.6	24%
Large City	14.8	17.6	31%	13.9	15.1	31%
Total	**20.0**	**18.0**	**100%**	**21.3**	**22.5**	**100**

Table 29: Teacher absenteeism by SES quintile

	Maths teacher			Reading teacher		
Quintile	**Days absent**	**SD**	**Percent**	**Days absent**	**SD**	**Percent**
1	23.6	12.5	20%	24.2	12.2	20%
2	23.7	23.5	21%	29.2	42.4	20%
3	21.8	15.5	20%	25.3	17.6	20%
4	21.8	21.4	20%	19.8	9.4	20%
5	7.8	9.2	20%	7.3	6.7	20%
Total	**20.0**	**18.0**	**100%**	**21.3**	**22.5**	**100%**

Table 30: Frequency of Missed Meals by School Socioeconomic Quintile

Frequency of Missed Meals by School Socioeconomic Quintile

	Quintile 1			**Quintile 2**			**Quintile 3**			**Quintile 4**			**Quintile 5**			**Total**		
	Breakf.	Lunch	Supper	Breakf.	Lunch	Supper	Breakf.	Lunch	Supper	Breakf.	Lunch	Supper	Breakf.	Lunch	Supper	**Breakf.**	**Lunch**	**Supper**
Not at all	9%	5%	3%	8%	2%	3%	9%	3%	5%	10%	4%	3%	5%	2%	1%	**8%**	**3%**	**3%**
1-2 days/week	10%	11%	7%	11%	8%	3%	12%	7%	5%	12%	10%	5%	11%	4%	2%	**11%**	**8%**	**5%**
3-4 days/week	7%	9%	5%	10%	9%	4%	9%	11%	7%	11%	10%	6%	9%	6%	3%	**9%**	**9%**	**5%**
Every day	73%	75%	85%	71%	81%	89%	70%	79%	83%	68%	77%	86%	75%	88%	94%	**72%**	**80%**	**87%**
Total	100 %	100 %	100 %	100 %	100 %	100 %	100 %	100 %	100 %	100 %	100 %	100 %	100 %	100 %	100 %	100 %	100 %	100 %

Appendix D

Regression Output - *Standard errors left aligned. Significance levels (*10%) (**5%)(***1%)*

	Learner Reading Score (Mean: 498 / SD: 115)					
Sample:	*Only those with Teacher test scores*	*Full Sample*	*Top Quartile (25%) of School SES*	*Bottom 3 Quartiles (75%) of School SES*	*Top Quintile (20%) of School SES*	*Bottom 4 Quintiles (80%) of School SES*
Young (<11y3m)	-11.3272**	-10.2311*	-1.6709	-10.3439*	-18.0156	-8.4601
	5.67	5.24	12.78	6.17	11.37	6.03
Old (>11y3m-12y8m)	-17.8731***	-17.2983***	-22.7698***	-16.8636***	-27.6640***	-16.6586***
	2.35	2.3	6.02	2.46	6.66	2.48
Very old (14y+)	-22.5645***	-21.8776***	-58.0512***	-21.8897***	-94.4145***	-20.9751***
	3.39	3.24	13.08	3.39	14.12	3.41
Male	-12.7621***	-11.8880***	-15.5935***	-10.3751***	-9.8167***	-12.7066***
	1.86	1.73	3.71	2.03	3.51	2.07
> 5 Days absent	15.6247	13.8835	2.8593	21.6067	32.8600***	12.4144
	11.54	10.77	10.35	14.94	11.32	13.01
Preschool - months	5.1058	6.7047	14.5074	5.6689	24.3637**	4.9687
	4.5	4.12	12	4.34	10.71	4.85
Preschool - 1 year	10.2733***	10.1128***	19.4733**	6.9585*	21.9514***	9.6837**
	3.73	3.57	7.72	3.85	8.13	3.97
Preschool - 2 years	16.6904***	15.6015***	24.7914***	11.1783***	28.2491***	14.7637***
	3.43	3.35	7.86	3.61	7.01	3.84
Preschool - 3 years or more	11.4327***	12.8022***	19.6639***	7.0959**	26.2906***	7.8281**

	3.18	3.17	6.65	3.44	6.78	3.67
SES	3.5845***	4.2431***	11.6224**	0.2211	2.0416	3.7068**
	1.37	1.3	4.75	1.5	5.18	1.74
SES squared	1.405	1.2847	-1.8926	-1.7096	0.4665	1.0455
	0.97	0.95	2.19	1.21	2.65	1.4
Lived with parents	-5.9337**	-5.1507**	-4.6688	-4.9476*	-7.2928	-4.7082*
	2.53	2.39	6.3	2.71	6.68	2.79
3 or more siblings	-9.6481***	-10.3899***	-0.8986	-8.3145***	-3.2341	-9.8606***
	2.16	2.08	4.63	2.42	4.25	2.35
Missed breakfast	6.6681***	8.1990***	-3.4094	9.3696***	-0.6356	7.5011***
	2.24	2.25	4.02	2.37	4.95	2.37
Missed lunch	-7.9566**	-8.7397**	0.7668	-8.1192**	-5.3965	-6.9262*
	3.6	3.57	6.34	3.96	7.58	3.91
Missed dinner	-14.1919***	-16.3802***	-28.0364***	-12.5431***	-17.4336*	-13.3904***
	3.54	3.46	8.89	3.54	9.82	3.68
More than 10 books at home	10.0660***	10.3273***	14.2999***	6.0077**	19.3394***	7.3998***
	2.3	2.28	4.19	2.76	4.22	2.69
Used PC before	19.5469***	17.1185***	32.8606***	20.9466***	34.4547**	19.8040***
	3.98	3.7	11.21	3.8	13.34	4.06
Urban	-11.3143**	-10.6834**	-7.8148	-16.2707**	5.5478	-12.1995**
	4.64	4.3	5.83	6.52	5.63	5.89
Mother or father has matric	10.8959***	10.2208***	21.8341***	8.5646***	14.8329**	10.7080***
	2.3	2.27	5.77	2.35	6.33	2.49
Mother or father has degree	7.3581**	9.3076***	4.1668	10.9522**	8.1255	6.2496
	3.65	3.45	5.31	4.5	5.5	4.62
Speak Eng. at home sometimes	20.8862***	21.7547***	19.3853***	21.8548***	13.9596**	22.0091***

	3.29	3.38	6.29	3.24	6.6	3.3
Speak Eng. at home always	38.8100***	40.6304***	47.7345***	12.4325**	41.7362***	31.0147***
	5.45	5.13	7.44	5.73	7.48	7.15
Orphan (double-orphan)	-6.2484	-3.7301	-23.1080*	-2.4916	-31.1428**	-2.6283
	3.91	3.95	13.95	3.97	14.92	4.2
Orphanage or children's home	-34.5487***	-36.7403***	3.8411	-30.2093***	-23.16	-32.1277***
	8.61	8.41	19.84	9.79	28.78	9.26
School SES	33.6697***	31.2937***	84.3425*	20.4248	115.1838*	35.8774***
	7.17	7.43	43.82	13.41	62.01	12.15
School SES squared	19.6121***	21.3193***	-24.0735	8.8348	-35.3879	23.5633**
	3.65	3.86	18.9	11.33	26.08	9.52
Homework - 1 or 2 times a month	6.8581	5.8443	4.2079	8.9681	40.3292*	4.7321
	6.65	6.5	22.28	6.35	21.49	6.62
Homework - 1 or 2 times a week	21.5040***	20.6854***	9.235	23.7241***	23.2084	21.7732***
	6.2	5.9	20.22	5.95	22.48	6.13
Homework - Most days	22.5036***	22.9842***	15.6771	21.6927***	28.5351	19.8469***
	6.22	6.18	19.97	6.08	21.24	6.18
Repeated a grade once	-19.1254***	-19.6211***	-29.2185***	-16.1709***	-22.7590***	-17.8257***
	2.66	2.53	6.52	2.75	7.22	2.8
Repeated a grade twice	-22.6856***	-23.4965***	-7.1717	-21.6766***	11.1294	-22.4734***
	4.08	3.99	14.38	4.13	16.45	4.13
Repeated a grade three or more	-40.4311***	-38.6558***	-52.0713*	-35.1059***	-74.5612**	-37.1603***
	6.29	5.79	27.3	5.71	35.82	6.04
No class library	-0.8494	-0.2744	-5.9923	0.2103	-4.0013	0.7665

	4.4	4.15	5.03	4.84	5.16	5.11
Class-size > 40	-2.5503	-2.1672	-4.1329	-1.5321	-11.9823	-1.656
	4.2	4.19	8.23	4.95	9.23	4.89
Sanitation (Std)	-4.9452	-6.5523*	-11.0377	-9.0872***	10.747	-6.8856**
	3.3	3.34	8.24	2.82	14.22	3.28
Building Index (Std)	7.3835	8.6468**	14.1013**	-0.2567	18.4334***	3.0349
	4.74	4.33	6.89	5.79	6.73	5.41
Equipment Index (Std)	-0.2664	-0.1214	8.8308	-0.266	14.6733	0.6948
	3.2	3.25	7.14	3.25	9.89	3.27
Avg. Parent matric (school)	1.3717	3.8301	2.7296	2.6938	4.7177	1.2556
	3.75	3.73	6.37	4.34	7.75	4.48
Avg. Parent degree (school)	7.6445**	8.5778***	7.5236***	-3.1115	6.2660***	-0.2967
	3.06	2.98	2.7	5.82	2.35	5.59
R/M Textbook -Teacher only	-2.3717	-0.1919	9.1455	-2.7401	34.3202**	-7.3956
	6.34	6.05	15.95	6.67	14.29	6.46
R/M Textbook -Share 2+	-3.3519	-1.7652	-32.7349**	6.3791	-23.6368	0.036
	6.08	6.34	16.6	6.58	19.29	6.62
R/M Textbook -Share with 1	16.6963***	17.1299***	23.7109*	21.6567***	23.4188**	19.5311***
	5.56	5.34	12.69	5.95	11.01	6.26
R/M Textbook - own textbook	14.9086***	11.6911**	10.3138	17.6076***	17.5849*	14.8348***
	4.88	4.78	12.19	5.39	10.46	5.56
Extra tuition (English/Maths)	-16.3418***	-15.8735***	-40.8851***	-3.7701	-43.5471***	-8.0151
	5.83	5.8	8.13	7.06	8.35	6.69

Reading-teacher Reading score	5.6023**		4.5489	2.7252	6.2489**	5.2829**
	2.39		2.91	2.86	2.46	2.6
Teacher Male	-1.3571	-2.0487	4.4922	-1.3253	15.6666***	-3.7268
	3.97	3.81	7.23	4.1	5.89	4.24
Teacher Jnr. Secondary	8.6946	5.2422	-0.6419	-4.9567	15.235	-0.7445
	11.83	9.87	21.18	16.91	17.12	15.05
Teacher Snr. Secondary	4.9229	4.1723	-5.155	-1.8563	-11.8263	1.9604
	7.16	6.57	7.32	7.7	8.98	7.8
Teacher A-level / further study	-9.1556	-8.5941	-27.5660**	-3.8813	-31.4778***	-5.2876
	7.25	6.94	10.84	7.72	11.52	8.16
Teacher Degree	2.3619	2.0679	4.0222	-4.9283	-2.7677	0.0154
	6.77	6.41	7.39	8.6	8.77	7.8
Teacher training: 2yrs	-1.5284	-0.6684	-19.4729	-9.2771	20.0413	4.583
	12.02	11.87	12.79	12.65	14.02	15.17
Teacher training: 3yrs	2.9196	2.0864	6.2172	-5.5958	9.4492	6.5168
	11.1	10.99	8.87	11.76	9.27	14.5
Teacher training: >3yrs	0.0628	-0.6179	8.2531	-6.0701	5.5598	4.8722
	10.8	10.64	8.1	12.1	8.37	14.94
Eastern Cape	17.2822	21.5309*	29.6618***	15.3451	25.5808**	16.4395
	12.44	11.88	11.09	13.64	12.35	13.99
Free State	7.52	7.5929	22.7083	2.2508	8.1769	4.1664
	8.26	7.93	15.31	9.18	14.53	9.71
Gauteng	27.9387***	30.2628***	22.1400***	15.0905*	16.4937**	22.2408**
	6.97	6.8	6.52	8.69	7.72	9.44
KwaZulu Natal	17.4944**	20.8421***	13.6172	8.8163	20.0741**	15.1043*
	6.85	6.74	8.43	7.84	8.96	8.29
Limpopo	-30.6788***	-27.8728***	-48.1543***	-25.9915***	-3.4631	-27.2719***
	9.03	8.77	17.68	8.57	16.02	9.14

Mpumalanga	4.2908	8.1221	1.5033	4.0729	-2.3347	6.9131
	7.69	7.39	9.62	9.4	9.94	9.41
Northern Cape	22.3081***	24.1182***	16.4974	19.4091**	18.865	20.4598**
	8.51	8.23	15.69	9.56	13.21	9.8
Western Cape	29.8220***	37.9730***	11.1291	45.4434***	4.2235	43.0378***
	8.2	8.26	8.04	13.15	9.49	11.94
Constant	434.7860***	431.6523***	397.9002***	433.6762***	362.1882***	425.0056***
	15.27	15.61	31.62	16.79	51.21	18.07
N	**7974**	**8822**	**8807**	**8238**	**8882**	**8163**
F-stat	**81.62857**	**83.2893**	**407.90375**	**17.23041**	**307.13396**	**24.30282**
Prob > F	**0**	**0**	**0**	**0**	**0**	**0**
R-squared	**0.59966**	**0.60186**	**0.51788**	**0.26963**	**0.4169**	**0.34484**
Pre1=2	0.09498	0.12963	0.40754	0.34237	0.26683	0.2647
Pre1=3	0.74102	0.42709	0.9689	0.97466	0.4017	0.67127
Pre2=3	0.09751	0.36761	0.33948	0.27498	0.65691	0.09318
Repeat1=2	0.37752	0.31761	0.14052	0.17178	0.04038	0.25809
Repeat1=3	0.00064	0.00088	0.40404	0.00091	0.14948	0.00118
Repeat2=3	0.01794	0.02807	0.09187	0.05405	0.01594	0.04379
Homework1=2	0.00082	0.0003	0.73758	0.00033	0.34739	0.00006
Homework1=3	0.00036	0.00004	0.42942	0.002	0.50543	0.0006
Homework2=3	0.76837	0.50096	0.31535	0.56762	0.3819	0.6017
ses ses2 jointsig	0.01709	0.00223	0.01856	0.30284	0.60381	0.10007
sses sses2 jointsig	0	0	0.00289	0.2327	0.00116	0.01074

Notes

a. Missed breakfast/lunch/dinner is a dummy variable taking on the value one if the learner reported that they normally missed breakfast/lunch/dinner at least once a week.

b. Sanitation (Std) is a measure of sanitation quality. It is the total non-flushing toilets in the school as a proportion of total toilets. The variable is standardised to have a mean of zero and a standard deviation of one

c. Avg. Parent *matric* (school) - is the proportion of learners in that school whose parents (mother or father) have a matric

d. Avg. Parent *degree* (school) - is the proportion of learners in that school whose parents (mother or father) have a degree

	Learner Mathematics Score (Mean: 497 / SD: 99)					
Sample:	*Only those with Teacher test scores*	*Full Sample*	*Top Quartile (25%) of School SES*	*Bottom 3 Quartiles (75%) of School SES*	*Top Quintile (20%) of School SES*	*Bottom 4 Quintiles (80%) of School SES*
Young (<11y3m)	-9.6983**	-8.5172**	-13.5894	-6.5205	-20.6138	-6.5137
	4.53	4.17	12.95	4.7	13.73	4.88
Old (>11y3m-12y8m)	-15.7735***	-16.4977***	-23.0888***	-12.6901***	-28.8115***	-13.4257***
	2.17	2.3	4.97	2.22	6.63	2.17
Very old (14y+)	-15.4261***	-16.9347***	-36.5056***	-14.0166***	-42.6602**	-14.9072***
	3.09	2.98	10.56	3.26	19	3.18
Male	0.5837	1.0902	10.8126***	-2.2593	12.2403***	-1.8619
	1.81	1.72	3.38	1.91	3.89	1.84
> 5 Days absent	11.4619	12.7813*	4.0727	15.8071*	15.4374*	12.8153
	7.22	6.56	8.26	9.32	9.24	8.85
Preschool - months	4.0568	4.6621	11.58	4.6124	18.5978	3.5034
	4.4	3.98	13.47	4.6	13.14	4.91
Preschool - 1 year	7.5115**	6.3516*	14.8372*	5.3217	22.6480**	6.6484*
	3.76	3.79	7.58	4.05	10.81	3.97
Preschool - 2 years	5.1589	4.6988	12.9197	2.3039	21.1958**	4.1688
	3.24	3.37	7.84	3.59	10.54	3.59
Preschool - 3 years or more	9.8495***	8.9323***	20.9900***	3.8474	33.5469***	3.4976
	3.05	3.01	6.94	3.35	9.94	3.3
SES	1.7816	1.6805	4.8174	-0.8444	-2.0591	1.5827
	1.4	1.25	4.83	1.71	5.98	1.62
SES squared	2.4947**	2.3522**	0.9209	-0.2823	1.9789	1.9446
	1.05	0.94	3.27	1.36	3.98	1.35

Lived with parents	-2.2669	-2.7911	-0.8212	-2.7493	0.5535	-2.2094
	2.47	2.34	5.92	2.64	7	2.6
3 or more siblings	-6.6935***	-6.2078***	-6.8937*	-4.1766*	-7.6656**	-5.3525**
	2	1.86	4.16	2.4	3.85	2.28
Missed breakfast	6.3352**	3.9478*	-4.736	8.4890***	-9.4732**	8.8227***
	2.45	2.31	4.13	2.96	4.71	2.76
Missed lunch	-1.2328	-1.567	7.2794	-2.5054	10.6301	-2.2179
	3.73	3.36	6.13	4.18	7.18	4
Missed dinner	-15.7649***	-15.2104***	-34.0008***	-12.5922***	-26.0707***	-14.3204***
	3.47	3.29	8.12	3.44	7.97	3.6
More than 10 books at home	5.6707**	6.1272***	16.1069***	1.1881	23.1295***	1.7039
	2.24	2.21	3.86	2.65	4.55	2.48
Used PC before	14.4515***	14.2246***	25.1372***	15.5621***	22.0761**	17.0792***
	2.84	2.76	7.77	3.04	9.21	2.92
Urban	-9.0562*	-6.7933	-0.631	-14.3180**	0.8774	-8.5408
	4.91	4.51	6.53	6.26	8.66	5.76
Mother or father has matric	6.0541***	6.9556***	15.8183***	4.2664*	13.6605**	5.6988***
	2.02	1.88	4.87	2.18	5.57	2.2
Mother or father has degree	10.0776***	9.4479***	7.6963*	11.4263***	9.3100**	7.7655*
	2.99	2.73	4.12	4.13	4.36	4
Speak Eng. at home sometimes	17.5700***	17.5406***	7.1063	19.3173***	12.4255*	18.3088***
	3.47	3.28	5.29	3.86	6.74	3.72
Speak Eng. at home always	17.7458***	19.2298***	11.9300*	6.3359	11.5858	18.0742***
	4.88	4.55	6.72	5.7	8.27	6.04
Orphan (double-	-2.7517	-4.2617	-26.8970*	-0.7681	-36.1224**	-0.5914

orphan)						
	4.55	4.28	14	4.84	17.03	4.7
Orphanage or children's home	-35.9238***	-33.2360***	0.8114	-37.5693***	-20.1038	-34.6541***
	10.27	10.16	19.71	11.64	29.35	10.88
School SES	35.6150***	36.3667***	23.8625	7.3676	-9.3011	27.1024**
	7.98	7.16	53.76	10.88	77.34	10.74
School SES squared	24.3934***	25.2035***	18.3893	1.844	30.2338	16.2137*
	4.05	3.67	21.98	11.4	32.26	8.73
Homework - 1 or 2 times a month	12.1372*	4.5593	8.2087	14.0662**	46.9937**	9.9026
	6.41	7.12	20.19	6.37	19.8	6.46
Homework - 1 or 2 times a week	26.0539***	19.9130***	16.6108	29.2542***	36.1088**	26.1129***
	6.03	6.85	15.65	6.17	16.03	6.19
Homework - Most days	27.5690***	22.4579***	25.4335*	27.6081***	50.2023***	24.8046***
	6.23	6.78	14.64	6.61	14.16	6.5
Repeated a grade once	-12.4634***	-13.0981***	-21.9836***	-8.8728***	-21.3993***	-9.5331***
	2.69	2.48	6.47	2.86	8.23	2.77
Repeated a grade twice	-17.5875***	-14.8220***	-5.4513	-16.3590***	-8.5424	-17.1644***
	4.46	4.13	11.04	4.49	14.48	4.5
Repeated a grade three or more	-26.2723***	-23.9977***	-22.7456	-23.3383***	-22.5635	-25.2535***
	6.17	5.75	17.14	6.3	24.88	6.21
No class library	2.4781	1.1812	6.2821	0.6594	7.7567	3.623
	3.53	3.72	7.25	3.9	8.19	3.89
Class-size > 40	0.2441	0.6784	-1.9603	-1.0579	-3.9601	-1.0278
	3.58	3.53	6.34	4.14	8.27	4.12
Sanitation (Std)	-1.8894	-1.718	-5.048	-5.1324	-8.8322	-3.6049
	3.62	3.16	16.27	3.35	15.13	3.6

Building Index (Std)	-0.967	-0.5933	17.1467***	-4.2844	34.0515***	-3.9465
	4.67	4.27	6.53	5.43	7.65	5.03
Equipment Index (Std)	2.925	3.1074	-4.5036	2.1311	-2.6316	2.8608
	3.06	2.85	6.66	2.98	12.42	3.06
Avg. Parent matric (school)	3.275	2.4793	-7.4169	9.7803**	-12.6552	7.1679*
	3.53	3.35	7.17	4.11	10.52	3.97
Avg. Parent degree (school)	6.1323**	6.8375***	19.1289***	-2.0728	16.3250***	-1.6509
	2.4	2.33	3.38	3.42	3.34	3.65
R/M Textbook -Teacher only	-3.1155	-4.3552	10.77	-5.6006	17.415	-2.6859
	13.51	11.7	13.97	16.03	19.56	15.4
R/M Textbook -Share 2+	-15.3923	-15.7513	-13.8104	-15.0747	-1.3331	-15.5556
	13.98	12.05	13.75	16.29	18.09	15.92
R/M Textbook -Share with 1	0.6782	-1.2715	3.2392	-1.452	6.3299	0.0899
	13.07	11.12	9.69	15.62	11.03	15.2
R/M Textbook - own textbook	-0.5016	-2.0834	-0.6	-2.6499	1.695	-2.6106
	12.9	11.22	8.06	15.78	9.04	15.52
Extra tuition (English/Maths)	-14.1059**	-14.7752***	-24.3184***	-5.6408	-23.3553***	-8.6495
	5.58	5.3	7.77	6.99	7.56	6.56
Reading-teacher Reading score	5.4792**		0.5196	4.9273	-1.1897	6.3153**
	2.44		3.24	3.26	3.35	3.08
Teacher Male	-1.93	-1.6183	11.8503	-2.7415	5.3889	-1.4068
	3.88	3.49	7.75	4.05	9.53	4.21

Teacher Jnr. Secondary	-7.8708	-8.7454	28.7871	-18.2392		-12.9359
	15.74	15.36	23.71	17.38		16.75
Teacher Snr. Secondary	-3.3316	-2.0599	5.167	-7.7282	-4.1001	-7.451
	8.16	7.81	11.44	9	12.04	9.16
Teacher A-level / further study	-0.0529	2.1864	32.0143***	-3.6583	38.9034***	-2.915
	6.64	6.42	10.86	7.41	12.46	7.46
Teacher Degree	5.4415	5.7156	33.2170***	-2.5799	29.4812***	0.9114
	6.49	6.32	7.61	8.2	9.32	7.96
Teacher training: 2yrs	-2.8249	-6.1259	42.7861*	-0.2871		0.3668
	10.59	10.61	21.81	13.63		14.53
Teacher training: 3yrs	1.0444	-1.4491	-18.4020*	0.9891	-21.8763*	4.5884
	8.77	8.74	10.2	12.8	11.53	13.46
Teacher training: >3yrs	-5.3062	-6.1118	-6.4811	-4.1689	-8.4967	-3.1685
	8.73	8.85	9.69	12.56	10.19	13.28
Eastern Cape	23.1148*	24.4611**	7.2788	23.4526*	14.8569	23.7458*
	12.27	10.92	12.92	12.13	12.72	12.62
Free State	11.91	13.2263*	4.2198	8.1484	-1.7251	9.9774
	7.28	7.1	15.41	7.54	16.29	7.69
Gauteng	13.9548**	14.3170**	20.0508**	6.2175	16.5423	6.3429
	6.86	6.2	9.2	8.27	10.7	8.37
KwaZulu Natal	13.7447**	14.0305**	31.0944***	6.5238	31.3604***	12.8027*
	6.22	5.91	9.29	6.44	11.44	7.06
Limpopo	-20.1164**	-19.8564***	-11.8818	-19.8263**	3.3246	-19.6014**
	8	7.64	15.75	8.17	19.81	8.34
Mpumalanga	9.4063	4.6163	-15.2611	9.2641	-12.9806	10.9518
	7.7	7.33	9.83	7.97	12.17	7.77
Northern Cape	18.7242**	19.7025***	30.5694	15.1605*	29.7925	14.2679*
	7.93	7.05	19.63	8.53	23.37	8.55
Western Cape	29.5496***	33.2081***	42.2009***	32.9548***	37.9167**	34.4386***

	7.83	7.61	11.79	10.44	14.66	9.39
Constant	435.1325***	443.0338***	379.8548***	441.2975***	359.0068***	435.3143***
	20.79	19.39	39.66	26.53	53.87	26.58
N	**7740**	**8803**	**8644**	**8167**	**8729**	**8082**
F-stat	**39.39817**	**41.19787**	**.**	**9.68821**	**264.89121**	**13.31858**
Prob > F	**0**	**0**	**.**	**0**	**0**	**0**
R-squared	**0.49508**	**0.5041**	**0.43823**	**0.16837**	**0.40683**	**0.21952**
Pre1=2	0.57352	0.6674	0.74353	0.52553	0.8117	0.59523
Pre1=3	0.53154	0.45008	0.19126	0.76435	0.0468	0.5022
Pre2=3	0.16035	0.18371	0.12668	0.7024	0.0102	0.87002
Repeat1=2	0.26287	0.68197	0.12434	0.11128	0.32964	0.10333
Repeat1=3	0.03038	0.07157	0.96488	0.02782	0.96162	0.01433
Repeat2=3	0.27088	0.21358	0.30336	0.3925	0.55977	0.31503
Homework1=2	0.00059	0.00005	0.48949	0.00012	0.50623	0.00005
Homework1=3	0.00011	0	0.11813	0.00089	0.83571	0.00026
Homework2=3	0.63002	0.37452	0.16097	0.6336	0.04315	0.69716
ses ses2 jointsig	0.02186	0.01275	0.08111	0.88344	0.86293	0.28863
sses sses2 jointsig	0	0	0	0.74259	0	0.02976

Notes

a. Missed breakfast/lunch/dinner is a dummy variable taking on the value one if the learner reported that they normally missed breakfast/lunch/dinner at least once a week.
b. Sanitation (Std) is a measure of sanitation quality. It is the total non-flushing toilets in the school as a proportion of total toilets. The variable is standardised to have a mean of zero and a standard deviation of one
c. Avg. Parent *matric* (school) - is the proportion of learners in that school whose parents (mother or father) have a matric
d. Avg. Parent *degree* (school) - is the proportion of learners in that school whose parents (mother or father) have a degree

Appendix E

Summary statistics of variables included in regressions *Table 1*: **Diagnostic statistics – comparing means of variables across different samples**

	Full sample	Teacher-test only	Teach-test as % of full sample	Quart1-3	Quart4	Quint1-4	Quint5	Quart 4 as % of Quart 1-3	Quint 5 as % of Quint 1-4	QUINT1-4 as % of QUART1-3	QUNIT5 as % of QUART4
Variable	Mean	Mean		Mean	Mean	Mean	Mean				
Reading Score	494.949	493.867	0%	451.5	623.7	459.6	645.5	38%	40%	2%	3%
Mathematics Score	494.838	493.502	0%	461.6	593.8	467	613.9	29%	31%	1%	3%
Health Score	502.823	502.735	0%	475.8	583.2	480.6	597.7	23%	24%	1%	2%
R -teacher Reading Score	757.726	757.726	0%	734.2	827	739.4	834.4	13%	13%	1%	1%
M-teacher Maths Score	763.624	763.624	0%	731.3	863.5	740.6	864.9	18%	17%	1%	0%
Young (<11y3m)	0.034	0.035	2%	0.037	0.028	0.038	0.019	-24%	-51%	4%	-33%
Old (>11y3m-12y8m)	0.250	0.250	0%	0.283	0.151	0.272	0.156	-47%	-43%	-4%	3%
Very old (14y+)	0.153	0.149	-2%	0.197	0.02	0.186	0.013	-90%	-93%	-6%	-35%
Male	0.492	0.493	0%	0.498	0.475	0.497	0.471	-5%	-5%	0%	-1%
> 5 Days absent	0.028	0.028	0%	0.029	0.024	0.029	0.021	-17%	-30%	1%	-14%
Preschool - months	0.048	0.048	-1%	0.053	0.035	0.053	0.028	-33%	-48%	1%	-22%
Preschool - 1 year	0.331	0.333	0%	0.358	0.253	0.353	0.239	-29%	-32%	-1%	-6%
Preschool - 2 years	0.154	0.153	0%	0.134	0.213	0.134	0.241	58%	81%	-1%	13%
Preschool - 3 years +	0.202	0.201	0%	0.141	0.382	0.151	0.42	170%	178%	7%	10%
SES	-0.051	-0.053	4%	-0.41	1.005	-0.33	1.149	-347%	-446%	-18%	14%
SES squared	1.017	1.017	0%	0.811	1.629	0.817	1.869	101%	129%	1%	15%
Lived with parents	0.730	0.733	1%	0.688	0.854	0.696	0.874	24%	26%	1%	2%
3 or more siblings	0.542	0.544	1%	0.63	0.279	0.611	0.246	-56%	-60%	-3%	-12%
Missed breakfast	0.284	0.286	1%	0.287	0.274	0.291	0.253	-4%	-13%	1%	-8%
Missed lunch	0.203	0.201	-1%	0.227	0.134	0.224	0.118	-41%	-47%	-1%	-12%
Missed dinner	0.127	0.128	1%	0.149	0.062	0.144	0.056	-58%	-61%	-4%	-9%
> 10 books at home	0.341	0.341	0%	0.23	0.672	0.25	0.731	193%	193%	9%	9%
Used PC before	0.478	0.474	-1%	0.319	0.949	0.361	0.978	197%	171%	13%	3%
Urban	0.319	0.320	0%	0.179	0.733	0.209	0.785	310%	275%	17%	7%
Mother/father matric	0.485	0.489	1%	0.388	0.772	0.403	0.834	99%	107%	4%	8%
Mother/father degree	0.128	0.129	1%	0.074	0.287	0.077	0.345	288%	350%	4%	20%
Spk Eng. at home sometime	0.611	0.613	0%	0.643	0.516	0.629	0.535	-20%	-15%	-2%	4%

	Full sample	Teacher-test only	Teach-test as % of full sample	Quart1-3	Quart4	Quint1-4	Quint5	Quart 4 as % of Quart 1-3	Quint 5 as % of Quint 1-4	QUINT1-4 as % of QUART1-3	QUNIT5 as % of QUART4
Spk Eng. at home always	0.153	0.152	-1%	0.072	0.395	0.095	0.401	450%	321%	33%	1%
Orphan (double-orphan)	0.090	0.089	-1%	0.104	0.05	0.104	0.033	-52%	-68%	0%	-34%
Orphanage or children's ho	0.007	0.007	6%	0.008	0.004	0.008	0.003	-43%	-67%	2%	-41%
School SES	-0.051	-0.054	7%	-0.41	1.005	-0.33	1.149	-347%	-446%	-18%	14%
School SES squared	0.560	0.559	0%	0.364	1.142	0.361	1.41	214%	291%	-1%	23%
Homework - 1 or 2 times a	0.087	0.089	2%	0.102	0.04	0.098	0.037	-61%	-62%	-4%	-8%
Homework - 1 or 2 times a	0.316	0.315	-1%	0.359	0.189	0.353	0.161	-47%	-54%	-2%	-15%
Homework - Most days	0.561	0.563	0%	0.495	0.758	0.507	0.792	53%	56%	2%	4%
Repeated a grade once	0.204	0.202	-1%	0.229	0.128	0.22	0.132	-44%	-40%	-4%	3%
Repeated a grade twice	0.050	0.049	-2%	0.064	0.011	0.06	0.011	-83%	-82%	-6%	-2%
Repeated a grade three or	0.031	0.029	-4%	0.039	0.007	0.037	0.004	-82%	-90%	-5%	-44%
No class library	0.554	0.564	2%	0.619	0.36	0.605	0.333	-42%	-45%	-2%	-7%
Class-size > 40	0.551	0.550	0%	0.63	0.319	0.618	0.27	-49%	-56%	-2%	-15%
Sanitation (Std)	0.154	0.157	2%	0.454	-0.72	0.37	-0.75	-258%	-302%	-18%	5%
Building Index (Std)	-0.067	-0.076	13%	-0.48	1.157	-0.39	1.295	-341%	-435%	-19%	12%
Equipment Index (Std)	-0.168	-0.168	0%	-0.49	0.795	-0.42	0.908	-261%	-316%	-15%	14%
Avg. Parent matric (school)	0.008	0.022	"176%"	-0.38	1.173	-0.33	1.427	-405%	-539%	-15%	22%
Avg. Parent degree (school	-0.001	0.006	"-760%"	-0.36	1.053	-0.34	1.437	-396%	-525%	-5%	36%
R-Textbook -Teacher only	0.064	0.063	-2%	0.069	0.052	0.068	0.05	-25%	-26%	-1%	-4%
R-Textbook -Share 2+	0.161	0.162	0%	0.206	0.03	0.193	0.028	-86%	-85%	-6%	-5%
R-Textbook -Share with 1	0.282	0.284	1%	0.301	0.223	0.307	0.171	-26%	-44%	2%	-23%
R-Textbook - own textbook	0.450	0.449	0%	0.379	0.661	0.388	0.713	74%	84%	2%	8%
M-Textbook -Teacher only	0.174	0.176	1%	0.165	0.2	0.184	0.132	21%	-28%	11%	-34%
M-Textbook -Share 2+	0.118	0.119	1%	0.147	0.032	0.138	0.031	-78%	-78%	-6%	-4%
M-Textbook -Share with 1	0.238	0.237	-1%	0.256	0.185	0.254	0.17	-28%	-33%	-1%	-8%
M-Textbook - own textbool	0.364	0.365	0%	0.315	0.509	0.312	0.583	62%	87%	-1%	14%
Extra Eng tuition	0.096	0.090	-6%	0.104	0.072	0.102	0.07	-31%	-32%	-2%	-3%
Extra Math tuition	0.098	0.091	-7%	0.093	0.114	0.096	0.108	23%	13%	3%	-5%
R-Teacher Male	0.303	0.310	2%	0.329	0.227	0.316	0.248	-31%	-21%	-4%	10%
R-Teacher Jnr. Secondary	0.017	0.013	-22%	0.019	0.011	0.018	0.012	-44%	-38%	-4%	7%

	Full sample	Teacher-test only	Teach-test as % of full sample	Quart1-3	Quart4	Quint1-4	Quint5	Quart 4 as % of Quart 1-3	Quint 5 as % of Quint 1-4	QUINT1-4 as % of QUART1-3	QUNIT5 as % of QUART4
R-Teacher Snr. Secondary	0.123	0.127	3%	0.133	0.095	0.124	0.12	-28%	-3%	-7%	26%
R-Teacher A-level / further	0.154	0.159	4%	0.166	0.119	0.162	0.12	-28%	-26%	-2%	1%
R-Teacher Degree	0.458	0.453	-1%	0.413	0.589	0.426	0.591	43%	39%	3%	0%
R-Teacher training: 2yrs	0.083	0.083	0%	0.088	0.067	0.1	0.011	-24%	-89%	13%	-
R-Teacher training: 3yrs	0.437	0.437	0%	0.518	0.198	0.488	0.222	-62%	-54%	-6%	12%
R-Teacher training: >3yrs	0.436	0.434	0%	0.366	0.64	0.375	0.689	75%	84%	2%	8%
M-Teacher Male	0.387	0.400	3%	0.428	0.268	0.414	0.275	-37%	-34%	-3%	3%
M-Teacher Jnr. Secondary	0.017	0.018	4%	0.022	0.001	0.021	0	-95%	-100%	-6%	-
M-Teacher Snr. Secondary	0.091	0.094	4%	0.111	0.032	0.104	0.037	-71%	-64%	-6%	16%
M-Teacher A-level / further	0.168	0.173	3%	0.157	0.199	0.155	0.223	27%	44%	-2%	12%
M-Teacher Degree	0.491	0.484	-1%	0.448	0.618	0.467	0.589	38%	26%	4%	-5%
M-Teacher training: 2yrs	0.059	0.059	1%	0.073	0.016	0.07	0.009	-79%	-87%	-4%	-41%
M-Teacher training: 3yrs	0.400	0.405	1%	0.477	0.175	0.453	0.182	-63%	-60%	-5%	4%
M-Teacher training: >3yrs	0.508	0.501	-1%	0.433	0.728	0.461	0.701	68%	52%	7%	-4%
Eastern Cape	0.163	0.148	-9%	0.213	0.013	0.197	0.017	-94%	-91%	-8%	33%
Free State	0.050	0.052	5%	0.054	0.037	0.051	0.043	-30%	-16%	-5%	15%
Gauteng	0.172	0.173	0%	0.109	0.361	0.121	0.39	232%	223%	11%	8%
KwaZulu Natal	0.237	0.244	3%	0.248	0.204	0.255	0.161	-18%	-37%	3%	-21%
Limpopo	0.134	0.139	4%	0.159	0.06	0.157	0.038	-62%	-76%	-1%	-37%
Mpumalanga	0.085	0.090	5%	0.097	0.05	0.093	0.055	-49%	-41%	-5%	10%
Northern Cape	0.020	0.021	1%	0.022	0.016	0.021	0.019	-26%	-10%	-5%	16%
Western Cape	0.078	0.073	-7%	0.039	0.195	0.052	0.192	401%	273%	32%	-1%
North West	0.101	0.061	-40%	0.059	0.064	0.054	0.086	10%	58%	-8%	33%

Stellenbosch University http://scholar.sun.ac.za

Table 2: *Full sample*

Full sample					
Variable	Obs	Mean	Std. Dev.	Min	Max
Reading Score	9071	494.9494	116.2362	63.04475	965.6984
Mathematics Score	9051	494.8381	98.07328	12.35914	999.4773
Health Score	9053	502.8229	101.1873	78.98245	958.0318
Reading -teacher Reading Score	8088	757.7259	81.69571	289.8527	1090.262
Maths-teacher Maths Score	7872	763.6243	108.8479	469.2937	1204.372
Young (<11y3m)	9071	0.03449	0.182494	0	1
Old (>11y3m-12y8m)	9071	0.249557	0.432781	0	1
Very old (14y+)	9071	0.152737	0.359753	0	1
Male	9071	0.492202	0.499967	0	1
> 5 Days absent	9071	0.027772	0.164329	0	1
Preschool - months	9071	0.048383	0.214587	0	1
Preschool - 1 year	9071	0.331366	0.47073	0	1
Preschool - 2 years	9071	0.154045	0.361012	0	1
Preschool - 3 years or more	9071	0.202099	0.401588	0	1
SES	9071	-0.05056	1.007307	-2.22579	2.375511
SES squared	9071	1.017113	1.177636	6.34E-09	5.643051
Lived with parents	9071	0.729578	0.444202	0	1
3 or more siblings	9071	0.541525	0.4983	0	1
Missed breakfast	9071	0.284057	0.450989	0	1
Missed lunch	9071	0.203428	0.40257	0	1
Missed dinner	9071	0.126969	0.332957	0	1
More than 10 books at home	9071	0.341401	0.474206	0	1
Used PC before	9071	0.478232	0.499554	0	1
Urban	9071	0.31858	0.465951	0	1
Mother or father has matric	9071	0.485058	0.499804	0	1
Mother or father has degree	9071	0.127625	0.33369	0	1
Speak Eng. at home sometimes	9071	0.611109	0.487526	0	1
Speak Eng. at home always	9071	0.153277	0.360274	0	1
Orphan (double-orphan)	9071	0.090349	0.286697	0	1
Orphanage or children's home	9071	0.006715	0.081674	0	1
School SES	9071	-0.05056	0.746732	-2.00301	1.832755
School SES squared	9071	0.560104	0.693849	2.43E-07	4.012039
Homework - 1 or 2 times a month	9071	0.086697	0.281406	0	1
Homework - 1 or 2 times a week	9071	0.316198	0.465017	0	1
Homework - Most days	9071	0.560932	0.496301	0	1
Repeated a grade once	9071	0.203578	0.402681	0	1
Repeated a grade twice	9071	0.050363	0.218705	0	1
Repeated a grade three or more	9071	0.030679	0.172456	0	1
No class library	9071	0.553718	0.497133	0	1
Class-size > 40	8936	0.550697	0.497451	0	1
Sanitation (Std)	8957	0.153832	1.032807	-0.78092	1.372758
Building Index (Std)	9071	-0.0674	1.024177	-1.49432	1.689308

Stellenbosch University http://scholar.sun.ac.za

Equipment Index (Std)	9071	-0.16779	1.053493	-2.23241	1.13875
Avg. Parent matric (school)	9071	0.008053	0.992004	-1.96436	2.10198
Avg. Parent degree (school)	9071	-0.00096	0.971916	-0.84508	4.181613
R-Textbook -Teacher only	9071	0.064367	0.245419	0	1
R-Textbook -Share 2+	9071	0.161455	0.36797	0	1
R-Textbook -Share with 1	9071	0.28156	0.449785	0	1
R-Textbook - own textbook	9071	0.450188	0.49754	0	1
M-Textbook -Teacher only	9071	0.173952	0.379089	0	1
M-Textbook -Share 2+	9071	0.117922	0.322533	0	1
M-Textbook -Share with 1	9071	0.238227	0.426022	0	1
M-Textbook - own textbook	9071	0.363908	0.481149	0	1
Extra Eng tuition	9071	0.095511	0.293935	0	1
Extra Math tuition	9071	0.098237	0.297651	0	1
R-Teacher Male	9071	0.302993	0.459578	0	1
R-Teacher Jnr. Secondary	8936	0.017111	0.129693	0	1
R-Teacher Snr. Secondary	8936	0.123205	0.328691	0	1
R-Teacher A-level / further study	8936	0.153675	0.360657	0	1
R-Teacher Degree	8936	0.457676	0.498233	0	1
R-Teacher training: 2yrs	8936	0.083024	0.275934	0	1
R-Teacher training: 3yrs	8936	0.436709	0.496006	0	1
R-Teacher training: >3yrs	8936	0.435778	0.495886	0	1
M-Teacher Male	9071	0.387272	0.487154	0	1
M-Teacher Jnr. Secondary	8936	0.01695	0.129093	0	1
M-Teacher Snr. Secondary	8936	0.091106	0.287775	0	1
M-Teacher A-level / further study	8936	0.167928	0.373823	0	1
M-Teacher Degree	8936	0.490983	0.499947	0	1
M-Teacher training: 2yrs	8936	0.058532	0.234759	0	1
M-Teacher training: 3yrs	8936	0.400291	0.489985	0	1
M-Teacher training: >3yrs	8936	0.507869	0.499966	0	1
Eastern Cape	9071	0.16285	0.369249	0	1
Free State	9071	0.049548	0.217021	0	1
Gauteng	9071	0.17221	0.377584	0	1
KwaZulu Natal	9071	0.236756	0.425114	0	1
Limpopo	9071	0.134316	0.34101	0	1
Mpumalanga	9071	0.085412	0.279509	0	1
Northern Cape	9071	0.02048	0.141644	0	1
Western Cape	9071	0.078268	0.268608	0	1

Stellenbosch University http://scholar.sun.ac.za

Table 3: *Only learners with non-missing teacher test scores*

Only learners with non-missing teacher test scores					
Variable	Obs	Mean	Std. Dev.	Min	Max
Reading Score	8578	493.8666	115.9382	63.04475	965.6984
Mathematics Score	8559	493.5022	97.93608	12.35914	999.4773
Health Score	8560	502.7348	101.0356	78.98245	958.0318
Reading -teacher Reading Score	8088	757.7259	81.69571	289.8527	1090.262
Maths-teacher Maths Score	7872	763.6243	108.8479	469.2937	1204.372
Young (<11y3m)	8578	0.035343	0.184655	0	1
Old (>11y3m-12y8m)	8578	0.250232	0.433172	0	1
Very old (14y+)	8578	0.14946	0.356562	0	1
Male	8578	0.492953	0.49998	0	1
> 5 Days absent	8578	0.027724	0.16419	0	1
Preschool - months	8578	0.048097	0.213984	0	1
Preschool - 1 year	8578	0.332621	0.47118	0	1
Preschool - 2 years	8578	0.153476	0.360467	0	1
Preschool - 3 years or more	8578	0.201142	0.400878	0	1
SES	8578	-0.05275	1.006956	-2.22579	2.375511
SES squared	8578	1.016624	1.174512	6.34E-09	5.643051
Lived with parents	8578	0.733256	0.442284	0	1
3 or more siblings	8578	0.544394	0.498054	0	1
Missed breakfast	8578	0.286448	0.452128	0	1
Missed lunch	8578	0.201284	0.400983	0	1
Missed dinner	8578	0.127736	0.333816	0	1
More than 10 books at home	8578	0.340617	0.473944	0	1
Used PC before	8578	0.473926	0.499349	0	1
Urban	8578	0.320094	0.46654	0	1
Mother or father has matric	8578	0.48921	0.499913	0	1
Mother or father has degree	8578	0.128685	0.334871	0	1
Speak Eng. at home sometimes	8578	0.613074	0.487075	0	1
Speak Eng. at home always	8578	0.151598	0.358652	0	1
Orphan (double-orphan)	8578	0.089371	0.285295	0	1
Orphanage or children's home	8578	0.00709	0.08391	0	1
School SES	8578	-0.05414	0.745916	-2.00301	1.832755
School SES squared	8578	0.559257	0.688786	2.43E-07	4.012039
Homework - 1 or 2 times a month	8578	0.088527	0.284076	0	1
Homework - 1 or 2 times a week	8578	0.314556	0.464366	0	1
Homework - Most days	8578	0.562556	0.4961	0	1
Repeated a grade once	8578	0.20194	0.401471	0	1
Repeated a grade twice	8578	0.049321	0.216549	0	1
Repeated a grade three or more	8578	0.029322	0.168717	0	1
No class library	8578	0.564104	0.495903	0	1
Class-size > 40	8578	0.55037	0.497485	0	1
Sanitation (Std)	8464	0.156664	1.032838	-0.78092	1.372758
Building Index (Std)	8578	-0.07602	1.028298	-1.49432	1.689308

Stellenbosch University http://scholar.sun.ac.za

Equipment Index (Std)	8578	-0.1685	1.056692	-2.23241	1.13875
Avg. Parent matric (school)	8578	0.022232	0.991923	-1.96436	2.10198
Avg. Parent degree (school)	8578	0.006348	0.968314	-0.84508	4.181613
R-Textbook -Teacher only	8578	0.062884	0.242769	0	1
R-Textbook -Share 2+	8578	0.162031	0.368501	0	1
R-Textbook -Share with 1	8578	0.283569	0.450756	0	1
R-Textbook - own textbook	8578	0.449286	0.49745	0	1
M-Textbook -Teacher only	8578	0.175631	0.380528	0	1
M-Textbook -Share 2+	8578	0.118517	0.323238	0	1
M-Textbook -Share with 1	8578	0.236704	0.425084	0	1
M-Textbook - own textbook	8578	0.365048	0.481472	0	1
Extra Eng tuition	8578	0.089938	0.286109	0	1
Extra Math tuition	8578	0.091396	0.288188	0	1
R-Teacher Male	8578	0.310418	0.462692	0	1
R-Teacher Jnr. Secondary	8578	0.013429	0.115111	0	1
R-Teacher Snr. Secondary	8578	0.126688	0.332642	0	1
R-Teacher A-level / further study	8578	0.159079	0.365771	0	1
R-Teacher Degree	8578	0.453452	0.497858	0	1
R-Teacher training: 2yrs	8578	0.082837	0.275652	0	1
R-Teacher training: 3yrs	8578	0.436852	0.496025	0	1
R-Teacher training: >3yrs	8578	0.43399	0.495652	0	1
M-Teacher Male	8578	0.399584	0.489841	0	1
M-Teacher Jnr. Secondary	8578	0.017648	0.131677	0	1
M-Teacher Snr. Secondary	8578	0.094421	0.292431	0	1
M-Teacher A-level / further study	8578	0.173006	0.378274	0	1
M-Teacher Degree	8578	0.484366	0.499785	0	1
M-Teacher training: 2yrs	8578	0.058997	0.235633	0	1
M-Teacher training: 3yrs	8578	0.404979	0.490917	0	1
M-Teacher training: >3yrs	8578	0.501344	0.500027	0	1
Eastern Cape	8578	0.148415	0.355531	0	1
Free State	8578	0.051868	0.221774	0	1
Gauteng	8578	0.172679	0.377991	0	1
KwaZulu Natal	8578	0.244143	0.429603	0	1
Limpopo	8578	0.139216	0.346192	0	1
Mpumalanga	8578	0.089753	0.285844	0	1
Northern Cape	8578	0.020609	0.142079	0	1
Western Cape	8578	0.072529	0.259377	0	1

Stellenbosch University http://scholar.sun.ac.za

Table 4: *Only learners from quartiles 1-3 of school SES*

Only learners from quartiles 1-3 of school SES					
Variable	Obs	Mean	Std. Dev.	Min	Max
Reading Score	6816	451.5417	82.54769	63.04475	834.6495
Mathematics Score	6806	461.6164	72.97995	12.35914	769.1535
Health Score	6805	475.7821	87.03681	78.98245	827.7368
Reading -teacher Reading Score	6097	734.2255	68.40918	289.8527	958.6404
Maths-teacher Maths Score	6035	731.3492	84.58788	551.4316	1051.123
Young (<11y3m)	6816	0.03672	0.188087	0	1
Old (>11y3m-12y8m)	6816	0.282831	0.450408	0	1
Very old (14y+)	6816	0.197403	0.398068	0	1
Male	6816	0.497948	0.500033	0	1
> 5 Days absent	6816	0.029042	0.167938	0	1
Preschool - months	6816	0.05278	0.223611	0	1
Preschool - 1 year	6816	0.357766	0.479378	0	1
Preschool - 2 years	6816	0.134303	0.341003	0	1
Preschool - 3 years or more	6816	0.141423	0.348483	0	1
SES	6816	-0.40647	0.803553	-2.22579	2.375511
SES squared	6816	0.810819	0.964958	6.34E-09	5.643051
Lived with parents	6816	0.687711	0.463461	0	1
3 or more siblings	6816	0.630054	0.482825	0	1
Missed breakfast	6816	0.287299	0.452535	0	1
Missed lunch	6816	0.226859	0.418832	0	1
Missed dinner	6816	0.148883	0.356	0	1
More than 10 books at home	6816	0.229823	0.42075	0	1
Used PC before	6816	0.319423	0.466287	0	1
Urban	6816	0.178819	0.383228	0	1
Mother or father has matric	6816	0.388499	0.487445	0	1
Mother or father has degree	6816	0.073913	0.261649	0	1
Speak Eng. at home sometimes	6816	0.643026	0.479142	0	1
Speak Eng. at home always	6816	0.071787	0.258154	0	1
Orphan (double-orphan)	6816	0.103855	0.305095	0	1
Orphanage or children's home	6816	0.007538	0.086502	0	1
School SES	6816	-0.40647	0.445685	-2.00301	0.365661
School SES squared	6816	0.363823	0.530601	2.43E-07	4.012039
Homework - 1 or 2 times a month	6816	0.102376	0.303165	0	1
Homework - 1 or 2 times a week	6816	0.359048	0.479757	0	1
Homework - Most days	6816	0.494532	0.500007	0	1
Repeated a grade once	6816	0.228991	0.420214	0	1
Repeated a grade twice	6816	0.063619	0.244091	0	1
Repeated a grade three or more	6816	0.038729	0.192963	0	1
No class library	6816	0.619102	0.485643	0	1
Class-size > 40	6705	0.629926	0.48286	0	1
Sanitation (Std)	6702	0.453924	1.025624	-0.78092	1.372758
Building Index (Std)	6816	-0.48019	0.770959	-1.49432	1.689308

Stellenbosch University http://scholar.sun.ac.za

Equipment Index (Std)	6816	-0.49239	0.990923	-2.23241	1.13875
Avg. Parent matric (school)	6816	-0.38459	0.733295	-1.96436	1.939326
Avg. Parent degree (school)	6816	-0.35621	0.539508	-0.84508	2.065111
R-Textbook -Teacher only	6816	0.068611	0.25281	0	1
R-Textbook -Share 2+	6816	0.205822	0.40433	0	1
R-Textbook -Share with 1	6816	0.301174	0.458802	0	1
R-Textbook - own textbook	6816	0.379245	0.485235	0	1
M-Textbook -Teacher only	6816	0.16527	0.371451	0	1
M-Textbook -Share 2+	6816	0.146835	0.353967	0	1
M-Textbook -Share with 1	6816	0.256001	0.436455	0	1
M-Textbook - own textbook	6816	0.314897	0.464509	0	1
Extra Eng tuition	6816	0.103538	0.304682	0	1
Extra Math tuition	6816	0.09294	0.29037	0	1
R-Teacher Male	6816	0.328675	0.469766	0	1
R-Teacher Jnr. Secondary	6705	0.019281	0.137521	0	1
R-Teacher Snr. Secondary	6705	0.132843	0.339431	0	1
R-Teacher A-level / further study	6705	0.165582	0.371732	0	1
R-Teacher Degree	6705	0.412805	0.492375	0	1
R-Teacher training: 2yrs	6705	0.088382	0.283871	0	1
R-Teacher training: 3yrs	6705	0.518227	0.499705	0	1
R-Teacher training: >3yrs	6705	0.366043	0.481757	0	1
M-Teacher Male	6816	0.4276	0.494767	0	1
M-Teacher Jnr. Secondary	6705	0.022385	0.147942	0	1
M-Teacher Snr. Secondary	6705	0.111297	0.314523	0	1
M-Teacher A-level / further study	6705	0.157279	0.36409	0	1
M-Teacher Degree	6705	0.447614	0.497285	0	1
M-Teacher training: 2yrs	6705	0.073198	0.260481	0	1
M-Teacher training: 3yrs	6705	0.477272	0.49952	0	1
M-Teacher training: >3yrs	6705	0.432803	0.495501	0	1
Eastern Cape	6816	0.213325	0.409685	0	1
Free State	6816	0.05367	0.225382	0	1
Gauteng	6816	0.108738	0.311333	0	1
KwaZulu Natal	6816	0.247882	0.431815	0	1
Limpopo	6816	0.159395	0.366071	0	1
Mpumalanga	6816	0.097425	0.296558	0	1
Northern Cape	6816	0.02193	0.146466	0	1
Western Cape	6816	0.038927	0.193436	0	1

Stellenbosch University http://scholar.sun.ac.za

Table 5*: Only learners from quartile 4 of school SES*

Only learners from quartile 4 of school SES					
Variable	Obs	Mean	Std. Dev.	Min	Max
Reading Score	2255	623.7288	105.8593	122.0835	965.6984
Mathematics Score	2245	593.8205	96.55771	252.3595	999.4773
Health Score	2248	583.2168	97.58814	78.98245	958.0318
Reading -teacher Reading Score	1991	827.0005	78.28377	624.6281	1090.262
Maths-teacher Maths Score	1837	863.5033	114.7843	469.2937	1204.372
Young (<11y3m)	2255	0.027874	0.164649	0	1
Old (>11y3m-12y8m)	2255	0.150843	0.357975	0	1
Very old (14y+)	2255	0.020223	0.140795	0	1
Male	2255	0.475156	0.499493	0	1
> 5 Days absent	2255	0.024004	0.153094	0	1
Preschool - months	2255	0.035339	0.184675	0	1
Preschool - 1 year	2255	0.253043	0.434852	0	1
Preschool - 2 years	2255	0.212614	0.409247	0	1
Preschool - 3 years or more	2255	0.382108	0.486011	0	1
SES	2255	1.005317	0.786603	-2.22579	2.375511
SES squared	2255	1.629133	1.496225	2.27E-07	5.643051
Lived with parents	2255	0.853785	0.3534	0	1
3 or more siblings	2255	0.278881	0.448549	0	1
Missed breakfast	2255	0.274439	0.44633	0	1
Missed lunch	2255	0.133911	0.340633	0	1
Missed dinner	2255	0.061956	0.24113	0	1
More than 10 books at home	2255	0.672424	0.469433	0	1
Used PC before	2255	0.949379	0.219272	0	1
Urban	2255	0.733215	0.442377	0	1
Mother or father has matric	2255	0.771525	0.419943	0	1
Mother or father has degree	2255	0.286972	0.452449	0	1
Speak Eng. at home sometimes	2255	0.516419	0.499841	0	1
Speak Eng. at home always	2255	0.395035	0.488967	0	1
Orphan (double-orphan)	2255	0.050281	0.218572	0	1
Orphanage or children's home	2255	0.004273	0.06524	0	1
School SES	2255	1.005317	0.363063	0.444907	1.832755
School SES squared	2255	1.142419	0.788192	0.197942	3.358993
Homework - 1 or 2 times a month	2255	0.04018	0.196425	0	1
Homework - 1 or 2 times a week	2255	0.189072	0.391653	0	1
Homework - Most days	2255	0.757925	0.428434	0	1
Repeated a grade once	2255	0.128182	0.334366	0	1
Repeated a grade twice	2255	0.011035	0.10449	0	1
Repeated a grade three or more	2255	0.006796	0.082177	0	1
No class library	2255	0.359741	0.480031	0	1
Class-size > 40	2231	0.318839	0.466131	0	1
Sanitation (Std)	2255	-0.71651	0.301545	-0.78092	1.372758
Building Index (Std)	2255	1.157249	0.626256	-1.49432	1.689308

Stellenbosch University http://scholar.sun.ac.za

Equipment Index (Std)	2255	0.795209	0.499446	-1.48326	1.13875
Avg. Parent matric (school)	2255	1.172924	0.702847	-1.43397	2.10198
Avg. Parent degree (school)	2255	1.052976	1.182658	-0.84508	4.181613
R-Textbook -Teacher only	2255	0.051776	0.221623	0	1
R-Textbook -Share 2+	2255	0.02983	0.170155	0	1
R-Textbook -Share with 1	2255	0.223371	0.416597	0	1
R-Textbook - own textbook	2255	0.660655	0.473592	0	1
M-Textbook -Teacher only	2255	0.199709	0.39987	0	1
M-Textbook -Share 2+	2255	0.032143	0.17642	0	1
M-Textbook -Share with 1	2255	0.185496	0.388786	0	1
M-Textbook - own textbook	2255	0.509309	0.500024	0	1
Extra Eng tuition	2255	0.071696	0.258041	0	1
Extra Math tuition	2255	0.113953	0.317824	0	1
R-Teacher Male	2255	0.226802	0.418856	0	1
R-Teacher Jnr. Secondary	2231	0.010762	0.103203	0	1
R-Teacher Snr. Secondary	2231	0.094999	0.29328	0	1
R-Teacher A-level / further study	2231	0.118832	0.323663	0	1
R-Teacher Degree	2231	0.588984	0.492129	0	1
R-Teacher training: 2yrs	2231	0.067345	0.250674	0	1
R-Teacher training: 3yrs	2231	0.198157	0.398701	0	1
R-Teacher training: >3yrs	2231	0.639848	0.480152	0	1
M-Teacher Male	2255	0.26763	0.442822	0	1
M-Teacher Jnr. Secondary	2231	0.001048	0.032362	0	1
M-Teacher Snr. Secondary	2231	0.032018	0.176087	0	1
M-Teacher A-level / further study	2231	0.199091	0.399406	0	1
M-Teacher Degree	2231	0.617899	0.48601	0	1
M-Teacher training: 2yrs	2231	0.015611	0.123994	0	1
M-Teacher training: 3yrs	2231	0.175011	0.380062	0	1
M-Teacher training: >3yrs	2231	0.727543	0.445324	0	1
Eastern Cape	2255	0.013106	0.113753	0	1
Free State	2255	0.037318	0.189582	0	1
Gauteng	2255	0.360515	0.480257	0	1
KwaZulu Natal	2255	0.203746	0.402872	0	1
Limpopo	2255	0.059914	0.23738	0	1
Mpumalanga	2255	0.049771	0.21752	0	1
Northern Cape	2255	0.016179	0.126191	0	1
Western Cape	2255	0.194983	0.396276	0	1

Stellenbosch University http://scholar.sun.ac.za

Table 6*: Only learners from quintiles 1 - 4 of school SES*

Only learners from quintiles 1 - 4 of school SES					
Variable	Obs	Mean	Std. Dev.	Min	Max
Reading Score	7271	459.6275	90.71092	63.04475	834.6495
Mathematics Score	7258	466.9693	76.41124	12.35914	894.4982
Health Score	7260	480.6357	89.53737	78.98245	958.0318
Reading -teacher Reading Score	6477	739.3882	72.55563	289.8527	991.4503
Maths-teacher Maths Score	6407	740.5964	92.59928	551.4316	1051.123
Young (<11y3m)	7271	0.03819	0.191668	0	1
Old (>11y3m-12y8m)	7271	0.271594	0.444812	0	1
Very old (14y+)	7271	0.185513	0.38874	0	1
Male	7271	0.497069	0.500026	0	1
> 5 Days absent	7271	0.029444	0.169058	0	1
Preschool - months	7271	0.053265	0.224576	0	1
Preschool - 1 year	7271	0.353155	0.477983	0	1
Preschool - 2 years	7271	0.133575	0.340218	0	1
Preschool - 3 years or more	7271	0.150962	0.358037	0	1
SES	7271	-0.33206	0.840781	-2.22579	2.375511
SES squared	7271	0.817082	0.981452	6.34E-09	5.643051
Lived with parents	7271	0.695627	0.460173	0	1
3 or more siblings	7271	0.610817	0.487599	0	1
Missed breakfast	7271	0.29145	0.454461	0	1
Missed lunch	7271	0.223559	0.416658	0	1
Missed dinner	7271	0.143594	0.350701	0	1
More than 10 books at home	7271	0.249934	0.433005	0	1
Used PC before	7271	0.360985	0.480319	0	1
Urban	7271	0.209161	0.406737	0	1
Mother or father has matric	7271	0.403149	0.490564	0	1
Mother or father has degree	7271	0.076601	0.265976	0	1
Speak Eng. at home sometimes	7271	0.628941	0.483122	0	1
Speak Eng. at home always	7271	0.095177	0.29348	0	1
Orphan (double-orphan)	7271	0.103734	0.304936	0	1
Orphanage or children's home	7271	0.007696	0.087397	0	1
School SES	7271	-0.33206	0.500449	-2.00301	0.663683
School SES squared	7271	0.360682	0.510326	2.43E-07	4.012039
Homework - 1 or 2 times a month	7271	0.098324	0.297773	0	1
Homework - 1 or 2 times a week	7271	0.352518	0.477787	0	1
Homework - Most days	7271	0.5068	0.499988	0	1
Repeated a grade once	7271	0.220363	0.41452	0	1
Repeated a grade twice	7271	0.05965	0.236854	0	1
Repeated a grade three or more	7271	0.036985	0.188737	0	1
No class library	7271	0.605461	0.488785	0	1
Class-size > 40	7136	0.618121	0.485881	0	1
Sanitation (Std)	7157	0.370249	1.036888	-0.78092	1.372758
Building Index (Std)	7271	-0.38708	0.838122	-1.49432	1.689308

Stellenbosch University http://scholar.sun.ac.za

Equipment Index (Std)	7271	-0.42031	1.008613	-2.23241	1.13875
Avg. Parent matric (school)	7271	-0.32502	0.761808	-1.96436	1.939326
Avg. Parent degree (school)	7271	-0.33843	0.563731	-0.84508	2.858799
R-Textbook -Teacher only	7271	0.067774	0.251376	0	1
R-Textbook -Share 2+	7271	0.192663	0.394418	0	1
R-Textbook -Share with 1	7271	0.307418	0.461456	0	1
R-Textbook - own textbook	7271	0.388463	0.487434	0	1
M-Textbook -Teacher only	7271	0.183684	0.387253	0	1
M-Textbook -Share 2+	7271	0.13832	0.345259	0	1
M-Textbook -Share with 1	7271	0.254206	0.435444	0	1
M-Textbook - own textbook	7271	0.312462	0.463529	0	1
Extra Eng tuition	7271	0.10161	0.302155	0	1
Extra Math tuition	7271	0.095943	0.294533	0	1
R-Teacher Male	7271	0.315797	0.464865	0	1
R-Teacher Jnr. Secondary	7136	0.018454	0.134596	0	1
R-Teacher Snr. Secondary	7136	0.123975	0.329576	0	1
R-Teacher A-level / further study	7136	0.161815	0.368307	0	1
R-Teacher Degree	7136	0.425669	0.494479	0	1
R-Teacher training: 2yrs	7136	0.10019	0.300274	0	1
R-Teacher training: 3yrs	7136	0.488133	0.499894	0	1
R-Teacher training: >3yrs	7136	0.375045	0.484169	0	1
M-Teacher Male	7271	0.413696	0.492529	0	1
M-Teacher Jnr. Secondary	7136	0.021017	0.14345	0	1
M-Teacher Snr. Secondary	7136	0.104063	0.305364	0	1
M-Teacher A-level / further study	7136	0.154739	0.361681	0	1
M-Teacher Degree	7136	0.467354	0.498968	0	1
M-Teacher training: 2yrs	7136	0.070362	0.255773	0	1
M-Teacher training: 3yrs	7136	0.452758	0.497798	0	1
M-Teacher training: >3yrs	7136	0.461423	0.498545	0	1
Eastern Cape	7271	0.196989	0.397751	0	1
Free State	7271	0.051087	0.22019	0	1
Gauteng	7271	0.121021	0.326174	0	1
KwaZulu Natal	7271	0.254626	0.435681	0	1
Limpopo	7271	0.157008	0.363833	0	1
Mpumalanga	7271	0.092624	0.289925	0	1
Northern Cape	7271	0.020873	0.14297	0	1
Western Cape	7271	0.051561	0.221154	0	1

Stellenbosch University http://scholar.sun.ac.za

Table 7*: Only learners from quintile 5 of school SES*

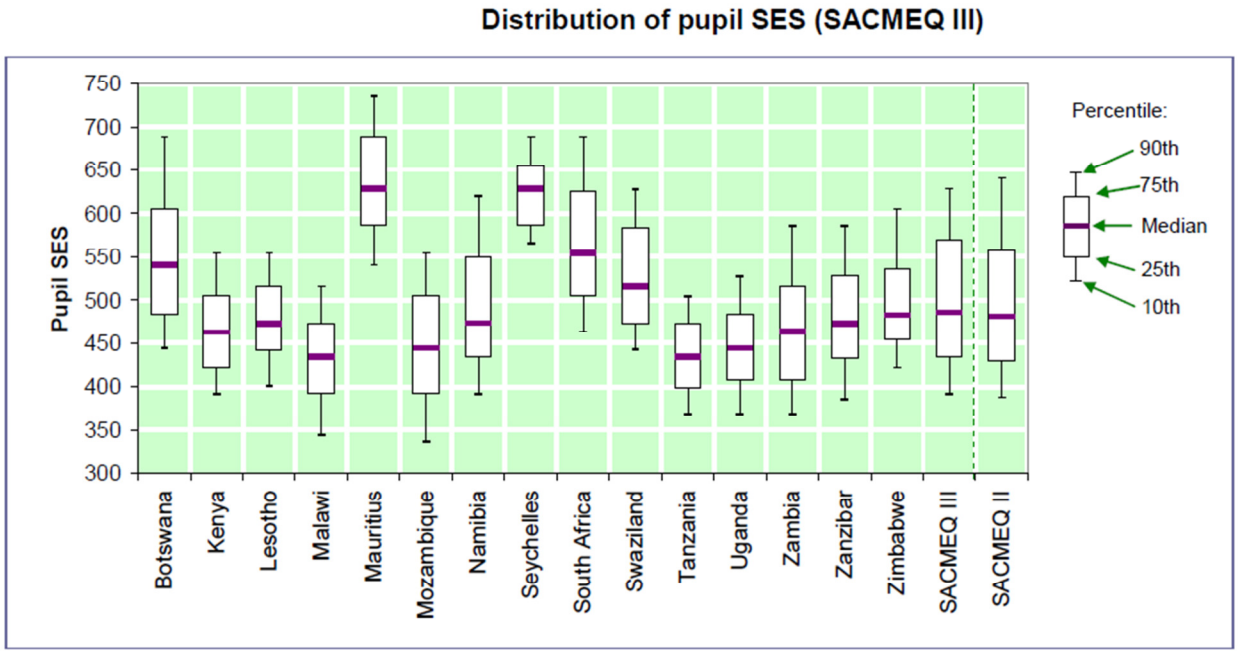
Only learners from quintile 5 of school SES					
Variable	Obs	Mean	Std. Dev.	Min	Max
Reading Score	1800	645.4604	89.7289	315.3185	965.6984
Mathematics Score	1793	613.9309	90.94288	252.3595	999.4773
Health Score	1793	597.6835	92.94615	290.7856	958.0318
Reading -teacher Reading Score	1611	834.3716	72.72778	624.6281	1090.262
Maths-teacher Maths Score	1465	864.8945	116.8848	469.2937	1204.372
Young (<11y3m)	1800	0.018724	0.135586	0	1
Old (>11y3m-12y8m)	1800	0.155655	0.362629	0	1
Very old (14y+)	1800	0.013073	0.113621	0	1
Male	1800	0.471463	0.499324	0	1
> 5 Days absent	1800	0.02065	0.142249	0	1
Preschool - months	1800	0.027583	0.16382	0	1
Preschool - 1 year	1800	0.238522	0.426298	0	1
Preschool - 2 years	1800	0.241271	0.427973	0	1
Preschool - 3 years or more	1800	0.419997	0.493695	0	1
SES	1800	1.148948	0.741415	-2.22579	2.375511
SES squared	1800	1.869471	1.51518	2.27E-07	5.643051
Lived with parents	1800	0.874245	0.331665	0	1
3 or more siblings	1800	0.246262	0.430953	0	1
Missed breakfast	1800	0.252557	0.4346	0	1
Missed lunch	1800	0.117645	0.322277	0	1
Missed dinner	1800	0.056131	0.230239	0	1
More than 10 books at home	1800	0.731153	0.443483	0	1
Used PC before	1800	0.977835	0.14726	0	1
Urban	1800	0.784828	0.411056	0	1
Mother or father has matric	1800	0.834084	0.372109	0	1
Mother or father has degree	1800	0.345041	0.475514	0	1
Speak Eng. at home sometimes	1800	0.535124	0.498903	0	1
Speak Eng. at home always	1800	0.400847	0.490206	0	1
Orphan (double-orphan)	1800	0.033315	0.179508	0	1
Orphanage or children's home	1800	0.002533	0.050282	0	1
School SES	1800	1.148948	0.299731	0.673315	1.832755
School SES squared	1800	1.409869	0.729187	0.453353	3.358993
Homework - 1 or 2 times a month	1800	0.037151	0.189185	0	1
Homework - 1 or 2 times a week	1800	0.161435	0.368034	0	1
Homework - Most days	1800	0.791597	0.406279	0	1
Repeated a grade once	1800	0.132051	0.338641	0	1
Repeated a grade twice	1800	0.010789	0.103335	0	1
Repeated a grade three or more	1800	0.003809	0.06162	0	1
No class library	1800	0.333237	0.471502	0	1
Class-size > 40	1800	0.269633	0.443893	0	1

Stellenbosch University http://scholar.sun.ac.za

Sanitation (Std)	1800	-0.74926	0.142989	-0.78092	0.122233
Building Index (Std)	1800	1.29483	0.483909	-0.58471	1.689308
Equipment Index (Std)	1800	0.908198	0.273397	0.015031	1.13875
Avg. Parent matric (school)	1800	1.427308	0.46612	0.068811	2.10198
Avg. Parent degree (school)	1800	1.437047	1.030869	-0.58052	4.181613
R-Textbook -Teacher only	1800	0.049847	0.21769	0	1
R-Textbook -Share 2+	1800	0.028473	0.166366	0	1
R-Textbook -Share with 1	1800	0.171374	0.37694	0	1
R-Textbook - own textbook	1800	0.713205	0.452391	0	1
M-Textbook -Teacher only	1800	0.132483	0.339109	0	1
M-Textbook -Share 2+	1800	0.031001	0.173369	0	1
M-Textbook -Share with 1	1800	0.170141	0.375861	0	1
M-Textbook - own textbook	1800	0.583127	0.493179	0	1
Extra Eng tuition	1800	0.069521	0.254408	0	1
Extra Math tuition	1800	0.108012	0.310482	0	1
R-Teacher Male	1800	0.248434	0.432225	0	1
R-Teacher Jnr. Secondary	1800	0.011514	0.106711	0	1
R-Teacher Snr. Secondary	1800	0.119997	0.325049	0	1
R-Teacher A-level / further study	1800	0.119742	0.324749	0	1
R-Teacher Degree	1800	0.591098	0.491768	0	1
R-Teacher training: 2yrs	1800	0.011467	0.1065	0	1
R-Teacher training: 3yrs	1800	0.222347	0.415938	0	1
R-Teacher training: >3yrs	1800	0.688943	0.463055	0	1
M-Teacher Male	1800	0.274674	0.446474	0	1
M-Teacher Jnr. Secondary	1800	0	0	0	0
M-Teacher Snr. Secondary	1800	0.037091	0.189037	0	1
M-Teacher A-level / further study	1800	0.222904	0.41631	0	1
M-Teacher Degree	1800	0.589484	0.492064	0	1
M-Teacher training: 2yrs	1800	0.009218	0.095595	0	1
M-Teacher training: 3yrs	1800	0.181575	0.385601	0	1
M-Teacher training: >3yrs	1800	0.701483	0.457735	0	1
Eastern Cape	1800	0.017382	0.130727	0	1
Free State	1800	0.04299	0.202891	0	1
Gauteng	1800	0.390333	0.487961	0	1
KwaZulu Natal	1800	0.160608	0.367271	0	1
Limpopo	1800	0.037624	0.190337	0	1
Mpumalanga	1800	0.05468	0.227418	0	1
Northern Cape	1800	0.018805	0.135874	0	1
Western Cape	1800	0.192073	0.39404	0	1

Stellenbosch University http://scholar.sun.ac.za

Figure 58: Distribution of Socioeconomic Status across SACMEQ countries



Source: SACMEQ (2010)

Stellenbosch University http://scholar.sun.ac.za

Table 31: Mean SACMEQ mathematics scores for participating countries for sub-groups

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

Table 32: Mean SACMEQ reading scores for participating countries for sub-groups

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

Appendix F

Design Effects

Variable	Province	Mean	Linearized SE	DEFF	DEFT	MEFF	MEFT	Sample size	
								Actual	Effective
Reading score	ECA	447.74	10.15	13.31	3.65	13.15	3.63	1068	293
	FST	491.17	12.48	14.73	3.84	15.59	3.95	958	250
	GTN	573.07	14.39	14.49	3.81	14.02	3.74	1020	268
	KZN	485.19	10.51	13.44	3.67	12.87	3.59	1492	407
	LMP	425.23	7.72	8.43	2.90	8.46	2.91	917	316
	MPU	473.61	11.14	11.90	3.45	11.73	3.43	869	252
	NCA	505.86	12.57	12.46	3.53	12.25	3.50	926	262
	NWP	506.26	14.19	14.36	3.79	14.45	3.80	914	241
	WCA	583.71	11.16	10.14	3.18	10.07	3.17	907	285
Maths score	ECA	468.77	10.31	14.41	3.80	14.57	3.82	1068	281
	FST	491.57	10.08	13.80	3.71	14.62	3.82	958	258
	GTN	545.01	11.99	13.93	3.73	13.96	3.74	1020	273
	KZN	485.23	8.22	11.84	3.44	11.28	3.36	1492	434
	LMP	446.72	5.25	5.24	2.29	5.25	2.29	917	400
	MPU	476.12	8.19	8.88	2.98	8.38	2.89	869	292
	NCA	498.72	10.83	11.59	3.40	11.28	3.36	926	272
	NWP	503.06	13.14	15.43	3.93	15.94	3.99	914	233
	WCA	565.69	12.01	11.83	3.44	11.19	3.35	907	264
Male	ECA	0.49	0.02	1.49	1.22	1.49	1.22	1068	874
	FST	0.49	0.01	0.71	0.84	0.71	0.84	958	1141
	GTN	0.49	0.02	1.41	1.19	1.41	1.19	1020	860

	KZN	0.49	0.01	1.21	1.10	1.21	1.10	1492	1354
	LMP	0.50	0.01	0.50	0.70	0.50	0.70	917	1303
	MPU	0.52	0.01	0.65	0.81	0.66	0.81	869	1074
	NCA	0.51	0.02	1.00	1.00	1.00	1.00	926	924
	NWP	0.49	0.02	0.84	0.92	0.84	0.92	914	997
	WCA	0.48	0.01	0.66	0.81	0.66	0.81	907	1114
≥ 1yr Preschool	ECA	0.67	0.03	5.59	2.36	5.63	2.37	1068	452
	FST	0.71	0.03	3.80	1.95	3.80	1.95	958	491
	GTN	0.81	0.02	3.80	1.95	3.78	1.95	1020	524
	KZN	0.61	0.02	3.76	1.94	3.78	1.95	1492	769
	LMP	0.66	0.04	7.03	2.65	6.94	2.63	917	346
	MPU	0.68	0.03	3.98	2.00	4.04	2.01	869	436
	NCA	0.62	0.03	4.69	2.17	4.68	2.16	926	427
	NWP	0.68	0.03	4.34	2.08	4.30	2.07	914	439
	WCA	0.75	0.03	3.06	1.75	3.14	1.77	907	518
Sometimes spoke English at home	ECA	0.63	0.03	4.74	2.18	4.83	2.20	1068	491
	FST	0.68	0.02	1.60	1.27	1.62	1.27	958	757
	GTN	0.64	0.03	4.95	2.22	4.90	2.21	1020	458
	KZN	0.54	0.04	7.35	2.71	7.39	2.72	1492	550
	LMP	0.55	0.05	7.71	2.78	7.75	2.78	917	330
	MPU	0.71	0.03	3.17	1.78	3.18	1.78	869	488
	NCA	0.59	0.03	3.17	1.78	3.15	1.77	926	520
	NWP	0.73	0.03	3.93	1.98	3.87	1.97	914	461
	WCA	0.58	0.04	4.78	2.19	4.79	2.19	907	415

Stellenbosch University http://scholar.sun.ac.za

Appendix G

Teacher Content Knowledge: An overview of the South African literature

While teacher content knowledge is only one of the many variables assessed in this thesis, the complete lack of nationally representative data on this variable, and the fact that SACMEQ III was the first time such data became available warrants a brief discussion of the literature on this topic in South Africa. This is included below.

Few would contest the assertion that teacher content knowledge is important for effective teaching. Taylor (2008, p. 24) states the obvious, but important, reality: "teachers cannot teach what they do not know." There is unequivocal consensus in the research community that teacher content knowledge in the majority of South African schools is unacceptably low. Some studies which address the topic of teacher content knowledge in South Africa are summarised below:

1. A 2008 study conducted by the HSRC in collaboration with Stanford University found that of the 49 Grade 6 teachers in Gauteng that were tested, the average teacher scored only 60% on both the mathematical content knowledge and pedagogical content knowledge parts of the test (Carnoy *et al.*, 2008). This is particularly disturbing given that the test was at a *Grade 5* level.
2. The Khanyisa Baseline Project conducted in 2004, assessed Grade 3 teachers in 24 schools testing them on Grade 6 mathematics and literacy items. The study found that the average score for the 23 teachers (writing a Grade 6 test) was 55% for literacy and 67% for maths (Taylor & Moyana, 2005)
3. Based on Stols *et al*'s (2007) analysis of 27 secondary school teachers, Taylor (2008, p. 12) questions whether the majority of South African high school teachers would be able to pass the Senior Certificate, the very exam their *learners* are expected to pass. He recommends that this hypothesis be tested on a larger more representative sample in order to draw definitive conclusions regarding teacher content knowledge, or lack thereof.

Fleisch (2008, p. 123) identifies three more studies highlighting the dearth of teacher content knowledge:

4. Bertram (2006) who found that a substantial portion of teachers enrolled in an honours level programme are not proficient or fluent readers.
5. Webb *et al.* (1998) who found that teachers did not score better than their own learners on a Grade 6 test on an electricity topic.
6. Van der Sandt and Niewoult (2003) who found weak teacher content knowledge in former Model C schools.

In addition to the above studies, the Department of Education itself has also identified the problem of inadequate teacher competencies: "The fact that learning outcomes among children and youth

Stellenbosch University http://scholar.sun.ac.za

remains low, while the percentage of qualified educators has increased, raises questions about the value of these qualifications as a measure of the competency of teachers" (DBE, 2011a, p. 55).

The quality of an education system cannot exceed the quality of its teachers, and the only way to improve outcomes is to improve instruction (Barber & Mourshed, 2007; in Taylor, 2008: 11). The findings of this thesis seem to suggest that teacher content knowledge may be only one of a variety of factors that determine teacher quality.

Stellenbosch University http://scholar.sun.ac.za

- END -