Do education inputs influence education outcomes differently for females and males in primary school in Malawi? A production function analysis.

Anna Downs

Research report submitted in partial fulfilment of the requirement for the degree of Masters in Education (Policy, Planning and Management)

In the

School of Education

University of the Witwatersrand

Supervisors: Helen Perry

February 2014

Declaration

I, Anna Downs, do declare that this dissertation is my own unaided work. It is being submitted for the degree of Masters of Education (Policy Planning and Management) at the University of Witwatersrand, in Johannesburg. It has not been submitted for any degree or examination at any other university.

Anna Downs Date

Copyright Notice

Acknowledgements

I would like to express my sincere appreciation and gratitude to the following people: my supervisor Helen Perry for her support and advice; Martin Gustafsson for his advice and constructive comments; SACMEQ for providing the data and information on the data; the Joint Education Trust and in particular to Double-Hugh Marera; Ed Downs and Maggi Jones for proofreading; and finally, to Simon Gates for his help formatting the report as well as his patience, understanding, and support.

Dedication

This research is dedicated to Malawi's girls who are disadvantaged both in the education system and in society. May the education system and society allow you to learn and live without discrimination or fear of violence. May the education system and society allow you to reach your full potential, be listened to and be valued. May you strive to be happy.

Abstract

Despite the challenges for females' education in Malawi, there is limited quantitative research which suggests what needs to be done in order to improve gender equality in primary education in Malawi. This report contributes to the literature on education inputs' effectiveness for female pupils at the primary school level and can be used to see if the results back up, and give weight to, some of the qualitative studies that have tried to answer this question. The analysis is unique in two main respects: firstly, the education production function analysis divides the sample into separate groupings of female and male pupils, and then compares the results of the separate regressions; secondly, it is an education production function analysis that focuses solely on Malawi using the SACMEQ III data.

The analysis shows that females in Malawi face serious educational disadvantages even after observable education inputs have been controlled for. This points to other unobserved factors such as cultural attitudes, household preferences, or whether a female is: married, due to be married, a mother or pregnant.

The analysis finds that the variables that are significantly correlated with the academic achievement of both females and males are: class size; hours of teaching; school resources; problems with the community; homework given; community paying bonus to school staff; years of teacher training; head teachers' number of years' experience; pupil/teacher ratio; and access to a computer. In addition to the factors that were found to correlate with male and female pupil test scores similarly, there were several factors that were associated with female pupils' but not with male pupils' test scores. The analysis found that female pupils at Grade 6 tend to do worse academically if they: are older than their peers; have more household chores; and have further to travel to school. There was some weaker evidence that a safe learning environment was associated with female pupils' academic achievement.

Key Words: Production function, Malawi, SACMEQ, primary school, education outcomes, education inputs, gender, female, age, economics of education.

Table of Contents

List of Figures

List of Tables

Acronyms

CSR Country Status Report

DfID Department for International Development

DHS Demographic Health Survey

EFA Education for All

EMIS Education Management Information System

EPDC Education Policy and Data Centre

GPI Gender Parity Index

MDGs Millennium Development Goals

MICS Multiple Indicator Cluster Survey

MOEST Ministry of Education, Science and Technology

MPRSP Malawi Poverty Reduction Strategic Paper

NER Net Enrolment Rate

NSO National Statistics Office

OECD Organisation for Economic Co-operation and Development

OLS Ordinary Least Squares

PIF Policy and Investment Framework

SACMEQ Southern and Eastern African Consortium for Monitoring Education Quality

SES Socio-economic Status

SWAp Sector Wide Approach

UNDP United Nations Development Programme

UNESCO United Nations Educational, Scientific and Cultural Organization

Terminology

Where possible the term ‘gender’ has been used is this report rather than the term ‘sex’. The term ‘sex’ has been used when the discussion refers to a particular study that uses that term. The term ‘gender’ is preferred over the term ‘sex’ because, rather than just referring to biological differences, it refers to the socially constructed roles and behaviours, which are important in this study.

When referring to pupils, the terms females and girls have been used interchangeably, as have the terms males and boys. As some of the pupils in Grade 6 can be as old as 25 years, the terms female pupil and male pupil are preferred. However, as most studies commonly refer to girls and boys, these terms are also used throughout.

Details on Computer Software and Dataset

Permission to use the SACMEQ III dataset was obtained from the UNESCO and IIEP SACMEQ Co-ordinating Centre in Paris (contact person there was Jan Maarse). Support from the office is hereby acknowledged and appreciated.

Two software packages were used to analyse the SACMEQ III data set: Stata 9 and Excel 2010.

1 Introduction

Using the Southern and Eastern African Consortium for Monitoring Education Quality (SACMEQ) data, this research report *aims* to identify which education inputs are correlated with education outcomes in Malawi, and to identify whether any of these education inputs have different relationships with male and female pupils education outcomes. This research report will aim to answer the following two research questions:

- Research question 1: Which education inputs have a significant and important relationship with education outcomes in Malawi? Is the gender variable one of these significant and important variables?

- Research question 2: If two sets of production functions are run, one set for males and one set for females in Malawi, how do education inputs vary in significance and importance? I.e. how do education inputs correlate with males' and females' test scores differently?

As will be shown in the literature review, and the analysis in this report, the SACMEQ III data shows that females' education attainment in Malawi at Grade 6 is lower than males'. Therefore, for research question one, there are two possible findings which, consequently, have differing policy implications:

- Scenario 1: After controlling for education inputs, the variable for gender is shown to be insignificant, i.e. the reason why females do not perform as well as males in primary school in Malawi is because they are disadvantaged in terms of the education inputs which are correlated with academic achievement. For example, females may tend to carry out more household chores than males, or have fewer meals than males. The policy implication of this finding is that if female pupils' educational attainment is to improve then the differences in the education inputs of males and females needs to be addressed.

- Scenario 2: After controlling for education inputs, the variable for gender is shown to be significant, suggesting that there are missing variables in the model that are associated with poor academic achievement for females. The policy implications of

this are that if female pupils' educational achievement is to improve, these missing variables need to be identified and addressed.

For research question two there are two possible findings, again with different policy implications:

- Scenario 1: Education inputs have the same or similar relationship with test scores of male and female pupils. If this is the case, then this finding does not provide any extra information on how to improve the situation for female students.

- Scenario 2: Education inputs are correlated with male and female pupils' test scores differently. For example, the likelihood of abuse at a school may have a stronger relationship with female pupils' test scores more than males'. The policy implication of this is that an extra way to improve females' educational achievement is to focus on improving the education inputs that have the strongest relationship with females' academic achievement.

1.1 Rationale

Achieving gender equality in education is viewed as a key priority internationally. The Dakar Final Framework for Action in the Education for All (EFA) and the Millennium Development Goals (MDGs) both stress the importance of gender equality in education:

• Dakar Final Framework of Action, number 5**:** 'Eliminating gender disparities in primary and secondary education by 2005, and achieving gender equality in education by 2015,with a focus on ensuring girls' full and equal access to and achievement in basic education of good quality'(UNESCO, 2000).

• Millennium Development Goal, number 3:'Promote Gender Equality and Empower Women: Eliminate gender disparity in primary and secondary education preferably by 2005, and at all levels by 2015' (United Nations, 2006).

These goals are very important. As the Global Campaign for Education (GCE) (2003) report stresses: 'education not only provides basic knowledge and skills to improve health and livelihoods, but it empowers women to take their rightful place in society and the

development process'. Education is a key part of breaking the poverty cycle in developing countries. However, although these goals exist, in countries such as Malawi, gender equality in education is unlikely to be achieved by 2015, unless gender equality in education is measured solely on overall primary school enrolment figures.

Unlike many African countries in which the gender gap has decreased significantly, with females achieving more than males in some countries (Hungi 2011), in Malawi there are still significant inequalities in test scores. In addition to the SACMEQ test scores showing inequalities, the Education Management Information System (EMIS) data (2009) shows that after Grade 6, female pupils' drop-out rate was much higher than male pupils' (for example, in Grade 7 the drop-out rate for males was 10 per cent, whereas for females it was 18 per cent). As drop-out tends to correlate to performance[1], finding out what factors are associated with females' performance also provides useful information if female drop-out rates are to be reduced in later grades.

Despite the challenges for females' education in Malawi, there is limited quantitative research which suggests what needs to be done in order to achieve gender equality in primary education in Malawi. This research report aims to contribute to quantitative evidence on the individual, family and school factors that are correlated with education outcomes for females, and to find out if the results back up, and give weight to some of the qualitative studies that have tried to answer this question.

There are a few studies that carry out education production functions analysis on SACMEQ data and include some commentary on Malawi (Hungi, (2011), Lee & Zuze (2011), Lee, Zuze & Ross (2005); however, these do not run regressions for female and male samples separately. There have also been some studies carried out that look specifically at Malawi using SACMEQ data. These have looked at progress in gender equality, focusing on changes in enrolment and achievement (Milner et al 2011b), Learner and Teacher knowledge about HIV and AIDS in Malawi (Milner et al 2011a), and Quality of Primary School Inputs in Malawi (Milner et al 2011c). Although their findings are useful, and the literature review looks at these studies in more detail, they do not consider how education inputs are related to

[1] 'The better the scores are, the more likely it is that learners will stay within the schooling system right up to the exit point (Hanushek and Luque 2003:483, quoted in Gustafsson 2006:46).

achievement, or how inputs may result in differences in achievement between females and males.

In summary this research report will fill a research gap and produce useful findings to help inform future policy.

2 Literature Review

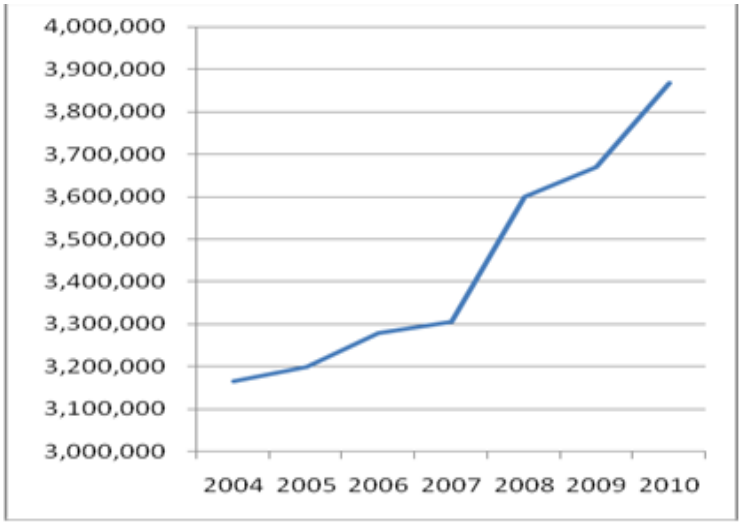
Three types of literature have been looked at. Firstly, the report examined literature and data that gives a background picture of Malawi's education system. Secondly, it examined studies that used education production functions. Thirdly, it examined studies that assessed why females do not do as well in education as males in Malawi and other developing countries.

2.1 Background

Since 1990, there have been significant achievements in education in the developing world, such as the increase in enrolment rates. However, many challenges still exist: for example, '…one in five girls of primary school going age is not at school in the developing world as a whole' (Gustafsson 2006). This section gives a brief summary of the education system in Malawi and then gives an overview of the SACMEQ results for Malawi.

Primary school education in Malawi was made compulsory in 1994 and, since free primary education was introduced, the number of students entering primary schools has increased enormously (see Figure 1). The influx of new pupils into schools is said to have put pressure on existing resources (Ministry of Education, Science and Technology (MOEST), (2001)).

Figure 1: Primary School Enrolment 2004-2010



Source: EMIS 2010

1

The Net Enrolment Ratio[2] (NER) given by EMIS (2010) is 100 which suggests that all children of primary school age attend school. However, it is likely that the enrolment figures are inflated[3] in EMIS and a more accurate estimation may be the NSO Population and Housing census (2010) which estimates the NER to be 75 per cent.

The EMIS (2010) data shows that between 2004 and 2010 the Gender Parity Index[4] (GPI) for enrolment in primary schools fluctuated between 0.99 and 1.03[5] in 2010, showing a slight movement from a small majority of males to a small majority of females in primary schools. However, the overall GPI masks the pattern of enrolment in different grades. As Figure 2 shows, GPI is higher than 1 in Grades 1 to 6 and lower than 1 for Grades 7 and 8, reaching a GPI of 0.89 in Grade 8 (EMIS 2010).

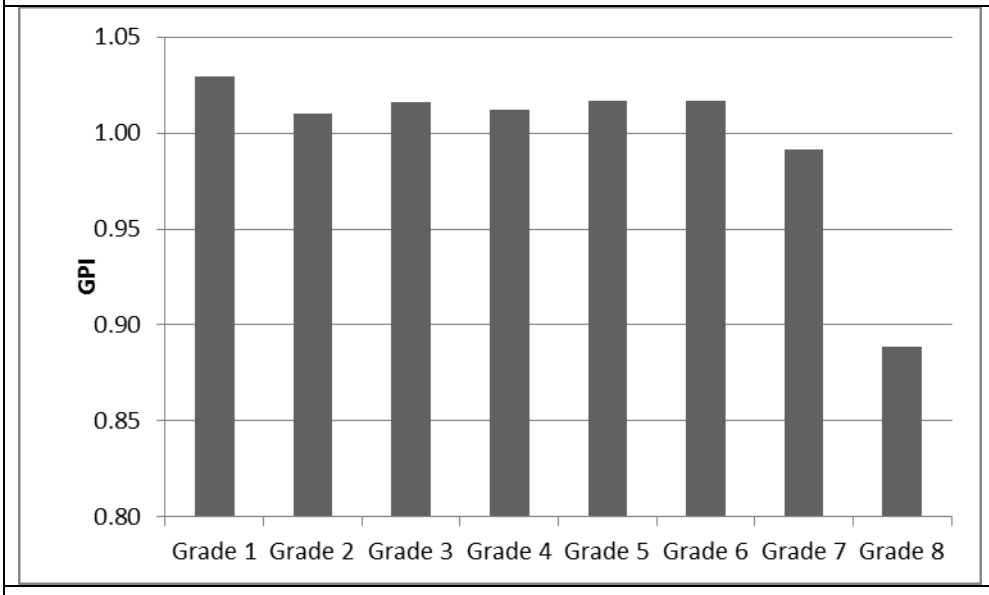
[2] The Net Enrolment Ratio (NER) is defined by the UNESCO Institute for Statistics as enrolment of the official age-group for a given level of education expressed as a percentage of the corresponding population.

[3] They may be inflated for two reasons. One: if schools believe the budget allocation they receive depends on the enrolment figures they provide, and two: EMIS 2010 uses an old census to base population estimates whilst the real figure is likely to be less than this

[4] The Institute for Statistics of UNESCO calculates GPI for any development indicator by dividing its value for females by its value for males.

[5] The GPI for net primary enrolment from other sources range from 1.12 (EPDC 2010) to a UNESCO 2009 statistic of 1.05.

Figure 2: Gender Parity Index for Female and Male Pupils in Primary School 2009/10



Source: EMIS 2010

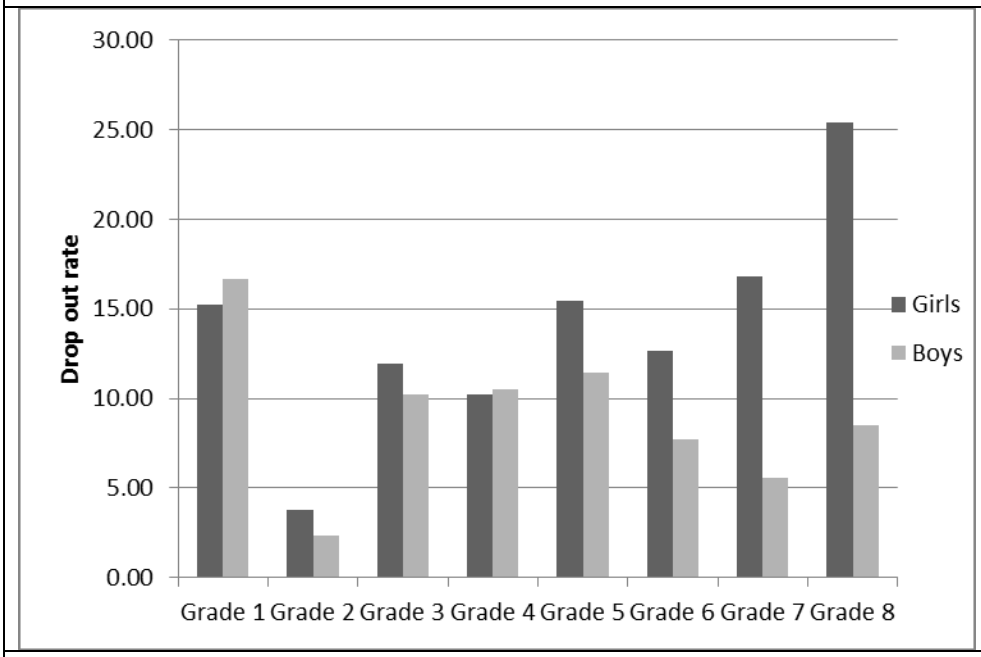
2

Note that GPI parity for enrolment does not necessarily translate into gender parity for other indicators, such as, drop-out, completion, and exam performance. Note also that females are 5.5 times more likely to be enrolled in non-formal education than males (Malawi Education Sector Implementation Plan (2009-13).

The drop-out rate for female pupils is higher than the drop-out rate for male pupils from Grade 5 upwards (see Figure 3). The female pupils' drop-out rate gets exponentially worse from Grade 6,[6] and by Grade 8, the drop-out rates is 8 per cent for males and 25 per cent for females (EMIS (2010)).

[6]EMIS (2010) data shows that by Grade 6 dropout rates were 8 per cent for boys and 13 per cent for girls; at Grade 7 they were 6 per cent for boys and 17 per cent for girls; finally, by Grade 8, the dropout rates were 8 per cent for boys and 25 per cent for girls.

Figure 3: Dropout Rate for Female and Male Pupils 2009/10



Source: EMIS 2010

[3]

As Table 1 shows, the most common reason for both female and male pupils to drop out of primary school in Malawi is family responsibilities. However, as the grades increase, female pupil drop-outs due to marriage and pregnancy increase. By Grade 6 marriage is the most common reason for female pupils to drop out of schooling and continues to be so in Grades 7 and 8. In Grade 6, 42 per cent of female pupils who drop out do so due to marriage or pregnancy; in Grade 7 this has increased to 54 per cent, and by Grade 8 it has increased to 63 per cent. As these are very important reasons for females dropping out, it is very likely they are also related to test scores. However, there is no SACMEQ data identifying whether a pupil is due to get married, is married, is pregnant or has a child. So unfortunately, these important issues could not be looked at in the variable analysis or regressions.

Table 1: Reasons for Pupils Dropping Out of School, 2009/10

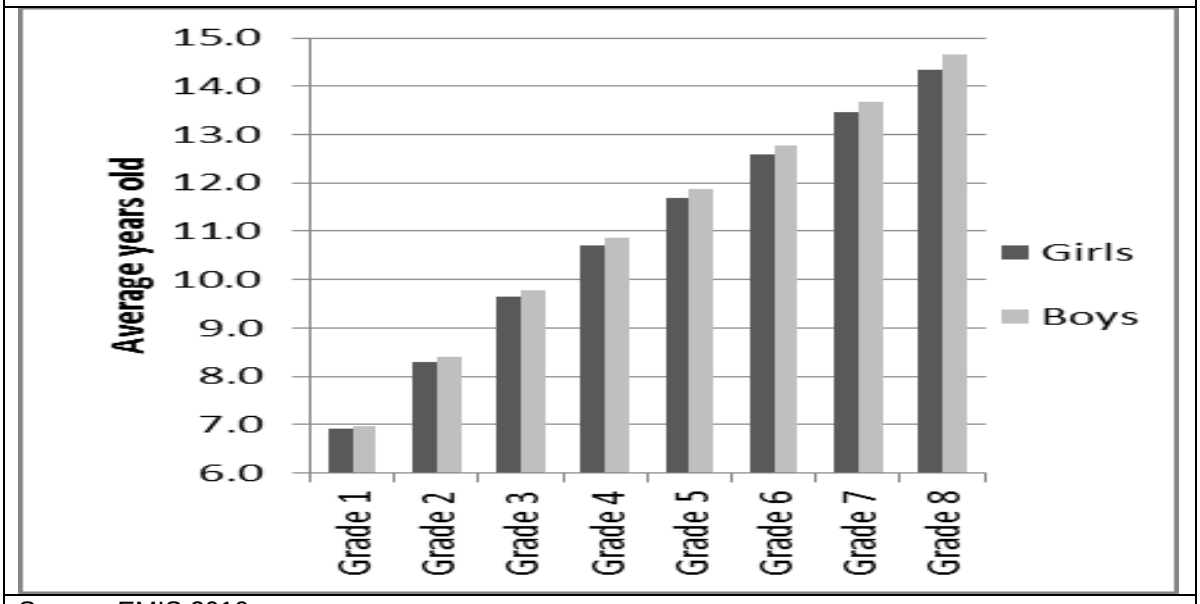
		Family Responsibilities	Pregnancy	Marriage	Fees	Employment	Sickness	Poor Facilities	Availability of Teachers	Long Distances	Violence	Other Reasons	Dropout rate
Grade 1	% Male	32%	0%	0%	1%	2%	4%	3%	2%	15%	0%	41%	16.66
	% Female	34%	0%	0%	1%	2%	4%	3%	2%	15%	0%	39%	15.25
Grade 2	% Male	35%	0%	0%	1%	4%	4%	3%	2%	10%	0%	40%	2.31
	% Female	37%	0%	0%	1%	3%	4%	3%	2%	11%	0%	38%	3.79
Grade 3	% Male	37%	0%	0%	1%	8%	3%	3%	2%	5%	1%	40%	10.23
	% Female	40%	0%	1%	1%	6%	3%	3%	3%	6%	1%	36%	11.97
Grade 4	% Male	37%	0%	1%	1%	10%	3%	3%	2%	4%	1%	37%	10.53
	% Female	38%	2%	6%	1%	8%	3%	3%	2%	4%	0%	33%	10.22
Grade 5	% Male	34%	0%	3%	2%	12%	4%	2%	2%	4%	0%	38%	11.42
	% Female	30%	7%	19%	1%	6%	2%	2%	2%	3%	0%	27%	15.48
Grade 6	% Male	32%	0%	5%	2%	12%	3%	3%	2%	3%	1%	37%	7.75
	% Female	24%	13%	29%	2%	5%	2%	2%	2%	2%	0%	21%	12.68
Grade 7	% Male	30%	1%	10%	2%	11%	2%	2%	2%	3%	1%	35%	5.58
	% Female	20%	18%	36%	1%	3%	1%	1%	1%	2%	0%	16%	16.82
Grade 8	% Male	31%	2%	17%	2%	11%	2%	2%	2%	2%	1%	30%	8.49
	% Female	16%	22%	41%	1%	2%	2%	2%	1%	1%	0%	13%	25.38
Total	% Male	34%	0%	2%	1%	6%	4%	3%	2%	9%	1%	39%	-
	% Female	32%	5%	10%	1%	4%	3%	3%	2%	8%	0%	32%	-

Source: EMIS, 2010

1

The UNFPA (2012) found that 50% of females are married by the time they are 18, and that one third of females aged 15-19 have begun childbearing (UNFPA (2013). There is no data in EMIS that says what age the drop-outs are, however, the EMIS (2010) data estimates that 13 per cent of Grade 6 females are of age 15 to 18. The data also shows that the average age of female pupils relative to male pupils in the same grade starts to decrease as the grade increases, see Figure 4. Female are consistently younger than males, and byGrade 8, the average age difference is just less than 4 months (0.3 of a year).

Figure 4: Difference in Age of Female and Male Pupils by Grades, 2009/10

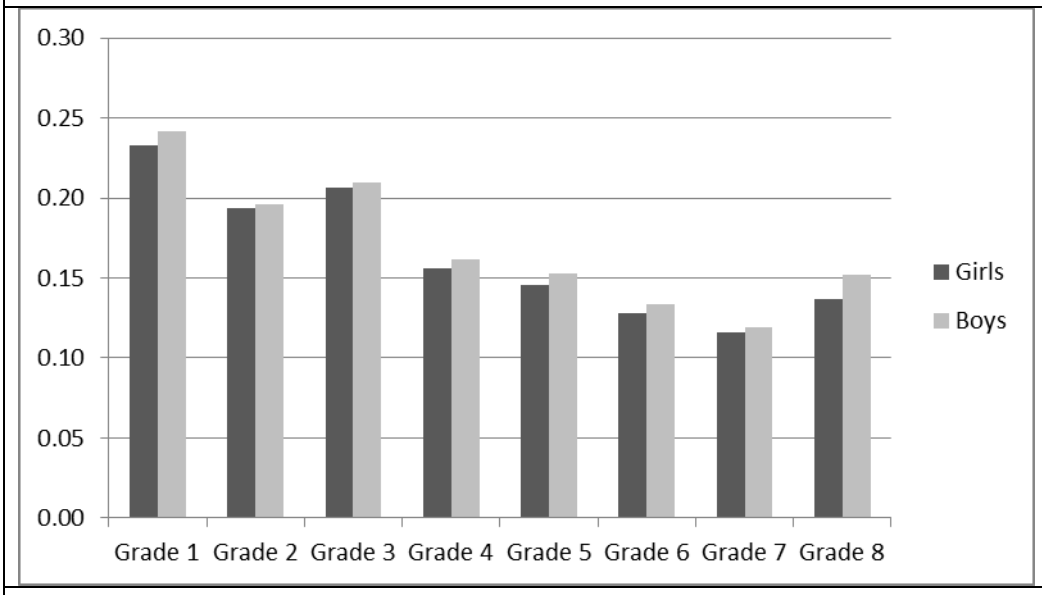


Source: EMIS 2010

[4]

However, the slight decrease in the relative average age of females in the later grades may be partially explained by the repetition statistics which show that female pupils repeat slightly less than males at every grade (see Figure 5).

Figure 5: Repetition Rate for Female and Male Pupils, 2009/10



Source: EMIS 2010

[5]

The World Bank (2010) Country Status Report (CSR) found that primary education repetition rates have increased to 20 per cent from 1999 -2006 resulting in Malawi having the highest repetition rates in the region. The World Bank (2010) CSR argues that reducing the repetition rate and high drop-out rates in primary education are two of the main challenges for Malawi.

The figures for primary to secondary transition rate range from 40 per cent to 33 per cent: the EMIS 2010 figure is 33 per cent, the DHS 2004 (2005) figure is 38 per cent, and the MICS 2006 figure is 40 per cent. This shows that the majority of primary school students do not go on to secondary school. The EMIS data (2010) shows that of those who do enrol in Form 1 of secondary school, the GPI is 0.87. Secondary education takes 4 years (Form 1 to Form 4) and requires students to have passed the primary school Grade 8 exams. They also have to pay fees.

Although students at secondary school and further education must pay fees, whereas primary education is free, it is still found that the government spend per pupil increases as the level of education increases. The World Bank (2010) estimates that, 'the 10 per cent most educated (those who study the longest) benefit from 73 per cent of the public resources in the education system. This makes the Malawian education system the most elitist system in Africa'. However, the Ministry of Education, Science and Technology's (MOEST) internal records show that the amount spent on primary education as a proportion of total education

spending has been increasing over the past 3 years up to a high of 55 per cent in 2010-11 and 56 per cent in 2011-12. One reason for the increase in allocation to primary education is the targets set by Development Partners who have recently started the Sector Wide Approach (SWAp) which states that, by 2011-12, 55 per cent of the education budget should be allocated to primary education. These targets follow Education for All (EFA) recommendations and aim to increase spend per pupil at primary school relative to the amount spent per pupil at higher levels of education.

In terms of the quality of education, the general consensus from the literature and the data on test scores, as shown in the next section, is that the quality in Malawi is low. For example, Banda (2003) stated that, high pupil/teacher ratios and high class sizes; poor infrastructure; high rate of wastage due to school drop-out before completion; rates of repetition and absenteeism; the absence of teaching and learning materials; pedagogical methods that encourage rote learning, as well as the use of untrained teachers, has lead to a low quality of education in Malawi. Banda (2003) also listed some reasons that cannot be measured by SACMEQ as leading to a low quality of education: poor teacher preparation, attitudes and motivation. A study by Milner et al (2011c) uses SACMEQ data to look at the inputs in the education system and finds that Malawi scored poorly in four key school input indicators: (a) basic learning materials, (b) mathematics textbooks, (c) pupil-teacher ratios, and (d) class size. Milner et al (2011c) suggests that these four indicators explain why Malawi does not score well compared to other countries. Where possible, the descriptive data section of the main report will look at these factors in more detail.

The next part of this background section will look at studies that analyse the test scores in SACMEQ by country, gender, SES, and urban or rural. Hungi et al (2010) show that for SACMEQ III Malawi achieved the lowest mean pupil reading score (433) of all countries[7] participating in SACMEQ. The mean score for all 15 countries was 512. For maths scores, Malawi had the third lowest score (447), with Swaziland, then Zambia, getting the lowest scores; the mean score for all countries was 510 (Hungi et al, 2010). This suggests that, compared to other SACMEQ countries, Malawi is a poor performer and provides a low quality of education for the majority.

[7] Malawi, Botswana, Kenya, Lesotho, Malawi, Mauritius, Mozambique, Namibia, Seychelles, South Africa, Swaziland, Tanzania (Mainland), Tanzania (Zanzibar), Uganda, Zambia, and Zimbabwe.

In industrialised countries and OECD countries female pupils tend to do better than male pupils in reading scores, and results for maths scores have become more mixed in recent years (OECD 2010, and Mullis 2008). In many African countries, females outperform males in some subjects. For example, Saito (2011) looked at SACMEQ scores for females and males and also found that, 'reading achievement was mostly in favour of females while the mathematics achievement was mostly in favour of males in 2007'. However, for Malawi the data does not follow this trend and male pupils significantly[8] outperform female pupils in both reading and maths. The only other country where females perform significantly worse than males in the reading test is Tanzania. Saito (2011) also looked at the number of students reaching level 4 for reading and maths. Saito (2011) found that, 'Malawi was the only country where statistical significance emerged with more boys reaching Level 4 for reading'.

As well as in SACMEQ tests, female pupils also perform worse in the national school leaving exams at both primary and secondary level. In 2010, only 62 per cent passed the Primary School Leaving Certificate of Education (PLSCE), 55 per cent passed the Junior Certificate of Examinations (JCE) and[9] 47 per cent passed the Malawi School Certificate of Examinations[10] (MSCE), compared to 74 per cent, 66 per cent and 58 per cent for male pupils (EMIS 2010).

These results show that females are significantly under performing in Malawi and confirm that gender is an important aspect to consider when carrying out education production function analysis for Malawi. The results also show that the gap between males' and females' results has widened since 2000 rather than narrowing, whilst the GPI for enrolment has improved (EMIS 2010). Saito (2010) finds that for most of the 15 SACMEQ countries there is a notable stability in gender differences. Saito (2010) concludes that, 'these results seem to suggest that there is a need to move the focus of the gender-related interventions beyond "access" and "participation" and concentrate more on "achievement", especially in less advantaged settings', such as Malawi.

Saito (2011) also looked at differences in test scores by urban or rural schools and found the unusual result that, for Malawi, the gender gap for reading was even larger at urban schools and higher SES groups, unlike in most other countries where the gender gap was worse in

[8] At the 95per cent confidence limit.
[9] The JCE is taken at the end of from two in secondary school
[10] The MCSE exam is taken in Form 4 at the end of secondary education.

rural schools and in lower SES groups. For maths, the gender gap in Malawi was also larger at urban schools and higher SES groups[11].

This background section so far has highlighted some of the challenges with Malawi's education system and the poor performance in education by Malawi compared to other South East African countries. Some of these challenges highlighted have included the gender inequality in the education system, for example, as shown by the significant gender gaps in educational achievement. Before moving on to look at production function studies, this section will briefly look at the gender related policies and interventions that are in place in Malawi. The Ministry of Education, Science and Technology's (MOEST) Policy and Investment Framework (PIF) for the period from 2000 to 2015 and the Malawi Poverty Reduction Strategic Paper (MPRSP) state that high priority would be given to the gender imbalance and inequity in the education system at all levels (Banda, 2003). The Ministry of Gender, Youth & Community, and the Gender Focal Point in Ministry of Education aim to:

- Promote gender mainstreaming throughout public sector
- Reduce repetition, absence, and drop-out of girls
- Implement gender-sensitive curriculum
- Increase enrolment of girls in science and technology
- Implement programmes on readmission of mothers, grants for orphans, child-friendly schools
- Increase number of mother groups
- Increase number of female teachers
- Increase number of girls' hostels at secondary schools.

However, although these policies exist, in practice gender does not receive the focus it should do. For example, the gender unit in MOEST receives the smallest budget allocation of any headquarters unit in the Ministry of Education (internal documents 2011). This unit has just two members of staff (2011), and is not integrated into the planning department. The UK Department for International Development (DfID) is planning to start a new girls' education programme in an attempt to improve the situation, whilst other Development Partners and MOEST have started to put together an overall strategy to improve female pupils'

[11]However, this time other countries also had larger gender gaps at urban schools and at higher SES groups (Saito 2011).

performance. However, it seems, at present, that the education sector in Malawi is guilty of what Global Campaign for Education (GCE) (2003) describes as having a, 'plethora of gender equity initiatives, which add up to a 'scatter-shot patchwork of 'girls' projects' rather than a comprehensive package of interventions backed by clear policy aims', and a budget to match.

2.2 Education Production Functions

This section will describe the rationale for using production functions, then briefly discuss the models used for some production functions, and then progress onto looking at the particular inputs that have been found important or not important in the literature.

The reason why this research report plans to carry out a production function analysis is because it allows analysis of which inputs have the most significant relationship with the education quality that children receive, i.e. which inputs produce the highest returns. One of the first production function studies that analysed the relationship between student outcomes, education inputs and family background was carried out by a sociologist called Coleman in 1996. The Coleman Report (1996) looked at 650,000 students in the United States and found that socio-economic status and pupil background had more of an effect on student outcomes than school resources. The Coleman report also found that differences in schools and teachers however, did have an effect on results. The findings of this report were considered highly controversial as at the time it was interpreted that it meant schools did not matter and family background was the only significant factor (Marshall 1998). Since this report, many other production functions have been carried out, many of which also found SES to be a key explanatory variable. For example, the PISA 2006 Survey found that, 'students' socio-economic differences accounted for a significant part of between school differences in some countries' (OECD (2007)).

In general, education production function studies of developed countries tend to find that marginal increases in inputs have only very small correlations with the outcomes; whereas studies for developing countries often find that small increases in inputs have larger increases in the outcomes. Boissiere (2004:p27) found that 'in contrast to the Coleman Report conclusion that schools are less important than SES in determining education outcomes, many

researchers conclude from the developing country evidence that schools matter much more in the setting of poor countries'. He argues that this is mainly due to the principle of diminishing returns, i.e. at low levels of inputs, having a bit extra makes more of a difference to education outcomes than having the same amount extra when you already have a high level of inputs. This is particularly true for inputs such as learning support materials (Gustafsson 2006:14). As Malawi is one of the world's lowest income countries (ranking 175 out 183 with $860 per capita PPP (IMF 2012)) and has low levels of most education inputs, then inputs are more likely to make a significant difference to education outputs, compared to a rich OECD country.

Gustafsson (2006: 12-13) says that production functions can answer some important policy questions:

1) Could output be improved with the current basket of inputs?
2) Could output be improved by changing the current mix of inputs within the budget?
3) On what should new education funds resulting from budgetary increases be spent?
4) What education inputs should be cut following budgetary cuts?
5) How should education inputs be rearranged following a change in the relative prices of inputs?
6) How much additional education funding is required by poor communities in order to achieve greater equality in the outputs?
7) What are education outputs likely to be in the future?

This research report will not answer all these questions, not least because it is beyond the scope of this report to look at costs of the inputs. However, the results of the production function will be useful for planners in MOEST in Malawi to enable them to start to answer, and address, these questions.

Gustafsson (2006:13) states that one of the many uses of education production functions is that they can be used to measure how much additional funding is required by poor communities in order to achieve greater equality in the outputs. He argues that production functions often find that higher SES is associated with higher learning outcomes, and therefore the results of production function can be used to inform 'what level of additional

school inputs are required to at least partly offset a learner's disadvantage'. If the analysis in this report finds, after controlling for education inputs, that being a female in Malawi still has a negative association with educational achievement, then if there was also evidence identifying likely causal mechanisms and levers not just association, the results of the production function could be used to calculate the extra education inputs required to offset this gender disadvantage. This could form the basis of a pro-female education financing approach.

It is important for the research report to choose the inputs carefully and make sure the ones in the model are the most important ones. This report selected the inputs that the literature review highlighted as being important in determining education achievement and used these to design the conceptual model.

This literature review will next look at some examples of how other researchers have grouped education inputs in their conceptual models. It will then look at particular inputs that have been found to be important or not important in the literature. After considering these inputs, this research reports' own conceptual model will be presented in the methodology section.

Gustafsson (2006:58) proposes a model which links the inputs to policy areas. He organises these inputs as:

- Educator inputs
- Curriculum inputs
- Learning support inputs
- Infrastructure inputs
- Management inputs
- Access promotion inputs
- Household inputs

Hungi (2011) looked at variables in terms of pupil level variables and school level variables. She split the pupil level variables into three categories:

- Individual characteristics (for example, gender and age)
- Personalized learning support (for example, preschool attendance, extra tuition, and homework help at home)
- Home environment (for example, pupil SES, number of siblings, and household tasks).

Hungi (2011) then split the school level categories into four areas:

- Teacher characteristics (for example, gender, education, and professional qualifications)
- Classroom environment (for example, class size and classroom resources).
- School head characteristics (for example, gender, education level, and experience)
- School environment (for example, school resources, type of school, pupils' behaviour problems, the contextual climate such as average pupil SES, and the proportion of female pupils in the school).

Hungi (2011) ran production functions for 15 African studies using the SACMEQ III data and found that at the pupil level, grade repetition, socio-economic background, pupil age, and pupil gender were found to be the most important factors affecting the variations in pupil achievement in these school systems; while at the school level, school resources and school location were identified as the important common factors. The table that follows summarises the variables that Hungi (2011) found significant for Malawi.

Table 2: Variables Correlated with Pupil Scores in Malawi	
For Reading Scores:	**For Maths Scores**
Pupil Level:	***Pupil Level:***
Pupil age	Pupil age
Pupil sex*	Pupil sex*
Days absent	Days absent
Learning culture (books at home)	Learning culture (books at home)
SES	Meals per week
Grade repetition	Living with parents/relatives
School Level:	***School Level:***
Class size*	Class size*
School head experience as head*	Teacher days absent
School resources*	Classroom resources*
School community problems*	SH advice teacher
Pupil-teacher ratio*	SH teaching hours per week
Pupil behaviour problems*	Speaking language of instruction
Proportion of female teachers*	
Teacher years of professional training*	
Notes: * means that the standardised regression coefficients were above 0.10. All results reported above were significant. Source: Hungi (2011)	

2

Table 2 shows, Hungi (2011) found that pupils' gender was a significant variable in education production functions and had a standardised regression coefficient higher than 0.10[12] for Malawi. Hungi (2011) also found that for Malawi, and the majority of other countries, younger pupils achieved better than their older counterparts in both reading and mathematics. Age is still significant when pupil repetition is in the model, so age is not simply a proxy for pupil repetition and may reflect other factors at the school level. For example, the report, 'Because I am Girl' (Plan International (2009:71)) states that 'The arrival of puberty, and the onset of adolescence, not only affects children's bodies; it is also the time when sex and gender differences become more sharply defined'.

As Table 2 shows, Hungi (2011) also found that class size was significant for Malawi whereas for most other countries it was not. This country difference is likely to be due to the large variance in class sizes in Malawi, and the existence of some very large (100+) classes.

[12] Hungi (2011) considers any standardised coefficient that is 0.10 or above to be 'important'.

In Boissiere's report (2004) he stated that there was no evidence that reducing pupil to teacher ratios below 40 pupils per teacher would have a significant effect on learning outcomes, however above this level PTRs can be detrimental to learning outcomes (Boissiere 2004). This is due to diminishing marginal returns discussed previously.

It is interesting to note that, as Table 2 shows, the pupil SES variable was not significant in the final mathematics model for Malawi; however, it was significant for most other countries (Hungi 2011:8). Hungi's (2011: 32) measure of SES included measures of home possessions, parental education, home quality and source of lighting at home. Finding that SES does not have a significant relationship with maths scores in Malawi is unusual. As previous literature has shown, SES is normally an important input to education outcomes, and it is often assumed that leaving SES out of a regression model would result in a significant proportion of variability in academic achievement being unexplained, i.e. a low R2.

For example Vinjevold and Crouch (2001) concluded that the reason for their low R2[13] in their production function models of Grade 3 students in South Africa was likely to be due to the fact that contextual factors such as poverty were not included in the model. Vinjevold and Crouch (2001) ran regressions to explain literacy and numeracy achievement using input measures such as pupil teacher ratio, the degree of training of the educators, their level of experience, the frequency of internal assessment, and the types of learning materials.

However, Perry (2002) followed up their study and added socio-economic status to the regression and found that, 'the variance in schools' overall assessment scores is hardly accounted for by school condition and socio-economic status'. Perry (2002) concluded that inputs such as educator and school manager know-how may be more important in determining learner achievement, however data is not available. Perry (2002) also noted that, 'if all schools in the country were measured (and not just schools in poorer socio-economic districts) it is likely that school inputs and socio-economic conditions might have more of an impact on learner performance'. This suggests that if most pupils in Malawi are of an equally poor social economic status and have equally low levels of other inputs, that the analysis may then find that there is insufficient variance in these factors to lead to statistically

[13]Their results show that, 'the variability in inputs is not strongly associated with variability in results', with 'only about 10 per cent of learning variation in numeracy, and about 7 per cent of learning variation in literacy' being explained by these inputs (Vinjevold and Crouch, 2001).

significant findings. As most pupils in Malawi have very low SES (as shown later in the descriptive data section), this may provide an explanation as to why SES was not found to be significant in Malawi. Note also that Hungi (2011) found in her study that Malawi had the smallest total variance in reading scores, and had the third smallest total variance in mathematics scores. The descriptive statistics in this research report will look in more detail at the distributions of test scores in Malawi and the inputs.

Nannyonjo (2007) analysed factors influencing the learning achievement of Grade 6 students in Uganda and found after running regressions that, 'if availability of inputs is accompanied by proper use and school processes that emphasize academic achievement, they do lead to better performance.' Nannyonjo (2007) found some pupil background characteristics had a positive influence on learning: parental education, number of books at home and language. However, the study also found some pupil background variables that had a negative influence on pupils' test scores: distance of school from pupil's home, family size and age. Nannyonjo (2007) also found that factors that contribute to longer learning time such as: pupil's punctuality, attendance, parental interest and presence of electricity or reliable lighting at home, had a positive influence on scores. The study also found funding per pupil, time spent on teaching a subject, and greater availability of textbooks to be positively correlated with pupil performance; whilst class size, pupils per bench and the number of pupils sharing a textbook were negatively correlated. Nannyonjo (2007) found that teacher experience and age also had a positive relationship with pupils' test scores, but that teacher qualifications, except for university education, had little influence on test scores in Uganda. Nannyonjo (2007) found that most of the coefficients for these significant variables were low and the overall R squared for the regression models explaining individual performance were 0.39 for literacy and 0.26 for maths. Nannyonjo (2007) suggested that the way schools are managed, classroom interaction, and use of school resources may be more strongly related to pupils' performance.

Gustafsson (2006:55) also argued that an important input in education outcomes is school management. He said that this is often excluded in production functions because it is hard to model and hard to measure, with little useful data existing. As SACMEQ does not contain data on the way schools are managed, it was expected that the R2 in this study would also be low. Gustafsson (2006) found in his regressions for South Africa using SACMEQ II data that the following inputs were also important: SES, repetition, classroom methodology, teacher

evaluation, in-service training of teachers, physical infrastructure of schools, education of the parents, access to textbooks, and whether learners had regular meals. Gustafsson (2006:53) argued that, 'two inputs that frequently appear in production function studies and are hardly ever rejected in the models focusing just on developing countries are: firstly, the quality of the human capital embodied in the teachers, and secondly, the time spent learning and teaching.'

As part of his study, 'The Failure of Input-based Schooling Policies', Hanushek (2002) presents meta-results from developing country production functions. He looked at the significance of certain variables that are tested in various studies: pupil teacher ratio, teacher education, teacher experience, teacher salary, expenditure per pupil, and facilities. Hanushek concluded that the meta-analysis gives very weak evidence that resource inputs have a significant effect on student outcomes[14]. However, Gustafsson (2006) looked at Hanushek's same study and rejected this conclusion[15], as does UNESCO (2005) who used Hanushek's 2002 meta-analysis results to discuss which resources do in fact make a difference in their EFA Global monitoring report. This UNESCO (2005) report also found that time spent by learners on learning activities and pupils' attendance at pre-school, to be important in influencing education outcomes. In a later paper, Hanushek summarised the findings on education production functions again and this time concluded that schools, money and resources may impact achievement but that 'there currently is no clear, systematic relationship between resources and student outcomes' (Hanushek 2007). He did however find teacher quality to be very important, but unfortunately, it is not clear which measurable attributes relate to teacher quality (Hanushek 2007).

[14]However, he does state that, 'the data hint that the importance of resources may vary with the level of resources' (Hanushek 2002) as Boissiere (2004) also argues.

[15]Gustafsson (2006) argues that even though some of the 96 studies Hanushek looked at have some insignificant results, if all the inputs were truly insignificant then a more equal mix of positive and negative coefficients would be seen. It should also be noted that there are a lot of factors specific to each study that Hanushek included in the meta analysis, that are neglected in this type of meta-analysis. Hanushek's meta-analysis also does not take into account other factors such as the methodology used in the study.

2.3 Explanations for Gender Differences in Education

This section summarises three qualitative studies and then two quantitative studies that offer explanations of why females underachieve in the education system. It also draws evidence from other studies but does not go into these studies in any depth.

Colclough et al (2000) looked at the relationship between poverty, schooling and gender inequality, and concluded that the gendered outcomes of under-enrolment are the product of cultural practice, rather than of poverty per se'. In their paper, Colclough et al (2000) mentioned the following situations that may give reasons why females' enrolment may be lower than males' in developing countries; however, there was no specific focus on Malawi so not all the findings may be relevant. The reasons they gave are summarised below:

Societal Reasons:

- Main leadership roles in local and national life are occupied by males
- Marriage of females occurs at a younger age than of males
- Religious or customary belief discourages social interaction between the sexes
- Conventional opinion encourages women to see their future as being centred on the home and family

School Environment Reasons:

- Male teachers may not provide sufficient support to female pupils
- Male teachers may be sexually threatening
- Toilet facilities for female pupils may be inadequate
- Other facilities such as a shortage or absence of chairs may be unfriendly to female pupils
- Harassment from male pupils at school may occur
- The journey to school may have greater risks for the safety of females than for males

Household reasons:

- Gendered division of labour can affect attendance of female pupils at school, in particular if females are expected to look after siblings and perform many household chores

- Where a female's allegiance after marriage is mainly to her future husband's family, then parents are likely to favour the education of sons over daughters.

- If schooling decisions are taken by men, then education of males may have greater preference.

Focusing on 9 countries including Malawi, the Global Campaign for Education (GCE) (2003) study looked at three areas of gender inequality in countries around the world. The areas of gender inequality it looked at were: enrolment, school level completed, and achievement. After drawing on various qualitative case studies, they concluded that the following factors can contribute to females' under enrolment, dropout and underachievement at school:

- Son preference.
- Early marriage.
- School is too expensive: even where fees do not exist there are still costs of school materials and donations required by school, etc.
- Female have a lot of chores at home.
- Government schools are too few and too far – can cause concern for abuse and sexual harassment on the way to school.
- Schools fail to motivate or encourage female pupils – in the study they state that not having enough female teachers as role models can contribute to this failure as can curriculum that are biased against females.
- Schools fail to protect the basic rights and dignity of female pupils – for example abuse from teachers and other pupils, bad sanitary facilities for females.

(GCE (2003))

For Malawi, some of the recommendations they give on how to improve gender equality in education are to: increase teaching and learning materials; provide more adult literacy and out of school youth programmes; strengthen Community Based Childcare Centres (CBCC) to

reduce pressure on females to look after siblings; increase the number of female teachers; reduce sexual harassment and early pregnancies; encourage females back to school after pregnancy or marriage; implement mechanisms to prevent transmission of HIV AIDS and mitigate impacts; increase support for orphans; change the curriculum so it is not biased against the poor and females (GCE (2003)).

Banda's (2003) paper on gender sensitive policy and practice looked at Malawi only. It draws on various studies[16] and gives a variety of factors which disadvantage females in education and in particular can cause them to drop out of school:

- Socio-economic factors include: family poverty; direct costs of the school, including cost of school materials and levies, and opportunity costs of schooling.
- Socio-cultural factors include: pregnancies and early marriages; initiation practices; parental attitudes and aspirations for children; household chores; puberty-related issues; death in the family; caring for sick parents or relatives.
- School-related factors include: distance to school; attitudes and perceptions of teachers; lack of female teachers to act as role models; academic performance; resource allocation within schools; curriculum length; quality of teaching time.

In addition to the reasons given above, Banda (2003) also referred to research that has shown that most schools in Malawi do not always guarantee the security of females, and that there were reports that in some schools females were subjected to different types of abuse and harassment from male students and male teachers. Banda (2003) also commented on the unequal representation of women in the labour markets and suggested that this may be one reason why females are not motivated to do well at school.

The explanations given by these three studies for why females may not achieve in education are very similar. Where possible, the factors suggested as important in determining the educational achievements of females were modelled in this report's analysis. However, data for many of the variables was unavailable.

[16] Davison and Kanyuka (1990), Kadzamira and Chibwana (1999), Maluwa-Banda and Kholowa (2002), Burchfield and Kadzamira (1996), Maluwa-Banda and Lunguzi (2002), Chimombo and Chonzi (2000), Mbilizi, Kanyongolo, Nzimpita, (2000).

The literature review will now look at two quantitative studies that offer explanations for gender differences. One quantitative analysis report is by Hungi (2010) in which SACMEQ data was used in a descriptive analysis on the number of children who had a female head teacher at primary school. Hungi (2010) found that in all countries (apart from the Seychelles and Lesotho) the percentage of head teachers who were female was less than the percentage of female primary teachers, implying that there was considerable bias in the allocation of school managerial positions in favour of males. Hungi (2010) stated that one of the implications of these biases is that 'these gender inequities could send the wrong signal to pupils (both boys and girls) - that female teachers are incapable of being leaders'. It is noted that although Hungi (2010) used the SACMEQ data to show the differences in proportions of male and female teachers and head teachers, she did not carry out research to see if these differences affected females' education outcomes.

Milner et al (2011b) also conducted a descriptive analysis which looked at gender difference in enrolment and achievement at Standard 6 using the SACMEQ data for Malawi and also looked at factors that they thought could affect gender equality in education. They found, as did Saito (2011), that although enrolment differences in general were minimal, females did significantly worse than males in the SACMEQ tests. After looking at the education inputs that they thought may explain why females were under-achieving, they concluded that, 'more efforts and new interventions need to be implemented to address the challenges, especially regarding construction of toilets, promotion of girls' achievement, and the allocation of female teaching staff and head teachers' (Milner et al 2011b). For example, they found the gender balance of staff unequal, and, in particular, that of head teachers, as well as finding a lack of safety and sanitary facilities at the schools.

The last two studies discussed were descriptive data studies which did not carry out analysis to see if the education inputs looked at were found to be statistically correlated with the test scores of females. This research report looks at these education inputs, where possible, to see if the results of the data analysis can back up the assertions of these studies and the other qualitative studies.

All of these studies have found at least one of the following factors to be important in explaining female underachievement in education: early marriage and pregnancy,

prioritisation of males' education, household poverty and family responsibilities, risk of violence and bullying, lack of female role models, biased curriculum and learning environment. Many other studies also support these findings in other developing countries which may offer some insights for Malawi. Some of these are mentioned briefly in the following paragraphs before moving on to the methodology section.

Plan International (2009) says that 'for an adolescent girl, entrance into puberty and the possibility of getting pregnant have one of the most significant impacts on her education'. As mentioned previously the UNFPA (2012) found that 50% of females are married by the time they are 18 and that one third of females aged 15-19 have begun childbearing (UNFPA (2013)). As seen in Table 1, in 2010 almost 9% of female pupils who dropped out of school did so because of early marriage and pregnancy, thus reducing the small number of females who make it into upper primary.

Plan International's (2009) report states various reasons why families can choose to prioritise sons' education over daughters' education. One of the reasons that parents may choose to withdraw daughters before they reach the end of primary school is if they are '… convinced that lack of post-primary places and the lack of employment prospects for girls do not justify the cost of keeping them in school'. The study also says that it may also be the case that if females are expected to take the role of mothers, wives and carers, then education can be perceived as unnecessary for these roles. This study also states that another reason for male education to be seen to give higher returns to a family than female education is in situations where a family know that '…their son will support them in their old age, whereas a daughter is likely to be living with her husband's family'. Empirical evidence for prioritisation of male over female education was found in a study in Peru which found that parents are more willing to pay for reduced travel time to secondary school for males than for females (Gertler & Glewwe 1992).

With the demands of household chores, family care and few employment opportunities in rural areas, parents must weigh the opportunity cost, which is significant for poor households, of keeping children in school. Gender and poverty combine to put females from poor households at a huge educational disadvantage (Oxaal (1997)). As Table 1 shows, 33% of females who dropped out in 2010 did so due to family responsibilities.

Plan International (2009:71) states that ‘Parental concerns about girls’ physical safety often forces girls to leave school before they have been able to acquire even minimal levels of learning’. Female can be vulnerable to violence and abuse on the way to school and at school; where the risk of this happening is high, families may keep females at home to protect them. According to one global report (Action Aid, (2010)) 60 million females were sexually assaulted at, or on their way to school every year. A study on Malawi showed that almost a quarter of children surveyed reported being forced to have sex against their will and 99 per cent of children reported being bullied (Burton (2005)).

The curriculum and environment has also been found to negatively affect the performance of female pupils if it enforces the message that females are inferior to males. For example, Plan International (2009) report states that Lewis and Lockheed (2006) found that ‘In Yemen researchers noticed that primary school girls were often seated at the back of a class, making it harder to participate’ which negatively affected their performance. Some studies have shown that the assignment of students to a same gender teacher significantly improves the achievement of both female and male pupils (Dee (2007)).

To put gender inequality in education in Malawi into context, it is also worthwhile mentioning that the UNDP gender equality index for Malawi was 0.374 in 2009 which makes it one of the lowest in the world (UNFPA (2013)). The UNFPA (2013) notes that ‘women and adolescent girls continue to meet unsurpassed challenges and limited opportunities which increase their vulnerability to reproductive health challenges and gender based violence’.

The next section of this report looks at the research methods used in this study.

3 Research Methods

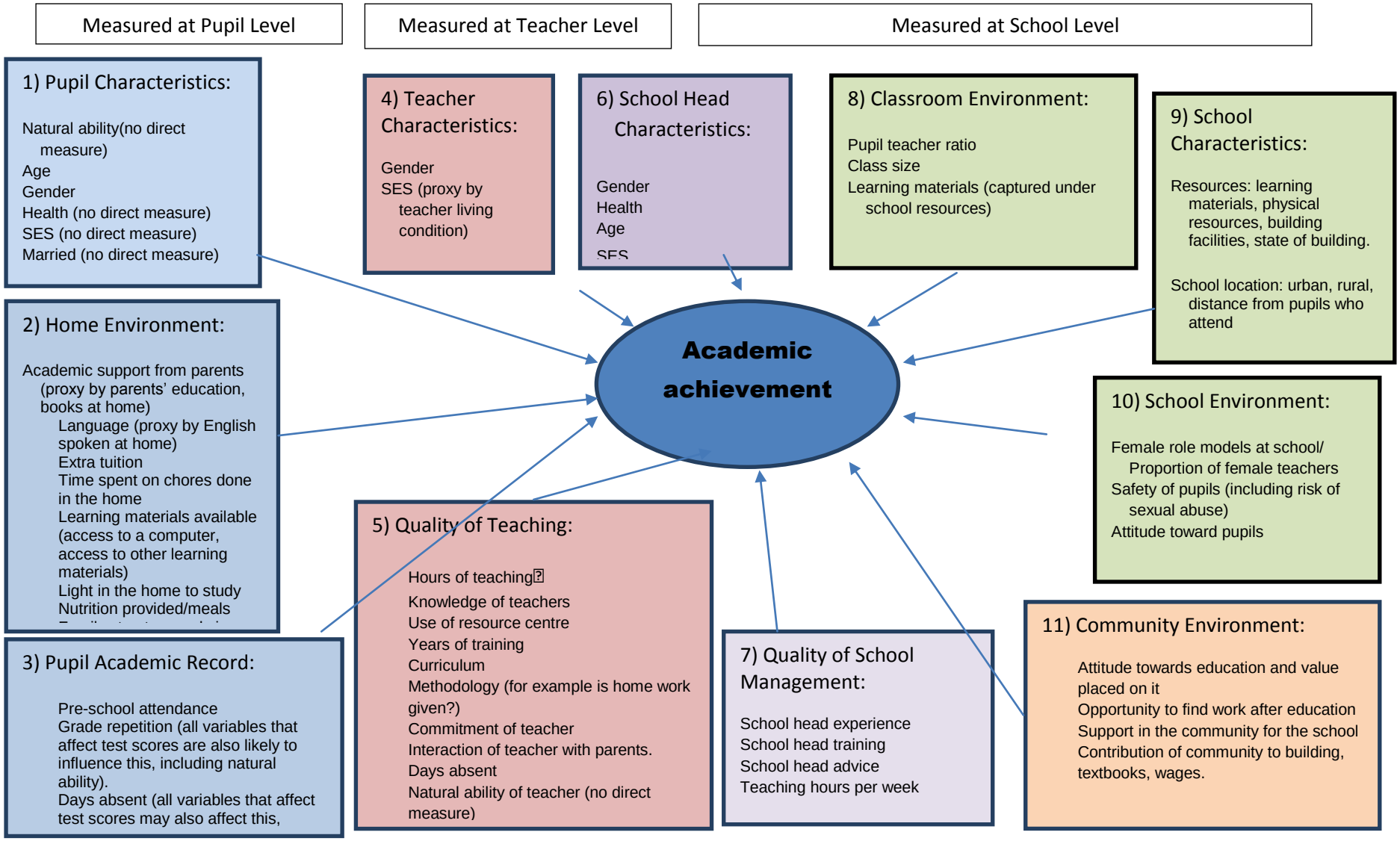
This research was carried out by, firstly, producing a conceptual model of the factors that are associated with educational attainment in Malawi; secondly, producing descriptive statistics and running basic analysis on the variables in the conceptual model; thirdly, carrying out regression analysis on the set of variables.

This section goes through each of the above steps, discussing any issues arising, and explaining decisions for methodology where appropriate.

3.1 Conceptual model

The variables that the literature review highlighted as being factors that are associated with test scores were captured and organised in a conceptual model. This conceptual model provided the theoretical basis of the analysis. The inputs were categorised into eleven fields, and the relationship between these inputs and educational attainment was examined (see Figure 6: Conceptual model). The mechanism through which these educational inputs (taken from the literature reviews) affect educational achievement has not been discussed due to the scope of the report. Therefore it is recommended throughout the discussion of the results that further investigation should be carried out to look at the mechanisms that explain the links between variables and test scores.

Figure 6: Conceptual Model



6

3.2 Descriptive Data and Basic Analysis of Variables

Each variable in the conceptual model was matched up with a suitable variable from SACMEQ, or, where appropriate, a new variable was created using the SACMEQ data (see Table 3 and Appendix 1 for details). For some variables, it was noted that no suitable data was available. Some of the SACMEQ variables were combined to create one variable. For example, in SACMEQ III there are 31 variables that measure whether a pupil had a certain possession at home; rather than include all of these 31 variables in the analysis, it was more helpful to combine a number of them into one single variable which measured pupils' possessions at home. Other variables have been adjusted so that they are more appropriate for the Malawi context, for example the measure of SES.

Table 3 contains a summary, in table form, of how the variables used in the analysis were constructed and some of the variables that were constructed but then rejected due to low correlation. Further detail is also presented in Appendix 1 to explain the decision making processes and analysis carried out which led to the rejection of certain variables or to the creation of new ones. However, the report will now summarise how SES was constructed given that this is a variable that it usually very significant in analysis and given that an alternative measure of SES was constructed rather than use the SACMEQ measurement of SES.

Ideally, social economic status would be measured by family income. However as SACMEQ cannot collect information on family income, a proxy variable for SES has been constructed in the SACMEQ data from roof type, light type, parents' education and some other variables, including those on possessions. This is useful if results between countries are to be compared; however, as this study is just looking at Malawi, a measure of SES that is more appropriate for Malawi (a very economically poor country) has been created. The SES variable this study used only included measures of possessions and home type: floor type and roof type in the home, if the house had electricity, a bed, a fridge, a TV and an electric fan. Many of the 31 possessions that SACMEQ measures if they are present in a home were very rare to have in a pupil's home in Malawi and therefore were excluded. For example, less than 1 per cent of pupils had the following possessions in their homes: vacuum cleaner, washing machine, computer, and internet. Therefore a measure of SES that includes these possessions becomes less relevant for Malawi. This variable used to measure SES can take possible

scores of 0 to 12, where 0 means the pupil has none of the house possessions included in the measure and has the most basic floor and roof type. For Malawi the median score is 1.75 which means that pupils at the 50th percentile have no electricity, no fridge, no electric fan and no TV at their house. A Malawian pupil scoring an SES of 1.75 may however have either a bed, and earth floor and an unsealed roof (for example thatch) or no bed in their house and a cement floor and metal roof.

It was decided that parents' education together with books at home would be excluded from the main measure of SES. This is because these two variables will be used instead to act as a proxy for the educational support a pupil can get at home so that the contribution of this to test scores can be separately assessed. It was decided as well as keeping academic support separate, to also keep the type of light a house had as separate, as this variable seemed correlated to pupil scores even after SES was accounted for. More details on how these variables were constructed is in Appendix 1. Note that the light type in a pupil's house, academic support: (parents' education and books at home) and SES are very highly correlated and this needs to be considered when interpreting results.

Following the creation of variables using the SACMEQ III data (more details in Appendix 1), analysis of these variables was subsequently carried out. Analysis included calculating the variables' distribution (mean, median, range, percentiles, variance), and calculating the correlations between the output variables and the input variables.

Table 3 : Computation of Variables from SACMEQ III data				
Conceptual variable	**Name of variable tested**	**Computation of variable from SACMEQ**	**Alternative variations of variables tried but rejected**	**SACMEQ variable meaning and notes**
1) Pupil Characteristics				
Gender	Zpsex	Straight from SACMEQ		0 = male 1= female
Age	Zpagemon	Straight from SACMEQ	Tried this variable squared but correlation was not as strong.	Pupil's age in months rounded to the nearest month.
SES	psesnolig12	psesnolig12= (zppos14*3) +((zproof*2)/4) + (zppos15*2) + (zppos17*2) + (zppos23)+ (zppos07) +(zpfloor/4) Added SACMEQ variables on possessions and house type together (floor type (zpfloor)and roof type (zproof) in the home, if the house had electricity (zppos14), a bed (zppos07), a fridge(zppos15), a TV (zppos23)and an electric fan (zppos17)). Possessions variables multiplied by 4 to give equal weighting to house type. Then the following weighting given: electricity*3, roof*2, fridge *2, fan *2. This weights the variables that have highest correlations with test scores higher: The variables that have a correlation of 0.14 get weighting of 3, those with a correlation of 0.10-0.11 get a weighting of 2.	zpsesscr - straight from SACMEQ which used the Rasch scaling approach. This had a lower correlation (0.11) with pupil test scores compared to the SES measure constructed for Malawi (0.16).	A measure for SES, better adapted to Malawi, and excluding parents ed and light type at home, so that these variables can be separately assessed. Takes values 0-12.
2) Home Environment				
Frequency of meals	Zpmealw	ZPMEALW = PMEAL1*0.5 + PMEAL2*1.5 + PMEAL3 where PMEAL1, PMEAL2, PMEAL3 (1=not at all; 2=1or2 /week; 3=3or4/week; 4=every day). PMEl1 1 =breakfast, pmeal2 = lunch, pmeal 3= evening meal.		Looked at correlation between the Pmeal variables and scores and found that lunch was most highly correlated, and breakfast least. So weighted lunch as 1.5, breakfast as 0.5 and evening meal as 1).
Academic support at home: parents education and books at home	zpparbook59	zpparbook59 = zpparbook5/300. Where: gen zpparbook5= zpparentsqsq+(zpbooksh*5) and gen zpparent= zpmother+zpfather gen zpparentsq=zpparent*zpparent gen zpparentsqsq=zpparentsq*zpparent	Mothers education on own, and fathers education on own. Father education var is stronger than mothers education variable, but even stronger when together.	Parents education^3 plus 5*number of books at home, (divided by 300 to make numbers more manageable).

Table 3 : Computation of Variables from SACMEQ III data				
Conceptual variable	**Name of variable tested**	**Computation of variable from SACMEQ**	**Alternative variations of variables tried but rejected**	**SACMEQ variable meaning and notes**
English spoken at home	Zpenglis	Straight from SACMEQ		If English is spoken at home: 0=never, 1=sometimes/all the time.
Extra tuition	Zpexttui	straight from SACMEQ		Whether or not a pupil has tuition outside school.
Chores done in the home	Ptasktotg	gen ptasktotg= ptask1+ ptask3+ ptask4+ ptask8+ ptask9+ ptask12+ ptask13		Gets point for each of the following jobs that they have to do: looking after younger relatives, taking care of sick relatives, cooking, fetching water, chopping firewood, gardening, taking care of livestock. 1 point = some days and 2= most days
Access to a computer	Zpcomptr	Straight from SACMEQ	Whether a pupil had the following learning materials available in the home: exercise books, notebooks, pens and pencils. These were not correlated to test scores so a variables measuring if these items were available in the home was rejected.	If pupil has ever used a computer
Light type in the home	Zplight	Straight from SACMEQ		The type of light a pupil has in the home, 1=no light, 4= electricity.
3) Pupil Academic Record				
Years of school repeated	Prepeatsq	prepeatsq=prepeat*prepeat	Prepeat was rejected as repetition squared had a higher correlation with pupil scores.	(1=never; 2=once; 3=twice; 4=three plus), then these values squared.
Days absent	Pabsent	Straight from SACMEQ	pabsent squared was not as significant so it was rejected.	Number of days pupil absent in a year
4) Teacher Characteristics				

Table 3 : Computation of Variables from SACMEQ III data				
Conceptual variable	**Name of variable tested**	**Computation of variable from SACMEQ**	**Alternative variations of variables tried but rejected**	**SACMEQ variable meaning and notes**
Teacher living conditions	Ycondliv	Straight from SACMEQ	Tried the same variable but for maths and reading teachers combined but it was not significantly correlated so rejected.	1=generally poor, 2=major repair, 3=minor repair, 4=generally good. (note this is just for the maths teacher)
5) Teaching Quality				
Hours of teaching (1= above 30 hours)	zhrteachdum4	Zxhrteac added to Zyhrteac together to make a variable for both math and reading. Then dummy variable created where 0= 30 hours and below and 1 = above 30 hours teaching for both maths and reading teachers combined.		A dummy variable has been created which takes the value 0 if 30 hours or less are taught and the value 1 if more than 30 hours are taught.
Knowledge of teachers	Zaloctmr	zaloctmr = zraloct+zmaloct		Test scores of the maths and reading teachers added together.
Homework given	Phmwk	Straight from SACMEQ		If a pupil gets homework. 1= never homework 2= 1-2 times a month 3= once or twice a week 4= everyday.
Math teacher interacts with parents	Ymeetpar	Straight from SACMEQ	How often the reading teacher meets with parents was rejected because it was significantly correlated.	How often the math teacher meets with pupils
Teachers days absent	Teachabs	genteachabs= xabsnt1+xabsnt2+xabsnt3+ xabsnt4 +xabsnt5 +xabsnt6 +xabsnt7 +xabsnt8 +xabsnt9 +xabsnt10 +xabsnt12 +yabsnt1 +yabsnt2 +yabsnt3 +yabsnt4 +yabsnt5 +yabsnt6 +yabsnt7 +yabsnt8 +yabsnt9 +yabsnt10 +yabsnt11 +yabsnt12		Total days teachers are absent (math and reading added)
Teachers years of training	Zqprodmr	zqprodmr= zxqprof+zyqprof		Combine these two for average.
6) School Head Characteristics				
Head teacher gender	Zssex	Straight from SACMEQ		Gender of head teacher
Head teacher age	Sage	Straight from SACMEQ		age of head teacher

Table 3 : Computation of Variables from SACMEQ III data				
Conceptual variable	**Name of variable tested**	**Computation of variable from SACMEQ**	**Alternative variations of variables tried but rejected**	**SACMEQ variable meaning and notes**
7) School Head Quality of Management				
Head teacher years experience (1= 13 years or above experience)	**Sexpalldum**	replace sexpalldum = 0 if sexpall<13, replace sexpalldum =1 if sexpall>=13		Takes the value 0 if the head teacher has worked less than 13 years and the value 1 if they have worked as a head teacher for 13 years or more
Head teacher teaching hours	Zstchhrs	**zstchhrs**= zstchmin/60		This is teaching minutes of head teacher per week divided by 60 and create a new variable for hours.
8) Classroom Environment				
Class size	**Clsize**	**clsize**=(xclsize+yclsize)/2		Average class size for reading and math
Pupil teacher ratio	**Zsptrati**	Straight from SACMEQ	Tried squaring this variable but not as significant.	Pupil teacher ratio from school data.
9) School Characteristics				
School resources	**resource 1**	Combines zreslev (School Resources (e.g. library, sports ground etc) level. It was created by classifying ZRESLOC into categories.) and tcres8 and tcres9 for reading and math teachers, to create a 1-8 level.		The level of classroom and school resources 1 is the lowest level and 8 is the highest.
State of school building	**Zsbldgco**	Straight from SACMEQ		0=only in need of minor repair/good, 1=in need of rebuilding/major repair
Distance to school	**Ptravel**	Straight from SACMEQ	Tried combining with ptravel 2 - modes of transport, but over 90% walk and it did not affect significance or results.	Distance traveled: where 1= 0-0.5 k, 2= 0.5-1 k etc up to 11 which =5k or more

Table 3 : Computation of Variables from SACMEQ III data				
Conceptual variable	**Name of variable tested**	**Computation of variable from SACMEQ**	**Alternative variations of variables tried but rejected**	**SACMEQ variable meaning and notes**
School location	**Zsloc**	Straight from SACMEQ		Location of school, rural or urban.
Free meals at school	**Zpmealsc**	Straight from SACMEQ		Free School Meals. 0 = none, 1= one free meal a day
Type of school (public/private)	**Zstype**	Straight from SACMEQ		0=government 1= private
10) School Environment				
Proportion of female teachers	**Spropf**	gen stchtotf=stchpf+stchtf and gen stchtotm=stchpm+stchtm and gen spropf=stchtotf/stchpm		Use STCHPM and STCHPF to create a variable to give the proportion of female teachers in a school.
Safe learning environment	**abusedum3**	gen abusedum3 =0 replace abusedum3=1 if zspupp13==1 & zstchp09==1 & zstchp06==1 & pupprobdum1==0 &zshriskt==0		abusedum3 takes the value of 1 if there is no sexual abuse from teachers or pupils, no alcohol abuse from teachers, a perceived low HIV risk for teachers and the pupils do have some behavior problems, and 0 otherwise
11) Community Environment				
Problems with community	**Sprobcom**	Straight from SACMEQ		Takes the value 1 if there is a school problem with the community. 0 otherwise.
Community contributes (building, textbooks)	**Zscommneg**	zscommneg=zscomm01+ zscomm02+ zscomm04+ zscomm08		Number of the following items that community contribute to school: building facility, maintenance of faculty, purchase of text books, payment of salaries of additional teachers.
Community pays bonus to school staff	zscomm11	Straight from SACMEQ		Takes the value 1 if the community pay a bonus to the non-teaching staff at the school

3

3.3 Regression Analysis

The first step of the regression analysis was to narrow down the number of variables to be included in the regressions. The following steps were taken in order to do this:

1) The conceptual model was built from the findings of the literature review.

2) For each education input in the conceptual model, a variable was built from the SACMEQ data where suitable data existed.

3) Descriptive data and basic analysis was produced for each of these variables

4) Variables were included from the descriptive data and basic analysis only if they were statistically significantly correlated with the Y variable.

5) Checks were carried out for multicollinearities[17], endogeniety[18], specification errors[19] and selectivity bias[20]; any of these identified were mostly dealt with by either combining closely related variables into single variables, excluding variables, finding an alternative variable, or generating a different form of variable e.g. squaring the variable, .

Once the set of variables to be included had been decided upon then three sets of regressions were run. The first set of regressions run were for both the female and male samples combined, in which special attention was given to whether or not the gender dummy was significant and important. Information was thereby obtained in order to answer research question one. The next

[17]Multi-collinearity is where the X variables are highly intercorrelated. 'Strictly speaking, multicollinearity refers to the existence of more than one exact linear relationship, and collinearity refers to the existence of a single linear relationship. But this distinction is rarely maintained in practice, and multicollinearity refers to both cases' (Gujarati (2003:342)). For example if one had one x variable as the SES of a pupil, another variable that looked at the type of lighting that a pupil had in their home, and another that looked at how many books a pupil had in their house, one may find that these three variables are strongly correlated. It is likely that the type of lighting a pupil has in their home and how many books they have in the home are strongly related to the SES of a pupil and having all three variables in the regression may result in these variables not showing up as significant in the model because they each have a similar effect on educational outcomes.

[18]Endogeniety is where the y values may affect the x values. For example if a country rewarded high performing schools with extra resources then having high performers in a school may result in more resources in a school. In this case it would be wrong to conclude that having more resources in a school resulted in higher performance as onewould not be able to separate out the direction of the two affects.

[19]'If the model is not "correctly" specified, one encounter the problem of model specification error or model specification bias' (Gujarati (2003:506)). One of the following errors can cause it: omission of relevant variables, inclusion of unnecessary variables, adopting the wrong functional form, errors of measurement, incorrect specification of stochastic error term.

[20]'In questionnaire-type surveys, the problem of non-response can be serious; a researcher is lucky to get a 40 per cent response to a questionnaire. Analysis based on such partial response may not truly reflect the behaviour of the 60 per cent who did not respond, thereby leading to what is known as 9sample) selectivity bias' (Gujarati (2003:30)).

two sets of regressions that were run were for the female and male samples separately. The results were then compared to see if the education inputs were correlated with outcomes differently for males and females. This regression analysis provided information to answer research question two.

The regression model of school production that was used is an Ordinary Least Squares (OLS) regression which has the following formula:

$$Y = a + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4 + b_5X_5 + u$$

The *Y* variable is the dependent variable which, in this analysis, was the educational outcome. This was measured by test scores which are a proxy for the quality of knowledge that is achieved in the education system. Note that test scores can sometimes be correlated to other education outcomes, for example, drop-out rates: 'The better the scores are, the more likely it is that learners will stay within the schooling system right up to the exit point.' (Hanushek and Luque 2003:483, quoted in Gustafsson 2006:46). However, no analysis has been carried out to see if the SACMEQ III test scores are correlated with drop-out rates or other education outcomes so it cannot assume this is the case for Malawi.

Lockheed and Longford (1989: 24) argued that it is important to take into account factors that affect how a pupil does in a test that are not in a regression, such as emotional state on the day, or health on that day. They referred to this as 'within pupil variance'. Ideally, in order to overcome this problem, a range of tests would be carried out on different days for the same pupil, and the average score would then be calculated. This has not been possible using SACMEQ as there is just one set of results. However, to minimise the risk of 'within pupil variance' the *Y* variable used was the average of the maths and reading scores for a pupil. This was thought to be appropriate as this research examined overall pupil performance rather than at maths or reading specifically.

The *X* variables shown in the above equation are the independent variables which, in this analysis, were the education input variables; for example, class size or text books. The symbol '*a*' is the intercept and 'b_1' and 'b_2', etc. are the slope coefficients. The coefficients measure the

relationship between the input variable and the output variable. Unlike the correlations, which only look at the association between one input variable and the output variable, the coefficients state the net relationship between an input variable and the output variable, while controlling for the correlations of the other variables. The symbol *'u'* is the error term which takes the value of the unexplained test scores, and, in OLS regressions, the coefficients are those that minimise the error terms.

The sample for the SACMEQ III data was not drawn at the individual student level, but at school level. It used complex two-stage cluster sampling including weighting adjustments to compensate for variations in the probability of selection. Since learners are clustered in schools, and schools are situated in regions, the sampling errors are therefore larger than those that would have been obtained by simple random sampling. An adjustment for the complex sampling method was made by using the SVY command in Stata and this showed that in general the coefficients that were only significant at the 0.05 level in the standard OLS regressions were no longer significant. The results of the SVY adjusted regression are considered when interpreting results of the OLS regressions and only coefficients that are statistically significant at 0.01 in the standard OLS regression are considered significant.

The analysis also looked at the magnitude of the change that could be expected in Yi, given a change in the value of the explanatory variable. For each regression the unstandardised (metric) coefficients were calculated and presented as well as the standardised (beta) coefficients. An unstandardised coefficient for a variable estimates how much the pupil score would increase or decrease if the variable increased by one unit, whilst controlling for the other variables. The standardised beta coefficients are the same coefficients but standardized so that their variances are equal to 1. Thus the absolute values of the standardised coefficients allow the ranking of variables by their relative impact on the outcome variable regardless of their original units of measurement. Hungi (2011) said that: 'In education, a standardized regression coefficient is considered important if its magnitude taken in absolute terms is ≥0.10'. Therefore, in the following discussions, particular attention has been paid to variables with beta coefficients that were 0.10 and higher.

This analysis did not use hierarchical linear models[21], as it was decided that given the research questions, and the scope of the research, the reasons for not carrying out hierarchical linear models outweighed the reasons for carrying them out, and therefore, OLS regression models were used instead. A key benefit of using the OLS models used in this research is that they are simpler and easier to interpret. As Johnes (2004) argued: 'The problem with multi-level models is that, in order to model the complexities of the educational production process, they must themselves become complex and as a consequence can become difficult for any but the expert to understand and interpret. Moreover, more advanced variants of multi-level modelling (for instance those involving limited dependent variables) can be computationally intractable.' Gustafsson (2007) carried out two types of regressions using SACMEQ 2000 data: firstly, he carried out a one level regression, i.e. Ordinary Least Squares; secondly, he carried out two level models, i.e. hierarchical linear model. It is interesting to note that he found that: 'The slope coefficients obtained using the two-level model are roughly equal to those obtained in the one-level models' (Gustafsson (2007)). Thus it may be that, even if this research had carried out hierarchical linear model regressions as well, it may have been found that the results were the same. This is an area for future research.

After running the regressions, conclusions were then made based on the significance of the coefficients and the size of the association that the variables had with outcomes as measured by the beta-standardised coefficients. The analysis evaluated the robustness of the variables in the model through comparison of a few variations of the models.

Personal knowledge of the Malawian education system, along with information from the literature review, was used to assist with the interpretation of results.

[21]Hierarchical linear models allow differentiated intercepts and slopes by school and district within one econometric model, so that the effects of the institution attended on the education outcomes are measured.

4 Research Data

The analysis used secondary data from SACMEQ III's data set which has variables for education outcomes, socio-economic status (SES), pupil and teacher characteristics, and school resources. SACMEQ III has appropriate data for most variables in the conceptual model presented earlier; however, there were some variables in the conceptual model for which no suitable SACMEQ III data exists. For example, there were no variables to measure: innate ability, management processes in a school, the value that the community places on education, whether a female pupil is married or whether a pupil is pregnant. There is a risk that these missing variables may have caused omitted variable bias. Low R2 was expected due to the missing explanatory variables.

SACMEQ uses questionnaires for the school principal, the educator and the learner, and has tests for the educator and the learner. The sample size of Malawi's SACMEQ III was 2781 pupils, 264 teachers, and 139 schools. Therefore, the data was large enough to run reliable regressions split into males and females. The data was collected for Grade 6 pupils over eight months of the school year. SACMEQ III had various sampling stages: firstly, regions within a country were used as a strata; twenty schools within a region were then selected; finally, 20 pupils were selected randomly from within each of these schools. The analysis presented in this report adjusted all the data by a probability weight so that the conclusions can be generalised for the whole country. However, the probability weight used assumed that the learners who were absent on the day of the test were absent at random. In reality this may not have been the case; it may be that those pupils who expected that they would do badly in the test, or who did not like tests, were the same pupils who did not attend school that day. This may mean that the test scores used had a slight upward bias.

There may also be other missing values in the SACMEQ data which typically are not random. Gustafsson (2006:31) stated that: 'Historically disadvantaged learners and schools have a disproportionately high number of missing values.' SACMEQ carries out an imputation process[22] for missing values if a variable only has a low number of missing values. However, if

[22]The SACMEQ imputation process was done on the basis of the results of the analysis of the missing values taking into consideration the level of missing values at different levels of analysis. The imputation process was only done for variables with levels of missing values that was less than 15 % at one of the levels of analysis. For example: level of education; (i) If the level of missing values for the 'level of education' (variable) at the first level of analysis (which is class level) was less than 15%, then the mode or mean value for the variable will be used for the missing values for this variable.(ii) If the level of missing values for the variable was more than 15%, then the level of missing values for this variable at

the variable has missing values of more than 15 per cent at each level (class, school, region, country), then no imputation is carried out. The analysis reports on how many non-imputed missing values were found for each variable. Where it was suspected that missing values may have been of some significance, the issues have been subsequently discussed and addressed where possible.

On the whole, the SACMEQ data was considered to be reliable and valid. Unlike national statistics, such as EMIS, the SACMEQ statistics are not viewed by schools as affecting the resources that a school receives; therefore, there were fewer incentives to give false information when questionnaires were being filled in.

This report will now go through each of the eleven types of variables identified in the conceptual model and give an overview of the data.

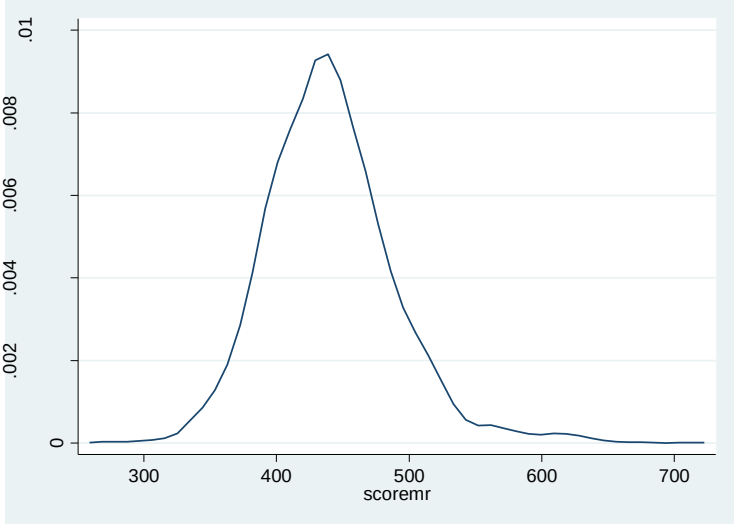
The next level of analysis (which is the school level) is checked. If the level is less that 15%, then the mode or mean value will be used for all missing values for this variable. (iii) If the level of missing values for the variable is more than 15% at all level of analysis (class, school, region, country), then no imputation is done. Examples are given below of the types imputation carried out:

1. School location: imputation was based on the information from the district or region (area where the school is located).
2. Community problem: imputation was based on the most common in the district or region (area where the school is located).
3. Maths teacher interact with parents: anything that is not ticked is assumed to be not discussed (which of the following issues you have discussed with parents).
4. Age and years of experience (teacher or school head): imputation was based on other administrative records/ other SACMEQ III instruments.
5. Years of training (teacher and school head): imputation was based on other administrative records
6. School resources: all un ticked options were assumed to be not available (indicate resources your school have)
7. State of the building: imputation was based on observation or other administrative records
8. Pupil tasks: imputation was based on the most common among other pupils in the same school.
9. Extra tuition: imputation was based on the mean of the class (for those who indicated that they had extra tuition).
10. Teaching hours: imputation was based on the group mean.
11. No imputation is done for pupil or teacher test scores. Any missing values are left as missing values.

5 Descriptive Data

5.1 Test Scores

Figure 7: Kernel Density Graph for Test Scores, 2007



SACMEQ III, own analysis

7

The descriptive statistics in Table 4 and kernel density graph show that the distribution is approximately normal apart from those scores above the 95 percentile which is 521, with the highest score at 714[23]. This skews the distribution to the right and explains why the mean is higher than the median.

[23]The score of 714 is an outlier. This is the highest score and the second highest score is 672. However, as this is likely to be a true score, correlations and regression will be more meaningful with it left in so it has not been removed.

Table 4: Descriptive Data for Test Scores, 2007									
Conceptual variable	**Name of variable tested**	**Non missing obs**	**10th percentile and 90th percentile**	**Median, mean**	**s.d.**	**Median, mean (Female)**	**Confidence Interval (CI) (95%) for Mean (Female)**	**Median, mean (Male)**	**Confidence Interval (CI) (95%) for Mean (Male)**
Pupil test scores	scoremr	2780	384, 500	436, 440	48	434, 435	432,437	442, 446	443,448
SACMEQ III, own analysis									

4

The data shows that on average females scored much lower than males. Female pupils have a median score of 434 whilst male pupils have a median score of 442. There is a significant gap between confidence intervals for females and males so it is almost certain that the true value of the parameter for females is lower than for males (the difference between the two values is significant at the corresponding value of α=0.05).

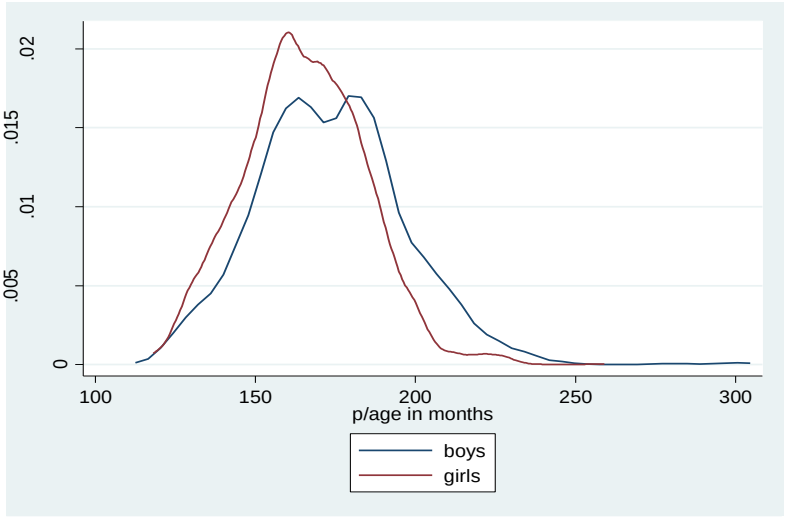
5.2 Pupil Characteristics

Table 5: Descriptive Data for Pupil Characteristics, 2007									
Conceptual variable	**Name of variable tested**	**Non missing obs**	**10th percentile and 90th percentile**	**Median, mean**	**s.d.**	**Median, mean (Female)**	**Mean CI (95%) (Female)**	**Median, mean (Male)**	**Mean CI (95%) (Male)**
Gender	zpsex	2781	0,1	0,0.49	50	1, 1	na	0, 0	na
Age	zpagemon	2781	142 (11yrs), 207 (17yrs)	168(14yrs) 169(14yrs)	21	165, 165 (13yrs)	164 to166	173, 174 (14yrs)	172 to175
SES	psesnolig12	2781	0.75,6.25	1.75,2.64	2.23	1.75,2.66	2.54 to 2.78	1.75,2.61	2.49 to 2.74
SACMEQ III, own analysis									

5

The statistics show that, in Malawi, approximately 49 per cent of pupils in Grade 6 are females. As shown previously, EMIS data (2010) shows that from Grade 6 the percentage of females dropping out of school starts to increase and the GPI for grade enrolment falls below 1 as the proportion of females to males falls. Thus if the data were for Grade 7 or 8, one would expect to find a lower percentage of females in the classes.

Figure 8: Kernel Density Graph for Age, 2007



SACMEQ III, own analysis

8

The analysis shows that pupils' ages range from approximately 11 years at the 10th percentile to 17 years old at the 90th percentile. Interestingly, the median for female pupils (13 ¾ years) is significantly lower than for male pupils (14 2/5 years) in 2007 as shown by the distance in confidence intervals. Note that age is higher for both males and females than the EMIS (2010) data suggested. The kernel density graph shows that there are more female pupils than male pupils who are under 15, and more males than females who are over 15. One reason that females tend to be younger than males may be because females repeat less (as shown by EMIS and analysis further on in the report). However, it may also be partially due to older females dropping out after puberty due to marriage or pregnancy, as suggested by some studies in the literature review (for example Colclough, 1997). However, there is no data available on drop out by age therefore it is not possible to say whether this theory is true in the case of Malawi.

The 10th percentile for SES is 0.75 and the 90th percentile is just 6.25. The median is 1.75 and the mean is 2.64 which demonstrate that the distribution is skewed to the left. The distribution of SES for males and females is very similar. Eighty per cent of the population only score 3 or less at this measure of SES which is what is expected given that 74 per cent of the population was

estimated to be below the poverty line in 2011 (UNDP, (2011)). Note that a pupil who scored 2.75 may live in a house with a bed, a cement floor and a metal roof but have no electricity, no fan, no fridge and no TV. A pupil who scored 1.75 (the median) would also have no electricity, no fan, no fridge and no TV however they would either live in a house with no bed, a cement floor and metal roof or more likely in a house with a bed, an earth floor and no permanent roof e.g. thatch. This data shows how economically poor the majority of the population is, and that for the majority of the population there is little variation in their SES.

Some variables that are in the conceptual model have been omitted due to lack of data. Ideally, there should be variables which measure whether or not a pupil is married, whether or not they are pregnant or have any children, and the level of health of a pupil. However, there are no direct measures of these variables in SACMEQ. Therefore, as mentioned in the methodology section, the conclusions drawn have considered the possibility of omitted variable bias.

5.3 Home Environment

Table 6: Descriptive Data for Home Environment, 2007

Conceptual variable	Name of variable tested	Non missing obs	10th percentile and 90th percentile	Median, mean	s.d.	Median, mean (Female)	Mean CI (95%) (Female)	Median, mean (Male)	Mean CI (95%) (Male)
Frequency of meals	Zpmealw	2781	9.5,12	12,11	1.4	12,11	11.16 to 11.31	11.5,11	10.99 to 11.14
Academic support:(parent education & books at home)	zpparbook59	2781	0.17,3.4	0.5,1.16	1.44	0.5,1.18	1.10 to 1.26	0.5,1.14	1.07 to 1.22
English spoken at home	Zpenglis	2781	0,1	1,0.54	0.5	1,0.50	0.48 to 0.53	1, 0.58	0.55 to 0.60
Extra tuition	Zpexttui	2781	0,1	0,0.14	0.35	0,0.16	0.14 to 0.178	0,0.13	0.11 to 0.15
Chores done in the home	Ptasktotg	2781	9,17	13,13	2.87	13,13.4	13.26 to 13.56	12,12.5	12.36 to 12.66
Access to a computer	Zpcomptr	2781	0,0	0,0.016	0.125	0,0.01	0.01 to 0.02	0, 0.02	0.01 to 0.028
Light type in the home	Zplight	2781	1,4	2,2.21	0.7	2, 2.24	2.20 to 2.27	2, 2.19	2.16 to 2.23

SACMEQ III, own analysis

6

A valuable resource for pupils who are trying to study outside school hours is a light in the house in which they live. The data shows that approximately 4 per cent of pupils lived in a house with no light; 82 per cent lived in a house with a paraffin lamp or candles; 2 per cent with a gas lamp; and 12 per cent with electricity.

The statistics also show that only 1.59 per cent of all pupils had ever used a computer. This once again demonstrates the very limited access that pupils in Malawi have to technology and higher-level learning materials.

The overlapping confidence intervals for males and females shows there are no significant differences by gender in the number of books or level of parents' education or in pupils' access

to a computer or the light type in the home. This suggests that pupils tended to come from similar home backgrounds, as the SES data in the previous table demonstrates.

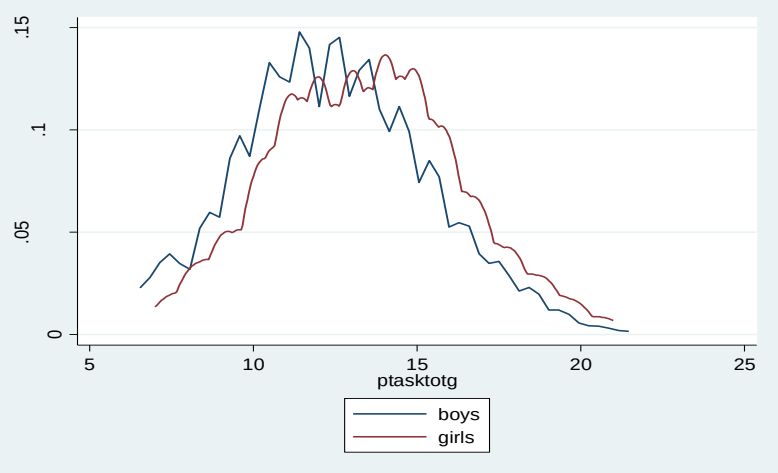
However, the data shows that frequency of meals was higher for females than for males. This finding is not supported by the literature and it would be interesting to look further into the reasons for this finding.

The data also shows that male pupils in Grade 6 were more likely to come from a home where English is spoken. Given that other variables have shown that males and females come from similar backgrounds, this finding is unexpected. This suggests that the reason why there are differences between the percentage of males and females who are from homes where English is spoken is representing something else as well as SES[24] or education of parents. For example, it may also be a proxy for parents' attitudes towards education. However, the literature review has not shown any evidence for this so, as with pupil meals, more analysis is required in order to further interpret these findings.

One difference that the data finds that is supported by the literature, is that females carry out significantly more household chores than males. The kernel density graph shows a similar shaped distribution for females and males, but the female graph is shifted slightly to the right of the male graph.

[24] As the data shows that approximately 48 per cent of females and 55 per cent of males reported that English is spoken in the home this suggests that it is not only in the middle class households in which English is spoken.

Figure 9: Kernel Density Graph of Chores Done Outside School, 2007



SACMEQ III, own analysis

9

Some of the variables that ideally would be looked at under 'home environment' such as the value that parents put on female and male education have been omitted due to lack of data.

5.4 Pupil Academic Record

Table 7: Descriptive Data for Pupil Academic Record, 2007

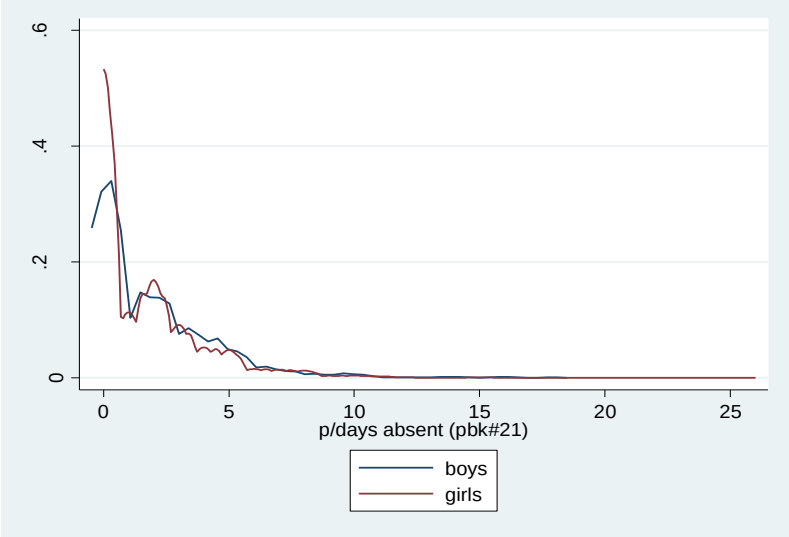
Conceptual variable	Name of variable tested	Non missing obs	10th percentile and 90th percentile	Median, mean	s.d.	Median, mean (Female)	Mean CI (95%) (Female)	Median, mean (Male)	Mean CI (95%) (Male)
Grade repetition (squared)	Prepeatsq	2781	1,9	4,4.15	3.8	4,4.03	3.83 to 4.23	4,4.26	4.06 to 4.47
Days absent	Pabsent	2781	0,5	1,1.66	2.31	0,1.52	1.41 to 1.64	1, 1.80	1.67 to 1.92

SACMEQ III, own analysis

7

The statistics show that in Malawi 40 per cent of pupils had never repeated a school year; 41 per cent of these had repeated once; 14 per cent had repeated twice; and 5 per cent had repeated three or more times. Slightly more males then females repeat (mean of 4.03 for females compared to 4.26 for males) however this is not a significant difference. Male tended to have more days absent with a mean of 1.8 days compared to 1.5 days for females; however, the pattern is similar, see kernel density graph. As the data shows females have significantly less absence than males. According to the literature and analysis presented later, this should have a positive relationship with their scores. However, it does not have a strong enough positive association to counteract the overall gender differences in scores discussed previously.

Figure 10: Kernel Density Graph for Days Absent, 2007



SACMEQ III, own analysis

10

5.5 Teacher Characteristics

Table 8: Descriptive Data for Teacher Characteristics, 2007									
Conceptual variable	**Name of variable tested**	**Non missing obs**	**10th percentile and 90th percentile**	**Median, mean**	**s.d.**	**Median, mean (Female)**	**Mean CI (95%) (Female)**	**Median, mean (Male)**	**Mean CI (95%) (Male)**
Teacher living conditions	ycondliv	2780	1, 3	1, 1.79	0.99	1, 1.77	1.72 to 1.82	1, 1.81	1.75 to 1.86
SACMEQ III, own analysis									

8

Data shows that the median score for teacher living conditions is 1, which suggests that most teachers were living in poor housing conditions. Additional tabulation of the data (see Table 9) also shows that approximately 52 per cent of maths teachers lived in poor housing conditions; 26 per cent in houses that needed major repairs; 12 per cent in houses that needed minor repairs; and 10 per cent in houses that were in good condition.

Table 9: Teacher Living Conditions for Math Teacher, 2007	
State of house:	**Per cent**
Poor	52.35
Need major repairs	26.07
Need minor repairs	11.86
Good condition	9.72
Total	**100**
SACMEQ III, own analysis	

9

5.6 Quality of Teaching

Table 10: Descriptive Data for Quality of Teaching, 2007

Conceptual variable	Name of variable tested	Non missing obs	10th percentile and 90th percentile	Median, mean	s.d.	Median, mean (Female)	Mean CI (95%) (Female)	Median, mean (Male)	Mean CI (95%) (Male)
Hours of teaching (1= above 30 hours)	zhrteachdum4	2781	0,1	1,0.695	0.46	1,0.693	0.67 to 0.72	1, 0.697	0.67 to 0.72
Knowledge of teachers	zaloctmr	2744	1302, 1656	1489, 1482	133	1489, 1481	1474.72 to 1488.94	1489, 1482	1475.53 to 1489.54
Homework given	phmwk	2781	2,4	3, 2.86	0.79	3, 2.86	2.82 to 2.90	3, 2.86	2.82 to 2.91
Maths teacher interacts with parents	ymeetpar	2781	1,4	3, 2.98	0.99	3, 2.97	2.91 to 3.02	3, 2.99	2.94 to 3.04
Teachers days absent	teachabs	2781	2,34	15,19	21.5	15,19	17.76 to 19.83	15,19	18.27 to 20.70
Teachers years of training	zqprodmr	2781	2, 5	4, 3.7	1.4	4, 3.7	3.62 to 3.77	4, 3.7	3.62 to 3.76

SACMEQ III, own analysis

10

The analysis shows that only 30.5 per cent of teachers taught for 30 hours or less. Knowledge of teachers varies from 1186 to 1803, with the median score at 1489 and the mean score at 1481. It has been noted, however, that there are only 37 missing observations for this variable. It is suspected that these may not be randomly distributed as it may have been that the teachers who expected to perform badly did not do the test. Although these missing values affect less than 2 per cent of the total observations, precautions have been taken and are explained later in the report.

The analysis shows that in Malawi approximately 5 per cent of Grade 6 students were never set homework; 24 per cent only received homework once or twice a month; 51 per cent got homework once or twice a week; 20 per cent were set homework every day.

The data also shows that the mean and median for teacher/parent contact is 3, suggesting that most pupils' parents met the maths teacher once a term. Additional tabulations in Table 11 show

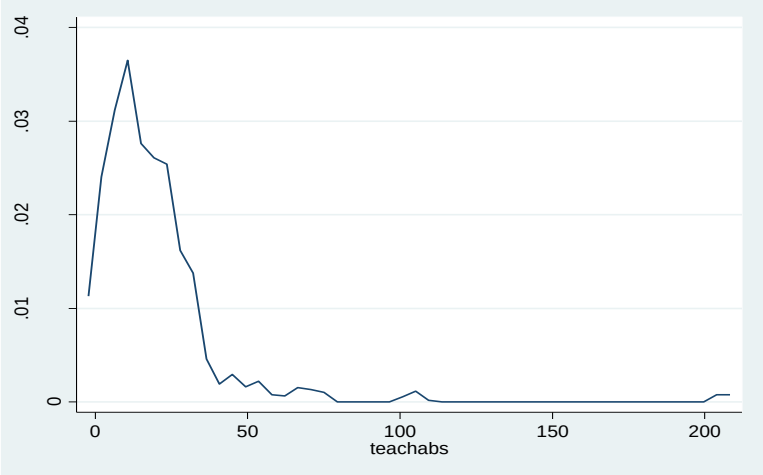
that 15 per cent of pupils' maths teachers and parents never met; 4 per cent met once a year; 48 per cent met once a term; 33 per cent met once or more a month.

Table 11: Teacher Parent Contact, 2007	
Frequency of contact teacher has with parent	**Per cent**
Never	15.64
Once per year	3.56
Once per term	47.95
Once or more per month	32.85
Total	**100**
SACMEQ III, own analysis	

[11]

The data illustrates that the mean days that maths and reading teachers were absent is a total of 19 days and the median is 15 days. The distribution is skewed to the right (see kernel density graph). Approximately 95 per cent of pupils had maths and reading teachers who were absent for less than 45 days a year, but 5 per cent of students had teachers who had between 45 and 206 days absence. Given that there are approximately 180 school days in a year and thus 360 days in total that maths and reading teachers should be teaching, this means that for 5 per cent of children between 12.5 per cent and 57 per cent of the time their teachers were supposed to be in front of a class they were in fact, absent.

Figure 11: Kernel Density Graph for Teacher Absenteeism, 2007



SACMEQ III, own analysis

11

The total years of training for maths and reading teachers ranges from 0 to 8 years, with the median being 4 years, (i.e. most teachers have had 2 years professional training). Approximately 63 per cent of students were being taught by maths and reading teachers who each had 2 years teaching training. The data shows that approximately 4 per cent of pupils were being taught by maths and reading teachers who had had no training.

All of these variables have a similar distribution for males and females suggesting that the quality of teaching that males and females receive is similar. However, again, some variables that may affect the quality of teaching have not been able to be captured by the data: for example, the quality of the curriculum and whether it reinforces negative stereotypes of females, or whether the teacher gives more attention or encouragement to males or females.

5.7 School Head Characteristics

Table 12: Descriptive Data for School Head Characteristics, 2007

Conceptual variable	Name of variable tested	Non missing obs	10th percentile and 90th percentile	Median, mean	s.d.	Median, mean (Female)	Mean CI (95%) (Female)	Median, mean (Male)	Mean CI (95%) (Male)
Head teacher gender	Zssex	2781	0,1	0, 0.13	0, 33	0, 0.14	0.12 to 0.16	0, 0.12	0.10 to 0.13
Head teacher age	Sage	2781	38, 52	45, 45	5.26	45,45	44.80 to 45.36	45,45	44.83 to 45.38

SACMEQ III, own analysis

12

The data shows that approximately 13 per cent of head teachers are female and 87 per cent are male. The mean and median age for head teachers in Malawi is 45 years old, with the 10th percentile at 38 and the 90th percentile at 52.

5.8 Quality of School Management

Table 13: Descriptive Data for Quality of School Management, 2007

Conceptual variable	Name of variable tested	Non missing obs	10th percentile and 90th percentile	Median, mean	s.d.	Median, mean (Female)	Mean CI (95%) (Female)	Median, mean (Male)	Mean CI (95%) (Male)
Head teacher years experience (1= 13 years or above experience)	sexpalldum	2781	0,1	0,0.13	0.34	0, 0.13	0.11 to 0.15	0, 0.13	0.12 to 0.15
Head teacher teaching hours	zstchhrs	2781	2.92, 26.25	12.5, 14	9.78	12.5, 14	13.46 to 14.50	12.25, 14	13.58 to 14.59

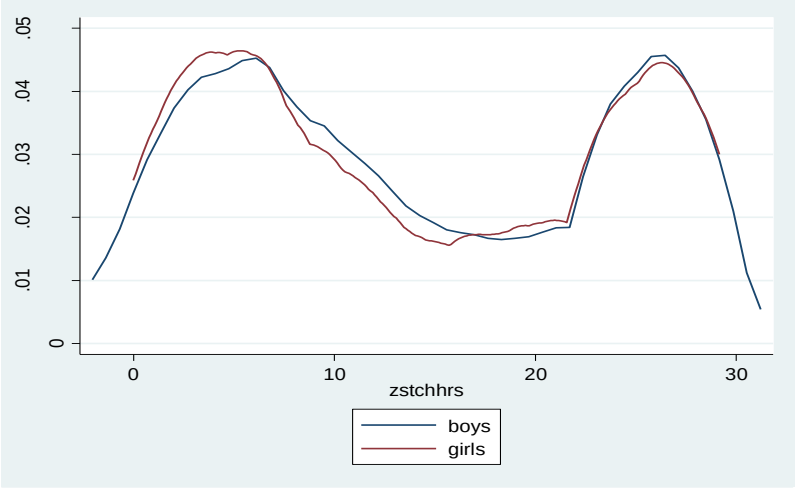
SACMEQ III, own analysis

13

The data shows that only 13 per cent of head teachers had worked as a head teacher for 13 years or more. The mean number of hours taught by the head teacher is 14 and the median number of hours is 12.5. The distribution is skewed to the right, with two peaks in the distribution (see

kernel density graph). The number of hours taught ranges from 2 hours a week to 30 hours a week. The reason for this is unclear.

Figure 12: Kernel Density Graph for Hours Taught by Head Teachers, 2007



SACMEQ III, own analysis

12

5.9 Classroom Environment

Table 14: Descriptive Data for Classroom Environment, 2007

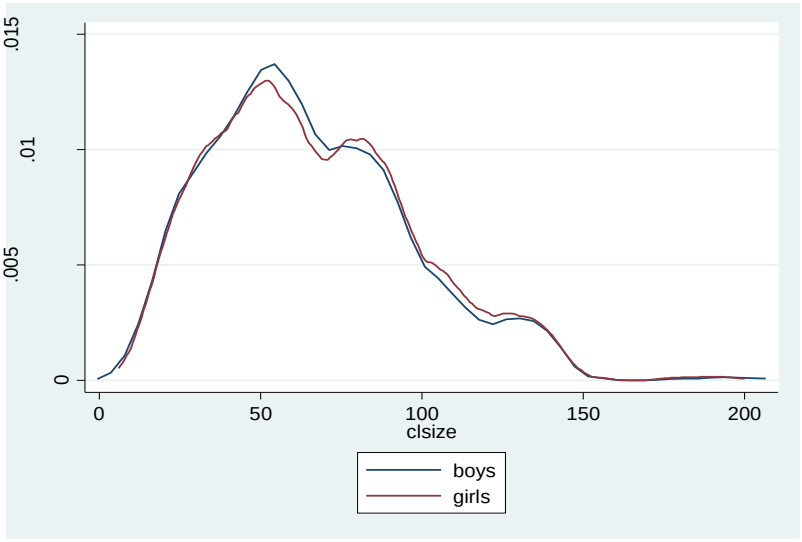
Conceptual variable	Name of variable tested	Non missing obs	10th percentile and 90th percentile	Median, mean	s.d.	Median, mean (Female)	Mean CI (95%) (Female)	Median, mean (Male)	Mean CI (95%) (Male)
Class size	Clsize	2781	28, 108	62,66	31.25	62, 66	65.24 to 68.59	60,65	64.10 to 67.33
Pupil teacher ratio	Zsptrati	2781	45, 141	74, 88	45.98	74,89	84.68 to 89.14	74,89	86.67 to 91.30

SACMEQ III, own analysis

14

Class size has a median of 62 pupils and a mean of 66 with the 10th percentile at 28 and the 90th percentile at 108 pupils. This highlights the frequency of large class sizes that exists in Malawi. However, class size does not explain how many teachers there are per class. In the case of a school which is short of staff, one teacher will teach two or more classes at the same time.

Figure 13: Kernel Density Graph for Class Size, 2007

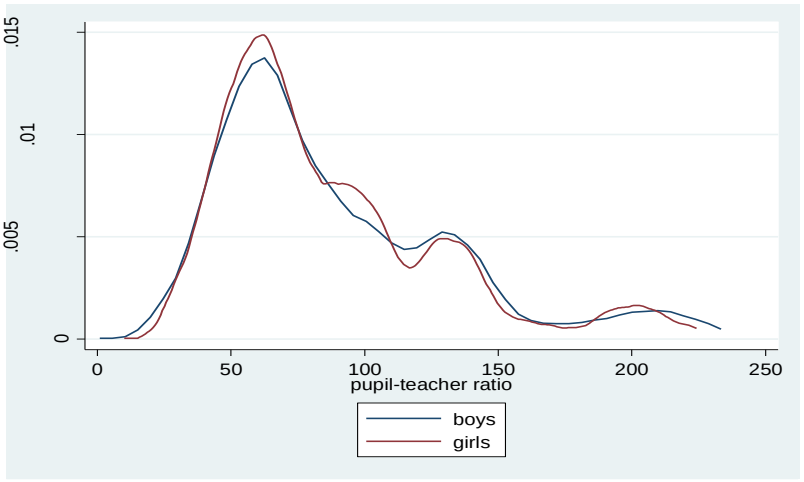


SACMEQ III, own analysis

13

The next variable (pupil/ teacher ratio) measures this. Male and females have a similar shaped distribution as shown by Figure 14. The analysis shows that, in Malawi's schools, pupil/teacher ratio for Grade 6 has a median of 74 pupils per teacher and a mean of 88 pupils per teacher. The pupil/teacher ratio at the 10th percentile is 45 and at the 90th percentile is 141. As these statistics are larger than those for class size, it is suggested that, as suspected, in some schools there is less than one teacher per class. Once again, these statistics show the large scale problem of high pupil/teacher ratio in the majority of pupils in Malawi's schools.

Figure 14: Kernel Density Graph for Pupil/Teacher Ratio, 2007



SACMEQ III, own analysis

14

For both of these variables there are no significant differences for female and male pupils and the distributions are similar (see kernel density graphs).

5.10 School Characteristics

Table 11: Descriptive Data for School Characteristics, 2007

Conceptual variable	Name of variable tested	Non missing obs	10th percentile and 90th percentile	Median, mean	s.d.	Median, mean (Female)	Mean CI (95%) (Female)	Median, mean (Male)	Mean CI (95%) (Male)
School resources	resource 1	2781	2,7	4, 4.34	1.99	4, 4.37	4.27 to 4.48	4, 4.32	4.22 to 4.43
State of school building	zsbldgco	2781	0,1	1, 0.54	0.5	1, 0.55	0.52 to 0.58	1, 0.54	0.511 to 0.563
Distance to school	ptravel	2781	1,9	3, 3.8	3.06	3,3.9	3.75 to 4.07	3,3.9	3.69 to 4.01
School location	zsloc	2781	0,1	0, 0.24	0.43	0, 0.24	0.22 to 0.27	0, 0.24	0.21 to 0.26
Free meals at school	zpmealsc	2781	0,1	0, 0,25	0.43	0, 0.26	0.23 to 0.28	0,2.4	0.22 to 0.26
Type of school (public/private)	zstype	2781	0,1	0, 0.31	0.46	0, 0.29	0.27 to 0.32	0, 0.32	0.30 to 0.34

SACMEQ III, own analysis

15

The data shows that of a possible score of 1-8 for school (whether there is a library, a staff room, water, electricity, a computer) and classroom resources (writing board, chalk, board duster, wall chart, cupboard, bookshelf, book corner, teacher chair and teacher table) the mean is 4.34. Approximately 50 per cent of pupils were in schools that were in need of rebuilding or major repair.

The median distance a pupil had to travel to school was between 1 and 1.5 kilometres, with approximately 10 per cent travelling more than 4 kilometres (see Table 16).

Table 16: Distance Pupil Lives From School, 2007		
Distance pupil lives from school	**Per cent**	**Cum.**
up to 0.5km	26.4	26.4
>0.5-1km	18.87	45.27
>1-1.5km	13.94	59.21
>1.5-2km	9.16	68.37
>2-2.5km	7.12	75.49
>2.5-3km	5.23	80.72
>3-3.5km	3.85	84.58
>3.5-4km	3.06	87.63
>4km-4.5km	2.71	90.34
>4.5-5km	3.48	93.82
>5km	6.18	100
Total	**100**	
SACMEQ III, own analysis		

16

The analysis found that 76 per cent of pupils went to a rural school and 24 per cent to an urban school. It was found that 25 per cent of pupils had free school meals at school. The data shows that approximately 31 per cent of students in Malawi attended private school.

There were no significant differences between the school characteristics of female or male pupils.

5.11 School Environment

Table 17: Descriptive Data for School Environment, 2007

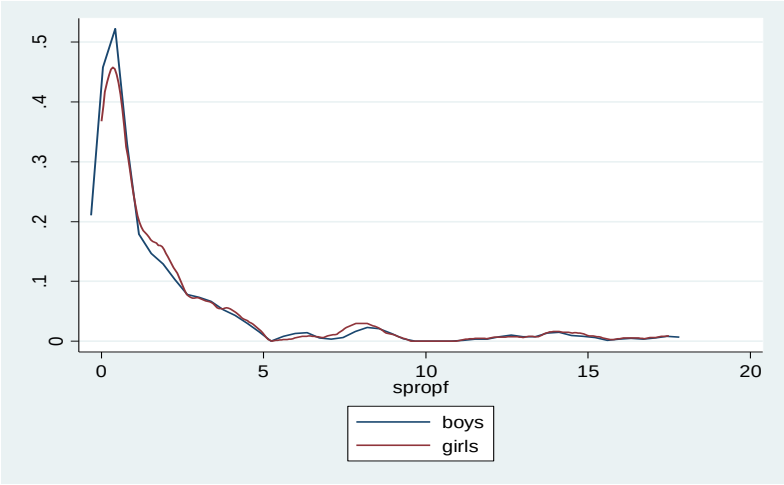
Conceptual variable	Name of variable tested	Non missing obs	10th percentile and 90th percentile	Median, mean	s.d.	Median, mean (Female)	Mean CI (95%) (Female)	Median, mean (Male)	Mean CI (95%) (Male)
Proportion of female teachers	spropf	2781	0, 6	0.75, 2.12	3.46	0.8, 2.25	2.06 to 2.44	0.67, 2.1	1.84 to 2.19
Safe learning environment	abusedum3	2781	0, 1	0, 0.22	0.41	0, 0.22	0.19 to 0.24	0, 0.22	0.19 to 0.24

SACMEQ III, own analysis

17

The proportion of female teachers ranges from 0 to 17.5; the 10th percentile is 0 and the 90th percentile is 6. The median is 0.75 and fifty four per cent of students were at schools that had more male than female teachers. The kernel density graph shows the distribution for this variable..

Figure 15: Kernel Density Graph for Proportion of Female Students, 2007



SACMEQ III, own analysis

15

The analysis finds that only 22 per cent of pupils were at schools which were classified as having a safe environment. If these statistics are broken down further, it can be found that that, 42 per cent of pupils were at schools which had a problem with sexual abuse of pupils by other pupils; 31 per cent were at schools where teachers had alcohol abuse difficulties; 17 per cent attended schools where teachers' sexual abuse of pupils had been identified as being a problem; 63 per cent went to schools with a high risk of teachers being infected with HIV; only 15 per cent of schools never had problems with pupil behaviour. This data highlights the high level of abuse which takes place in Malawian schools, and demonstrates that the majority of pupils are learning in unsafe school environments. Regardless of the correlation of this variable with pupil scores this data tells us that unsafe schools are a serious problem in Malawi and something which policy makers must address.

Ideally there would also be data that measured whether a pupil had or was being abused by pupils or teachers at school, or felt at risk from this. If this information existed it would be possible to analyse the correlation between abuse and test scores, however this data is not available in SACMEQ.

5.12 Community Environment

Table 18: Descriptive Data for Community Environment, 2007

Conceptual variable	Name of variable tested	Non missing obs	10th percentile and 90th percentile	Median, mean	s.d.	Median, mean (Female)	Mean CI (95%) (Female)	Median, mean (Male)	Mean CI (95%) (Male)
Problems with community	Sprobcom	2781	1,3	2, 1.77	0.62	2, 1.77	1.74 to 1.80	2, 1.76	1.73 to 1.80
Community contributes (building, textbooks)	zscommneg	2781	1,3	2,2	0.83	2, 1.98	1.94 to 2.03	2, 2.01	1.98 to 2.06
Community pays bonus to school staff	zscomm11	2781	0,0	0, 0.09	0.29	0, 0.09	0.07 to 0.10	0,0.1	0.08 to 0.12

SACMEQ III, own analysis

18

The analysis found that 34 per cent of pupils attended schools where there were no problems between the school and the community; 56 per cent were at schools which had minor problems; 10 per cent attended schools that experienced major difficulties between the school and the community. The data shows that most pupils went to schools where the community contributed to two out of four of the following: building facilities, maintenance of facilities, purchase of text books, payment of salaries of additional teachers. The data also shows that only approximately 9 per cent of pupils were at schools where the community paid a bonus to non-teaching school staff.

Again it should be noted that many variables that were highlighted in the conceptual model did not have data available and so are not included in the regressions models. For example, there is no data to measure the attitude in the community towards education and value placed on it for males and females, or the opportunity to find work after education for males and females. The many missing variables mean the R2 in the regressions is expected to be low.

In summary, the data has shown that male and female pupils come from very similar backgrounds, attend similar schools, learn in similar classrooms, with similar teachers, head teachers, school management, school environment and community environment. The only observed variables where there were significant differences between male and female pupils were for: test scores, age, household chores, frequency of meals, whether English is spoken in the home, and pupil absenteeism. The next section looks at how these variables are correlated with test scores and each other.

6 Correlations

6.1 Pupil Characteristics

Table 19: Correlations of Test Scores with Pupil Characteristics, 2007

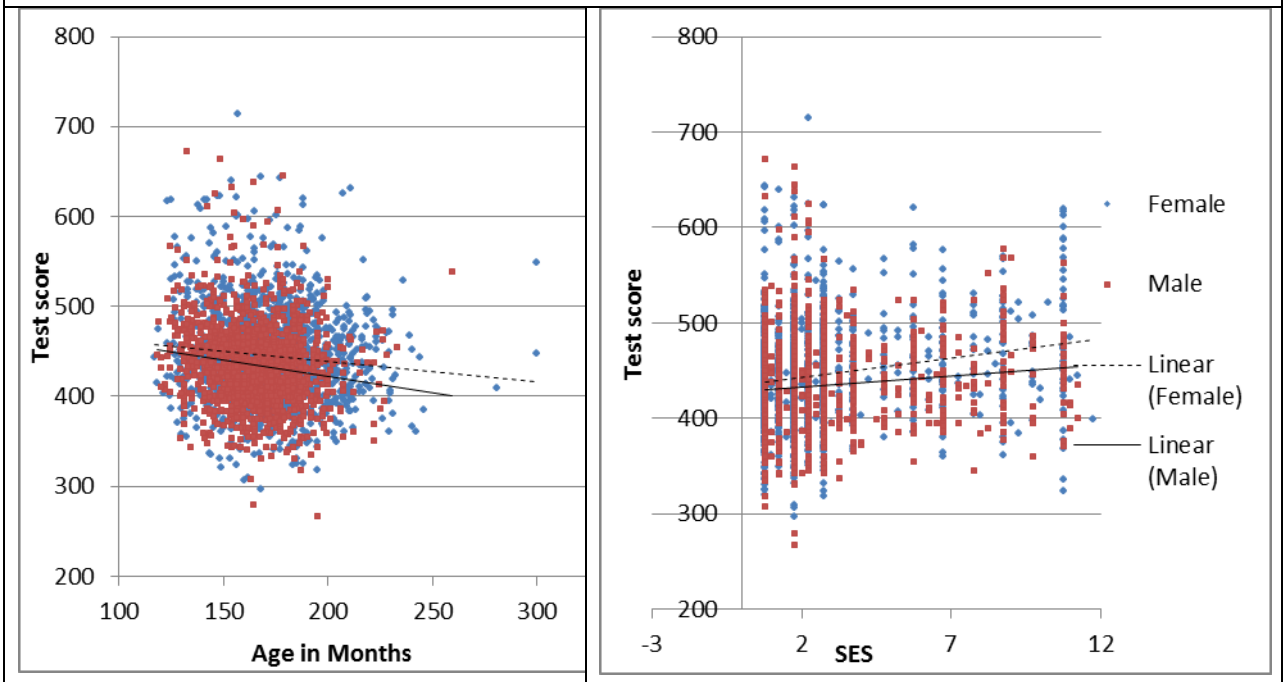
Conceptual variable	Name of variable tested	Non missing obs	Corr. for F and M	Corr. for F	Confidence Interval (CI) for F	Corr. for M	Confidence Interval (CI) for M
Gender	Zpsex	2781	-0.1113 (p 0.000)	Na		Na	
Age	zpagemon	2781	-0.1096 (p 0.000)	-0.1626 (p 0.000)	-0.214 to -0.110	-0.1149 (p 0.000)	-0.166 to -0.063
SES	psesnolig12	2781	0.1595 (p 0.000)	0.1156 (p 0.000)	0.063 to 0.168	0.1993 (p 0.000)	0.149 to 0.249

SACMEQ III, own analysis

19

As expected from the literature review, the correlation between pupil gender and test scores was significant with a correlation of (-0.11) suggesting that female pupils performed worse than male pupils. The correlations between test scores and age are significant: they are higher when the correlation was run for the sample of female pupils, however not significantly different as shown by the overlapping confidence intervals. SES has been found to be significantly correlated to test scores for both males and females (0.16). Here there is a stronger correlation for males (0.19) than for females (0.12) but again the difference is not significant.

Figure 16: Scatter Graphs of Test Scores and Age and SES of Pupils, 2007



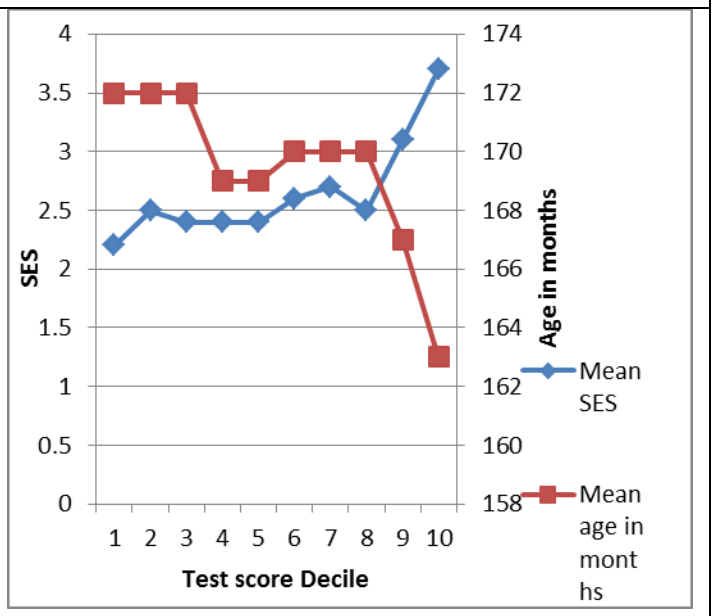
SACMEQ III data, own analysis

16

The scatter graphs in Figure 16 illustrate test scores' negative relationship with age and positive relationship with SES. However, the scatter plots do not show a strong relationship. It can be seen that the majority of pupils have SES of below 4 and for test scores over 600 it can be seen that most of these scores are from pupils with a low SES. Looking at the data by test score deciles is therefore useful.

Figure 17: Mean age and SES by Test Score Decile, 2007

Test score Decile	Mean Test score	Mean age in months	Mean SES
1	364	172	2.2
2	393	172	2.5
3	408	172	2.4
4	421	169	2.4
5	432	169	2.4
6	442	170	2.6
7	453	170	2.7
8	467	170	2.5
9	486	167	3.1
10	536	163	3.7



SACMEQ III data, own analysis

17

Figure 17 has the age and SES means for the weighted observations in each test score decile. This table and graph shows that the pupils who score the highest and are in the 9th and 10th test score deciles are on average younger and from a higher SES. However, for deciles 1-8 the relationship between age is not as clear and the mean age in deciles 4 and 5 is lower than mean age in deciles 7, 8, 9. The relationship between SES and test sores in test score deciles 2 to 8 is also not very clear and the mean SES for decile 8 is the same as the mean SES in decile 2. When the data is sorted into SES deciles it also shows no clear relationship with test scores until the 9th and 10th decile.

Table 20: Mean Score of Female and Male by Age Decile, 2007						
Age decile	**Age in months (decile ranked)**	**Ratio of female to male pupils in age decile**	**Mean score for females in Age ranked decile**	**Mean score for males in Age ranked decile**	**Difference in test scores**	**Standard Error of difference**
1	134 (11yrs)	1.28	453	464	-11	4.60
2	147 (12yr)	1.36	444	456	-12	4.43
3	155 (12 yrs)	1.44	435	448	-12	4.45
4	160 (13 yrs)	1.09	434	443	-9	3.74
5	165 (14yrs)	1.04	428	447	-19	4.24
6	171 (14yrs)	1.25	429	443	-14	3.97
7	177 (14yrs)	1.16	433	447	-15	3.94
8	183 (15 yrs)	0.71	430	438	-8	3.54
9	190 (15yrs)	0.65	422	443	-21	4.15
10	210 (17 yrs)	0.33	430	437	-8	3.54
SACMEQ III data, own analysis						

[20]

Table 20 shows that when the data is ranked by age, and sorted into age deciles, then the data shows that there are more females than males in the first 1-7 age deciles, and more males than females in the deciles 8-10, i.e. most of the older pupils are males, as shown previously by the kernel density graph. The data in Table 20 also shows that at all age deciles female students perform less well in the test than males, and that as pupils get older their scores tend to be lower. However, there is no evidence to suggest that older female students' do worse compared to older male students than the younger female students do compared to younger male students. Thus there is no evidence to suggest that it is older females with very bad test scores who are the cause in the overall gender differences in test scores.

Figure 17 shows that in SES decile 10 female pupil score is much higher than any other deciles by 11 points. However, from deciles 1 to 9 there is no apparent relationship between test score and SES. For males in deciles 1-8 there is no clear relationship; however, in decile 9 the average score is 8 points higher than other deciles, and decile 10 is then 15 points higher than decile 9. This supports the finding in Table 14 that males have a higher correlation with SES than females. The data also shows that in all SES deciles, apart from decile 4, male pupils' average score is higher than females'. The data also shows that the difference between males' and females' test

scores increases as SES increases. This backs up Hungi (2011) who found that unlike most other SACMEQ countries, in Malawi, as SES increases the gender inequality in test scores increases, Thus there is no evidence to suggest that gender differences in education in Malawi would reduce if there were an overall increase in living standards and SES.

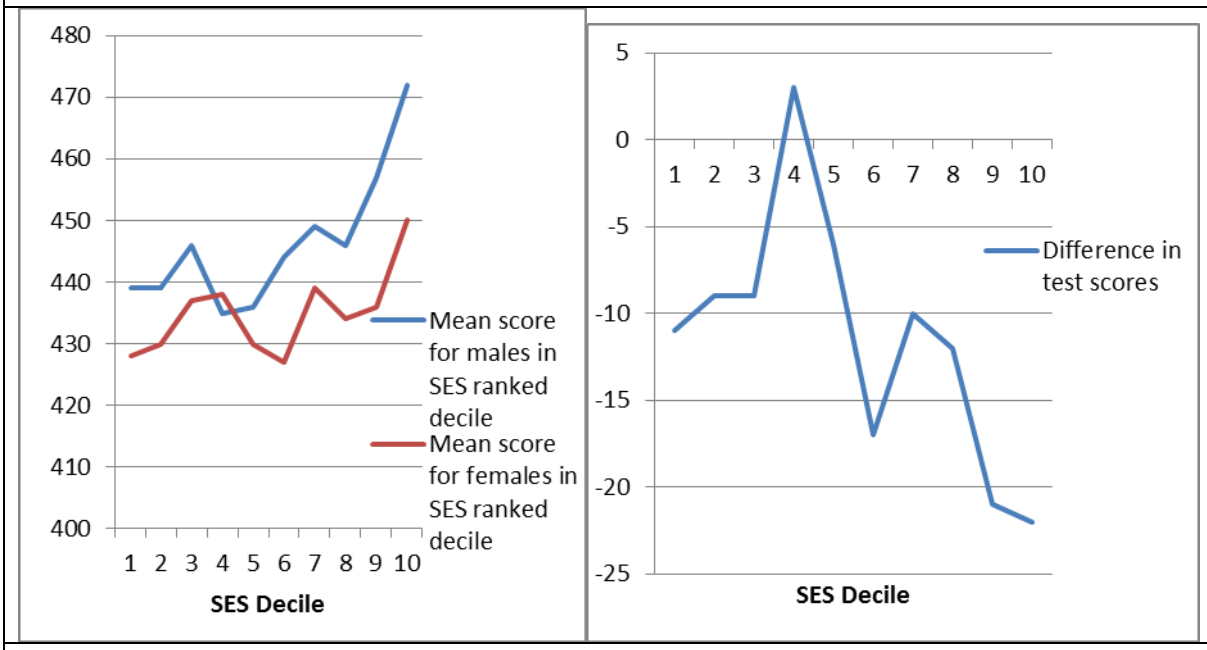
Table 21: Mean Score of Female and Male Pupils in SES Deciles, 2007

SES Decile	SES (decile ranked)	Ratio of females to males in decile	Mean score for females in SES ranked decile	Mean score for males in SES ranked decile	Difference in test scores	Standard Error
1	.8	0.9	428	439	-11	3.83
2	.8	0.9	430	439	-9	3.84
3	1.1	1.1	437	446	-9	3.99
4	1.7	0.9	438	435	3	3.93
5	1.8	0.9	430	436	-6	4.20
6	2.0	1.1	427	444	-17	4.15
7	2.6	1.0	439	449	-10	4.34
8	2.8	1.0	434	446	-12	3.71
9	4.4	1.0	436	457	-21	3.76
10	8.7	1.0	450	472	-22	4.55

SACMEQ III data, own analysis

21

Figure 18: Mean Scores for Female and Male Pupils by SES Deciles, 2007



SACMEQ III data, own analysis

18

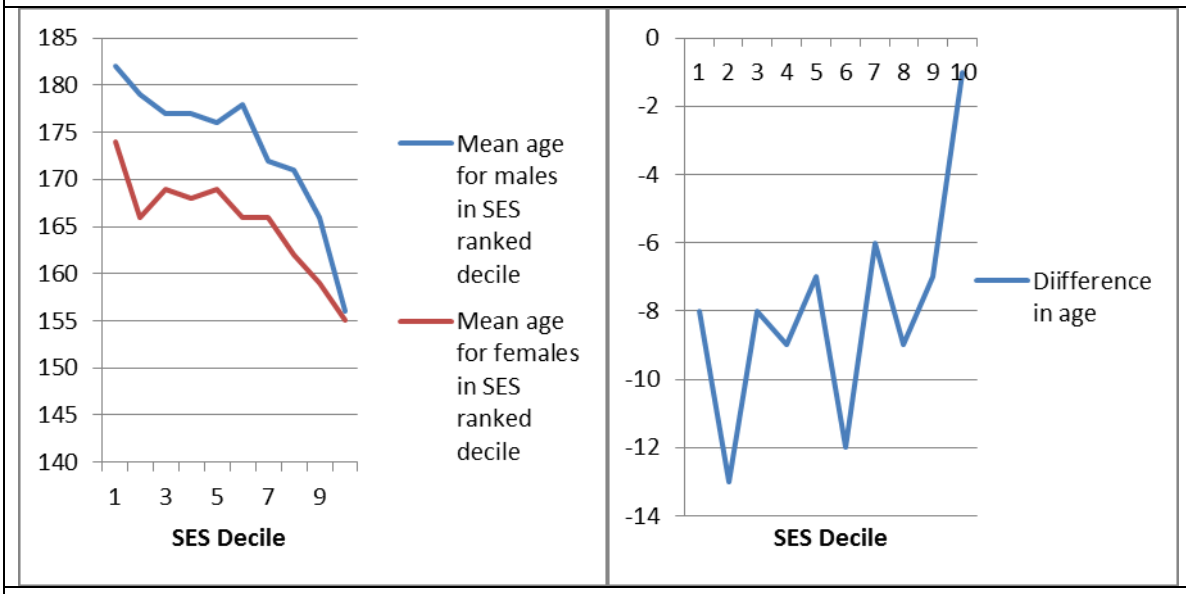
When the correlations between all variables were run (see Appendix 2 for the correlation matrixes), particular attention was shown to any correlations whose absolute values were 0.3 or above. It was found that SES had a significant negative correlation of approximately -0.3[25] with pupil age. When interpreting the results of regressions, this correlation will be noted. This negative relationship is also found when the data is split into SES deciles and the mean age is looked at, see Table 22.

[25]These factors were controlled for in the regressions so the affect age has on scores independently of these factors could be seen. Note that although this significant (at 0.001 level) correlation exists, correlations of this magnitude are not deemed large enough to result in multi-collinearity problems therefore both variables were kept in the regressions.

Table 22: Mean Age of Female and Male Pupils in SES Deciles, 2007					
SES (decile ranked)	**SES (decile ranked)**	**Ratio of females to males in decile**	**Mean age for females in SES ranked decile**	**Mean age for males in SES ranked decile**	**Difference in age**
1	0.8	0.9	174	182	-8
2	0.8	0.9	166	179	-13
3	1.1	1.1	169	177	-8
4	1.7	0.9	168	177	-9
5	1.8	0.9	169	176	-7
6	2	1.1	166	178	-12
7	2.6	1.0	166	172	-6
8	2.8	1.0	162	171	-9
9	4.4	1.0	159	166	-7
10	8.7	1.0	155	156	-1
SACMEQ III data, own analysis					

22

Figure 19: Mean Age for Female and Male Pupils by SES Deciles, 2007



SACMEQ III data, own analysis

19

Table 22 and Figure 19 show that at every SES decile the mean female age is lower than the mean male age, but that the gap narrows as SES increases and in the last decile it shows that there is only a very small, 1 month, difference. Even the highest difference is only just over a year difference, decile 2. The data in Table 22 shows also that there are almost the same ratios of female and male pupils in each SES decile. This again gives no evidence that the reason for female pupils doing relatively worse than males is because they are of a different SES composition.

SES was correlated with many other explanatory variables. However, for some of these variables the relationship between them and SES was bidirectional or may be due to a third common causation variable. For example, the direction of the relationship between academic support in the home and SES (0.48 correlation) is likely to run in both directions. Location and SES are correlated (0.41); however, it may be that a third factor, such as the availability of jobs in certain locations, could have resulted in the correlation between these two variables. Apart from academic support and location, there were some other variables that SES has a correlation with that are 0.3 or above: SES and proportion of females at school (0.44), SES and school head teaching hours (-0.28), and SES and school head gender (0.26). Thus pupils from a higher SES background are more likely to attend schools which have a higher ratio of female pupils, where the school head teaches less and the head teacher is more likely to be female.

It should be noted that, as the correlation matrixes in Appendix 2 show, SES is also significantly correlated with all of the variables that have been considered, apart from: pupil gender, repetition, head teacher age, free school meals, public or private school, and problems with the community. This highlights the important relationship of SES and other variables that can, in turn, have a correlation with test scores, even if SES is not significant in the regressions, as Hungi (2011) found, for Malawi.

6.2 Home Environment

Table 23: Correlations of Test Scores with Home Environment, 2007

Conceptual variable	Name of variable tested	Non missing obs	Corr. for F and M	Corr. for F	CI for F	Corr. for M	CI for M
Frequency of meals	Zpmealw	2781	0.0813 (p 0.000)	0.084 (p 0.000)	0.031 to 0.136	0.0931 (p 0.000)	0.041 to 0.145
Academic support:(parent education & books at home)	zpparbook59	2781	0.1277 (p 0.000)	0.1051 (p 0.001)	0.052 to 0.157	0.1542 (p 0.000)	0.101 to 0.203
English spoken at home	Zpenglis	2781	0.1414 (p 0.000)	0.0870 (p 0.0013)	0.034 to 0.139	0.1754 (p 0.000)	0.124 to 0.225
Extra tuition	Zpexttui	2781	0.0664 (p 0.0005)	0.0809 (p 0.0028)	0.028 to 0.133	0.0638 (p 0.0164)	0.012 to 0.115
Chores done in the home	Ptasktotg	2781	-0.1560 (p 0.000)	-0.1615 (p 0.0000)	-0.213 to -0.109	-0.1244 (p 0.0000)	-0.175 to -0.073
Access to a computer	Zpcomptr	2781	0.1494 (p 0.0000)	0.0975 (p 0.0003)	0.045 to 0.150	0.1775 (p 0.0000)	0.127 to 0.227
Light type in the home	Zplight	2781	0.1741 (p 0.000)	0.1414 (p 0.000)	0.089 to 0.193	0.2129 (p 0.0000)	0.163 to 0.262

SACMEQ III, own analysis

23

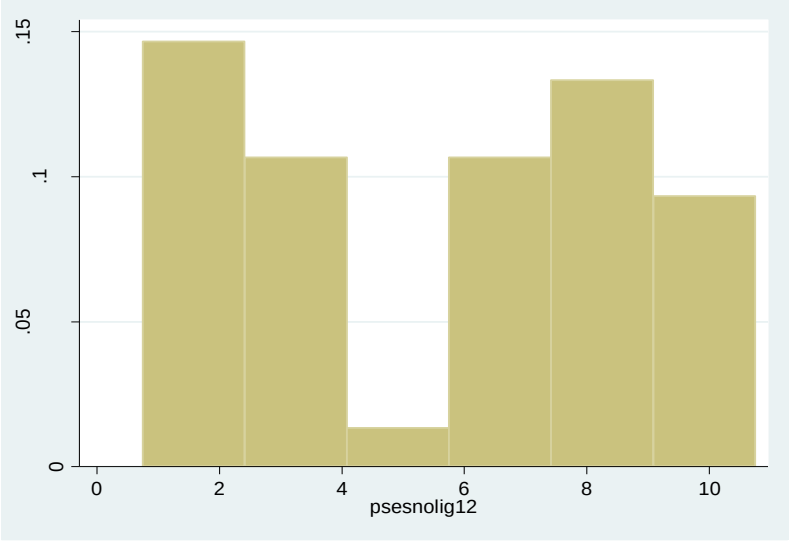
The analysis shows that meals are positively correlated with scores (0.08) but that for this variable as with all other variables in the table, there are no statistically significant differences between the correlations for males and females. However it is still interesting to note that for the variable for language spoken in the home the correlation for males alone (0.18) is much larger than that of females (0.09), and that the result has a higher level of significance for males. This may suggest that this variable is representing something other than a student's ability to understand and perform in the SACMEQ exams which are in English.

Tuition outside of school is significantly and positively correlated with test scores. Participation in household chores outside school is significantly and negatively correlated for both males and

females and again, with a higher correlation for females (-0.16) than for males (-0.12) although the difference is not significant. The correlation matrixes in Appendix 2 shows that chores at home have a stronger correlation with SES for females (-0.27) than for males (-0.19), suggesting that as pupils' SES gets lower, the need for females to complete household chores increases more than for males.

The variable for whether a pupil had used a computer before is correlated to pupil scores for both males and females (0.15), for females only (0.10) and for males only (0.18); these are all statistically significant at the 0.001 level. The correlation matrixes in Appendix 2 show that whether a pupil had used a computer before is significantly correlated to SES at a correlation of 0.17. Figure 20 shows the values that SES takes for the pupils who had used a computer before; it also shows that pupils from a range of SES backgrounds had access to a computer, not just those from a high SES background as was expected.

Figure 20: SES for Pupils Who Have Had Access to a Computer, 2007



SACMEQ III, own analysis

20

It is noted that, like SES, academic support (parents' education and books at home) and light type at home, have higher correlation coefficients for male pupils than female pupils, however they are not significantly different as the confidence intervals overlap. The correlations for academic support (parents' education and books at home) with test scores at home are

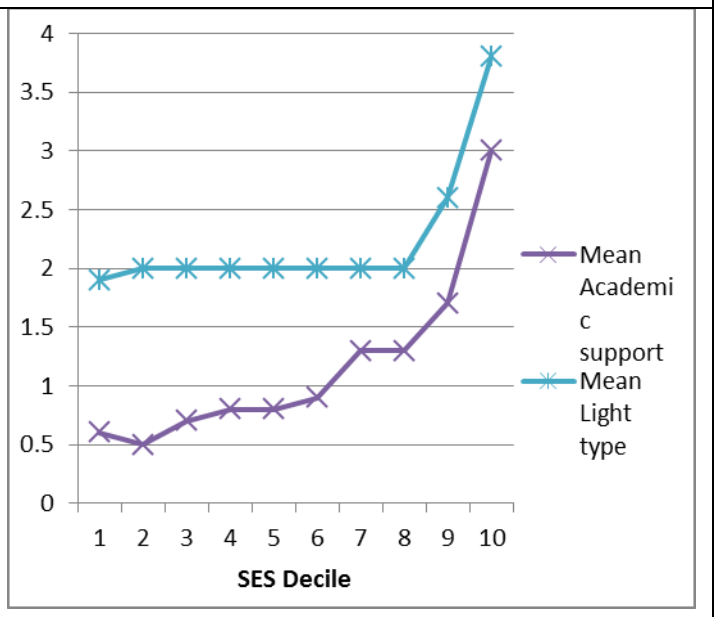
statistically significant at the 0.001 level. The variable for light type is also statistically significantly correlated at the 0.001 level, with correlations of 0.17 for males and females, 0.14 for females and 0.21 for males.

Academic support and light type both have a high correlation with SES, (see Appendix 2) and therefore they are given special attention. Academic support (parents' education and books at home) has a correlation of 0.48 with SES whilst the variable for light type in a pupil's home has an even correlation with SES of 0.77.

Figure 21 shows that light type in the house stays at a mean value of 2 (oil or paraffin lamp) and then at SES decile 9 and 10 the mean value for light type increase up to nearly 4 (electricity) at the 10th decile. Mean academic support (parents' education and books at home) increases gradually as SES decile increase and then increases sharply for the 9th and 10th decile.

Figure 21: Mean Academic Support and Light Type by SES Decile, 2007

SES Decile	Mean SES	Mean Academic support	Mean Light type
1	.8	.6	1.9
2	.8	.5	2.0
3	1.1	.7	2.0
4	1.7	.8	2.0
5	1.8	.8	2.0
6	2.0	.9	2.0
7	2.6	1.3	2.0
8	2.8	1.3	2.0
9	4.4	1.7	2.6
10	8.7	3.0	3.8



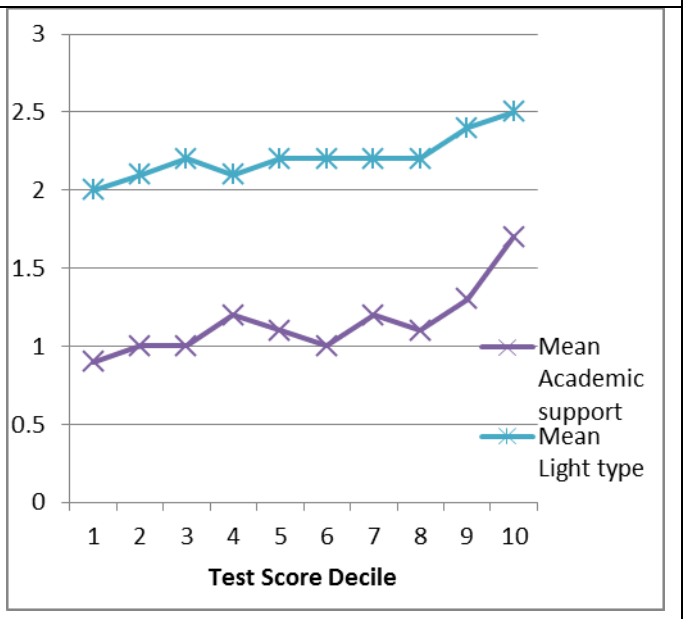
SACMEQ III data, own analysis

21

The graph and table show that for test scores deciles 9 and 10 there is a positive relationship with test scores and academic support and light type in the home.. This follows the same trend as seen for SES, as expected.

Figure 22: Mean Academic Support and Light Type by Test Score Decile, 2007

Test score Decile	Mean Test score	Mean Academic support	Mean Light type
1	364	0.9	2.0
2	393	1.0	2.1
3	408	1.0	2.2
4	421	1.2	2.1
5	432	1.1	2.2
6	442	1.0	2.2
7	453	1.2	2.2
8	467	1.1	2.2
9	486	1.3	2.4
10	536	1.7	2.5



SACMEQ III data, own analysis

22

The variables, for academic support (parents' education and books) and light type in the home, will be included in all the regressions, whilst being cautious of their relationship with SES. Therefore, there were checks for multi-collinearity of these variables in the regressions, and an awareness that, if the variable for SES is not included in the model then these two variables may pick up some of the effects of SES, rather than just measuring light available to study by and academic support available in the home.

6.3 Pupil Academic Record

Table 24: Correlations of Test Scores with Pupil Academic Record, 2007

Conceptual variable	Name of variable tested	Non missing obs	Corr. for F and M	Corr. for F	CI for F	Corr. for M	CI for M
Grade repetition	prepeatsq	2781	-0.0795 (p 0.0000)	-0.0926 (p 0.0006)	-0.145 to -0.040	-0.0757 (p 0.0044)	-0.127 to -0.024
Days absent	pabsent	2781	-0.0784 (p 0.000)	-0.0518 (p 0.0556)	-0.105 to 0.001	-0.1114 (p 0.000)	-0.163 to -0.060

SACMEQ III, own analysis

[24]

Grade repetition has a correlation of -0.08 for males and females, with a correlation of -0.08 for males and -0.09 for females. Days absent is statistically significantly correlated for males at the 0.001 level, with a correlation of -0.11. However, for females it is not even statistically significantly correlated at the 0.05 level. It is unclear why these differences in significance of correlations exist.

6.4 Teacher Characteristics

Table 25: Correlations of Test Scores with Teacher Characteristics, 2007

Conceptual variable	Name of variable tested	Non missing obs	Corr. for F and M	Corr. for F	CI for F	Corr. for M	CI for M
Teacher living conditions (maths)	Ycondliv	2780	0.0417 (p 0.0278)	0.0585 (p 0.0307)	0.005 to 0.111	0.0254 (p 0.3389)	-0.027 to 0.077

SACMEQ III, own analysis

[25]

In the analysis it was found that the living conditions of the maths teachers, a proxy for their SES, was correlated with test scores at the 0.05 level of significance but not the 0.01 level of significance, for the sample of female pupils and the sample of both female and male pupils. It was not significant at any level for the male only sample.

6.5 Quality of Teaching

Table 26: Correlations of Test Scores with Quality of Teaching, 2007

Conceptual variable	Name of variable tested	Non missing obs	Corr. for F and M	Corr. for F	CI for F	Corr. for M	CI for M
Hours of teaching (1= above 30 hours)	zhrteachdum4	2781	-0.1140 (p 0.0000)	-0.1120 (p 0.000)	-0.164 to -0.059	-0.1183, (p 0.000)	-0.169 to -0.067
Knowledge of teachers	zaloctmr	2744	0.0418, (p 0.0285)	0.0922 (p 0.0007)	0.039 to 0.145	-0.0005 (p 0.9856)	-0.053 to 0.052
Homework given	phmwk	2781	0.1231 (p 0.0000)	0.1594 (p 0.0000)	0.107 to 0.211	0.0929 (p 0.0005)	0.041 to 0.144
Maths teacher interacts with parents	ymeetpar	2781	0.0778 (p 0.0040)	0.0110 (p 0.6804)	0.025 to 0.130	0.0110 (p 0.6804)	-0.041 to 0.063
Teachers' days absent	teachabs	2781	-0.0547 (p 0.0039)	-0.0446 (p 0.0997)	-0.097 to 0.009	-0.0656 (p 0.0135)	-0.117 to -0.014
Teachers' years of training	zqprodmr	2781	0.0889 (p 0.000)	0.0716 (p 0.0082)	0.019 to 0.124	0.1052 (p 0.0001)	0.053 to 0.156

SACMEQ III, own analysis

26

The correlations for the hours taught dummy variable are statistically significant at the 0.0001 level, all with correlations of approximately 0.11. Interestingly, the knowledge of teachers has no statistical significance on the test scores of males (p value of 0.9856) but has a statistically significant correlation with females' scores (p value of 0.0007).

Frequency of homework is statistically significantly correlated with pupils' test scores; with the correlation for females at 0.16 compared to 0.09 for males, however there is a small overlap in confidence intervals so we cannot conclude the difference is significant.

The analysis looked at how often maths and reading teachers met with the parents or guardians of pupils and it has found that this variable is only correlated with the maths teacher meeting regularly with female pupils' parents or guardians but not the males.

Absenteeism of teachers is significantly correlated (0.07) for males at the 0.05 significance level but not the 0.01 level and not at all for females.

The years of teacher training variable has a positive correlation with test scores, with a correlation of 0.11 for males and 0.07 for females.

6.6 School Head Characteristics

Table 27: Correlations of Test Scores with School Head Characteristics, 2007

Conceptual variable	Name of variable tested	Non missing obs	Corr. for F and M	Corr. for F	CI for F	Corr. for M	CI for M
Head teacher gender	zssex	2781	0.0501 (p 0.0083)	0.0603 (p 0.0260)	0.007 to 0.113	0.0496 (p 0.0621)	-0.003 to 0.101
Head teacher age	sage	2781	0.0705 (p 0.0002)	0.0811 (p 0.0027)	0.028 to 0.134	0.0621 (p 0.0195)	0.010 to 0.114

SACMEQ III, own analysis

27

The gender of head teacher was found to be significant at the 0.05 level for females but not significant for males.

The data shows that there is a positive correlation (0.07) with the age of a head teacher and the scores of pupils that is statistically significant at the 0.001 level.

6.7 Quality of School Management

Table 28: Correlations of Test Scores with Quality of School Management, 2007							
Conceptual variable	**Name of variable tested**	**Non missing obs**	**Corr. for F and M**	**Corr. for F**	**CI for F**	**Corr. for M**	**CI for M**
Head teacher years experience (1= 13 years or above experience)	Sexpalldum	2781	-0.0645 (p 0.0007)	-0.0618 (p 0.0225)	-0.114 to -0.009	-0.0695 (p 0.0089)	-0.121 to -0.018
Head teacher teaching hours	Zstchhrs	2781	-0.1624 (p 0.000)	-0.1389 (p 0.0000)	-0.191 to -0.086	-0.1870 (p 0.0000)	-0.237 to -0.136
SACMEQ III, own analysis							

28

The dummy variable for head teacher experience (takes the value 1 if a head teacher has been a head teacher for 13 years or more) is significantly negatively correlated with pupil scores at the 0.05 level for both males and females. Experience is usually considered to be positive, so this correlation is unexpected. If this variable is also found to be negatively correlated with test scores in the regressions then more investigation would be required in order to make observations about why this correlation occurs.

The variable for hours taught by a head teacher is also negatively correlated with test score which corresponds to Hungi's (2011) findings. A plausible explanation for this is that the more hours a head teacher had to teach classes, the fewer hours they had for managing the school. However, there may be other explanations for this correlation, such as head teachers being required to teach when there is a shortage of teachers. This would support the finding that there is a correlation of 0.45 between pupil teacher ratio and head teacher teaching hours. The correlations are all statistically significantly correlated at the 0.0001 level.

6.8 Classroom Environment

Table 29: Correlations of Test Scores with Classroom Environment, 2007

Conceptual variable	Name of variable tested	Non missing obs	Corr. for F and M	Corr. for F	CI for F	Corr. for M	CI for M
Class size	Clsize	2781	-0.0313 (p 0.0991)	-0.0541 (p0.0456)	-0.107 to -0.001	-0.0079 (p0.7666)	-0.060 to 0.044
Pupil/teacher ratio	Zsptrati	2781	-0.1801 (p 0.0000)	-0.1736 (p 0.0000)	-0.225 to -0.122	-0.1926 (p 0.0000)	-0.242 to -0.142

SACMEQ III, own analysis

29

Class size is only statistically significantly correlated for females at the 0.05 level, whereas for males it is not at all statistically significant.

Pupil/teacher ratio is statistically significantly correlated for both males and females at the 0.001 level. Additional analysis found that for the SACMEQ III data for Malawi, pupil/teacher ratio stops being significantly correlated to test scores if pupil/teacher ratio is below 50; however, above this level it is still significant. This is consistent with Boissiere's (2004) report which stated that below a pupil teacher ratio of 40, pupil teacher ratio did not significantly affect pupil scores but over this level it could be influential.

6.9 School Characteristics

Table 30: Correlations of Test Scores with School Characteristics, 2007

Conceptual variable	Name of variable tested	Non missing obs	Corr. for F and M	Corr. for F	CI for F	Corr. for M	CI for M
School resources	resource 1	2781	0.1579 (p 0.0000)	0.1440 (p 0.0000)	0.092 to 0.196	0.1746 (p 0.0000)	0.124 to 0.225
State of school building	zsbldgco	2781	-0.1367 (p 0.0000)	-0.1554 (p 0.0000)	-0.207 to -0.103	-0.1198 (p 0.0000)	-0.171 to -0.068
Distance to school	ptravel	2781	-0.0586 (p 0.0020)	-0.0627 (p 0.0207)	-0.115 to -0.010	-0.0537 (p 0.0435)	-0.105 to -0.002
School location	zsloc	2781	0.1509 (p 0.0000)	0.1376 (p 0.0000)	0.085 to 0.189	0.1670 (p 0.0000)	0.116 to 0.217
Free meals at school	zpmealsc	2781	-0.0791 (p 0.0000)	-0.0691 (p 0.0107)	-0.122 to -0.016	-0.0859 (p 0.0012)	-0.137 to -0.034
Type of school (public/private)	zstype	2781	0.0632 (p 0.0000)	0.0379 (p 0.1618)	-0.015 to 0.091	0.0793 (p 0.0028)	0.027 to 0.131

SACMEQ III, own analysis

30

School resources has a correlation of 0.16 with pupil scores and is statistically significant at the 0.001 level. Another variable that has been found to be statistically significantly correlated with pupil scores, with a correlation of -0.14 for males and females, is the state of the school building.

The distance a pupil had to walk to school has a statistically significant correlation with pupil scores for males and females at the 0.05 level. Whether a pupil's school was classified as being in an isolated or rural place or in a town or city is highly significantly correlated for all samples. The variable 'school location' is also shown in Appendix 2's correlation matrixes to have strong and statistically significant relationships with various other factors whose association with pupil scores is being assessed. Correlation scores of 0.3 and above have been found between location and 9 variables, which are all also statistically significantly correlated with each other:

1. Academic support: parents' education & books. 0.3
2. SES 0.4
3. Proportion of female teachers 0.6
4. Teacher gender 0.5
5. School head gender 0.4
6. School head teaching hours -0.5
7. Teachers hours 0.4
8. Pupil teacher ratio -0.4
9. Class size 0.5

These correlations show that: firstly, pupils with higher SES and higher parent education tend to go to schools in urban locations; secondly, that urban schools have a higher proportion of female teachers and head teachers; thirdly, that there are higher teacher shortages (higher pupil teacher ratios) in rural schools despite the smaller class sizes. It is likely that these shortages explain the fourth trend, that both teachers and head teachers in rural schools teach longer hours than those who work in urban schools.

The analysis also found that, although the free school meals variable was not highlighted in the literature review, it is statistically correlated with pupil scores. However, the direction of the relationship was not as first expected. The purpose of the free school meals program in Malawi was to improve the performance of poor students who may not get enough nutrition at home. However, the correlation for males and females is -0.08 and is statistically significant at the 0.001 level. This may be because poor schools were selected to receive free meals. However, the correlation matrix in Appendix 2 does not show a statistically significant correlation for free school meals with pupils' SES, academic support (parents' education and books at home), or school resources. As the interpretation of this variable is unclear, this variable is added in one set of regressions and excluded in another set. Where this variable is included in the regressions, it

is noted that this variable may not be measuring the effectiveness of school meals but instead be measuring whether or not a school had been selected to receive free school meals. Further investigation is needed to find out how a school is selected to receive free school meals, what effect this has on the pupils who attend school and how these factors correlate with test scores.

The data shows that attending a private school has a positive correlation (0.06) with test scores on average for males and females that is statistically significant at the 0.001 level. However, although the relationship is also significant for males there is no statistically significant relationship for females. The correlation matrix in Appendix 2 shows that there are positive relationships between private schools and teacher living conditions, teacher years of training, urban location, free school meals, problems with the community, and community paying bonuses to non-teaching staff at the school. It also shows that there are negative correlations with the community contributing to buildings and text books, the state of the school, and head teacher gender. However, it should be noted that all of these correlations are less than 0.20 in size and all of these variables are controlled for in the regressions.

6.10 School Environment

Table 31: Correlations of Test Scores with School Environment, 2007

Conceptual variable	Name of variable tested	Non missing obs	Corr. for F and M	Corr. for F	CI for F	Corr. for M	CI for M
Proportion of female teachers	spropf	2781	0.1201 (p 0.0000)	0.1201 (p 0.0000)	0.067 to 0.172	0.1301 (p 0.0000)	0.079 to 0.181
Safe learning environment	abusedum3	2781	0.1085 (p 0.0000)	0.0707 (p 0.0000)	0.103 to 0.206	0.1550 (p 0.0000)	0.019 to 0.122

SACMEQ III, own analysis

[31]

The proportion of female teachers at a school is statistically significantly correlated with both female and male performance. However, once again, until other variables that have correlations

with the proportion of female teachers, such as location and SES are controlled for, no conclusions can be made.

The variable to measure whether a school had a safe learning environment has a positive correlation with pupil scores and is statistically significant at the 0.001 level for males and females. The strength of the relationship for females (0.16) is larger than that of males (0.07) however there is a small overlap in the confidence intervals.

6.11 Community Environment

Table 32: Correlations of Test Scores with Community Environment, 2007

Conceptual variable	Name of variable tested	Non missing obs	Corr. for F and M	Corr. for F	CI for F	Corr. for M	CI for M
Problems with community	sprobcom	2781	-0.0655 (p 0.0005)	-0.0931 (p 0.0006)	-0.145 to -0.040	-0.0417 (p 0.1170)	-0.094 to 0.010
Community contributes (building, textbooks)	zscommneg	2781	-0.1849 (p 0.0000)	-0.1821 (p 0.0000)	-0.233 to -0.130	-0.1943 (p 0.0000)	-0.244 to -0.144
Community pays bonus to school staff	zscomm11	2781	0.0778 (p 0.0000)	0.0745 (p 0.0059)	0.022 to 0.127	0.0779 (p 0.0034)	0.026 to 0.129

SACMEQ III, own analysis

32

Schools which had problems with the community were significantly negatively correlated with pupil scores for both males and females at a correlation of -0.07; however, the correlation is not significant for males.

The variable which measured whether the community contributed to certain aspects of the school (building facility, maintenance of facility, purchase of text books, and payment of salaries of additional teachers) has a negative correlation of -0.18 which is statistically significant at the 0.001 level. However, the factors causing this result are not straight forward. The literature review suggested that it would be beneficial if the community was involved in the running of a school. The correlations in Appendix 2 show a statistically significant correlation with the community contributing to a school (building facility, maintenance of facility, purchase of text

books, and payment of salaries of additional teachers) and a school being in a bad state of repair. This suggests that the schools in the worst state of repair are the schools that receive very little contributions from the government (in terms of funds to repair buildings and other resources such as text books) and therefore the community may contribute what they can to try and improve the school. There is also a statistically significant correlation (0.22) between the community contributing to certain aspects of the school (building facility, maintenance of facility, purchase of text books, and payment of salaries of additional teachers) and teaching hours.

In contrast, the variable that measured whether the community paid an additional amount on top of the normal salary of non-teaching staff, i.e. a staff bonus, has a statistically significant positive correlation (0.08). This may be due to the community being able to measure how hard staff in the school work and thus the bonus worked as a positive incentive to staff. However, it may be that this variables positive relationship with test scores is caused by other factors. For example, unlike community contribution to building and textbooks, the community paying bonuses to staff is correlated with urban schools and higher SES. Interestingly, there is a strong significant correlation between communities who paid bonuses to school staff and whether or not a school was in receipt of free school meals. As school meals have had a negative association with scores, and this variable has had a positive relationship with scores, it suggests there are some other variables that have not been taken into account.

More investigation is required in order to discover what type of school is likely to benefit from these types of community contributions before any conclusions can be drawn. As interpretation of these variables is very unclear, like the variable for free school meals, these two variables have also been excluded from half of the regressions.

The next section of this report will consider whether there are still correlations between test scores and the variables when multivariate analysis is carried out.

7 Regression Analysis

OLS regressions were run for three sets of observations: the first set of regressions were for males and females together; the second set included the female only sample; the third set included the male only sample. The first regression for each set consisted of running a regression for all of the 37 variables and removing the variable with the highest p-value until all the variables in the regression were significant at the 5 per cent significance level. This resulted in the same variables being left in model as when both forward and backward stepwise regressions were run. Due to the risk of artificially narrow confidence intervals being created by the multi-level nature of the SACMEQ III data, a coefficient will be viewed as significant if $p< 0.01$ (rather than $p < 0.05$).

For the samples where teacher knowledge was significant in the model (female sample only), another regression was then run where this variable was excluded. As discussed previously, this variable was excluded because it had 37 missing observations which were suspected to not be randomly distributed observations. Regressions were also run for each sample, with the exclusion of three variables: community contributes (to building facility, maintenance of facility, purchase of textbooks, and payment of salaries of additional teachers), community contributes to non-teaching staff bonuses, and free school meals. As discussed in the correlation section, these three variables were excluded because the reason behind their correlation with test scores was not clear. The exclusion of these variables also acted as a robustness[26] check on other variables in the model.

As discussed in the methodology section, the standardised beta coefficients were also calculated for each regression in order to compare the impact of variables in pupils' scores. Particular attention was paid to the beta coefficients that were 0.10 and above.

[26] "A common exercise in empirical studies is a "robustness check," where the researcher examines how certain "core" regression coefficient estimates behave when the regression speciation is modified by adding or removing regressors. If the coefficients are plausible and robust, this is commonly interpreted as evidence of structural validity." (White & Lu (2010)).

7.1 Results

Table 33: Regression Results for Male and Female Samples				
	Regression 1		**Regression 2**	
Excluded regardless of significance:	None		Community contributes, community pays bonus, free meals at school.	
Gender of sample	Male and female		Male and female	
N	2780		2780	
R^2	0.202		0.19	
adj. R^2	0.195		0.184	
Output=	Scoremr		Scoremr	
Coefficient reported.		beta		beta
1) Pupil Characteristics				
Gender	-9.648***	-0.099***	-10.49***	-0.108***
Age			*-0.0887*	*-0.040*
SES				
2) Home Environment				
Frequency of meals	*1.581***	*0.046***	*1.521*	*0.044*
Academic support at home: parent education and books in the home.				
English spoken at home	9.205***	0.095***	8.828***	0.091***
Extra tuition	*5.857*	*0.042*		
Chores done in the home	-1.040**	-0.062**	-0.991**	-0.059**
Access to a computer	29.65***	0.077***	31.25***	0.081***
Light type in the home	4.404**	0.063**	4.505**	0.065**
3) Pupil Academic Record				
Grade repetition	-0.698**	-0.055**	-0.660**	-0.052**
Days absent	-1.519***	-0.072***	-1.515***	-0.072***
4) Teacher Characteristics				
Teacher living conditions (maths)				
5) Teacher Quality				
Hours of teaching (1= above 30 hours)	-11.07***	-0.105***	-11.87***	-0.113***
Knowledge of teachers				
Homework given	6.044***	0.099***	5.786***	0.094***
Maths teacher interacts with parents	*1.963*	*0.040*		
Teachers days absent	*-0.142****	*-0.063****	*-0.137****	*-0.061****
Teachers years of training	*3.479****	*0.100****	*3.493****	*0.101****
6) School Head Characteristics				
Head teacher gender				
Head teacher age	*0.797****	*0.086****	*0.631****	*0.068****
7) School Head Quality Of Management				
Head teacher years experience (1= 13 years or above experience)	*-12.61****	*-0.088****	*-11.08****	*-0.077****

Table 33: Regression Results for Male and Female Samples

	Regression 1		Regression 2	
Head teacher teaching hours				
8) Classroom Environment				
Class size	-0.266***	-0.172***	-0.266***	-0.171***
Pupil teacher ratio	*-0.0817****	*-0.073****	*-0.104****	*-0.093****
9) School Characteristics				
School resources	2.525***	0.104***	3.104***	0.127***
State of school building	*-5.902***	*-0.061***	*-7.901****	*-0.081****
Distance to school				
School location				
Free meals at school	-8.337***	-0.074***		
Type of school (public/private)	*6.057***	*0.058***	*7.783****	*0.074****
10) School Environment				
Proportion of female teachers				
Safe learning environment			*4.567*	*0.039*
11) Community Environment				
Problems with community	-7.799***	-0.100***	-7.859***	-0.101***
Community contributes (building, textbooks)	*-3.473***	*-0.060***		
Community pays bonus to school staff	17.93***	0.108***		
Constant	**406.7*****		**428.9*****	

Notes: ** denotes that the coefficient is significant at p < 0.01 and *** denotes that the coefficient is significant at p < 0.001. If there is no star then the coefficient is only significant at p < 0.05.

The font is in italics for variables which did not show as significant when SVY command was used in the regression to adjust for the complex sampling design.

SACMEQ III, own analysis

33

For the sample of males and females, the variable, knowledge of teachers, was not found to be so significant for males and females in the regressions, therefore, the number of observations in the regressions is 2780, i.e. only 1 observation for which there was no data[27].

The first regression found that the following variables were not significant at the 0.01 level: age, SES, academic support from parents, extra tuition, teacher living conditions, teacher interaction with parents head teacher gender, head teacher teaching hours, distance to school, school location, proportion of female teachers, and safe learning environment. The following additional variables (shown in italics in the table) were not found significant in the regressions if an

[27] See methodology section for more details on SACMEQ imputation process of some missing values.

adjustment for the complex sampling design was made: frequency of meals, teacher days absent, head teacher age, and community contributes. The second regression found similar results as regression one. However, as expected when statistically significant variables were removed from the regressions, the R2 of this second model was lower than the first regression.

The R2 for both of these regression models is very low, at 0.20 and 0.19. This highlights that there are important explanatory variables omitted from the model, as noted throughout this report. One reason why the R2 of all models may be low is that interestingly, these regressions show that, although there were significant correlations with test scores and SES and Academic support: (parents' education and books at home), once other variables were controlled for then these variables have no statistically significant relationship with test scores. This is unusual; however, it backs up Hungi's (2011) finding which showed that SES was not statistically significant for explaining pupils' maths scores in Malawi. Light type in the home however is showing as significant and this is highly correlated with SES and academic support. As previously discussed, if this variable is the only one of the three variables showing as significant, it may be working as a proxy for SES and academic support, as well as measuring the effect of different light type in the home on test scores through a pupil's ability to study after dark.

The most significant and important variables in the regression are pupil gender, class size, whether English is spoken at home, access to a computer, how much homework is given, hours of teaching and pupil absenteeism,

Table 34: Regression Results for Female Samples

	Regression 1		Regression 2		Regression 3		Regression 4	
Excluded regardless of significance:	None		Teacher knowledge		Community contributes, community pays bonus, free meals at school.		Teacher knowledge, Community contributes, community pays bonus, free meals at school.	
Gender of sample	Female		Female		Female		Female	
N	1342		1364		1342		1364	
R^2	0.206		0.19		0.195		0.185	
adj. R^2	0.193		0.179		0.182		0.173	
Output=	scoremr		Scoremr		scoremr		Scoremr	
Coefficient reported.		beta		Beta		beta		Beta
1) Pupil Characteristics								
Gender								
Age	-0.198**	-0.084**	-0.230***	-0.097***	-0.217**	-0.092**	-0.208**	-0.088**
SES								
2) Home Environment								
Frequency of meals	*1.923*	*0.059*	*1.734*	*0.053*	*2.076*	*0.063*	*1.880*	*0.057*
Academic support:(parent education & books at home)								
English spoken at home	*4.848*	*0.054*	*5.746*	*0.064*			*4.890*	*0.054*
Extra tuition								
Chores done in the home	-1.347**	-0.085**	-1.477***	-0.094***	-1.509***	-0.095***	-1.365**	-0.087**
Access to a computer	22.58**	0.053**	21.98**	0.051**	23.17**	0.054**	19.75*	0.046*
Light type in the home								
3) Pupil Academic Record								
Grade repetition	*-0.660*	*-0.055*			*-0.775*	*-0.065*	*-0.682*	*-0.057*
Days absent								
4) Teacher Characteristics								
Teacher living conditions (maths)	*2.940*	*0.063*			*2.857*	*0.061*		
5) Teacher Quality								
Hours of teaching (1= above 30 hours)	-9.827**	-0.100**	-14.21***	-0.146***	-11.01***	-0.113***	-11.78***	-0.121***
Knowledge of teachers	*0.0323****	*0.096****			*0.0331****	*0.098****		
Homework given	5.940***	0.106***	6.284***	0.111***	6.335***	0.113***	6.841***	0.121***
Maths teacher interacts with parents	*2.960*	*0.065*			*2.699*	*0.060*		
Teachers days absent								
Teachers years of training	*2.935***	*0.091***	*2.884***	*0.089***	*2.650***	*0.083***	*3.034***	*0.094***
6) School Head Characteristics								
Head teacher Gender								
Head teacher age	*0.586*	*0.069*	*0.814****	*0.095****	*0.572*	*0.067*	*0.638***	*0.074***
7) School Head Quality Of Management								

Table 34: Regression Results for Female Samples

	Regression 1		Regression 2		Regression 3		Regression 4	
Head teacher years experience (1= 13 years or above experience)	-12.70***	-0.092***	-12.10***	-0.090***	-13.65***	-0.099***	-11.05**	-0.082**
Head teacher teaching hours								
8) Classroom Environment								
Class size	-0.247***	-0.174***	-0.308***	-0.216***	-0.232***	-0.164***	-0.268***	-0.188***
Pupil teacher ratio	*-0.0811***	*-0.076***			*-0.097****	*-0.091****	*-0.0809***	*-0.076***
9) School Characteristics								
School resources	*1.795***	*0.079***	2.229***	0.098***	2.254***	0.100***	2.422***	0.107***
State of school building	*-7.898***	*-0.088***	-9.205***	-0.102***	-8.987***	-0.100***	-10.39***	-0.115***
Distance to school			-0.853	-0.058			-0.818	-0.055
School location								
Free meals at school								
Type of school (public/private)			*5.760*	*0.058*	*6.070*	*0.061*	7.970**	0.081**
10) School Environment								
Proportion of female teachers								
Safe learning environment	*9.120***	*0.084***	*8.989***	*0.082***	*8.953***	*0.083***	*8.368***	*0.077***
11) Community Environment								
Problems with community	-8.278***	-0.114***	-9.507***	-0.131***	-7.801***	-0.108***	-8.869***	-0.122***
Community contributes (building, textbooks)	*-4.302***	*-0.079***	*-3.072*	*-0.057*				
Community pays bonus to school staff	13.44**	0.086**	16.78***	0.107***				
Constant	**396.1*****		**453.1*****		**393.0*****		**451.3*****	**0.185**

Notes: ** denotes that the coefficient is significant at p < 0.01 and *** denotes that the coefficient is significant at p < 0.001. If there is no star then the coefficient is only significant at p < 0.05.

The font is in italics for variables which did not show as significant when SVY command was used in the regression to adjust for the complex sampling design.

SACMEQ III, own analysis

34

When the regressions were run for the female sample it was found that the following variables were not significant at the 0.01 level: SES, frequency of meals, English spoken at the home, academic support from parents, extra tuition, light type in the home, grade repetition, teacher living conditions, teacher interaction with parents, teacher days absent, head teacher age, head teacher gender, head teacher teaching hours, distance to school, school location, free meals at school, proportion of female teachers and type of school. The following additional variables (shown in italics in the table) were not found significant in the regressions if an adjustment for the complex sampling design was made: knowledge of teacher, teacher years training, pupil

teacher ratio, safe learning environment, and community contribution. The R2 for this regression was 0.21. Teachers' knowledge was noted to be significant for this female sample. Therefore, because of the missing values for this variable[28], the regressions were run again but with this variable excluded. When teacher knowledge was excluded the R2 for this model dropped to 0.19.

When regression three for the female sample was run again, including the variable teacher knowledge, but excluding the variables for community contributions, community pays teacher bonus, free meals at school, then it was found that compared to the first regression for females, there were no differences in the variables that were significant at the 0.01 level. The R2 dropped, as expected, due to omitted variables, from 0.21 to 0.20.

When the regression was run excluding teacher knowledge, and also excluding three other variables, the same variables as in regression 2 were shown to be significant. The R2 for this model was not surprisingly, the lowest for the female sample at 0.19.

In summary the most important and significant variables that are correlated with female pupil test scores are: age, chores done in the home, access to a computer, hours of teaching, homework given, head teacher years of experience, class size, school resources, state of school building, problems with the community.

Again note that the R2 in all of these regressions is low, once more suggesting that important variables have been omitted. It is also interesting to note that in all these regressions for the female sample, light type in home as well as SES and academic support (parents' education and books) are not significant. As mentioned previously, it is unusual for education production functions to find no correlation between any measures of SES and test scores. Therefore although the correlation section found that the access to a computer is not strongly correlated to SES, it seems likely that some aspect of SES is being picked up by this variable. Further investigation is needed to fully understand the mechanisms behind the results here.

[28] Suspected to be non-randomly distributed.

Table 35: Regression Results for Male Samples

	Regression 1		Regression 2	
Excluded regardless of significance:	None.		Community contributes, community pays bonus, free meals at school.	
Gender of sample	Male		Male	
N	1416		1416	
R^2	0.194		0.186	
adj. R^2	0.183		0.176	
Output=	Scoremr		Scoremr	
Coefficient reported.		Beta		Beta
1) Pupil Characteristics				
Gender				
Age				
SES				
2) Home Environment				
Frequency of meals				
Academic support: (parent education & books at home)				
English spoken at home	12.62***	0.122***	12.28***	0.119***
Extra tuition				
Chores done in the home				
Access to a computer	37.99***	0.106***	38.52***	0.107***
Light type in the home	8.011***	0.107***	8.164***	0.109***
3) Pupil Academic Record				
Grade repetition	*-0.698*	*-0.053*	*-0.663*	*-0.051*
Days absent	-1.886***	-0.088***	-1.854***	-0.087***
4) Teacher Characteristics				
Teacher living conditions (maths)				
5) Teacher Quality				
Hours of teaching (1= above 30 hours)	-12.44***	-0.112***	-11.63***	-0.104***
Knowledge of teachers				
Homework given	6.257***	0.096***	5.641***	0.087***
Maths teacher interacts with parents				
Teachers days absent	*-0.186***	*-0.085***	*-0.175***	*-0.080***
Teachers years of training	3.810***	0.105***	3.909***	0.108***
6) School Head Characteristics				
Head teacher Gender				
Head teacher age	*0.658*	*0.068*	*0.678***	*0.070***
7) School Head Quality of Management				
Head teacher years experience (1= 13 years or above experience)	*-14.07***	*-0.094***	*-12.26***	*-0.082***
Head teacher teaching hours				
8) Classroom Environment				
Class size	-0.258***	-0.157***	-0.252***	-0.153***
Pupil teacher ratio	*-0.0985***	*-0.086***	-0.122***	-0.106***

Table 35: Regression Results for Male Samples

	Regression 1		Regression 2	
9) School Characteristics				
School resources	3.140***	0.123***	3.554***	0.139***
State of school building	*-5.851*	*-0.057*	*-6.677*	*-0.065*
Distance to school				
School location				
Free meals at school	*-8.159*	*-0.068*		
Type of school (public/private)	*6.730*	*0.061*	*7.480*	*0.068*
10) School Environment				
Proportion of female teachers				
Safe learning environment				
11) Community Environment				
Problems with community	*-5.890***	*-0.072***	*-6.041***	*-0.074***
Community contributes (building, textbooks)				
Community pays bonus to school staff	17.45***	0.102***		
Constant	**402.5*** **	**0.194**	**401.8*** **	**0.186**

Notes: ** denotes that the coefficient is significant at p < 0.01 and *** denotes that the coefficient is significant at p < 0.001. If there is no star then the coefficient is only significant at p < 0.05.

The font is in italics for variables which did not show as significant when SVY command was used in the regression to adjust for the complex sampling design.

SACMEQ III, own analysis

35

When the regression for the male sample was run it was found that the following variables were not significant: age, SES, extra tuition, time spent on household chores, grade repetition, teacher living conditions, maths teacher interacts with parents, head teacher age, head teacher gender, head teacher teaching hours, school location, distance to school, state of school building, type of school, free meals at school, proportion of female teachers, safe learning environment, community contributes towards building and textbooks. The following additional variables (shown in italics in the table) were not found to be significant in the regressions if an adjustment for the complex sampling design was made: teacher days absent, head teacher years' experience, problems with the community. In the second regression, where the variables community contributes, community pays teacher bonus, free meals were removed, the results were similar.

The regressions for males show that the most significant and important variables are: English spoken at home, access to a computer, light type in the home, days absent, hours of teaching,

homework given, teacher years of training, class size, school resources and whether the community pays bonus to school staff.

Note that light type in home, which is strongly correlated to SES, and academic support in the home (parents' education and books) is important and significant but that SES and academic support are not. Access to a computer is again very significant. However, even though light type is an important and significant variable in these regression models, note that the R2 is still very low at just 0.19.

7.2 Discussion of Results

7.2.1 Pupil Characteristics

The regressions for the combined female and male samples show that being female has a negative correlation with test scores, with the normal coefficient showing that females score, on average, 10 points less than males. The standardised beta coefficients for pupil gender are between 0.10 and 0.11, demonstrating that this correlation is of a relatively large magnitude. Thus, it is concluded that this variable has both a significant and important relationship with test scores in Malawi.

The regressions for females, but not for males, find that the older a person is the lower the test scores are likely to be. The results show that, on average, an increase of female pupil's age by one month is associated with a reduction in test scores of 0.2 marks. The beta coefficients are between 0.8 and 0.10 thus showing that age is an important variable for females. One reason why age has a strong relationship with female educational achievement in Malawi may be that as a female gets older the more likely they are to get married, get pregnant, be at risk from sexual abuse, have more household chores, enter puberty, etc. (Plan International 2009). All of these factors may have a negative effect on their education; therefore, age may be a proxy for these factors which have not been measured and included in the model. However, more analysis is required before it can be concluded what the mechanisms behind this correlation is.

The measure for SES used in the regressions is not found to be significant in any of the regressions. However, it is noted that SES is strongly correlated to many other variables in the model and it may be that other variables such as light type at home or access to a computer are a better measure of SES in Malawi.

7.2.2 Home Environment

The average number of meals that a pupil consumes is not significant at the 0.01 level for all regressions except the first regression for males and females. Academic support in the home, measured by parents' education[29] and the number of books in the home, is not found to be significant in any of the regressions. One of the variables that academic support in the home is statistically significantly correlated with (see Appendix 2) is whether English is ever spoken in the home. This variable is statistically significant in all the regressions apart from one of the regressions for the female sample. English being spoken in the home has the largest and most significant relationship with male pupils' test scores and has beta coefficients of approximately 0.12 and normal coefficients of between 12 and 13. When English is spoken in the home, the test scores, on average, are 12-13 points higher for the male sample. For the female sample the associations are less significant. This suggests that some other factors are affecting this relationship with test scores in addition to this variable having a positive correlation with test scores due to the tests being conducted in English. Once again, further investigation would be needed before firm conclusions can be made concerning the reasons for these findings.

The variable measuring whether a pupil has extra tuition outside school is not significant at the 0.01 significance level in any of the regressions. It also has the second lowest beta coefficient at 0.04, thereby suggesting that this is one of the least important variables in the model.

The variable which considered time spent on household chores is found to be significant in all the regressions, apart from those for the male sample. The beta coefficient for the female sample is between 0.09 and 0.10 which shows that this is an important factor in influencing females' test scores. One reason why this variable has a stronger correlation with females' scores may be because females are required to spend more hours on these tasks than males; however, time spent on chores, has not been measured by the variable. To understand this finding in more depth would involve measuring the number of hours spent on the various household chores that make up this variable. What can be stated however, is that this finding supports studies in the literature review which found that, if females have to carry out a lot of household duties, their education is affected in negative ways.

[29] As shown in Appendix 2 father's education had stronger correlation than mother's education with pupil scores, and the strongest correlation was when mother's and father's education was combined.

Whether a pupil has ever had access to a computer is significant and positive in all of the regressions. Computer access has a particularly strong correlation in the regressions for the male sample, with beta coefficients of around 0.11. The beta coefficients for females are approximately 0.5 and for the female and male combined sample regressions they are around 0.8. However as the confidence intervals overlap (see Appendix 3) it cannot be concluded that there is any significant difference.

The type of lighting used in a pupil's home is found to be significant and important in all of the regressions for males, with a beta coefficient of 0.11; however, in the regressions for the female sample, it is not significant. This variable is highly correlated and significantly correlated to SES and thus may be partially working as a proxy for SES. However, it may also have an effect in its own right on test scores, as the type of lighting in a house can determine a pupil's ability to study after the sun goes down. The reasons why this variable is important for males but not for females is unclear and this important finding requires further investigation.

7.2.3 Pupil Academic Record

Grade repetition is significant at the 0.01 significance level only for the regressions for the male and female sample. It has a negative relationship with the test scores, with a beta coefficient of between -0.5 and -0.7. As it is expected that weaker students are more likely to repeat grades, no conclusions can be made about the usefulness of the grade repetition policy in Malawi. All that can be said is that there is a negative relationship between test scores and grade repetition.

The variable for days absent is significant for the male sample regressions but not for the female sample regressions. This variable has a beta coefficient of -0.09 for males and normal coefficient of approximately -2, showing that a pupil being absent for an additional day, is associated with a test scores 2 points lower. Whereas an increase in the number of days a female pupil is absent does not appear to be associated with a decrease in test scores. Again, the reasons for these differences require further investigation.

7.2.4 Teacher Characteristics

The living condition of maths teachers is not significant at the 0.01 significance level in any regressions.

7.2.5 Quality of Teaching

The number of hours a teacher spends teaching is a significant and important variable in all of the regressions. This variable seems to be correlated with males' and females' scores with a similar magnitude: all regressions have a beta coefficient between -0.10 and -0.11, apart from one of the regressions for females which has a particularly important beta coefficient of -0.15. The regressions show that if a pupil has a teacher who teaches more than 30 hours a week, then this is associated with a pupil test score that is lower by approximately 10-14 points. A rational explanation for the result is that if teachers are not over-worked and have less than 30 hours teaching per week then this has a positive effect on both female and male pupils' academic achievement. Presumably, this is so that teachers can prepare properly for classes and have time to mark and assess pupils' work, etc.; however, it may be due to other factors such as being a proxy for whether the school has double shifts. More investigation would be needed if the mechanisms behind the correlation were to be understood.

Knowledge of teachers, as measured by test scores, was only significant in the regressions for females where it had a beta coefficient of approximately 0.10. However it was not significant when the effects of the complex sampling design had been taken into account. This is in contrast to what was expected as it makes sense that teachers who score higher on these tests are better teachers.

Whether a pupil is given homework is also found to be a factor that has an important and significant relationship with both female and male pupil scores. All of the regressions find that if a pupil is given homework regularly, the better their test scores are. This may be because practice is imperative for student learning and therefore homework can improve learning. The beta coefficients for the male sample and for the female and male sample together, are between

and 0.9 and 0.10, whereas the beta coefficient for females is 0.11. However as the confidence intervals overlap (see Appendix 3), it cannot be concluded that there are any significant differences between the results for males and females for this variable.

Whether the maths teacher interacts with pupils' parents or guardians is not found to be significant at the 0.01 significance level in any regressions. It also has the lowest beta coefficient in the model at 0.4. Therefore, it can be concluded that there is weak evidence that this variable is important for predicting pupil scores.

The number of days teachers are absent for is significant at the 0.001 level in all the regressions, apart from those for the female sample. The beta coefficients are around 0.8 for the male sample and 0.6 for the mixed sample. However note that when the complex sampling design is taken into account this variable becomes less significant, thus one should be cautious not to conclude that this variable has a significant association with test scores.

The number of years of professional training that teachers undergo is significant for both female and male pupils. As the descriptive data section discussed, the average number of years training teachers in Malawi receive is 2 years. However, some teachers receive no training at all, whilst a small proportion of teachers have 4 years training. This variable has a beta coefficient for females of around 0.9 and a beta coefficient for males of around 0.11. For the female and male samples together the beta coefficient is 0.10. This is consistent with Hungi's (2011) finding for reading test scores and emphasises the need to have trained teachers in Malawi's education system if pupils are to perform well at school.

7.2.6 School Head Characteristics

The gender of the head teacher has not been found to be significant in any of the models. Therefore, it is concluded that, once other factors are taken into account, this variable is no longer associated with test scores. The age of the head teacher is positively associated with test scores at the 0.01 level in 5 out of 8 regressions and has beta coefficients between 0.07 and 0.10. However note that when the complex sampling design is taken into account this variable becomes less significant, thus one should be cautious when interpreting this result.

7.2.7 Quality of School Management

The dummy variable for school head experience is found to be significant in all of the regressions with the beta coefficient being between -0.08 and -0.10. However when the complex sampling design is taken into account the association is not significant for male pupils. The results suggest that if a teacher has been working as a head teacher for 13 years or more, then female pupils' score, on average, 11-14 marks less than if the head teacher has been a head teacher for less than 13 years. The possible reasons for this relationship are discussed in the descriptive data section. Although it showed relatively high correlations with test scores the results show that, once other variables have been controlled for, the number of hours taught by head teachers does not have an important correlation with pupil scores. This suggests that the correlation was perhaps being driven by a third variable (such as pupil teacher ratio[30]).

7.2.8 Classroom Environment

Class size is statistically significant (at the 0.001 level) and very important (beta coefficients are between -0.15 and -0.22) in all of the regressions, including when the complex sampling design is taken into account. It is interesting to note that this variable is so strongly significant, because when the correlations were analysed this variable was only significant for females. One interpretation may be that it is only once other important variables have been controlled for that the negative impact that class size has on pupil performance can be seen. It is also interesting to note that class size has a stronger relationship with scores than pupil/teacher ratio. Pupil/teacher ratio is significant in all but one of the regressions (for the female sample), however once the complex sampling design is adjusted for it is no longer significant. One interpretation of this may be that large class sizes, regardless of the number of teachers per class, have negative effects on pupil scores. It is possible that by including both class size and pupil teacher ratio there are multicollinearity problems. However although for most other SAQMEQ countries class size and

[30] PTR has a statistically significant correlation of 0.45 with head teacher hours.

pupil teacher ratio are highly correlated, in Malawi there are not strong or significant correlations between these two variables (correlation of 0.02).

7.2.9 School Characteristics

School resources are significant and important in all the regressions with beta coefficients varying from 0.08 to 0.14. Based on beta coefficients, this variable is one of the most important variables in the regressions for both males and females. These findings support studies that cite resources as being an important input into educational achievement (Banda 2003, Milner et al 2011c, Hungi 2011, Willms and Somers 2001, Gustafsson, Hungi 2011, Nannyonjo 2007, Coates 1998, Colclough 2000). The result suggests that whilst it may be the case that internationally 'there currently is no clear, systematic relationship between resources and student outcomes' (Hanushek 2007), in Malawi, a country with very low levels of resources, they are associated with student outcomes.

The state of the school building is also significant at the 0.01 level in the regressions for females but not for males. For females it has beta coefficients between 0.09 and 0.11 and the normal coefficients show that if the building is in need of major repair, then female pupil scores are, on average, 8 to 10 points lower.

The distance a pupil has to travel to school is significant in two of the regressions for females, in which the variable for teacher knowledge has been removed. Although it is only significant at the 0.05 level, it is still significant when the complex sampling design is taken into account unlike most other variables that are only significant at the 0.05 level in the standard regressions. Therefore in this case we can conclude there is some evidence that distance for school is significantly related to female pupil test scores but not for males. The beta coefficients are 0.06 for these regressions. This suggests that this variable may affect female pupil scores, as the literature suggests. The reason why it is only significant in the regressions where the variable for teacher knowledge has been excluded is unclear.

The variable for school location, which indicates whether the school is urban or rural, is not significant in any of the regressions. This suggests that once other variables correlated with

school location have been controlled for, the school location has no relationship with pupil scores.

Whether a school provides free school meals to pupils is significant in the regressions for males and females however at the 0.01 significance level it is not significant in the male only or female only samples. It has a negative relationship with pupil scores with beta coefficient of -0.07 and normal coefficient of -8 thereby demonstrating that the pupils who go to schools with free school meals score on average 8 points less. It may be that the schools that are selected to receive free meals have characteristics that are negatively correlated with pupil performance and it may be that it is these other characteristics that are causing this negative relationship. It may also be that the meals are keeping poorer pupils from attending school. However, without further research, the reason for this result is unclear and therefore this variable was left out of half of the regressions. It was then found that removing this variable did not seem to have any significant effect on the models.

Whether the school is a public or private school is significant at the 0.01 level in the regressions for the female and male sample, and in one of the regressions for the female sample. There are no significant differences in the coefficients for males and females, and the beta coefficients vary from 0.06 to 0.08. This result suggests that in Malawi going to a private school rather than a public school has a positive relationship with pupils' scores. As characteristics of the school have been controlled for, such as pupil/teacher ratio and resources, this finding suggests that there are some unmeasured processes, or factors that differ between public and private schools. This merits further investigation to see what these differences may be.

7.2.10 School Environment

Once other variables are controlled for, the proportion of female teachers in a school is not found to be important, and does not appear to be significant in any of the regressions, not even for the female sample.

The variable that measures whether pupils are being taught in a safe learning environment is significant (at the 0.01 level) in all of the regressions for females and none of the regressions for

males. The beta coefficients for the female regressions are 0.08 and the normal coefficients are on average 9, which shows that females who attend schools which have a safe environment score, on average, 9 points more than females studying in an unsafe environment. This variable is not as significant once the complex sampling design has been taken into account. However the fact that this variable is significant at all in a sample representative of all females in Malawi where the variable is not directly measuring whether or not a pupil has experienced any abuse highlights that safety in a school is likely to have a relationship with female pupils' scores. Further investigation is needed here before concluding that abuse of females in the education system may be a contributory factor to the under achievement of females in Malawi's education system. However if this was concluded after further analysis it would support qualitative literature which argues that one reason why females do not do as well at school or drop-out is due to abuse and fear of abuse in school. (Colclough 2000, GCE 2003, Banda 2003). It would beneficial for further study to look at the individual SACMEQ III variables that constitute a 'safe' school separately, as well as try and find any data that measures abuse more directly. Sexual abuse and fear of abuse, particularly in school, as well as being a human rights issue, is also an educational achievement issue that urgently needs to be addressed.

7.2.11 Community Environment

The variable measuring problems between the community and the school is significant in all of the regression models. It has a beta coefficient between 0.11 and 0.12 for the female sample and 0.07 for the male sample. When the complex sampling design is taken into account, this variable is no longer significant for males. However, note that the confidence intervals (see Appendix 3) overlap for the male and female samples so it cannot be concluded that this variable has a significantly different relationship with test scores for females and males. On average, the results show that if a pupil is at a school which has problems with the community, then this is associated with pupil scores of 8 points less.

Whether the community contributes to the school building and textbooks is not significant for the male sample and is only significant at the 0.01 level in one of the female regressions. The variable that measures whether a pupil is at a school where the community contributes to school

staff bonuses is significant for males and females with higher beta coefficients of between 0.09 and 0.11. More investigation and understanding into the common characteristics of schools for which the community contributes to buildings, textbooks and/or school staff bonuses is needed before conclusions are made about why these variables are correlated with pupil scores.

8 Conclusion

As mentioned throughout the report, for many of the results to be fully understood, further research is needed. Although guesses and assumptions could be made about many of these relationships, until further investigations have been carried out, reliable conclusions cannot be reached. Thus one recommendation from the report is that, in order to enrich the research, some of the findings of this study should be explored further through the use of qualitative as well as quantitative methods.

The research could also be enriched if further quantitative analysis is carried out, such as:

- the results from the OLS regressions results are compared to the results for regressions using the hierarchical linear model,
- tests are carried out for heteroskedasticity,
- some variables are constructed differently and the results compared, and
- the method for choosing variables to add into the regression is altered and the results compared.

Although these additions to the research would enhance it, the findings of this study are still a very useful starting point for information on how education inputs correlate with female and male pupils differently and which factors should be looked at in more detail if pupils' education outcomes are to be improved in Malawi.

The report has shown that females in Malawi face serious educational disadvantages. They start primary school equitably with a GPI for enrolment in Grade 1 of more than 1. However, although they repeat less, are absent less, and are younger than males, the SACMEQ III data shows that Grade 6 female pupils of all SES backgrounds, are not achieving as well academically as male pupils.

The first research question asked which education inputs had a significant and important influence on education outcomes in Malawi and if the gender variable is one of these significant

and important variables. This research question has been answered using production function analysis and found that the gender variable is still both significant and important after various education inputs have been controlled for. As this study rejects the concept that females are innately less intelligent than males, this finding suggests that there are other important variables which disadvantage female pupils that have not been included in the model. Some of the variables which the literature review and conceptual model suggest should be included in the production function could not be included due to lack of data. For example, SACMEQ III does not contain information on key variables such as: management processes inside schools, innate ability, cultural attitudes to female education, whether a pupil is pregnant or has children, whether a pupil is married or due to get married, or whether a pupil is being abused at school. If data existed on all of the factors which the qualitative studies suggest are very important in determining education outcomes, then this report could have provided more detailed insights into the key determinants of education outcomes.

The low R squares of between 0.19 and 0.21 for all regressions also indicate that there are still variables missing from the regressions that might explain variance in pupil performance. The fact that SES was not shown to be significant may also partially explain the low R2. However, it is possible that missing variables may not further explain the variance in pupil performance. This may be the case if there is little variance in variables, as is the case for SES which was found to be at very low levels for the majority of pupils. It is interesting that it appears to be the case that Malawi is one of the few documented country cases where gender differences are not mitigated by improving SES.

The second research question has been answered by running the regressions again but separating the male and female samples and comparing the results. This analysis finds that some education inputs are correlated with male and female pupils' academic achievement very differently, whereas others have similar relationships. It is very important for policy makers to understand which inputs have a relationship with females' test scores if academic achievements of females in Malawi are to improve. The study shows that when education production function analysis is carried out with the male and female samples together, then the results may hide the importance and significance of certain education inputs for female pupils' academic achievement; for

example, the significance and importance, for female pupils, of carrying out chores outside school.

The regression analysis finds that certain education inputs are important to female pupils' education outcomes but not for males' education outcomes. These variables are: age of pupil; chores done in the home; distance to school. Some of these variables that are significant in the female but not the male regressions show up as significant in the regressions for both females and males but with a smaller coefficient and significance, whereas other variables do not feature in the mixed sample regressions. The finding that female pupils at Grade 6 do less well academically if they: are older; have more household chores and have further to travel to school, support the qualitative literature studies which explain why females may not achieve as much as males in the education system[31]. This report, therefore, gives quantitative evidence that certain factors need to be addressed and improved if female pupils are to perform better in Malawi's primary education system.

The analysis shows that approximately 42 per cent of pupils were at schools which had a problem with sexual abuse of pupils by other pupils and 17 per cent attended schools where teachers' sexual abuse of pupils had been identified as being a problem. There was also some weak evidence that female pupils' test scores were associated with how safe a school was. However regardless of the strength of the association between a safe learning environment and female pupil test scores, this scale of the risk of sexual abuse is very troubling and policy makers should address this urgently.

The regression also finds that certain education inputs are important to male pupils' education outcomes but not for females' education outcomes. These variables are: English spoken at home; light type in the home; teacher days absent; and pupil days absent. All of these variables are found to have a negative relationship with male pupils' education apart from light type in home which has strong correlations with SES and has a positive correlation. It is unclear why this variable is correlated with males' but not females' academic achievement. Further investigation is required if this and other gender differences for variables are to be explained.

[31]The finding that females' but not males' academic achievement is affected by teacher knowledge, and the amount the community contributes to the school building and text books has not been explained by the literature reviewed in this report and requires further investigation if this is to be explained.

The analysis also finds that some variables are significantly correlated with the academic achievement of both females and males. The variables are: class size; hours of teaching; school resources; problems with the community; homework given; community paying bonus to school staff; years of teacher training; head teachers' number of years' experience; pupil/teacher ratio; access to a computer. In other studies, most of these variables, or similar variables to these, have been found to have significant effects on education outcomes. Thus if the mechanisms behind these factors being correlated with academic achievement are understood and addressed this could improve both female and male pupils' education outcomes in Malawi.

In summary, the analysis has suggested some factors which policy makers may wish to consider, looking at how these factors are associated with academic achievement, in order to reach a better understanding of how to improve academic achievement at primary school level in Malawi. The analysis has shown which factors are associated with both male and female pupils' academic achievement, and which are only associated with either female or male pupils. The analysis has shown that if policy makers want to improve female pupils' academic achievements, as well as addressing the factors that are associated with academic achievement for both females and males, they should also look to address factors that associated with lower academic achievement for females only. If further investigation was carried out and there was strong evidence identifying likely causal mechanisms and levers not just association, then the results of the production function could be used to calculate the extra education inputs required to offset this gender disadvantage. Policy makers32 could put costs to the education inputs and work out the most efficient ways to improve education outcomes according to the regressions results. This could form the basis of a pro-female education financing approach.

[32] It would also be advisable for policy makers to consider any externalities to improving certain education inputs. For example, if schools are made safer for females, and in particular, they are less at risk of sexual abuse, this may also have many positive externalities outside an improvement in their education achievements.

Reference List

ActionAid. (2010). *Destined to fail? How violence against women is undoing development:* http://www.actionaid.org.uk/doc_lib/destined_to_fail.pdf [Accessed 6 July 2012]

Banda, D. (2003). *Gender sensitive educational policy and practice: the case of Malawi*, Background paper prepared for the Education for All Global Monitoring Report 2003/4 Gender and Education for All: The Leap' to Equality

Boissiere, M. (2004) *Determinants of Primary Education Outcomes in Developing Countries.* Background Paper for the Evaluation of the World Bank's Support for Primary Education. The World Bank Operations Evaluation Department. http://www.worldbank.org/oed/education/documents/education_primary_determinants_paper.pdf [Accessed 10 October 2012]

Burton, P. (2005). *Suffering At School. Results of the Malawi Gender-Based Violence in Schools Survey.* National Statistical Office

Coates. (1998). *Education Production Functions Using Instructional Time as an Input.* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=98308 [Accessed 6 July 2012]

Coleman, J, et al. (1966). *Equality of educational opportunity*. Washington, D.C.: U.S. Government Printing Office.

Colclough, C., Rose, P. & Tembo, M. (2000). *Gender Inequalities in Primary School: The role of poverty and adverse cultural Practice*, IDS Working Paper 78. http://www.sciencedirect.com/science/article/pii/S0738059399000462 [Accessed 20 May 2012]

Dee, Thomas S. (2007). *Teachers and the Gender Gaps in Student Achievement. Journal of Human Resources* 42, 3 (528-554)

DHS 2004, Malawi Demographic Health Survey 2004. (2005). http://www.measuredhs.com/publications/publication-FR175-DHS-Final-Reports.cfm [Accessed 6 May 2010]

EMIS, 2009. http://www.nso.malawi.net/ [Accessed 6 May 2010]

EMIS. (2010). http://www.nso.malawi.net/ [Accessed 6 May 2010]

Education Policy and Data Centre. (2010). http://epdc.org/ [Accessed 10 May 2012]

Gertler, P. & Glewwe, P. (1992). *The Willingness to Pay for Education for Daughters in Contrast to Sons: Evidence from Rural Peru*, World Bank Economic Review, World Bank Group, vol. 6(1), pages 171-88, January. http://ideas.repec.org/s/oup/wbecrv.html [Accessed 6 July 2012]

Global Campaign for Education (GCE). (2003). A fair chance - United Nations Girls' Education Initiative. http://www.ungei.org/resources/index_598.html [Accessed 6 July 2012]

Goldstein, H. (1997). *Methods in school effectiveness research*, within School Effectiveness and School Improvement, 8(4), 369–95.

Gujarati, D, N. (2003). *Basic Econometrics*. 4th Edition. McGraw Hill.

Gustafsson, M. (2006). *School Production Modelling to Strengthen government monitoring programmes in developing countries*, Thesis for University of Stellenbosch. http://scholar.sun.ac.za/handle/10019.1/2683 [Accessed 12 May 2012]

Gustafsson, M. (2007). Using the hierarchical linear model to understand school production in South Africa. South African Journal of Economics, 75(1): 84-98

Hanushek. (2002). *The Failure of Input-based Schooling Policies*, The Economic Journal, 113, pp 64, 98. http://hanushek.stanford.edu/sites/default/files/publications/Hanushek%202003%20EJ%20113%28485%29.pdf [Accessed 2 November 2012]

Hanushek. (2007). *Education Production Functions*. http://hanushek.stanford.edu/sites/default/files/publications/Hanushekper cent202008per cent20PalgraveDict_0.pdf [Accessed 6 July 2012]

Harbison, R. & Hanushek, E. (1992). *Educational Performance of the Poor: Lessons from Rural Brazil*, Oxf. U. P, New York.

Hungi, N. (2010). Is there gender equity in school managerial positions? , Policy Issues Series Nr 8. http://www.sacmeq.org/downloads/policy/008-SACMEQPolicyIssuesSeries-Schoolmngpositions.pdf [Accessed 5 July 2012]

Hungi, N., Makuwa, D., Ross, K., Saito, M., Dolato, S., Capelle, F., Paviot, L. &Vellien, J. (2010). SACMEQ III Project Results: Pupil Achievement Levels in reading and mathematics, SACMEQ.

Hungi, N. (2011). Accounting for the Variations in the Quality of Primary School Education. http://www.sacmeq.org/downloads/Workingper cent20Papers/07_Multivariate_Final.pdf [Accessed 5 July 2012]

International Monetary Fund. (2012). World Economic Outlook Database. http://www.imf.org/external/pubs/ft/weo/2012/01/weodata/weorept.aspx?pr.x=74&pr.y=16&sy=2009&ey=2012&scsm=1&ssd=1&sort=country&ds=.&br=1&c=676&s=NGDPDpercent2CNGDPDPCper cent2CPPPGDPper cent2CPPPPCper cent2CLP&grp=0&a=. [Accessed 19 April 2012]

Johnes, J. (2004). *'Efficiency measurement',* in International Handbook of the Economics of Education. Edward Elgar Publishing Ltd. USA.

Lee, V. E., & Zuze, T. L. (2011). *School resources and academic performance in sub-saharan Africa*. Comparative Education Review

Lee, V.E., Zuze, T.L., and Ross, K.N. (2005). *School Effectiveness in 14 Sub-Saharan African Countries*: Links With the Reading Achievement of Grade 6 Pupils. *Studies in Educational Evaluation*

Lewis, Maureen and Marlene Lockheed. (2006). *Inexcusable Absence: Why 60 Million Girls Still Aren't In School and What To Do About It*. Washington, DC: Center for Global Development.

Lockheed, M.E. & Longford, N.T. (1989). *A Multilevel Model of School Effectiveness in a Developing Country*. World Bank Discussion paper No. 69. Washington, DC: World Bank

Marshall, G. (1998). 'Coleman Report.' A *Dictionary of Sociology*. <http://www.encyclopedia.com>. [Accessed 9 July 2012]

Milner, G., Mulera, D., Chimuzu Banda, T., Matale, E., &Chimombo, J. (2011a). *Learner and Teacher knowledge about HIV and AIDS in Malawi.* http://www.sacmeq.org/downloads/policy/MALAWIper cent20per cent20HIV_AIDSper cent20Finalper cent20Versionper cent2016per cent20Juneper cent202011per cent20KR.pdf [Accessed 8 July 2012]

Milner, G., Mulera, D., Chimuzu Banda, T., Matale, E., & Chimombo, J. (2011b). *Progress in Gender Equality in Education: Malawi*. http://www.sacmeq.org/downloads/Genderpolicybriefs/06_MAL_Policy_briefper cent20-per cent20Gender_NewFormat_ver12_Final_111004.pdf [Accessed on 8th July 2012]

Milner, G., Mulera, D., Chimuzu Banda, T., Matale, E., &Chimombo, J. (2011c). Quality *of Primary School Inputs in Malawi.* http://www.sacmeq.org/downloads/Schoolper cent20inputs/MAL_School_Inputs_15Oct2011_FINAL.pdf [Accessed 8 July 2012]

Ministry of Education, Science & Technology. (2009). *Education Sector Implementation Plan 2009-2013, ESIP*

MOEST. (2001). *Policy and investment framework for the education sector*. Lilongwe: Government of Malawi.

Mullis, I., Martin, M., & Foy, P. (2008). *TIMSS 2007 International Mathematics Report Findings from IEA's Trends* in International Mathematics and Science Study at the Fourth and Eighth Grades. Chestnut Hill: TIMMS & PIRLS International Study Centre.

Multiple Indicator Cluster Survey (UNICEF). (2006). http://www.unicef.org/infobycountry/malawi_statistics.html [Accessed 12 May 2012]

Nannyonjo, H. (2007) *Education Inputs In Uganda: An Analysis of Factors Influencing Learning Achievement in Grade Six*. World Bank Working Paper No. 98.

National Statistics Office. (2004). Demographic Household Survey. http://www.nso.malawi.net/index.php?option=com_content&view=article&id=17:2004-mdhs&catid=8:reports&Itemid=68 [Accessed 10 May 2012]

National Statistics Office. (2001). *Population and Housing Census 2008: Analytical Report Volume 5: Education and Literacy, Government of Malawi.*

OECD. (2001). *Equally prepared for life? How 15-year-old boys and girls perform in school. Programme for International Student Assessment (PISA*). Paris : OECD.

OECD, (2007). *Executive Summary PISA 2006: Science Competencies for Tomorrow's World.* The Programme for International Student Assessment (PISA). http://www.oecd.org/pisa/pisaproducts/pisa2006/39725224.pdf [Accessed 10 October 2012]

Oxaal, Z. (1997). *Education and Poverty: A Gender Analysis, Report prepared for the Gender Equality Unit*. Swedish International Development Cooperation Agency (Sida)

Perry, H. (2002). *The impact of socio-economic status and school condition on Grade 3 learners' literacy and numeracy results*. Unpublished document prepared for the Joint Education Trust.

Plan International. (2009). Because I am a Girl, *The State of the World's Girls 2009, Girls in the Global Economy: Adding It All Up.* http://plan-international.org/girls/reports-and-publications-detail-3.php?lang=en [Accessed 5 July 2012]

Saito, M. (2010). Has gender equality in reading and mathematics achievement improved?, SACMEQ Policy series nr 4. http://www.sacmeq.org/downloads/policy/004-SACMEQPolicyIssuesSeries-Genderdifferences.pdf [Accessed 5 July 2012]

Saito, M. (2011). *Trends in the Magnitude and Direction of Gender Differences in Learning Outcomes.* http://www.sacmeq.org/downloads/Workingper cent20Papers/04_Gender_Final_21-02b.pdf [Accessed 5 July 2012]

UNDP (2011). *Sustainability and Equity: A Better Future for All.*
Explanatory note on 2011 HDR composite indices. Malawi
HDI values and rank changes in the 2011 Human Development Report

UNESCO. (2009/2010). Data centre. http://www.uis.unesco.org/Pages/default.aspx [Accessed 10 May & 2 November 2012]

UNESCO. (2005). *Education For All Global Monitoring Report*.
http://www.unesco.org/new/en/education/themes/leading-the-international-agenda/efareport/reports/2005-quality/ [Accessed May 20 2012]

UNESCO. (2000). *The Dakar framework for action: Education for All – meeting our collective commitments.* World Education Forum, Dakar, Senegal, 26-28 April. Paris: UNESCO.

United Nations. (2006). *The millennium development goals report 2006*. New York: United Nations

UNFPA, (2012). *Marrying Too Young. End Child Marriage.* Chapter 3, Global & Regional Levels and Trends.
http://www.unfpa.org/webdav/site/global/shared/documents/publications/2012/ChildMarriage_4_chapter3.pdf [Accessed 15 March 2013]

UNFPA, (2013). UNFPA in Malawi.
http://unfpamalawi.org/index.php?option=com_content&view=article&id=89&Itemid=88 [Accessed 15 March 2013]

Vinjevold P & Crouch L. (2001). *Literacy, Numeracy and Policy: Towards Assessing Learning in Grade 3. District Development Support Programme,* Johannesburg. USAID

White, H., Lu, X. (2010). *Robustness Checks and Robustness Tests in Applied Economics*. Department of Economics, University of California, San Diego

Willms, J. D., & Somers, M.-A. (2001). *Family, classroom, and school effects on children's educational outcomes in Latin America*. International Journal of School Effectiveness and Improvement

World Bank. (2010). *The education system in Malawi.* World Bank Working Paper, No. 182. Washington DC: The International Bank for Reconstruction and Development/The World Bank.

Appendices

Appendix 1: Methodology behind the creation of some variables and the rejection of others

Pupil Characteristics

The analysis tested using age squared, rather than just age which presumes a linear relationship, however the correlations were weaker so this variable was rejected.

Ideally, social economic status would be measured by family income. However as SACMEQ cannot collect information on family income, a proxy variable for SES has been constructed in the SACMEQ data from roof type, light type, parents' education and some other variables, including those on possessions. This is useful if results between countries are to be compared; however, as this study is just looking at Malawi, a measure of SES that is more appropriate for Malawi (a very economically poor country) and the purposes of this study has been created. The SES variable this study used only included measures of possessions and home type. In order to choose which possessions and house features to use in this measure of SES each home possessions and home type was correlated with test scores and the significance and size of the correlations was assessed. From this analysis it was decided that the best measure for SES for Malawi would include information on floor type and roof type in the home, if the house had electricity, a bed, a fridge, a TV and an electric fan. Possession of electricity, fridge and fan as well as roof type were given higher weightings, because they had higher correlations. Many of the 31 possessions that SACMEQ measures if they are present in a home were excluded as they were not significantly correlated with test scores and many of these possessions were very rare to have in a pupil's home. For example, less than 1 per cent of pupils had the following possessions in their homes: vacuum cleaner, washing machine, computer, internet. Therefore a measure of SES that includes these possessions becomes less relevant for Malawi.

This variable used to measure SES can take possible scores of 0 to 12, where 0 means the pupil has none of the house possessions included in the measure and has the most basic floor and roof type. For Malawi the median score is 1.75 which means that pupils at the 50th percentile have no electricity, no fridge, no electric fan and no TV at their house. A Malawian pupil scoring an SES

of 1.75 may however have either a bed, and earth floor and an unsealed roof (for example thatch) or no bed in their house and a cement floor and metal roof.

It was decided that parents' education together with books at home would be excluded from the main measure of SES. This is because these two variables were used instead to act as a proxy for the educational support a pupil received at home so that the contribution of this to test scores could be separately assessed. It was decided as well as keeping academic support separate, to also keep the type of light a house had as separate, as this variable seemed correlated to pupil scores even after SES was accounted for. Note that the light type in a pupil's house, academic support: (parents' education and books at home) and SES are very highly correlated and this needs to be considered when interpreting results.

Home Environment

To construct the variable to measure nutrition the analysis first looked at the three SACMEQ variables that measured how often pupils had breakfast, lunch and dinner. As the following tables show, the most common meal to have every day was the evening meal, followed by lunch, and lastly breakfast.

Table 36: Per Cent of Pupils Having Meals Each Week

	Morning meal (per cent)	Lunch (per cent)	Evening meal (per cent)
Not at all	14.31	1.77	1.44
1-2 days a week	13.85	4.80	2.60
3-4 days a week	11.67	5.41	4.14
Every day	60.17	88.03	91.83

SACMEQ III, own analysis

36

The analysis then looked at which meals were most correlated to test scores and then generated a new variable to sum weight the meals accordingly. The analysis showed that in Malawi lunch is the most highly correlated with test scores and breakfast has the weakest, this variable weights lunch as 1.5, breakfast as 0.5 and the evening meal as 1.0. . Therefore range of this variable was

from 3.5 to 12, where a score of 12 means that a pupil had breakfast, lunch and dinner every day, and a score of 3.5 means that a child in general had no meals but occasionally had breakfast.

To measure the academic support available at a pupil's home, a variable was created that combines parents' education and the number of books at a pupil's house. Correlations were firstly run for mothers' education, fathers' education and thirdly for mothers' and fathers' education combined. It was found that parents' education was correlated with males' scores more than females' scores, and for both males and females it was found that fathers' education has a higher correlation with test scores more than mothers' education did, but that mothers' and fathers' education added together had higher correlations than they did separately. This finding is contrary to what Gustafsson (2006:p107) found for South Africa where the mothers' education had more of a correlation with scores than the fathers' education. Therefore it was decided to use mothers' and fathers' education combined. The analysis tried squaring and cubing parents' education and found that for Malawi, parents' education cubed was most significantly correlated with pupil scores, suggesting that parents' education has an exponential correlation with scores up until a certain point. This variable was then added to the number of books at a pupil's home after being weighted by 5, and then the whole variable was divided by 300 so that the numbers were more manageable. The combination of these two variables creates a variable that has a higher correlation with test scores than either of the variables on their own, which supports the use of this combined variable to measure academic support (parents' education and books at home).

In order to create a variable to measure the household chores a pupil does, correlations for the various chores measured by the SACMEQ III data and test scores were assessed. The chores that had the largest negative relationship with pupil scores were then added to make one variable. The variable that resulted from this combined analysis ranges from 7 to 21. A score of 7 means pupils never have to carry out the following activities: looking after younger relatives, taking care of sick relatives, cooking, fetching water, chopping firewood, gardening, taking care of livestock, and a score of 21 would mean they have to carry out all of the activities almost every day.

A variable was created to measure whether pupils owned the following learning materials: exercise books, notebooks, pens and pencils. The data found that 63 per cent owned none of

these items, 36 per cent owned 1 or 2 of these items and less than 1 per cent owned three of four of these items. This again displays the level of poverty that exists in Malawi and the lack of any educational resources in most pupils' homes. This variable, however, does not show up as correlated with scores at the 0.01 level for males, females or both together, therefore the regressions do not include this variable. Therefore, an alternative measure of access to learning resources was used, which was if a pupil had ever used a computer before. This was found to be a more important measure. This measured whether a pupil had used a computer, not necessarily just at home, but also included use of a computer anywhere, such as at a friend's or relative's, or at an internet café, etc. This variable takes the value 1 if a computer had ever been used by the pupil and 0 if a computer had never been used.

The literature review suggested that whether or not a pupil lived with their biological parents may have an effect on academic achievement. However, the SACMEQ data for Malawi did not find any significant correlations and therefore this variable has not been added to the regressions. There was also no suitable variable to look at family size so it was not possible to account for this in the analysis.

Pupil Academic Record

The SACMEQ III statistics show that 72 per cent of pupils never attended pre-school, and 28 per cent only attended for a few months or more. A study in the literature review (UNESCO, 2005) selected this variable as affecting academic achievement. However, for Malawi this variable did not have a statistically significant correlation at the 0.01 level and therefore was not included in the regressions.

The variable that has been used for grade repetition is the frequency a person has repeated, squared, so that a higher weight is given to the higher values. The variable takes a value of 1 if a person had never repeated, 4 if they had repeated once, 9 if they had repeated twice and 16 if they had repeated three times or more.

Teacher Characteristics

A variable was created that takes the value 0 if a pupil had no female teachers, 1 if either their maths or reading teacher was female and 2 if both the maths and reading teachers were female. The data suggests that 66 per cent of pupils had no female teacher for maths or reading classes, 17 per cent had one female teacher and 17 per cent had 2 female teachers. This variable is statistically significantly positively correlated for males and females at the 0.001 level at 0.09 correlation and surprisingly there is a slightly stronger correlation for males than females with males having a correlation of 0.10 and females of 0.08. However, this variable also has a strong (0.7) statistically significant correlation with the proportion of female teachers at a school and therefore it has been decided not to include this variable in the regressions to avoid problems with multi-collinearity. The proportion of female teachers at a school is discussed under the section 'school environment'.

Teacher Quality

Analysis was carried out for the number of hours taught and it was found that on average if more than 30 hours or more were taught between the maths and reading teacher a week, i.e. more than 3 hours a day on each subject, then this had a negative relationship with pupil scores. Therefore a dummy variable has been created which takes the value 0 if 30 hours or less were taught and the value 1 if more than 30 hours were taught.

In order to measure the knowledge of teachers, SACMEQ collected data on how well teachers do on a subject test they are given. As the report is interested in both maths and science teachers the scores from both tests have been added together. Note that this variable has 37 missing observations. Ideally analysis should be carried out to look at the 37 missing observations and to see if these are randomly distributed or if they are correlated with any other variables, as it is likely they are not. It may be that the teachers that expected to do badly in the test refused to carry out the test or did not show up at school on the day they were expected to carry out the test. Therefore, in some of the regressions this variable has been excluded and results compared to check the robustness of the results and for possible bias.

A variable was considered that measured if a teacher used a school resource centre, and takes the value 0 if resource centres were used for both subjects, 1 the resource centre was used for only one subject and 2 if the resource centre was not used for any subjects. The analysis shows that 80 per cent of teachers used resource centres for both subjects. This variable is statistically significantly correlated at the 0.001 level for males and females, and for males at the 0.05 statistical significance level. However, there are only 2466 observations for this variable, i.e. 315 missing observations. These missing variables are for schools that had no resource centre and therefore are not randomly distributed. Thus it has been decided not to use this variable as it would result in a non-random section of the sample being excluded which would mean the results were no longer representative of the population.

There are no variables in SACMEQ to measure the type of curriculum taught and whether or not there were any gender or SES biases in it.

School Head Characteristics

SES, health, natural management ability characteristics of a school head teacher may be correlated with scores; however, none of the literature directly points to these as important variables for Malawi and SACMEQ does not collect information on these variables therefore they have not been included.

Quality of School Management

The data indicates that 31 per cent of head teachers have worked as a head teacher for 5 years or less, 34 per cent have worked between 10 and 5 years, and the remaining 35 per cent have worked over 10 years as a head teacher. The median is 5 years and the mean is 5.8 years. The data shows that experience as a head teacher was negatively correlated with pupil performance. This is contrary to what was expected from the literature review, which showed that Hungi (2011) found that school head experience had a positive correlation with pupil scores. As the analysis found this negative correlation, it was decided to transform this variable into a dummy variable, and to find the point at which more experience has a negative influence on scores. After

some variable analysis it was found that if a head teacher has worked as a head teacher for 13 or more years then this can have a negative impact on students' scores. This may be due to a lack of energy and enthusiasm of a head teacher who has been in the role for a long time compared to a head teacher who has not been in the position as long and may be more enthusiastic and motivated to have a positive influence at the school. However, more research into this would be needed before firm conclusions were made.

No correlations were found between pupil scores and the amount of advice given to teachers from head teachers so this variable was left out of all the regressions. The variable in SACMEQ measures whether teachers received any advice from the head teacher or not. If however, SACMEQ contained information on the frequency of advice given, then it may have been possible to have constructed a more useful variable which could have avoided the issue of the possibility of under-performing teachers being given more advice, and thus this effect biasing the results.

Seventy five per cent of school heads had 2 years training and the others had either 4 years training or between 2 years and 0 years training. However, no correlations were found for this variable and pupil scores, therefore this variable was excluded from further analysis.

School Characteristics

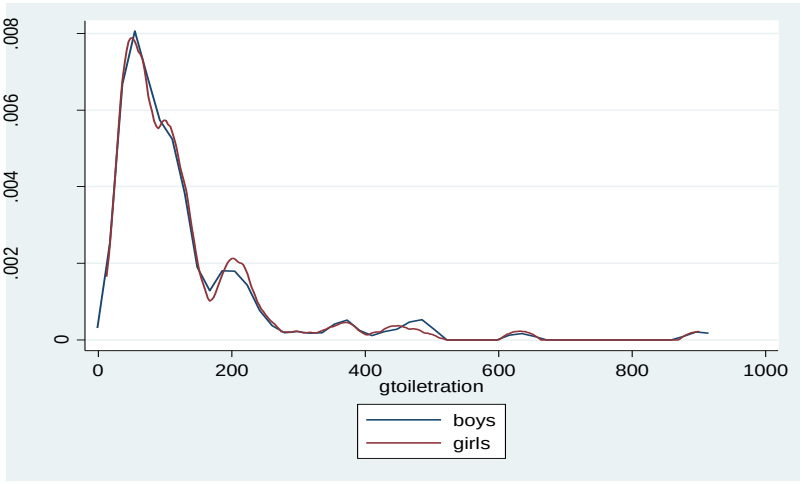
Many studies (Banda 2003, Milner et al 2011c, Hungi 2011, Willms and Somers 2001, Gustafsson, Hungi 2011, Nannyonjo 2007, Coates 1998, Colclough 2000, Banda 2003) found that resources available at a school have a positive influence on academic achievement especially in poor countries such as Malawi. There are however, various different types of school resources that could make a difference to pupils' learning, and some of these are highly correlated. SACMEQ collects data on the resources that are in the classrooms of maths and reading teachers such as writing board, chalk, board duster, wall chart, cupboard, bookshelf, book corner, teacher chair and teacher table. Most of these resources in Malawi are strongly and significantly correlated with each other. The classroom resources also tend to be strongly and significantly

correlated to the school resources, such as whether there is a library, a staff room, water, electricity, a computer, etc. Therefore, rather than have many variables in the regressions measuring each of these things individually, a variable has been constructed, which is a proxy measure for overall classroom and school resources. This variable (resource1) ranges from 1-8, where 8 is the highest level of resources and 1 is the lowest level.

One variable that the literature review pointed at as being significant was text book to pupil ratio. Shockingly, the data shows that 92 per cent of Grade 6 pupils in Malawi had no text books in a class at all. 4 per cent of students had at least one book in the class but no more than one between 2 students and 4 per cent had between 1.6 and 1.8 books per person in a class. This lack of variance may explain why there is no relationship found between pupil scores and text book to pupil ratio. A dummy variable was tried which takes the value 0 if there were no text books in a class at all and 1 if there were any text books in a class, but this also was not found to be significant.

The data shows that the median number of females to each toilet was 90. The 10^{th} percentile was 30 female pupils per toilet and at the 90^{th} percentile the data shows there were 390 female pupils per toilet, see kernel density graph in figure 23.

Figure 23: Kernel Density Graph for Ratio of Girls to Toilets, 2007



SACMEQ III, own analysis

23

Approximately 1.5 per cent of students were at schools that had no girls' toilets at all. This again displays the lack of crucial resources at schools in Malawi. Unfortunately, there was no information on whether female toilets were sufficiently separate to male toilets. However, the females to female toilet ratio was not statistically significant in these correlations, therefore it has not been included in the regressions.

The proportion of female teachers at a school has been chosen to be added into the regressions rather than teacher gender. This is because the two variables are highly correlated to each other, and because it was thought that the proportion of female teachers at a school is a better measure of the influence of female role models at a school than the gender of the reading and maths teachers.

School Environment

In order to create a variable for safety of environment, five different variables that measure levels of abuse and safety in a school were looked at: sexual abuse of pupils by teachers, sexual abuse of pupils by pupils, alcohol abuse by teachers, HIV risk, and pupil behaviour problems.

These variables were all found to statically significantly correlate with each other at the 0.01 significance level. This suggests that if there is a problem of sexual abuse of pupils by teachers there is also likely to be problem of sexual abuse of pupils by pupils, a problem of alcohol abuse by teachers, higher risks of teachers contracting HIV and problems with pupil behaviour (cheating, bullying staff and vandalism).

Note that the pupil behaviour used in this variable (pupprobdum1) takes the value 0 when there was sometimes a problem with pupil behaviour (cheating, vandalism, and bullying staff) and 1 when there was never a problem of these things. This variable is negatively correlated with scores, i.e. if there were no pupil problems reported at all this had a negative impact on pupil scores. It may be that if pupils are too obedient, and passive, it can make them more vulnerable to other possible negatives such as abuse from teachers. However, further research is needed before firm conclusions are made. Note that there are positive and significant correlations with pupil behaviour problems and the variables for sexual abuse from pupils and teachers, and alcohol abuse from teachers, i.e. if there is never a problem with alcohol or sexual abuse then there is more likely to never be a problem with all three of the pupil behaviours.

The variable constructed (abusedum3) takes the value of 1 if the school was deemed to be a safe environment, i.e. there is no sexual abuse from teachers or pupils, no alcohol abuse from teachers, a perceived low HIV risk for teachers and the pupils do have some behaviour problems. The variable takes the value 0 if these conditions did not exist and the school is therefore classified as an unsafe environment.

Attitudes and perceptions towards students was highlighted in the literature review (Banda 2003) as effecting pupils' academic achievement, particularly in relation to the attitudes towards female students in the classroom. However, there is no suitable measurement in SACMEQ of this variable so this has not been included in the regressions.

Community Environment

The analysis found that the following contributions by the community had a negative relationship with pupil scores: building facility, maintenance of facility, purchase of text books, and payment of salaries of additional teachers. These were combined into a single variable.

Appendix 2: Correlation Tables

Table 37: Correlation Matrix for Male and Female Sample (Part A)

		Maths and reading scores	Pupil gender	Pupil age (months)	SES	Frequency of meals	Academic support: (parent education & books at home)	English spoken in home	Extra tuition	Chores done in the home	Used a computer	Light type in the home	Grade repetition
		Scoremr	zpsex	zpagemon	psesno~12	zpmealw	zpparb~59	zpenglis	zpexttui	ptasktotg	zpcomptr	zplight	prepeatsq
Maths and reading scores	scoremr	1											
Pupil gender	Zpsex	-0.111***	1										
Pupil age (months)	zpagemon	-0.112***	-0.193***	1									
SES	psesnol~12	0.159***	0.000856	-0.277***	1								
Frequency of meals	zpmealw	0.0801***	0.0619**	-0.131***	0.159***	1							
Academic support: (parent education & books at home)	zpparbo~59	0.128***	0.00118	-0.271***	0.483***	0.162***	1						
English spoken in home	zpenglis	0.141***	-0.0776***	-0.0737***	0.168***	0.0338	0.173***	1					
Extra tuition	zpexttui	0.0662***	0.0357	-0.0219	0.123***	0.0156	0.115***	0.113***	1				
Chores done in the home	ptasktotg	-0.155***	0.159***	0.152***	-0.224***	-0.0504**	-0.148***	-0.0665***	-0.0241	1			
Used a computer	zpcomptr	0.150***	-0.0388*	-0.0979***	0.170***	0.0485*	0.186***	0.102***	0.0802***	-0.0779***	1		
Light type in the home	zplight	0.173***	0.0255	-0.232***	0.763***	0.158***	0.400***	0.117***	0.119***	-0.224***	0.155***	1	
Grade repetition	prepeatsq	-0.0807***	-0.0286	0.186***	-0.0218	-0.0157	-0.0504**	-0.0589**	-0.0195	0.0824***	-0.0253	-0.0241	1

Days absent	pabsent	-0.0781***	-0.0512**	0.176***	-0.0960***	-0.118***	-0.0863***	-0.0678***	0.00212	0.0904***	-0.0490*	-0.0886***	0.0281
Maths teachers living conditions	ycondliv	0.0438*	-0.0168	0.0471*	-0.0459*	-0.0109	-0.0864***	-0.0185	-0.0324	0.117***	-0.0436*	-0.0259	0.0566**
Hours taught	zhrteach~4	-0.112***	-0.00394	0.101***	-0.208***	-0.0857***	-0.161***	-0.140***	-0.0653***	0.0662***	-0.0606**	-0.127***	-0.00166
Knowledge of teachers	zaloctmr	0.0418*	-0.00289	-0.0316	-0.0436*	-0.0119	-0.00327	-0.00512	-0.00516	0.0307	-0.00494	0.00325	0.00775
Homework given	phmwk	0.124***	-0.00512	-0.0514**	0.0871***	0.0334	0.0362	-0.0315	0.160***	-0.025	-0.000743	0.0842***	0.000586
Maths teacher interacts with parents	ymeetpar	0.0442*	-0.0158	-0.0301	0.0535**	-0.00516	0.0455*	-0.0301	0.00684	-0.000254	0.0371	0.0181	-0.0235
Teachers days absent	teachabs	-0.0546**	-0.0155	0.0304	-0.0585**	-0.00162	-0.0274	0.0186	-0.0279	-0.000525	-0.0408*	-0.0490*	0.018
Teachers years of training	zqprodmr	0.0884***	0.00269	-0.0454*	0.0764***	0.0514**	0.0387*	-0.0855***	-0.0443*	-0.0872***	0.0313	0.0936***	0.00967
Head teacher gender	zssex	0.0507**	0.0366	-0.127***	0.257***	0.0646***	0.216***	0.0584**	0.113***	-0.111***	0.0122	0.201***	0.02
Head teacher age	sage	0.0712***	-0.000879	-0.0443*	0.013	-0.0466*	-0.0147	0.0482*	-0.106***	0.00534	-0.0023	-0.00659	-0.0506**
Head teacher years' experience	sexpalldum	-0.0724***	-0.00986	-0.0104	-0.0682***	-0.00409	-0.0449*	-0.0361	-0.0789***	0.0878***	-0.0249	-0.0809***	-0.00513
Head teacher teaching hours	zstchhrs	-0.167***	-0.00567	0.166***	-0.279***	-0.106***	-0.220***	-0.152***	-0.152***	0.143***	-0.0884***	-0.227***	0.0383*
School resources	resource1	0.158***	0.0149	-0.0282	0.112***	0.0803***	0.0594**	0.0195	-0.0589**	-0.0394*	0.140***	0.156***	0.0237
State of school building	zsbldgco	-0.139***	0.0137	0.0202	-0.0567**	-0.036	-0.0137	-0.0463*	0.0685***	0.0930***	-0.0654***	-0.0767***	0.00988
Distance to school	ptravel	-0.0574**	0.012	0.0699***	-0.0906***	-0.0573**	-0.0494**	0.00498	-0.0431*	0.118***	0.0044	-0.0736***	0.0283
School location	zsloc	0.152***	0.012	-0.210***	0.412***	0.115***	0.319***	0.183***	0.0830***	-0.198***	0.178***	0.319***	-0.0410*

Free school meals	zpmealsc	-0.0826***	0.0118	0.0228	0.0294	0.0395*	0.00108	0.00657	0.120***	0.0339	-0.0394*	0.034	-0.0189
Public or private school	zstype	0.0629***	-0.032	0.00749	0.0268	0.0611**	0.0129	0.0580**	-0.0109	-0.00579	-0.00199	0.014	-0.00548
Proportion of female teachers	spropf	0.121***	0.0357	-0.168***	0.435***	0.111***	0.292***	0.158***	0.115***	-0.199***	0.0652***	0.341***	-0.0118
Safe learning environment	abusedum3	0.110***	0.00217	-0.0750***	0.119***	0.012	0.0905***	0.0512**	0.0558**	-0.0710***	0.00446	0.0892***	-0.0690***
Problems with community	sprobcom	-0.0681***	0.00827	0.031	0.0222	0.0883***	0.0132	0.0854***	0.0564**	0.0125	0.0243	-0.00434	-0.00118
Community contributes (building, textbooks)	zscommneg	-0.185***	-0.0228	0.0829***	-0.178***	-0.0821***	-0.107***	-0.0737***	0.00706	0.131***	-0.104***	-0.181***	0.0378*
Community pays bonus to school staff	zscomm11	0.0785***	-0.0172	-0.0534**	0.0821***	0.0668***	0.0553**	-0.00765	-0.0295	-0.0184	0.00371	0.0730***	-0.0267
Class size	clsize	-0.0315	0.0204	-0.148***	0.236***	0.120***	0.142***	0.0982***	-0.0427*	-0.114***	0.0749***	0.190***	0.00881
Pupil to teacher ratio	zsptrati	-0.182***	-0.026	0.158***	-0.211***	-0.0622**	-0.204***	-0.156***	-0.0732***	0.136***	-0.0885***	-0.159***	0.0162

Notes: * p<0.05, ** p<0.01 *** p<0.001

37

Table 38: Correlation Matrix for Male and Female Sample (Part B)

		Days absent	Maths teachers living conditions	Hours taught	Knowledge of teachers	Homework given	Maths teacher interacts with parents	Teachers days absent	Teachers years of training	Head teacher gender	Head teacher age	Head teacher years' experience	Head teacher teaching hours
		pabsent	ycondliv	zhrteac~4	zaloctmr	phmwk	ymeetpar	teachabs	zqprodmr	zssex	sage	sexpall~m	Zstchhrs
Days absent	pabsent	1											
Maths teachers living conditions	ycondliv	0.0235	1										
Hours taught	zhrteach~4	0.0121	0.0144	1									
Knowledge of teachers	zaloctmr	0.00155	-0.102***	0.169***	1								
Homework given	phmwk	0.0215	0.0466*	-0.00679	0.0734***	1							
Maths teacher interacts with parents	ymeetpar	0.00102	0.0423*	-0.115***	-0.117***	0.0275	1						
Teachers days absent	teachabs	-0.0141	0.127***	-0.141***	-0.139***	-0.0287	-0.105***	1					
Teachers years of training	zqprodmr	-0.0368	0.0552**	0.100***	0.0356	-0.0175	-0.00324	0.0242	1				
Head teacher gender	zssex	-0.0512**	-0.0744***	-0.254***	-0.140***	0.0480*	0.117***	-0.0677***	0.0513**	1			
Head teacher age	sage	-0.0309	0.0993***	0.0332	0.0415*	0.0591**	-0.0646***	-0.0101	-0.0302	-0.118***	1		
Head teacher years' experience	sexpalldum	0.016	0.0448*	-0.00417	0.0922***	0.018	0.0231	0.0655***	-0.0743***	-0.146***	0.237***	1	
Head teacher teaching hours	zstchhrs	0.0034	0.149***	0.390***	0.0287	-0.131***	-0.0454*	0.0692***	-0.0276	-0.154***	-0.0256	-0.00427	1

School resources	resource1	-0.0107	0.140***	0.0696***	-0.0382*	0.0903***	-0.0510**	0.116***	0.151***	0.0336	0.0409*	0.0296	-0.145***
State of school building	zsbldgco	-0.0440*	-0.156***	-0.00266	0.0307	0.0155	-0.150***	0.0536**	-0.148***	0.0563**	0.0424*	0.143***	0.0395*
Distance to school	ptravel	0.0804***	0.013	-0.0312	-0.00486	-0.0366	-0.0311	0.0305	-0.0515**	-0.00592	-0.0106	-0.0422*	0.0437*
School location	zsloc	-0.0214	-0.178***	-0.298***	-0.0810***	0.0636***	0.117***	-0.0820***	0.108***	0.398***	-0.0204	-0.0682***	-0.478***
Free school meals	zpmealsc	-0.101***	-0.0649***	0.104***	0.00259	0.0928***	0.110***	0.188***	-0.0318	-0.0708***	-0.0663***	-0.021	0.00405
Public or private school	Zstype	0.0173	0.0905***	0.00909	0.0412*	-0.00875	0.0328	-0.0169	0.0857***	-0.0999***	-0.156***	0.0137	-0.0141
Proportion of female teachers	Spropf	-0.0550**	-0.0390*	-0.360***	-0.186***	0.0389*	0.140***	-0.0144	0.111***	0.532***	-0.0778***	-0.127***	-0.420***
Safe learning environment	abusedum3	-0.0151	-0.0628**	-0.0473*	0.0406*	0.132***	-0.0308	-0.114***	0.0302	0.0508**	0.105***	-0.150***	-0.0928***
Problems with community	sprobcom	-0.0728***	0.107***	0.000356	-0.162***	-0.115***	-0.116***	0.0285	0.0337	0.131***	-0.0252	-0.195***	0.0708***
Community contributes (building, textbooks)	zscommneg	0.115***	0.0660***	0.216***	0.0136	-0.0814***	0.125***	-0.00762	-0.111***	-0.114***	0.108***	0.215***	0.322***
Community pays bonus to school staff	zscomm11	-0.0584**	0.0987***	0.0683***	0.00611	-0.023	0.0879***	0.170***	0.0109	-0.0723***	-0.0306	0.153***	-0.101***
Class size	Clsize	-0.122***	-0.0848***	-0.230***	-0.0813***	0.0202	0.112***	0.0219	0.207***	0.161***	-0.0371	-0.0231	-0.322***
Pupil to teacher ratio	Zsptrati	0.0332	0.0915***	0.360***	0.00773	0.0416*	0.0769***	-0.0239	0.139***	-0.181***	-0.0952***	-0.0609**	0.449***

Notes: * p<0.05, ** p<0.01 *** p<0.001

Table 39: Correlation Matrix for Male and Female Sample (Part C)

Correlations for female and male sample c)		School resources	State of school building	Distance to school	School location	Free school meals	Public or private school	Proportion of female teachers	Safe learning environment	Problems with community	Community contributes (building, textbooks)	Community pays bonus to school staff
		resource1	zsbldgco	ptravel	zsloc	zpmealsc	zstype	spropf	abusedum3	sprobcom	zscommneg	zscomm11
School resources	resource1	1										
State of school building	zsbldgco	-0.200***	1									
Distance to school	ptravel	0.00438	0.0367	1								
School location	zsloc	0.222***	-0.141***	0.00349	1							
Free school meals	zpmealsc	-0.0303	0.0595**	-0.0498**	-0.0127	1						
Public or private school	zstype	-0.0376*	-0.145***	0.0284	0.0857***	0.110***	1					
Proportion of female teachers	spropf	0.114***	-0.0727***	-0.0223	0.665***	0.0701***	0.0657***	1				
Safe learning environment	abusedum3	-0.0970***	-0.0239	-0.0111	0.151***	-0.200***	0.0273	0.0649***	1			
Problems with community	sprobcom	-0.0522**	0.0461*	0.0434*	0.0058	0.0537**	0.142***	0.0757***	0.0860***	1		
Community contributes (building, textbooks)	zscommneg	-0.135***	0.236***	0.0541**	-0.217***	-0.00607	-0.192***	-0.227***	-0.131***	-0.154***	1	
Community pays bonus to school staff	zscomm11	0.181***	-0.0397*	-0.00996	0.115***	0.389***	0.183***	0.256***	-0.122***	-0.0371	-0.0348	1
Class size	clsize	0.180***	-0.141***	-0.0431*	0.459***	0.177***	0.0600**	0.393***	-0.0919***	-0.0914***	-0.0241	0.190***
Pupil to teacher ratio	zsptrati	-0.0982***	-0.0728***	-0.0188	-0.354***	0.0765***	-0.0125	-0.318***	-0.0228	0.0498**	0.147***	-0.127***

Notes: * p<0.05, ** p<0.01 *** p<0.001

39

Table 40: Correlation Matrix for Female Sample (A)

Correlations for female sample		Math and reading scores	Pupil age (months)	SES	Frequency of meals	Academic support: (parent education & books at home)	English spoken in home	Extra tuition	Chores done in the home	Used a computer	Light type in the home	Grade repetition
		scoremr	zpagemon	psesno~12	zpmealw	zpparb~59	zpenglis	zpexttui	ptasktotg	zpcomptr	zplight	Prepeatsq
Math and reading scores	scoremr	1										
Pupil age (months)	zpagemon	-0.166***	1									
SES	psesnol~12	0.111***	-0.264***	1								
Frequency of meals	zpmealw	0.0834**	-0.141***	0.137***	1							
Academic support: (parent education & books at home	zpparbo~59	0.102***	-0.261***	0.443***	0.147***	1						
English spoken in home	zpenglis	0.0861**	-0.129***	0.190***	0.0372	0.214***	1					
Extra tuition	zpexttui	0.0807**	-0.0332	0.131***	-0.00623	0.120***	0.126***	1				
Chores done in the home	ptasktotg	-0.160***	0.204***	-0.270***	-0.0484	-0.212***	-0.0931***	-0.0566*	1			
Used a computer	zpcomptr	0.0984***	-0.0609*	0.0979***	0.0533	0.143***	0.106***	0.0624*	-0.0842**	1		
Light type in the home	zplight	0.137***	-0.236***	0.783***	0.126***	0.371***	0.153***	0.140***	-0.239***	0.139***	1	
Grade repetition	prepeatsq	-0.0941***	0.180***	-0.00404	0.0136	-0.0321	-0.0791**	-0.0126	0.0688*	-0.0128	-0.0106	1
Days absent	pabsent	-0.0513	0.204***	-0.0836**	-0.132***	-0.117***	-0.0218	0.0221	0.102***	-0.0269	-0.0845**	0.000773
Math teachers living conditions	ycondliv	0.0619*	0.0977***	-0.0403	0.00828	-0.0745**	0.00039	-0.00814	0.131***	-0.0107	-0.0309	0.0481
Hours taught	zhrteach~4	-0.108***	0.111***	-0.216***	-0.0834**	-0.170***	-0.136***	-0.0803**	0.0868**	-0.0649*	-0.108***	0.00564
Knowledge of teachers	zaloctmr	0.0922***	-0.0457	-0.0715**	-0.0398	-0.011	0.0223	-0.0164	0.0226	-0.0482	-0.0215	-0.00186
Homework given	phmwk	0.161***	-0.0443	0.0868**	0.00365	0.0555*	0.0184	0.194***	-0.00257	-0.00768	0.0909***	0.0157
Math teacher interacts with parents	ymeetpar	0.0804**	-0.0293	0.0817**	-0.0103	0.0552*	-0.0154	0.0299	-0.0221	0.0352	0.0448	0.00456

Teachers days absent	teachabs	-0.0443	0.0253	-0.038	-0.00503	-0.0395	0.00411	0.0067	-0.016	-0.0437	-0.0325	-0.0167
Teachers years of training	zqprodmr	0.0707**	-0.0148	0.0718**	0.0695*	0.00611	-0.154***	-0.0484	-0.051	0.0231	0.0716**	0.0266
Head teacher gender	zssex	0.0613*	-0.114***	0.281***	0.0648*	0.219***	0.0756**	0.146***	-0.162***	0.000961	0.229***	0.00743
Head teacher age	sage	0.0821**	-0.0629*	0.000285	-0.0112	-0.00275	0.0595*	-0.109***	0.0129	-0.0149	0.00672	-0.0811**
Head teacher years' experience	sexpalldum	-0.0729**	0.00359	-0.0668*	0.0306	-0.0314	-0.0521	-0.114***	0.120***	-0.0394	-0.0982***	0.00204
Head teacher teaching hours	zstchhrs	-0.145***	0.181***	-0.269***	-0.0544*	-0.213***	-0.112***	-0.174***	0.215***	-0.0464	-0.226***	0.0281
School resources	resource1	0.144***	-0.0515	0.0969***	0.0910***	0.0677*	0.000954	-0.0208	-0.0236	0.163***	0.125***	0.0168
State of school building	zsbldgco	-0.159***	0.0133	-0.0958***	-0.0132	-0.0078	-0.0476	0.00826	0.0812**	-0.0852**	-0.0711**	0.0223
Distance to school	ptravel	-0.0604*	0.0495	-0.0943***	-0.0678*	-0.0425	0.0111	-0.0587*	0.123***	-0.0146	-0.0585*	0.0315
School location	zsloc	0.140***	-0.234***	0.416***	0.0850**	0.301***	0.184***	0.113***	-0.268***	0.172***	0.340***	-0.0564*
Free school meals	zpmealsc	-0.0732**	0.0212	0.0219	0.00131	-0.00212	0.0315	0.137***	0.0610*	-0.0459	0.0145	0.0141
Public or private school	zstype	0.0369	0.0225	-0.0325	0.0484	-0.00917	0.0107	0.00689	0.0213	0.0109	-0.0579*	0.0281
Proportion of female teachers	spropf	0.122***	-0.157***	0.423***	0.0931***	0.265***	0.167***	0.162***	-0.246***	0.0451	0.330***	-0.00808
Safe learning environment	abusedum3	0.157***	-0.120***	0.119***	0.018	0.0992***	0.0702*	0.105***	-0.129***	0.0102	0.124***	-0.0856**
Problems with community	sprobcom	-0.0975***	0.0452	0.0119	0.0798**	0.0285	0.0719**	0.0635*	-0.00462	0.0393	-0.00667	-0.00652
Community contributes (building, textbooks)	zscommneg	-0.182***	0.0802**	-0.158***	-0.0885**	-0.104***	-0.0256	-0.037	0.154***	-0.0976***	-0.153***	0.0191
Community pays bonus to school staff	zscomm11	0.0755**	-0.05	0.0382	0.0599*	0.0309	-0.0264	-0.000367	-0.0207	-0.0337	0.0355	-0.0505
Class size	clsize	-0.0547*	-0.163***	0.244***	0.121***	0.123***	0.112***	-0.00935	-0.130***	0.0607*	0.180***	0.00589
Pupil to teacher ratio	zsptrati	-0.176***	0.154***	-0.190***	-0.0682*	-0.209***	-0.0993***	-0.0860**	0.217***	-0.0690*	-0.158***	0.00927
Notes: * p<0.05, ** p<0.01 *** p<0.001												

40

Table 41: Correlation Matrix for Female Sample (B)

Correlations for female sample		Days absent	Math teachers living conditions	Hours taught	Knowledge of teachers	Homework given	Math teacher interacts with parents	Teachers days absent	Teachers years of training	Head teacher sex	Head teacher age	Head teacher years' experience
		pabsent	ycondliv	zhrteac~4	zaloctmr	Phmwk	ymeetpar	teachabs	zqprodmr	zssex	sage	sexpall~m
Days absent	pabsent	1										
Math teachers living conditions	ycondliv	0.0562*	1									
Hours taught	zhrteach~4	0.00753	0.0580*	1								
Knowledge of teachers	zaloctmr	0.0107	-0.119***	0.163***	1							
Homework given	phmwk	-0.00211	0.033	-0.0243	0.0944***	1						
Math teacher interacts with parents	ymeetpar	-0.0379	0.0363	-0.192***	-0.150***	0.0533	1					
Teachers days absent	teachabs	-0.0434	0.0896**	-0.130***	-0.122***	0.000617	-0.0703**	1				
Teachers years of training	zqprodmr	-0.00103	0.0798**	0.112***	-0.031	-0.0277	-0.0139	0.000207	1			
Head teacher gender	zssex	-0.0798**	-0.0558*	-0.283***	-0.170***	0.0739**	0.118***	-0.0664*	0.0276	1		
Head teacher age	sage	-0.0447	0.102***	0.0015	0.0693*	0.0792**	-0.0321	-0.0491	-0.0843**	-0.0892**	1	
Head teacher years' experience	sexpalldum	0.0596*	0.0472	-0.0276	0.110***	0.037	0.0338	0.0199	-0.0712**	-0.151***	0.221***	1
Head teacher teaching hours	zstchhrs	-0.0245	0.146***	0.415***	0.0408	-0.118***	-0.0767**	0.100***	-0.00574	-0.159***	-0.0324	0.0024
School resources	resource1	0.0178	0.172***	0.0841**	-0.0533	0.120***	0.0025	0.0893**	0.153***	0.0386	0.0252	0.0146
State of school building	zsbldgco	-0.035	-0.177***	0.00373	0.0282	0.00755	-0.179***	0.0678*	-0.151***	0.0411	0.0579*	0.137***
Distance to school	ptravel	0.123***	-0.0277	-0.0224	0.0476	-0.0362	-0.0543*	0.0338	-0.000742	-0.0387	-0.00835	-0.0489
School location	zsloc	-0.0196	-0.181***	-0.325***	-0.103***	0.0564*	0.157***	-0.0762**	0.0995***	0.397***	-0.0144	-0.0782**

Free school meals	zpmealsc	-0.0989***	-0.0902***	0.0786**	0.0139	0.120***	0.122***	0.201***	-0.0165	-0.0445	-0.0425	-0.0383
Public or private school	zstype	0.0277	0.0694*	0.0681*	0.044	-0.0251	0.0426	-0.00119	0.114***	-0.0863**	-0.148***	0.0421
Proportion of female teachers	spropf	-0.0542*	-0.0379	-0.392***	-0.195***	0.0575*	0.163***	-0.017	0.103***	0.554***	-0.0621*	-0.137***
Safe learning environment	abusedum3	-0.0534	-0.114***	-0.0780**	0.0594*	0.115***	-0.0386	-0.0993***	0.0212	0.0377	0.0761**	-0.166***
Problems with community	sprobcom	-0.122***	0.122***	0.0297	-0.179***	-0.125***	-0.102***	0.0429	0.0348	0.146***	-0.031	-0.206***
Community contributes (building, textbooks)	zscommneg	0.122***	0.0783**	0.189***	0.0241	-0.0747**	0.0951***	0.0054	-0.136***	-0.137***	0.0940***	0.205***
Community pays bonus to school staff	zscomm11	-0.0443	0.101***	0.0751**	0.00665	0.00124	0.118***	0.139***	0.001	-0.0696*	-0.0497	0.141***
Class size	clsize	-0.107***	-0.103***	-0.264***	-0.0880**	0.0174	0.115***	0.0205	0.187***	0.163***	-0.0382	-0.0296
Pupil to teacher ratio	zsptrati	-0.00734	0.0512	0.341***	0.0319	0.0144	0.036	0.015	0.147***	-0.203***	-0.135***	-0.0303

Notes: * p<0.05, ** p<0.01 *** p<0.001

41

Table 42: Correlation Matrix for Female Sample (C)

		Head teacher teaching hours	School resources	State of school building	Distance to school	School location	Free school meals	Public or private school	Proportion of female teachers	Safe learning environment	Problems with community	Community contributes (building, textbooks)	Community pays bonus to school staff	Class size	Pupil to teacher ratio
		zstchhrs	resource1	zsbldgco	ptravel	zsloc	zpmealsc	Zstype	spropf	abusedum3	sprobcom	zscommneg	zscomm11	clsize	zsprati
Head teacher teaching hours	zstchhrs	1													
School resources	resource1	-0.144***	1												
State of school building	zsbldgco	0.0485	-0.209***	1											
Distance to school	ptravel	0.042	0.0145	0.00585	1										
School location	zsloc	-0.469***	0.233***	-0.183***	-0.00588	1									
Free school meals	zpmealsc	0.0234	-0.0445	0.0439	-0.0528	0.00664	1								
Public or private school	zstype	0.0103	-0.0264	-0.122***	0.0518	0.0881**	0.101***	1							
Proportion of female teachers	spropf	-0.424***	0.109***	-0.0803**	-0.0532	0.668***	0.0838**	0.0653*	1						
Safe learning environment	abusedum3	-0.116***	-0.103***	-0.00793	-0.00639	0.159***	-0.206***	0.0199	0.0729**	1					
Problems with community	sprobcom	0.0796**	-0.0557*	0.0199	0.0211	-0.0136	0.0759**	0.148***	0.0715**	0.0946***	1				
Community contributes (building, textbooks)	zscommneg	0.323***	-0.116***	0.220***	0.0144	-0.235***	-0.0112	-0.179***	-0.234***	-0.131***	-0.156***	1			

Community pays bonus to school staff	zscomm 11	-0.0903***	0.162***	-0.0451	0.0046	0.113***	0.362***	0.202***	0.244***	-0.111***	-0.0406	-0.0336	1		
Class size	clsize	-0.335***	0.150***	-0.144***	-0.0568*	0.478***	0.174***	0.0793**	0.394***	-0.0918***	-0.0761**	-0.0172	0.192***	1	
Pupil to teacher ratio	zsptrati	0.467***	-0.104***	-0.0541*	0.00572	-0.356***	0.0918***	0.000437	-0.328***	-0.0642*	0.0332	0.135***	-0.105***	0.0334	1
Notes: * p<0.05, ** p<0.01 *** p<0.001															

42

Table 43: Correlation Matrix for Male Sample (A)

		Math and reading scores	Pupil age (months)	SES	Frequency of meals	Academic support: (parent education & books at home)	English spoken in home	Extra tuition	Chores done in the home	Used a computer	Light type in the home	Grade repetition	Days absent	Math teachers living conditions
		scoremr	zpagemon	psesno~12	zpmealw	zpparb~59	Zpenglis	zpexttui	ptasktotg	zpcomptr	zplight	prepeatsq	pabsent	ycondliv
Math and reading scores	scoremr	1												
Pupil age (months)	zpagemon	-0.116***	1											
SES	psesnol~12	0.199***	-0.296***	1										
Frequency of meals	zpmealw	0.0913***	-0.108***	0.178***	1									
Academic support: (parent education & books at home)	zpparbo~59	0.153***	-0.289***	0.519***	0.177***	1								
English spoken in home	zpenglis	0.176***	-0.0606*	0.150***	0.0403	0.134***	1							
Extra tuition	zpexttui	0.0622*	-0.000187	0.116***	0.0331	0.111***	0.106***	1						
Chores done in the home	ptasktotg	-0.124***	0.177***	-0.188***	-0.0729**	-0.0895***	-0.0177	-0.00267	1					
Used a computer	zpcomptr	0.178***	-0.135***	0.221***	0.0503	0.221***	0.0965***	0.0991***	-0.0661*	1				
Light type in the home	zplight	0.213***	-0.230***	0.746***	0.185***	0.430***	0.0846**	0.0959***	-0.223***	0.172***	1			
Grade repetition	prepeatsq	-0.0767**	0.188***	-0.0373	-0.0385	-0.0675*	-0.0447	-0.0244	0.106***	-0.0361	-0.0355	1		
Days absent	pabsent	-0.110***	0.144***	-0.106***	-0.101***	-0.0594*	-0.118***	-0.0134	0.0983***	-0.0674*	-0.0905***	0.0489	1	
Math teachers living conditions	ycondliv	0.0266	0.00493	-0.0506	-0.0257	-0.0974***	-0.0386	-0.0550*	0.112***	-0.0672*	-0.0205	0.0632*	-0.00521	1
Hours taught	zhrteach~4	-0.117***	0.0952***	-0.201***	-0.0877**	-0.152***	-0.144***	-0.0495	0.0491	-0.0592*	-0.146***	-0.00868	0.0158	-0.0254

Knowledge of teachers	zaloctmr	-0.000482	-0.0228	-0.0184	0.014	0.00419	-0.0322	0.00659	0.0401	0.0252	0.0277	0.0164	-0.0068	-0.0881***
Homework given	phmwk	0.0929***	-0.0613*	0.0874**	0.0620*	0.0171	-0.0817**	0.125***	-0.0458	0.00385	0.0778**	-0.0138	0.0422	0.0590*
Math teacher interacts with parents	ymeetpar	0.0112	-0.0375	0.0282	0.00138	0.0362	-0.0469	-0.0158	0.0257	0.0381	-0.00737	-0.0503	0.0338	0.0472
Teachers days absent	teachabs	-0.0654*	0.0298	-0.0743**	0.00271	-0.0176	0.0285	-0.0575*	0.0165	-0.0402	-0.0624*	0.0442	0.0062	0.155***
Teachers years of training	zqprodmr	0.105***	-0.0707**	0.0807**	0.0348	0.0702**	-0.0192	-0.0403	-0.125***	0.0379	0.115***	-0.00574	-0.0681*	0.0333
Head teacher gender	zssex	0.0502	-0.132***	0.235***	0.0605*	0.214***	0.0465	0.0735**	-0.0735**	0.0238	0.170***	0.0349	-0.0204	-0.0919***
Head teacher age	Sage	0.0630*	-0.0317	0.0246	-0.0791**	-0.0262	0.0374	-0.104***	-0.00145	0.00649	-0.0196	-0.0227	-0.0191	0.0968***
Head teacher years' experience	sexpalldum	-0.0750**	-0.0253	-0.0694**	-0.0343	-0.0578*	-0.0224	-0.0429	0.0622*	-0.0159	-0.0639*	-0.0122	-0.0226	0.0425
Head teacher teaching hours	zstchhrs	-0.190***	0.159***	-0.289***	-0.155***	-0.227***	-0.195***	-0.129***	0.0765**	-0.121***	-0.227***	0.0477	0.0282	0.151***
School resources	resource1	0.175***	-0.00558	0.126***	0.0691**	0.0514	0.0399	-0.0995***	-0.0602*	0.127***	0.186***	0.0308	-0.0344	0.112***
State of school building	zsbldgco	-0.122***	0.0315	-0.0213	-0.0588*	-0.0195	-0.0432	0.130***	0.102***	-0.0519	-0.0829**	-0.000864	-0.0509	-0.136***
Distance to school	ptravel	-0.0532*	0.0933***	-0.0874**	-0.0493	-0.0561*	0.000969	-0.028	0.111***	0.0186	-0.0890***	0.0261	0.0447	0.0498
School location	zsloc	0.168***	-0.196***	0.409***	0.142***	0.337***	0.185***	0.0503	-0.138***	0.187***	0.297***	-0.0259	-0.022	-0.176***
Free school meals	zpmealsc	-0.0896***	0.0294	0.0364	0.0742**	0.00423	-0.0163	0.101***	0.00439	-0.0348	0.0531*	-0.0494	-0.102***	-0.0413
Public or private school	zstype	0.0787**	-0.0152	0.0794**	0.0766**	0.034	0.0985***	-0.0267	-0.0214	-0.0131	0.0848**	-0.0376	0.00541	0.108***
Proportion of female teachers	spropf	0.131***	-0.173***	0.450***	0.125***	0.321***	0.156***	0.0595*	-0.168***	0.0850**	0.351***	-0.0134	-0.0527*	-0.0391
Safe learning environment	abusedum3	0.0716**	-0.0408	0.119***	0.00615	0.0820**	0.0331	0.00468	-0.0175	0.000641	0.0550*	-0.0535*	0.0192	-0.0163
Problems with community	sprobcom	-0.0426	0.0239	0.0316	0.0955***	-0.00174	0.100***	0.0485	0.0267	0.0148	-0.00246	0.00422	-0.0285	0.0936***

Community contributes (building, textbooks)	zscommneg	-0.195***	0.0804**	-0.197***	-0.0739**	-0.109***	-0.124***	0.0538*	0.120***	-0.112***	-0.208***	0.0537*	0.108***	0.0543*
Community pays bonus to school staff	zscomm11	0.0785**	-0.0641*	0.120***	0.0752**	0.0782**	0.00726	-0.0574*	-0.0114	0.0279	0.109***	-0.00653	-0.0722**	0.0966***
Class size	clsize	-0.00771	-0.134***	0.230***	0.118***	0.162***	0.0882***	-0.0797**	-0.107***	0.0884***	0.199***	0.0127	-0.134***	-0.0672*
Pupil to teacher ratio	zsptrati	-0.194***	0.158***	-0.229***	-0.0543*	-0.200***	-0.214***	-0.0591*	0.0747**	-0.105***	-0.158***	0.021	0.0648*	0.125***

Notes: * p<0.05, ** p<0.01 *** p<0.001

43

Table 44: Correlation Matrix for Male Sample (B)

		Hours taught	Knowledge of teachers	Homework given	Math teacher interacts with parents	Teachers days absent	Teachers years of training	Head teacher gender	Head teacher age	Head teacher years' experience	Head teacher teaching hours	School resources	State of school building
		zhrteac~4	zaloctmr	phmwk	ymeetpar	teachabs	zqprodmr	zssex	sage	sexpall~m	zstchhrs	resource1	zsbldgco
Hours of teaching (1= above 30 hours)	zhrteach~4	1											
Knowledge of teachers	zaloctmr	0.176***	1										
Homework given	phmwk	0.0103	0.0530*	1									
Math teacher interacts with parents	ymeetpar	-0.0405	-0.0854**	0.00243	1								
Teachers days absent	teachabs	-0.151***	-0.154***	-0.0529*	-0.135***	1							
Teachers years of training	zqprodmr	0.0892***	0.0987***	-0.00774	0.00693	0.0436	1						
Head teacher gender	zssex	-0.225***	-0.110***	0.0212	0.117***	-0.0689**	0.0759**	1					
Head teacher age	sage	0.0637*	0.0152	0.0397	-0.0955***	0.021	0.0211	-0.147***	1				
Head teacher years experience (1= 13 years or above experience)	sexpalldum	0.018	0.0755**	-0.000267	0.0128	0.102***	-0.0771**	-0.140***	0.253***	1			
Head teacher teaching hours	zstchhrs	0.364***	0.0169	-0.145***	-0.0149	0.0439	-0.0489	-0.149***	-0.0191	-0.0108	1		
School resources	resource1	0.0558*	-0.0237	0.0621*	-0.101***	0.139***	0.149***	0.0274	0.0559*	0.044	-0.145***	1	
State of school building	zsbldgco	-0.00873	0.0333	0.0233	-0.122***	0.043	-0.146***	0.0714**	0.0276	0.149***	0.0308	-0.192***	1

Distance to school	ptravel	-0.0396	-0.0545*	-0.0369	-0.00874	0.0284	-0.0996***	0.0271	-0.0127	-0.0357	0.0455	-0.00555	0.0658*
School location	zsloc	-0.271***	-0.0596*	0.0707**	0.0796**	-0.0872**	0.116***	0.398***	-0.0261	-0.0584*	-0.487***	0.212***	-0.101***
Free school meals	zpmealsc	0.129***	-0.00837	0.0661*	0.0987***	0.179***	-0.0467	-0.0999***	-0.0894***	-0.00423	-0.0152	-0.0167	0.0745**
Public or private school	zstype	-0.0468	0.0385	0.00643	0.0227	-0.0303	0.0600*	-0.112***	-0.163***	-0.0131	-0.0378	-0.0473	-0.166***
Proportion of female teachers	spropf	-0.328***	-0.177***	0.02	0.119***	-0.0113	0.121***	0.506***	-0.0938***	-0.116***	-0.416***	0.119***	-0.0661*
Safe learning environment	abusedum3	-0.0173	0.0226	0.148***	-0.0233	-0.127***	0.0388	0.0647*	0.133***	-0.135***	-0.0701**	-0.0916***	-0.0394
Problems with community	sprobcom	-0.0279	-0.147***	-0.106***	-0.129***	0.0174	0.0326	0.115***	-0.0197	-0.185***	0.0622*	-0.0492	0.0711**
Community contributes (building, textbooks)	zscommneg	0.242***	0.00355	-0.0882***	0.153***	-0.0186	-0.0863**	-0.0893***	0.122***	0.224***	0.322***	-0.153***	0.252***
Community pays bonus to school staff	zscomm11	0.0619*	0.00553	-0.0457	0.0598*	0.195***	0.02	-0.0741**	-0.0132	0.163***	-0.111***	0.200***	-0.0342
Class size	clsize	-0.196***	-0.0747**	0.0232	0.109***	0.024	0.227***	0.157***	-0.0359	-0.0164	-0.308***	0.208***	-0.138***
Pupil to teacher ratio	zsptrati	0.379***	-0.0143	0.0663*	0.113***	-0.0543*	0.133***	-0.159***	-0.0594*	-0.0887***	0.433***	-0.0919***	-0.0892***

Notes: * p<0.05, ** p<0.01 *** p<0.001

44

Table 45: Correlation Matrix for Male Sample (C)

		Distance to school	School location	Free school meals	Public or private school	Proportion of female teachers	Safe learning environment	Problems with community	Community contributes (building, textbooks)	Community pays bonus to school staff	Class size	Pupil to teacher ratio
		Ptravel	zsloc	zpmealsc	zstype	Spropf	abusedum3	sprobcom	zscommneg	zscomm11	clsize	Zsptrati
Distance to school	Ptravel	1										
School location	Zsloc	0.0123	1									
Free school meals	zpmealsc	-0.0472	-0.0323	1								
Public or private school	Zstype	0.00741	0.0842**	0.119***	1							
Proportion of female teachers	Spropf	0.00814	0.663***	0.0552*	0.0687*	1						
Safe learning environment	abusedum3	-0.0157	0.143***	-0.195***	0.0345	0.0565*	1					
Problems with community	sprobcom	0.0644*	0.0246	0.0319	0.138***	0.0796**	0.0777**	1				
Community contributes (building, textbooks)	zscommneg	0.0920***	-0.199***	-0.000565	-0.205***	-0.219***	-0.130***	-0.151***	1			
Community pays bonus to school staff	zscomm11	-0.0229	0.117***	0.416***	0.165***	0.271***	-0.133***	-0.0336	-0.0366	1		
Class size	clsize	-0.0303	0.441***	0.180***	0.0429	0.392***	-0.0920***	-0.107***	-0.0298	0.190***	1	
Pupil to teacher ratio	zsptrati	-0.0403	-0.352***	0.0630*	-0.0256	-0.309***	0.0152	0.0653*	0.157***	-0.147***	0.00203	1

Notes: * p<0.05, ** p<0.01 *** p<0.001

45

Appendix 3: Regression Tables

Table 46: Table of Regressions for All Samples

	Regression 1		**Regression 2**		**Regression 1**		**Regression 2**		**Regression 3**		**Regression 4**		**Regression 1**		**Regression 3**	
Excluded regardless of significance:	None.		zscommneg, zscomm11 &zpmealsc		None.		zaloctmr		zscommneg, zscomm11 &zpmealsc		zaloctmr, zscommneg, zscomm11 &zpmealsc		None.		zpmealsc&zscommneg zcomm11	
Gender of sample	Male and female		Male and female		Female		Female		Female		Female		Male		Male	
N	2780		2780		1342		1364		1342		1364		1416		1416	
R^2	0.202		0.19		0.206		0.19		0.195		0.185		0.194		0.186	
adj. R^2	0.195		0.184		0.193		0.179		0.182		0.173		0.183		0.176	
Output=	Scoremr		Scoremr		scoremr		scoremr		scoremr		Scoremr		scoremr		Scoremr	
Coefficient reported.		beta		Beta		beta		beta		beta		beta		beta		Beta
1 Pupil Characteristics																
Gender	**-9.648*****	**-0.099*****	**-10.49*****	**-0.108*****												
CI	-13.131 to -6.164		-14.106 to -6.874													
CI SVY	-13.56 to -5.735		-14.70 to -6.279													
Age			**-0.0887***	**-0.040***	**-0.198****	**-0.084****	**-0.230*****	**-0.097*****	**-0.217****	**-0.092****	**-0.208****	**-0.088****				
CI			-0.174 to -0.004		-0.330 to -0.066		-0.358 to -0.102		-0.349 to -0.084		-0.338 to -0.078					
CI SVY			-0.191 to 0.0136		-0.357 to -0.0380		-0.384 to -0.0750		-0.380 to -0.0538		-0.373 to -0.0430					

Table 46: Table of Regressions for All Samples

	Regression 1		Regression 2		Regression 1		Regression 2		Regression 3		Regression 4		Regression 1		Regression 3	
SES																
CI																
CI SVY																
2 Home Environment																
Frequency of meals	**1.581****	**0.046****	**1.521***	**0.044***	**1.923***	**0.059***	**1.734***	**0.053***	**2.076***	**0.063***	**1.880***	**0.057***				
CI	0.390 to 2.772		0.334 to 2.709		0.250 to 3.597		0.118 to 3.350		0.393 to 3.759		0.254 to 3.507					
CI SVY	-0.219 to 3.382		-0.273 to 3.315		-0.150 to 3.997		-0.245 to 3.713		0.0296 to 4.123		-0.0950 to 3.856					
Academic support: parent education & books at home																
CI																
CI SVY																
English spoken at home	**9.205*****	**0.095*****	**8.828*****	**0.091*****	**4.848***	**0.054***	**5.746***	**0.064***			**4.890***	**0.054***	**12.62*****	**0.122*****	**12.28*****	**0.119*****
CI	5.758 to 12.653		5.335 to 12.321		0.097 to 9.599		1.028 to 10.464				0.114 to 9.666		7.455 to 17.792		7.103 to 17.452	
CI SVY	3.702 to 14.71		3.008 to 14.65		-2.105 to 11.80		-0.936 to 12.43				-1.998 to 11.78		5.717 to 19.53		5.416 to 19.14	

Table 46: Table of Regressions for All Samples

	Regression 1		Regression 2		Regression 1		Regression 2		Regression 3		Regression 4		Regression 1		Regression 3	
Extra tuition	**5.857***	**0.042***														
CI	0.332 to 11.383															
CI SVY	-4.083 to 15.80															
Chores done in the home	**-1.040****	**-0.062****	**-0.991****	**-0.059****	**-1.347****	**-0.085****	**-1.477*****	**-0.094*****	**-1.509*****	**-0.095*****	**-1.365****	**-0.087****				
CI	-1.660 to -.420		-1.621 to -0.361		-2.186 to -0.509		-2.332 to -0.622		-2.350 to -0.667		-2.236 to -0.494					
CI SVY	-2.024 to -0.0560		-1.981 to -0.0011		-2.567 to -0.128		-2.682 to -0.272		-2.737 to -0.280		-2.629 to -0.101					
Access to a computer	**29.65*****	**0.077*****	**31.25*****	**0.081*****	**22.58****	**0.053****	**21.98****	**0.051****	**23.17****	**0.054****	**19.75***	**0.046***	**37.99*****	**0.106*****	**38.52*****	**0.107*****
CI	14.292 to 45.001		15.910 to 46.590		5.999 to 39.168		5.931 to 38.022		6.114 to 40.227		3.247 to 36.247		17.648 to 58.340		18.218 to 58.822	
CI SVY	16.80 to 42.50		18.55 to 43.95		11.77 to 33.40		12.21 to 31.75		10.27 to 36.08		7.935 to 31.56		23.30 to 52.68		23.57 to 53.47	
Light type in the home	**4.404****	**0.063****	**4.505****	**0.065****									**8.011*****	**0.107*****	**8.164*****	**0.109*****
CI	1.619 to 7.189		1.747 to 7.263										3.776 to 12.247		3.910 to 12.419	
CI SVY	0.722 to 8.086		0.879 to 8.132										3.366 to 12.66		3.436 to 12.89	
3 Pupil Academic Record																
Grade repetition	**-0.698****	**-0.055****	**-0.660****	**-0.052****	**-0.660***	**-0.055***			**-0.775***	**-0.065***	**-0.682***	**-0.057***	**-0.698***	**-0.053***	**-0.663***	**-0.051***

Table 46: Table of Regressions for All Samples

	Regression 1		Regression 2		Regression 1		Regression 2		Regression 3		Regression 4		Regression 1		Regression 3	
CI	-1.143 to -0.253		-1.108 to -0.211		-1.279 to -0.041				-1.382 to -0.168		-1.304 to -0.060		-1.344 to -0.052		-1.306 to -0.021	
CI SVY	-1.203 to -0.193		-1.183 to -0.136		-1.387 to 0.0675				-1.485 to -0.0654		-1.395 to 0.0309		-1.454 to 0.0579		-1.411 to 0.0840	
Days absent	**-1.519*** **	**-0.072*** **	**-1.515*** **	**-0.072*** **									**-1.886*** **	**-0.088*** **	**-1.854*** **	**-0.087*** **
CI	-2.257 to -0.781		-2.262 to -0.768												-2.913 to -0.794	
CI SVY	-2.694 to -0.345		-2.775 to -0.255												-3.313 to -0.395	
4 Teacher Characteristics																
Teacher living conditions math					**2.940***	**0.063***			**2.857***	**0.061***						
CI					0.285 to 5.595				0.264 to 5.450							
CI SVY					-1.438 to 7.317				-1.535 to 7.249							
5 Teaching Quality																
Hours of teaching 1= above 30 hours	**-11.07*** **	**-0.105*** **	**-11.87*** **	**-0.113*** **	**-9.827** **	**-0.100** **	**-14.21*** **	**-0.146*** **	**-11.01*** **	**-0.113*** **	**-11.78*** **	**-0.121*** **	**-12.44*** **	**-0.112*** **	**-11.63*** **	**-0.104*** **
CI	-15.652 to -6.494		-16.193 to -7.552		-16.045 to -3.609		-19.970 to -8.441		-16.906 to -5.119		-17.684 to -5.874		-18.819 to -6.055		-17.869 to -5.381	
CI SVY	-20.53 to -1.619		-21.13 to -2.613		-20.44 to 0.785		-24.45 to -3.958		-20.86 to -1.168		-21.85 to -1.710		-22.47 to -2.404		-21.56 to -1.690	

Table 46: Table of Regressions for All Samples

	Regression 1		Regression 2		Regression 1		Regression 2		Regression 3		Regression 4		Regression 1		Regression 3	
Knowledge of teachers					**0.0323*** **	**0.096*** **			**0.0331*** **	**0.098*** **						
CI					0.013 to 0.051				0.014 to 0.052							
CI SVY					-0.0066 to 0.0712				-0.0070 to 0.0731							
Homework given	**6.044*** **	**0.099*** **	**5.786*** **	**0.094*** **	**5.940*** **	**0.106*** **	**6.284*** **	**0.111*** **	**6.335*** **	**0.113*** **	**6.841*** **	**0.121*** **	**6.257*** **	**0.096*** **	**5.641*** **	**0.087*** **
CI	3.844 to 8.245		3.613 to 7.958		3.127 to 8.752		3.458 to 9.110		3.50 to 9.172		3.981 to 9.701		2.990 to 9.523		2.380 to 8.901	
CI SVY	1.580 to 10.51		1.210 to 10.36		1.707 to 10.17		1.791 to 10.78		1.977 to 10.69		2.229 to 11.45		1.004 to 11.51		0.300 to 10.98	
Math teacher interacts with parents	**1.963***	**0.040***			**2.960***	**0.065***			**2.699***	**0.060***						
CI	0.165 to 3.761				0.557 to 5.364				0.340 to 5.059							
CI SVY	-1.354 to 5.280				-0.607 to 6.528				-0.917 to 6.316							
Teachers days absent	**-0.142*** **	**-0.063*** **	**-0.137*** **	**-0.061*** **									**-0.186*** **	**-0.085*** **	**-0.175** **	**-0.080** **
CI	-0.223 to -0.061		-0.217 to -0.057										-0.292 to -0.080		-0.281 to -0.070	
CI SVY	-0.307 to 0.0231		-0.314 to 0.0390										-0.363, to 0.00927		-0.373 to 0.0225	
Teachers years of training	**3.479*** **	**0.100*** **	**3.493*** **	**0.101*** **	**2.935** **	**0.091** **	**2.884** **	**0.089** **	**2.650** **	**0.083** **	**3.034** **	**0.094** **	**3.810*** **	**0.105*** **	**3.909*** **	**0.108*** **

Table 46: Table of Regressions for All Samples

	Regression 1		Regression 2		Regression 1		Regression 2		Regression 3		Regression 4		Regression 1		Regression 3	
CI	2.137 to 4.821		2.125 to 4.860		1.030 to 4.839		0.972 to 4.796		0.751 to 4.549		1.082 to 4.986		1.888 to 5.732		1.975 to 5.843	
CI SVY	0.0880 to 6.869		-0.0748 to 7.061		-0.860 to 6.730		-1.129 to 6.897		-1.361 to 6.661		-1.074 to 7.143		0.329 to 7.291		0.310 to 7.507	
6 School Head Characteristics																
Head teacher Gender																
CI																
CI SVY																
Head teacher age	**0.797*** **	**0.086*** **	**0.631*** **	**0.068*** **	**0.586***	**0.069***	**0.814*** **	**0.095*** **	**0.572***	**0.067***	**0.638****	**0.074****	**0.658***	**0.068***	**0.678****	**0.070****
CI	0.445 to 1.148		0.286 to 0.976		0.105 to 1.067		0.345 to 1.283		0.098 to 1.046		0.164 to 1.111		0.148 to 1.169		0.168 to 1.187	
CI SVY	-0.118 to 1.711		-0.235 to 1.497		-0.378 to 1.550		-0.157 to 1.785		-0.347 to 1.491		-0.330 to 1.606		-0.298 to 1.615		-0.268 to 1.624	
7 School Head Quality Of Management																
Head teacher years' experience	**-12.61*** **	**-0.088*** **	**-11.08*** **	**-0.077*** **	**-12.70*** **	**-0.092*** **	**-12.10*** **	**-0.090*** **	**-13.65*** **	**-0.099*** **	**-11.05****	**-0.082****	**-14.07****	**-0.094****	**-12.26****	**-0.082****
CI	-18.273 to -6.942		-16.566 to -5.601		-19.488 to -5.918		-18.943 to -5.265		-20.305 to -6.987		-17.703 to -4.391		-23.004 to -5.134		-21.032 to -3.485	
CI SVY	-24.87 to -0.346		-22.65 to 0.485		-25.14 to -0.270		-23.49 to -0.714		-25.23 to -2.064		-21.83 to -0.263		-28.80 to 0.666		-26.43 to 1.909	
Head teacher teaching hours																

Table 46: Table of Regressions for All Samples

	Regression 1		Regression 2		Regression 1		Regression 2		Regression 3		Regression 4		Regression 1		Regression 3	
CI																
CI SVY																
8 Classroom Environment																
Class size	**-0.266*** **	**-0.172*** **	**-0.266*** **	**-0.171*** **	**-0.247*** **	**-0.174*** **	**-0.308*** **	**-0.216*** **	**-0.232*** **	**-0.164*** **	**-0.268*** **	**-0.188*** **	**-0.258*** **	**-0.157*** **	**-0.252*** **	**-0.153*** **
CI	-0.331 to -0.202		-0.330 to -0.202		-0.332 to -0.161		-0.394 to -0.222		-0.316 to -0.147		-0.354 to -0.182		-0.355 to -0.162		-0.346 to -0.157	
CI SVY	-0.435 to -0.0981		-0.432 to -0.100		-0.410 to -0.0837		-0.484 to -0.131		-0.396 to -0.0673		-0.443 to -0.0920		-0.445 to -0.0717		-0.435 to -0.0686	
Pupil teacher ratio	**-0.082*** **	**-0.073*** **	**-0.104*** **	**-0.093*** **	**-0.0811** * **	**-0.076** **			**-0.097*** **	**-0.091*** **	**-0.081** **	**-0.076** **	**-0.099** **	**-0.086** **	**-0.122*** **	**-0.106*** **
CI	-0.125 to -0.039		-0.146 to -0.062		-0.139 to -0.024				-0.154 to -0.041		-0.138 to -0.0238		-0.160 to -0.037		-0.183 to -0.061	
CI SVY	-0.176 to 0.0123		-0.195 to -0.0128		-0.174 to 0.0119				-0.187 to -0.0078		-0.170 to 0.0081		-0.208 to 0.0109		-0.225 to -0.0194	
9 School Characteristics																
School resources	**2.525*** **	**0.104*** **	**3.104*** **	**0.127*** **	**1.795** **	**0.079** **	**2.229*** **	**0.098*** **	**2.254*** **	**0.100*** **	**2.422*** **	**0.107*** **	**3.140*** **	**0.123*** **	**3.554*** **	**0.139*** **
CI	1.547 to 3.502		2.103 to 4.106		0.530 to 3.061		0.990 to 3.468		0.977 to 3.530		1.141 to 3.703		1.650 to 4.629		2.070 to 5.037	
CI SVY	0.285 to 4.764		0.717 to 5.492		-0.372 to 3.963		0.0140 to 4.443		0.0750 to 4.432		0.159 to 4.685		0.382 to 5.897		0.785 to 6.322	
State of school building	**-5.902** **	**-0.061** **	**-7.901*** **	**-0.081*** **	**-7.898** **	**-0.088** **	**-9.205*** **	**-0.102*** **	**-8.987*** **	**-0.100*** **	**-10.39*** **	**-0.115*** **	**-5.851***	**-0.057***	**-6.677***	**-0.065***

Table 46: Table of Regressions for All Samples

	Regression 1		Regression 2		Regression 1		Regression 2		Regression 3		Regression 4		Regression 1		Regression 3	
CI	-9.615 to -2.188		-11.480 to -4.321		-12.866 to -2.931		-13.907 to -4.503		-13.883 to -4.090		-15.094 to -5.677		-11.273 to -0.430		-12.106 to -1.247	
CI SVY	-13.59 to 1.788		-15.67 to -0.132		-15.96 to 0.166		-17.12 to -1.288		-17.08 to -0.892		-18.43 to -2.345		-14.99 to 3.291		-15.90 to 2.546	
Distance to school							**-0.853***	**-0.058***			**-0.818***	**-0.055***				
CI							-1.619 to -0.088				-1.590 to -0.048					
CI SVY							-1.637 to -0.0691				-1.619 to -0.0180					
School location																
CI																
CI SVY																
Free meals at school	**-8.337*****	**-0.074*****											**-8.159***	**-0.068***		
CI	-12.751 to -3.923												-14.523 to -1.794			
CI SVY	-16.24 to -0.436												-16.82 to 0.505			
Type of school public/ private	**6.057****	**0.058****	**7.783*****	**0.074*****			**5.760***	**0.058***	**6.070***	**0.061***	**7.970****	**0.081****	**6.730***	**0.061***	**7.480***	**0.068***
CI	2.157 to 9.958		3.959 to 11.607				0.538 to 10.982		0.806 to 11.333		2.834 to 13.105		0.935 to 12.525		1.780 to 13.180	
CI SVY	-1.342 to 13.46		0.158 to 15.41				-1.895 to 13.42		-2.294 to 14.43		0.0475 to 15.89		-2.159 to 15.62		-1.492 to 16.45	
10 School Environment																

Table 46: Table of Regressions for All Samples

	Regression 1		Regression 2		Regression 1		Regression 2		Regression 3		Regression 4		Regression 1		Regression 3	
Proportion of female teachers																
CI																
CI SVY																
Safe learning environment			**4.567***	**0.039***	**9.120****	**0.084****	**8.989****	**0.082****	**8.953****	**0.083****	**8.368****	**0.077****				
CI					3.273 to 14.967		3.071 to 14.907		3.118 to 14.787		2.483 to 14.252					
CI SVY					-0.876 to 19.12		-1.897 to 19.87		-1.044 to 18.95		-2.454 to 19.19					
11 Community Environment																
Problems with community	**-7.799*****	**-0.100*****	**-7.859*****	**-0.101*****	**-8.278*****	**-0.114*****	**-9.507*****	**-0.131*****	**-7.801*****	**-0.108*****	**-8.869*****	**-0.122*****	**-5.890****	**-0.072****	**-6.041****	**-0.074****
CI	-10.599 to -4.999		-10.644 to -5.075		-12.047 to -4.509		-13.266 to -5.748		-11.540 to -4.063		-12.585 to -5.154		-9.976 to -1.804		-10.142 to -1.940	
CI SVY	-14.20 to -1.400		-14.29 to -1.432		-14.63 to -1.927		-16.58 to -2.437		-13.95 to -1.650		-15.72 to -2.020		-13.02 to 1.244		-13.32 to 1.234	
Community contributes building, textbooks	**-3.473****	**-0.060****			**-4.302****	**-0.079****	**-3.072***	**-0.057***								
CI	-5.722 to -1.224				-7.285 to -1.319		-5.975 to -0.169									
CI SVY	-8.208 to 1.261				-9.121 to 0.517		-7.836 to 1.693									

Table 46: Table of Regressions for All Samples

	Regression 1		Regression 2		Regression 1		Regression 2		Regression 3		Regression 4		Regression 1		Regression 3	
Community pays bonus to school staff	**17.93*** **	**0.108*** **			**13.44** **	**0.086** **	**16.78*** **	**0.107*** **					**17.45*** **	**0.102*** **		
CI	11.439 to 24.413				5.437 to 21.434		8.908 to 24.645						7.872 to 27.025			
CI SVY	8.052 to 27.80				2.889 to 23.98		6.529 to 27.02						6.245 to 28.65			
Constant	406.7***		428.9***		396.1***		453.1***		393.0***		451.3***	0.185	402.5***	0.194	401.8***	0.186
CI	380.765 to 432.688		398.810 to 458.921		342.444 to 449.724		412.239 to 493.937		339.742 to 446.355		410.915 to 491.779		373.302 to 431.757		372.442 to 431.197	
CI SVY	355.9 to 457.6		374.3 to 483.4		297.1 to 495.0		392.7 to 513.4		293.5 to 492.6		392.1 to 510.6		349.2 to 455.9		347.9 to 455.8	

Notes: * p<0.05, ** p<0.01 *** p<0.001

46