CLASS ATTENDANCE AND STUDENT PERFORMANCE: A CASE STUDY OF KAFUE SECONDARY SCHOOLS

By

SYDNEY CHISHIMBA (BA)

A dissertation submitted to the University of Zambia in partial fulfilment of the requirements of the Degree of Master of Arts in Economics

THE UNIVERSITY OF ZAMBIA

LUSAKA

2016

DECLARATION

I, **Sydney Chishimba,** declare that this dissertation**:**

(a) Represents my own work;

(b) Has not previously been submitted for a degree at this or any other University; and

(c) Does not incorporate any published work or material from another dissertation.

Signed …….

Date…….

APPROVAL

This dissertation of Sydney Chishimba has been approved as a partial fulfilment of the requirements for the award of the degree of Master of Arts in Economics by the University of Zambia.

Signed: **Date:**

Examiner: …… ……

Examiner: …… ……

Examiner: …… ……

Supervisor: …… ……

ABSTRACT

Student class attendance is not only important to the performance of the education sector in Zambia, but also to the country's economic productivity. Despite the removal of examination fees at junior secondary level, student performance in junior secondary examinations is low in Zambia. The purpose of this study is to determine the extent to which class attendance of students in public day secondary schools is linked to their performance.

The study utilised 2014 Grade 9 final examinations data on composite scores from ECZ and data from school administrative records to analyse the effect of class attendance on student performance in public day secondary schools of Kafue district. The data were also used to assess students' performance by geographical location and gender. The study applied Instrumental Variables (IV) regression strategy to assess the effect of attendance on performance, while t-tests were used to analyse the difference in student performance by school location and gender.

Study findings show that attending an extra day of class increases student average composite score by 0.670. The results also show significant differences in student performance with students in urban schools having higher mean scores than those from the rural schools at 1 percent level of significance. Similarly, males performed better than females with a significant difference at 1 percent level.

In view of these findings, Government and communities need to work together to build more secondary schools to reduce inter school distances. It is vital for policy makers to consider providing all schools with enough resource allocations for educational materials and libraries. The findings also call for the need to have Government and school authorities sensitise communities to support the education of girls and invest more resources to make schools more "girl friendly".

DEDICATION

This work is dedicated to my nephew Alex and my nieces Joyce and Precious. I salute them for the emotional support rendered to me, particularly during the period of my study. They gave me strength when I needed it most.

ACKNOWLEDGEMENTS

I am sincerely grateful to the Almighty God for the favour granted to me to embark on this overwhelming task of writing this dissertation. I wish to acknowledge with gratefulness the contributions of the following individuals and friends:

Special thanks go to my supervisor Dr. M.C. Chama-Chiliba who patiently read through my work and provided guidance, support, constructive criticism, technical and editorial expertise which shaped this study. Her commitment and devotion to duty stands unchallenged. She has left a mark in my life I will live to remember.

The Statistician at Ministry of General Education headquarters, Mr. Tresfore Chanda, for granting me permission to access the statistical data base. My supervisor at the place of work, Mr. Thomson Lungu, for granting me leave during the period of my study.

Special thanks go to Mr. Willy Chigoma, for his material support and his being there for me whenever I called upon him. Working with him brought profound experiences that we shared.

I sincerely thank all Head teachers, Deputy Head teachers, school Heads of Department and teachers of secondary schools for providing data needed for the study.

All my course mates especially Mr. Lazarus Nyirenda, Mr. Chilimboyi and Miss. Alliness Siwila for their support.

Finally, my deepest gratitude goes to my brother Mr. Alfred T. Kasapo for his support and encouragement throughout my study, my best friend and nephew Alex, my nieces Joyce and Precious who gave me strength to go on with my work. To them I owe this dissertation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND ACCRONYMS

ASCQ - Annual School Census Questionnaire

CSO - Central Statistical Office

DEBS - District Education Board Secretary

ECE - Early Childhood Education

ECZ - Examinations Council of Zambia

ED*ASSIST. - Education Automated Statistical Software Information System Toolkit

EFA - Education for All

GA - Grant Aided

GPI - Gender Parity Index

GRZ - Government of the Republic of Zambia

IODP - Infrastructure Operational Development Plan

MDG - Millennium Development Goal

MOE - Ministry of Education

MOGE - Ministry of General Education

MOHE - Ministry of Higher Education

NGO - Non-Governmental Organisation

OVC - Orphans and Vulnerable Children

PEO - Provincial Education Officer

RSNDP - Revised Sixth National Development Plan

TTL - Time to Learn

UNZA - University of Zambia

CHAPTER ONE

INTRODUCTION

1.0 Study Background

In this era of increasing scientific and technical knowledge and movement towards globalization, education of a country's citizens is vital. Besides it being a basic right of every child, education is widely considered to be a means of poverty reduction, improving health outcomes, quality of life and minimizing gender and social disparities. It plays an important role in the development and accumulation of human capital stock that is essential for a nation's social and economic development. "It is estimated that an additional year of education per person, results in 6 percent average increase in per capita GDP" (UNECOSOC, 2011: 4). As such, academic performance of students is a key variable of interest to all institutions and policy makers mandated with the education function as it serves as an indicator of both institutional and policy success. The implication of this is that positive or negative student outcomes are associated with the way students get to the learning institutions and the corresponding frequencies of student class contact.

In facilitating achievement of positive student outcomes, a successful national policy on education ensures that student and teacher interaction is maximized. It is generally believed that maximizing teacher and student interaction in the transfer of knowledge is important, especially in developing countries where e-learning technology is still low. In developing countries, productivity of school resources is also viewed to be low in relation to performance of students, thus creating a new demand for national governments to

formulate deliberate policies designed to increase student class attendance rates (Farooq, Chaudhry, Shafiq and Berhanu, 2011).

Zambia and many other developing nations have recognized the importance of student access to institutions of learning by passing education policy pronouncements that advocate for free education. This has been done as an attempt to meet the demands of the second Millennium Development Goal (MDG) and the goals on Education for All (EFA) that a number of countries adopted in 2000 (IBRD and World Bank, 2014). According to the second MDG, every child is entitled to universal access to free education (MOF, 2008). In the Zambian education system not only is education at primary level from Grade 1-7 free, but also public examinations at primary and junior secondary levels are free (ECZ, 2013). In Sub-Sahara Africa access to education had been increasing as many countries were on their path to attaining universal primary education and gender parity by 2015.

In addition to making education free at primary school level in 2002, the Zambian Government in 2012 made further policy pronouncements to extend free access to schooling up to Grade 9. This was done in line with the current national development plan; the Revised Sixth National Development Plan (RSNDP). The main focus of the RSNDP on education is to develop rural infrastructure that should facilitate equitable access to free education to help alleviate rural poverty and build human capacity for sustainable national development (MOF, 2014).

Despite the above national efforts, class attendance patterns in Zambia have been a matter of policy concern. According to the 2006-2010 Living Conditions Monitoring Survey

(CSO, 2012), there have been variations in class attendance patterns at both primary and secondary school levels between male and female students and between rural and urban schools.

Class attendance rates for primary school female students increased by 4.8 percentage points from 79 percent in 2006 to 83.8 percent in 2010 while for males it increased from 77 percent to 81.7 percent during the same period. At this level, more females than males attended classes. The attendance patterns were reversed at secondary school level, where attendance rates for females between 2006 and 2010 increased by 4.3 percentage points from 69 percent to 73.2 percent. Male attendance rate for the same period rose from 79 percent to 81.5 percent. Comparison of rural and urban student attendance rates between 2006 and 2010 also shows variations. At secondary school level, class attendance rates in rural schools increased only by 2.2 percentage points from 73 percent in 2006 to 75.2 percent in 2010. In urban schools it increased from 77 percent to 80.8 percent during the same period (CSO, 2012; MOE, 2006).

A number of studies have assessed the role of human capital acquisition in an economy. Relatively few of these studies have examined the mechanism of human capital acquisition. This has raised important questions on how people gain knowledge and what the relationship is between the learning environment and educational achievement of those receiving education (Dolton, Marcenaro and Navarro, 2001).

From the economic point of view, education can be considered as a production process in which a number of school inputs are used to generate output for current and next period utility. From the side of an educational institution, the manner in which school resources

are utilized to transform students into well-qualified individuals is cardinal. This leads to yet other questions regarding whether more taxpayers' money should be used to finance public schooling and whether or not the resulting gains from additional public expenditure are worthwhile (Dolton et al., 2001).

Most of the literature that estimates how students achieve their examination scores and grades pay little attention to taking into account the means by which students access schools, how much time they spend learning each school day, how their access to such schools is affected by school geographical location and exactly how frequency of class attendance translates into examination performance.

Regardless of the availability of existing studies of education production, the evidence would suggest that we are still far from comprehending how education is produced in terms of how class attendance is transformed into knowledge and subsequently performance in examinations. This therefore, serves as a justification to have a new empirical study which attempts to provide further insight into the process by which student class attendance is related to academic performance.

1.1 Statement of the Problem

Since the 2012 free education policy pronouncement, primary school enrolments have been increased. This increase was driven by Government expansion of more primary school infrastructure in the country that began in 2008. However, little attention was paid on secondary school infrastructure development and expansion. This led to more students accessing and completing primary school education, thereby exerting excessive pressure and demand for few available secondary school learning spaces.

Kafue district was not an exception. Given the spatial nature of school distribution, the district has only 9 GRZ secondary schools against 34 GRZ primary schools. The majority of the schools are rural schools (ED*ASSIST, 2014). To accommodate a large number of students from primary schools, most GRZ day secondary schools attract students from wide catchment areas where some students cover long distances between home and school. This results in occasional loss of learning time and class absences.

The problem with this situation has been that student performance at junior secondary in the district has been declining. District analysis of Grade 9 performance indicates that on average, district pass rates declined from 34 percent in 2009 to 28 percent in 2010 and from 30 percent in 2011 to 26 percent in 2012. Performance in 2013 and 2014 showed a slight improvement from 27 percent to 38 percent (MOGE, 2015). While much research has been done to explain the performance- attendance linkage, we have little comprehension of how precisely class attendance influences student performance in both rural and urban public day secondary schools of Kafue district. A clear understanding of this relationship will provide relevant insight for new policy considerations.

1.2 General Objective

The general objective of this study was to assess the effect that class attendance has on the performance of students in junior secondary schools in Kafue district as measured by Grade 9 composite scores.

1.3 Specific Objectives

The specific objectives of the study were to:

i. Determine the magnitude of class attendance effect on student performance in public day secondary schools

ii. Ascertain if there is a significant difference in student performance between students in rural and students in urban day secondary schools

iii. Ascertain if there is a significant difference in student performance between male and female students

1.4 Statement of Hypotheses

i. Class attendance does not influence student performance

ii. There is no significant difference in performance between students in rural and urban day secondary schools

iii. There is no significant difference in performance between male and female students in day secondary schools

1.5 Justification and Significance of the Study

The rationale for this study arises from the limited empirical evidence on how class attendance influences student performance in day secondary schools of Zambia and Kafue district in particular. Empirical evidence from other countries indicates a mixture of findings. Some studies report a positive association between student class attendance and academic performance while others do not support this relationship. In most countries worldwide, student attendance has been taken to be so important that attendance policies have been passed and implemented through student class attendance registers. This therefore, demands that we understand the benefit that such attendance policies have on student academic welfare.

In an effort to increase student access to learning facilities country wide, Zambia has implemented the second goal of the MDG 2015 and the 2000 EFA convention agreements (EFA, 2008). This has meant the removal of school fees at primary education level and examination fees at both primary and junior secondary school levels (ECZ, 2013). The purpose was to facilitate equal access to schooling for both children in rural and urban communities (MOF, 2014). The result has been a boost in attendance rates at primary school level with a large number of primary school students exiting the level in demand for secondary education. Investment in infrastructure development at secondary school level has been lower than in primary, resulting in some secondary schools over enrolling students from wider catchment areas. The study is therefore, meant to provide an assessment of how well the attendance patterns fit with the policy.

This study is also intended to benefit other stakeholders such as Non-Governmental Organisations (NGOs) who have committed financial and material resources in districts like Kafue for the purpose of increasing student attendance rates. Kafue district has since 2013, been receiving financial support for Orphans and Vulnerable Children (OVC) in 9 GRZ secondary schools from Time to Learn (TTL) for payment of their school fees. In addition, World Bicycle Relief has donated relief bicycles to students in GRZ day secondary schools that walk long distances to schools (World Bicycle Relief, 2011).

The study findings are, thus, critical to both Ministry of General Education and these NGOs in having an informed evaluation of how well Government and donor resources translate into student performance in the final examinations.

1.6 Study Limitations

The results of this study need to be interpreted with caution because the analysis of data was restricted only to variables that were available from school records which included individual student class attendance. Other important characteristics such as measures of student ability, motivation and effort were not included. A measure of family socio-economic status, proxied by income of parent or guardian, was only captured as a dummy, but a better measure would be the actual income figures of parents or guardians to obtain more precise estimates; school records only indicated parents as either working or not working. Another limitation was the reported distances from schools. Since the student distances reported in school records were approximated, it implied that such measurements need to be interpreted in light of this limitation. The other limitation in this study was non-inclusion of a measure of teacher quality. It was difficult to determine the standard against which teacher quality would be measured. However, it would have been better to include this aspect in the analysis since teacher quality is important in student performance.

1.7 Organization of the Remainder of the Study

Following presentation of chapter one, the remainder of the study is organized as follows: Chapter two presents a description of Zambia's education sector and education policy developments. Review of related literature and empirical research related to the assessment of the effect of class attendance on student performance is presented in the third chapter. The methodology and procedures used to gather data for the study are presented in chapter four. The results of analyses and findings emerging from the study

are in chapter five. Chapter six contains a discussion of findings while conclusions drawn from the findings and recommendations are in chapter seven.

CHAPTER TWO

A DESCRIPTION OF THE EDUCATION SECTOR IN ZAMBIA AND EDUCATION POLICY DEVELOPMENTS

2.0 A Description of the Education Sector in Zambia

The current formal education in Zambia is based on a four-tier system. It consists of academic learning at Early Childhood Education (ECE), primary education from Grades1-7, secondary education from Grades 8-12 and tertiary education.

The current administrative structure of the education sector in the country provides for two ministries. There is a Ministry of General Education (MOGE) and Ministry of Higher Education (MOHE). The Ministry of Higher Education is responsible for management of tertiary education institutions in Zambia. The Ministry of General Education has been mandated to manage the affairs of other non-tertiary educational institutions.

The Ministry of General Education headquarters is in charge of the overall management and direction of the formal education system from ECE to secondary school level in the country. The Provincial Education Officers (PEOs) are responsible for coordinating smooth delivery of education services in provinces and supervise District Education Board Secretaries (DEBS) who actually run ECE, primary and secondary schools in their respective districts.

2.1 Education Delivery under Ministry of General Education

The basic level of education under this ministry is ECE. Zambia recognized the importance of this educational level as cardinal in the early development of a child. It is one of the oldest goals that Zambia and other African countries agreed to fulfill at the

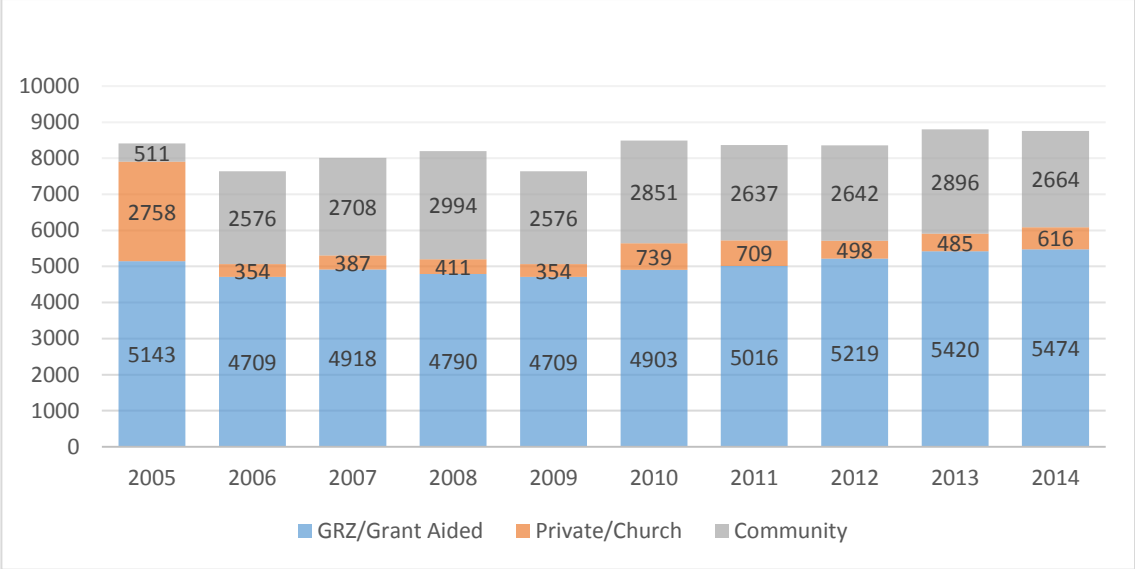
Education for All (EFA) conference held in Dakar, Senegal in 2000 (EFA,2008). In spite of rapid increase in student enrolments at primary school level, expansion of ECE facilities and enrolments have been slow. Statistics indicate that between 2010 and 2014, the percentage of first grade entrants with pre-grade experience fluctuated between 17.3 percent in 2010, 15.1 percent in 2011, 18.7 percent in 2012 and 15.4 percent in 2014 (MOGE, 2014).

2.1.1 Primary Education Delivery

Primary education in Zambia is principally provided by GRZ institutions of learning. Nonetheless, there are two other providers of education services at this level; the private and community schools. Student enrolments at primary level have been on an increase in GRZ primary schools countrywide as a result of Government investment in primary school infrastructure constructions. Between 2005 and 2014, learning spaces at primary education level exhibited a 1.5 percent average annual growth rate. At national level, the number of primary schools in 2005 was 8,412 classified as GRZ combined with Grant Aided (GA), private and community. These numbers increased by 4.1 percent in 2014 to 8,754. Of significance is the finding that in 2014 community schools in the country still constituted about 30.4 percent of the primary schools (MOGE, 2014). The number of community primary schools showed a decline between 2013 and 2014. This was because some community schools either acquired a GRZ status or stopped operating. For Kafue district, the number of GRZ primary schools in 2005 was 53. The private and community schools were 27 and 62 respectively (ED*ASSIT, 2010). However, these numbers decreased in 2011. From 2011 to date, the district has 34 GRZ primary schools, 9 private schools and 18 community schools (MOGE, 2016). The decline in the number of schools

in Kafue was a result of the creation of Chilanga district out of Kafue district. Figure 1 below depicts school distribution in Zambia by running agency.

Figure 1: Number of primary schools by agency from 2005-2014



Source: MOGE 2014 Statistical Bulletin

2.1.2 Distribution of Primary Schools by Grades Offered and Location

In 2014 MOGE had 551 schools that offered Grades 1-4 only, 3,864 schools that offered Grades 1-7 only, while 2,821 schools offered Grades 1 to 9. A large number of primary schools (81.8 percent) in 2014 were found in rural areas and only a small percentage (18.2) was in urban areas. In Kafue district 71 percent of GRZ primary schools (24 schools) are found in rural areas and only 29 percent (10 GRZ primary schools) are in urban. However, it is in the rural communities where much of the infrastructure is poor and a large proportion of students have to walk long distances daily to access education. In terms of quality of staff accommodation in these rural areas, it was estimated that over 50 percent of teachers' housing units were built out of sub-standard materials (MOE, 2005; MOGE,

2014). This inadequacy and poor quality of staff accommodation has compromised the quality of education service provision especially in rural communities as it acts as a demoralizing factor for the teachers serving in rural areas. To show the seriousness of teachers' accommodation problem, MOGE statistical bulletin reports that there were 78,395 primary school teachers against 16,676 permanent housing units in 2014. This implies that only about 21 percent of primary school teachers nationwide had decent shelter (MOGE, 2014).

2.1.3 Primary School Enrolment

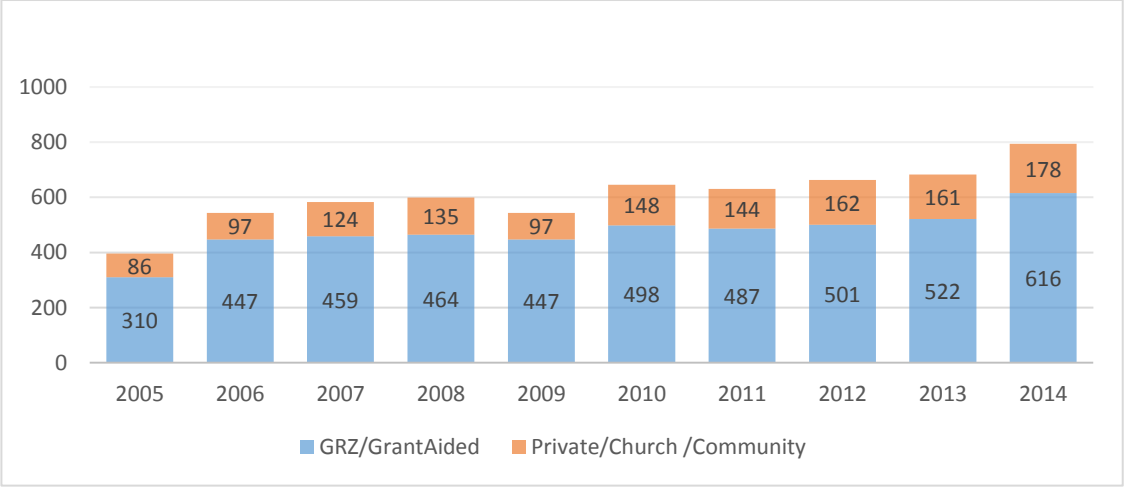
Total primary school enrolment in 2014 was 3,217,872 segregated as 1,613,748 males and 1,604,124 females compared to 2,821,018 (1,407,483 males and 1,413, 535 females) in 2010. This represents an increase in enrolment of 14.1 percent (MOGE, 2014). The increase was to a greater extent influenced by the introduction of free primary education (UNECOSOC, 2011).

2.1.4 Secondary Education Delivery

The transition from high school system to secondary system in 2011 brought with it a number of areas of immediate concern. One of the major issues at this level of education is the absorption capacity of secondary schools. The rise in the population of school-going children and expansion of primary school infrastructure, have jointly created excess demand for secondary education in the country. This has created a need for Government to shift emphasis from primary school expansion to creation of more learning spaces at secondary education level.

In 2014, the total number of schools nationwide that offered secondary education was 794 compared to 683 in the previous year. Out of these schools, 616 were run by a combination of GRZ and GA and 178 by a combination of private, church and community. In Kafue district there were only 4 GRZ secondary schools in 2005. In 2008 the number increased to 5 secondary schools. The current number of the GRZ schools in the district is now 9 (MOE, 2005; MOE, 2008; MOGE, 2016). Figure 2 depicts the 2014 national distribution of secondary schools by running agency between 2005 and 2014.

Figure 2: Number of secondary schools by agency from 2005-2014



Source: MOGE 2014 Statistical Bulletin

2.1.5 Distribution of Secondary Schools by Grades Offered

Countrywide, the number of secondary schools that offered Grades 8-12 in 2014 was 403. For the tenth to twelfth grade range, there were 25 secondary schools. The schools that offered Grades 1-12 were 66 countrywide. These schools that offered classes from Grade 1-12 are largely private schools (MOGE, 2014).

There are a number of issues that need to be resolved at this level of education also. One of the major concerns has been the problem of inadequate availability of teachers with appropriate professional qualifications. This situation is worrisome because teacher qualification is positively correlated with performance of students. In 2014 out of the 20,710 teachers assigned to teach in secondary schools, surprisingly 2,465 teachers had only a primary teacher's certificate and 13,300 had either a primary diploma or a secondary diploma. There were only 23.88 percent of teachers (4,945) with degree qualifications. Most of the teachers assigned to handle senior grades had diploma qualifications. The situation was even worse in the area of mathematics and science. In 2011 the country had approximately 885 teachers with degrees in social science out of the required 6,508 (MOF, 2014; MOGE, 2014). The demand for teachers of Science and Mathematics at this education level is so high that serious Government intervention is needed. To address this situation, MOGE embarked on the programme of teachers' skills development in these two subjects by facilitating the training of teachers of science and mathematics through the "fast-track" training initiative offered by the University of Zambia (UNZA).

Government has also made more effort to increase the supply of secondary school teachers to reduce the student-teacher ratio. Over the years the ratio of secondary school students to teachers has declined favourably from 56.84 in 2005 to 35.75 in 2014 (MOGE, 2014). Between 2005 and 2014, student-teacher ratio had been fluctuating. It was 56.84 in 2005 and declined to 37.04 in 2006. In 2007 it slightly increased to 38.67. For the years 2008 and 2009, the ratio stood at 40.46 and 40.81 respectively. It dropped to 33.83 in 2010, but

increased substantially to 48.16 in 2011. Nevertheless, it declined to 39.91 in 2012 and 38 in 2013. It declined further to 35.75 in 2014.

Regardless of such efforts by Government, national transition rates from Grade 9 to Grade 10 have still been lower than that from Grade 7 to Grade8. Student transition rate was 59.3 percent in 2014 at Grade 7 compared to 42.3 percent at Grade 9. In fact the transition rate from Grade 9 to Grade10 dropped from 46.7 percent in 2012 to 42.3 percent in 2014 (MOGE, 2014). In Kafue district, the student transition rate in 2014 at Grade 9 was only 38 percent (MOGE, 2015).

One of the major factors responsible for stagnation at secondary level has been low budgetary allocation to the development of school infrastructure over the years. Unbalanced focus was directed at the development of infrastructure at primary school sub-sector at the expense of secondary school sub-sector thus, consistently driving down the transition rates (MOF, 2014). For instance, the Infrastructure Operational Development Plan (IODP) of the Ministry of Education, indicates that Government constructed 247 new primary schools against 45 new secondary schools between 2008 and 2010. During the same period, 3 new primary schools were constructed in Kafue district, but no secondary school. This means excess demand for secondary school places was created (MOE, 2010).

Gender parity has been also another area of concern at secondary school level. Comparison of national gender parity indices in 2014 between primary and secondary levels of education indicated differences. While Gender Parity Index (GPI) at primary level for Grades 1-4 was favorable (1.01), that at secondary level showed that more female students had dropped out than males, giving the GPI of 0.83. For the years 2011, 2012 and 2014, the GPI varied between 0.82, 0.83 and 0.85 respectively. The continuing high

number of female students dropping out at secondary level is partly explained by pregnancies from 2,100 recorded in 2012 to 3,103 in 2014. The pregnancy cases for Kafue in 2012 and 2014 were 53 and 71 respectively (MOF, 2014; MOGE, 2014; MOGE, 2015)

2.1.6 Secondary School Enrolment

In 2014 secondary school enrolment for male students was 421,591 and 380,003 females, giving a total of 801,594. This was a 40.85 percent increase from 569,099 in 2010. These figures however, do not correspond to the available leaning spaces in secondary schools. This condition has led to some secondary schools to overenroll students. Not only is secondary school infrastructure inadequate in the country, it is also inadequate in districts. Kafue district is one such district where the ratio of GRZ primary to secondary schools is 34:9 (ED*ASSIST, 2014).

2.1.7 Situation Analysis of Secondary Schools in Kafue District

The district has both GRZ and private secondary schools. There are nine GRZ secondary schools of which three are boarding and six are day schools. There are two private secondary schools of which one is a boarding and the other one is a day school. Among the GRZ secondary schools, four are in rural areas while the other five are in urban. The distribution of the two private schools is that one is in a rural area while the other one is in town (MOGE, 2016).

Two of the rural GRZ schools are surrounded with village clusters and commercial farms and the other two are bordered with villages, commercial farms and tour operator lodges on Kafue and Zambezi rivers. The four GRZ secondary schools in the rural areas service sixteen primary schools in their catchment areas, while the five GRZ schools in the urban

areas service eighteen primary schools. This spatial distribution of secondary schools imply accessibility challenges faced by students.

The wider catchment areas of these secondary schools entail that some students walk long distances to these schools. The average distance to the nearest day secondary school covered by students is 2km with the longest being approximately10km (MOGE, 2015). The implication of such long distances is that student daily class attendance is affected. Most affected are students who live furthest from school who may not attend classes consistently due partly to walking fatigue. Even when such students strive to attend classes daily, concentration in class is likely to be affected due to fatigue. To mitigate the negative impact of distances on daily attendance of students, some schools in Kafue have been given relief bicycles by World Bicycle Relief (World Bicycle Relief, 2011) to enable students who walk long distances attend classes regularly. However, this relief has been limited to students in rural areas on the premise that children in such areas come from families with low incomes and live in places with poor transportation systems compared to their urban counterparts.

Despite the above intervention, secondary education in the district has been affected by low performance rates. For instance, statistics of student performance at junior secondary school obtained from Kafue District Education Office indicate that student achievements have been declining over the years (MOGE, 2015). Table 1 shows the declining trend in the Grade 9 examination composite scores for Kafue district with a slight improvement in 2014.

Table 1: 2009-2014 Grade 9 pass rates

Year	No. Registered for Exams	No. Sat	No. Passed	Pass Percentage
2009	3013	2627	894	34%
2010	2256	2144	600	28%
2011	2932	2311	693	30%
2012	3508	3095	805	26%
2013	4196	3920	1058	27%
2014	4892	4405	1668	38%

Source: 2015-2018 Kafue District Strategic Plan.

Apart from school distances there are other factors that might impinge on class attendance of pupils in Kafue schools. MOGE (2016) identifies these factors to include social, cultural and economic. Socio-cultural practices in rural areas such as cattle herding, subsistence farming, hunting and fishing might significantly reduce student time allocation for school. Girls in both rural and urban communities are more involved than boys with domestic responsibilities such as cooking, cleaning, fetching water and general family care. These activities also leave girls with little time for school and self-study and might also affect their concentration in class. The economic factor also might affect student class attendance in both rural and urban schools. To supplement family income, older school children in rural areas most likely work in nearby commercial farms and lodges. Such part-time jobs affect their school performance. In urban schools some families in non-formal employment send their children to sell at the markets. This also could contribute to reduction in student annual attendance days.

School infrastructure and modern learning equipment are a feature that differentiates the status of rural schools with that of urban schools in the district. While all four secondary schools in rural areas have permanent classrooms, two of these have no electricity, and have no waterborne toilets. Instead, they have pit latrines. In addition, only one out of four rural secondary schools has laboratories for science subjects (MOGE, 2016). None of these rural secondary schools has modern teaching and learning facilities such as computers, projectors and internet connectivity necessary for research by teachers. In terms of school library, only one rural school has a separate library stocked with books. This is in contrast with secondary schools in Kafue urban where all schools are electrified and enable students extend hours of self-study. In addition, urban schools have ablution blocks with running water to provide convenience to students, especially females. They are all equipped with science and computer laboratories and are connected to the internet. Each school has a well-stocked library (MOGE, 2016).

2.2 Educational Policy Developments and Reforms in Zambia

Provision of education in Zambia has been guided by a number of educational policies. Beginning with the educational reforms of 1977, the country has recognized the importance of providing opportunities to all individuals to receive an education based on their interests, abilities and needs in order to acquire intellectual and skills development (Education Reforms, 1977). This goal was revisited in 1992 by concentrating on student learning (Focus on Learning, 1992).

In 1996 the Government developed a national education policy which focused on restructuring the then existing organizational structure of the Ministry of Education (MOE) to make it "more responsive to the changing needs and requirements of society"

(Educating our Future, 1996: 131). This involved the change of the curricular in 2006 that would be responsive in making the country attain its long term vision. The Zambia Vision 2030 stipulates how the education sector has to contribute to the achievement of the vision (MOF, 2006).

Further policies and reforms were pronounced by Government in 2012. Key among them was free education policy which involved the abolition of school fees at primary level and scraping of examination fees at both primary and junior secondary levels. This was done in fulfilment of the 2000 EFA convention agreement. The other policy reform has been a further change in the school curricular at both primary and secondary levels by introducing "two-career path ways" where students now learn both academic and practical skills which, in line with the Revised Sixth National Development Plan (MOF,2014), are intended to prepare them to be future entrepreneurs.

CHAPTER THREE

LITERATURE REVIEW

There exists vast literature world over that contains empirical evidence on the existence of a relationship between class attendance and student performance. There are also studies that have shown how school geographical location and student gender are associated with student performance. Conclusions of these empirical studies have, to a greater extent, depended on the type of research methodologies employed in the investigation of these relationships. To effectively compare study findings, this section starts by examining studies that explain the relationship between student academic performance and factors other than class attendance. This is then followed by the review of literature specific to the relationship between class attendance and student academic performance.

3.0 Factors Influencing Student Performance other than Class Attendance

Numerous studies have been conducted in both developed and developing countries on student academic performance. Nizamettin and Bekir (2015) used correlation analysis method to assess the impact of student teacher ratio on student performance and performance ranking of cities in Turkey. The authors utilised secondary data from Ministry of National Education of Turkey and Student Selection and Placement Centre. Data were obtained for 81 cities. The study finding was that student performance in promotion to higher education examinations was negatively correlated with student-teacher ratio. Another finding of the study was that achievement ranking of cities was negatively correlated with student-teacher ratio, implying that a city with the lowest average number of students per teacher ranked higher in performance in examinations.

The finding on the effect of student-teacher ratio by Nizamettin and Bekir (2015) is similar to Idowu and Oluwole (2014) who also reported a significant correlation between student-teacher ratio and student academic achievement in mathematics. The study also applied a Pearson product moment correlation to analyse data for 120 students randomly selected from 3 senior secondary schools in Port Harcourt Local Government area of River State, Nigeria. Nizamettin and Bekir's (2015) study however, deviated from that by Idowu and Oluwole (2014) in three ways. Firstly, the Turkish study was based on the analysis of standardised final examinations whereas in the Nigerian study, the researchers prepared and administered the test items. Secondly, more subjects in the standardised examination in the Turkish study were considered in the analysis in contrast with one subject in a non-standardised test in the study in Nigeria. Thirdly, the study in Turkey on the effect of student-teacher ratio on student performance was implicit. The explicit focus was on city ranking in performance based on this ratio. In contrast, the study in Nigeria focused on individual student performance in relation to student teacher ratio. While the independent variable was student-teacher ratio in both studies, the study in Turkey considered the average student-teacher ratio in each of the 81 cities as the predictor of city ranking based on student performance. On the other hand, the individual student performance in the Nigerian study was associated with student-teacher ratio in classes from three senior secondary schools.

The above findings on the significance of student-teacher ratio are contrary to Ramous, Duque and Nieto (2012) who report a non-statistically significant influence of student-teacher ratio and class size on student performance in Columbia. The authors, instead find

mother's education to have a significant influence on the academic performance of their children.

School, family and socio-economic factors can equally be essential in explaining academic achievement of students as reported by Mbugua, Kibet, Muthaa and Nkonke (2012) who examined factors contributing to poor performance of secondary school students in school certificate examinations and proposed strategies that students could adopt to improve performance in mathematics in Baringo County in Kenya. The study employed a descriptive survey research design based on a sample of 1,876 respondents of whom 1,718 were form 3 students, 132 teachers of mathematics and 26 Head teachers. The study results show that low staffing levels and insufficient teaching and learning resources were responsible for students performing poorly. Other contributing factors were level of parents' education, income of parent or guardian and cultural factors. Parents' education and income are also found to be important determinants of student performance in studies done in Pakistan by Farooq et al. (2011) and Melack (2014). Melack (2014) equally finds inadequate teaching and learning materials and low staffing levels in Sumbawanga in Tanzania, similar to Mbugua et al. (2012), to negatively affect the performance of students. Other findings by Melack (2014) were that long walking distances by students and absence of library and laboratory facilities in schools impacted negatively on the performance of students. The author recommends to education stake holders to provide learning materials and to Sumbawanga District Council to construct school laboratories and student dormitories. An earlier study in Mbeya Municipality, in Tanzania by Mlozi, Kaguo and Nyamba (2013) that assessed factors influencing academic performance of students in Community and Government built secondary schools,

confirmed findings on the positive effect of teaching and learning materials on student performance.

Results from the above studies require careful handling when drawing inferences. Although these results may be a reflection of environments in which the studies were conducted, findings may not be generalised to larger populations as the respective samples on which they were based were not large enough to warrant such inferences. On the other hand, conclusions from Melack (2014) and Idowu and Oluwole (2014) were based on respondents' subjective views which may not hold similar conclusions if replicated in other settings. As for Mbugua et al. (2012) despite the sample size being fairly large (1718), the author did not indicate at what level of statistical significance the reported factors responsible for poor performance were. This suggests the need to exercise caution in the interpretation and application of these study findings.

Other studies have reported qualification of teachers and experience in the teaching of specific subjects as important factors contributing to student performance. Among these Idowu and Oluwole (2014) find a significant positive relationship between teacher qualification and student performance in mathematics. However, these findings are contradicted by Musau and Abere (2015) who studied a sample of 600 students who sat for the Kenya Certificate of Secondary Education, 8 Head teachers and 40 teachers of science, mathematics and technology (SMT) and found no significant relationship between teacher qualification and student performance in SMT subjects in Kitui County in Kenya.

3.1 Specific Studies on Student Performance in Relation to Class Attendance

The studies dealt with in the previous section, considered the number of factors that explained performance but were not specific to class attendance. Nonetheless having studies directed at a specific factor to explain student performance is preferable for well-informed policy.

In a study of the effect of school days and absences on test score performance, Aucejo and Romano (2014) used longitudinal administrative data from Carolina public schools to study a sample of 382,335 students and 29,202 teachers in 1,305 schools. The robust findings from the Instrumental Variables regression analysis show that extending a school calendar by 10 days increased maths and reading scores only by 0.8 percent. Alternatively, the study finds that reducing absences leads to a 5.8 percent gain in maths scores and 3 percent in reading. The authors suggest that given policy choices of financing between programmes to extend a school calendar and reduce absenteeism, the absenteeism programme would yield more benefits than the other. The study also suggests that observed student heterogeneity in student ability would imply that targeting absenteeism of low performing students would assist in lessening the current gaps in performance.

These findings contradict an earlier finding in Maryland by Caviglia-Harris (2006) who used two-stage least squares (2SLS) regression analysis and found that the number of days a student attended an economics class at Salisbury University had no significant impact on student performance in the course. The author, however, explains that course absences are likely to impact grades, but after controlling for endogeneity and simultaneity bias, the claim of an OLS positive relationship of attendance with student grades is not supported. The study instead, finds students' previous knowledge of a subject to be the most

significant predictor of performance. The two studies were further dissimilar in the targeted population and instrumentation. The study by Aucejo and Romano (2014) though utilising a robust dataset, focused on the primary grades from grades 3 to grade 5 whereas Caviglia-Harris (2006) targeted university students. Flu data was used to instrument student attendance rates in the 2014 study (Aucejo and Romano, 2014), but Caviglia-Harris (2006) instrumented number of times the student withdrew from courses on student course attendance. Extending results of these two studies to the secondary school education level may either support or depart from the class attendance argument. The present study therefore, intends to investigate the influence of class attendance in public secondary schools. Additionally the present study deviates from the two studies by using distance covered by students and student state of orphan hood as instruments for attendance. Nevertheless, both studies were interested in assessing the effectiveness of attendance policies. The attendance policy in the 2014 study was found to be more effective (Aucejo and Romano, 2014) compared to the study by Caviglia-Harris (2006) that did not find policies on attendance beneficial.

The study conducted in Nigeria underscores the findings of (Aucejo and Romano, 2014) over Caviglia-Harris (2006) regarding the influence of class attendance on student performance. Oghuvbu (2010) investigated the correlation between attendance and academic performance of secondary school students in Delta State by using Pearson product moment correlational analysis on a sample of 2,860 students. The researcher compared attendance rates and performance of students in Delta regions of the North, Central and South. The study findings were that student attendance was positively

correlated with performance. Further findings indicate that increasing attendance by 1 percent increased average performance by 0.192.

The findings from Oghuvbu (2010) need careful interpretation. Although the study found a positive correlation between attendance and performance, the findings from the applied OLS regression on the linear association of attendance with performance was not statistically significant. In addition, the linearity of the attendance relationship with performance was threatened by endogeneity problems. Since the study failed to control for the endogenous nature of attendance, there is a possibility that the parameter estimate of attendance was biased, thus affecting the drawing of inferences. Departing from the study by Oghuvbu (2010), we acknowledge the endogeneity problem of attendance in support of Aucejo and Romano (2014) and Caviglia (2006) and control for this endogeneity by performing a 2SLS regression with distance and orphan hood as instruments for attendance.

Other studies have also applied an Instrumental Variables strategy to model student achievement on attendance. Andrietti, D'Addazio and Gómez (2008) sought to establish the relationship between class attendance and academic performance among Spanish university students of economics using a sample of 316. The study used a combination of survey data and student administrative records. To address the problem of endogeneity of lecture attendance, the authors used student distances to the university and working status as instruments for attendance. The results indicate that attending an extra percentage point of economics lectures increased test scores by about 0.13 percent at 1 percent level of statistical significance. The authors found that students who attended more lectures scored significantly better than those with lower attendance rates.

The findings by Andrietti et al. (2008) are supported by Gottfried (2009) who hypothesised that the number of days a student was present in school positively affected the learning outcomes. The conclusions by Gottfried (2009) were based on a comprehensive longitudinal dataset of elementary and middle school students in the Philadelphia school district. The study was robust and used a full sample of 332,924 secondary data observations. The study results showed a consistent positive and statistically significant relationship between student attendance and academic achievement for both elementary and middle school students. The study also reported a negative correlation between distance and attendance. The author justified the use of distance as an instrument by stating that distance was an exogenous measure of attendance that minimized the confounding influence of unobservable student characteristics such as ability and motivation. The results of the two studies diverged on the reported measures of student performance; whereas Andrietti et al. (2008) used student scores in economics assessments, Gottfried (2009) used students Grade Point Averages (GPA). The two studies however, converged on the use of student distance to instrument for attendance, though the 2008 study had used two instruments for attendance compared to Gottfried's (2009) use of single instrument.

3.2 Significance of School Geographical Location in Student Academic Performance

Unequal development between urban and rural areas of most developing countries has attracted considerable amount of research. There is no consensus on the conclusions drawn from study findings on the relationship between school location and student performance. There are studies that posit that students in rural settings perform poorer

than those in urban areas. There are also those that do not find a significant difference in student performance by location.

In Zambia a report by the National Assessment Survey covering a sample of 8,520 Grade 5 students drawn from the 2007 Annual School Census Questionnaires (ASCQ), found rural and urban differences. The report findings were that male and female students in urban schools performed better than their counterparts in rural areas. The survey, further, shows that urban females performed better than rural males in mathematics, english, life skills and Zambian language (MOE, 2008: 18). Scholars like Ramous, Duque and Nieto (2012) attribute student poor performance in rural areas to the low socio-economic conditions that students in rural setups experience. Using micro data from 2006 and 2009 Programme for International Student Assessment (PISA), Ramous et al. (2012) applied a combination of Oaxaca-Blinder and Juhn Murphy-Pierce methods in the analysis. The study reveals that educational mean performance in reading and mathematics was poorer for students in rural areas than those in urban setups. Similar results in support of the rural-urban performance disparity are those by Usaini and Abu Bakar (2015). The study investigated the influence of school environment on secondary school students' academic performance in Kuala Terrengganu in Malaysia. The school environment was determined by presence or absence of modern learning equipment such as computers, internet facility, libraries and availability or lack of competent intelligent teachers. The sample for analysis was 377 secondary school students drawn from a population of 26,569 students. The study, employing linear regression analysis, confirmed a positive relationship between student performance and school environment. This meant that schools with poor learning conditions were worse off in student results.

Similar conclusions are drawn by a study of secondary schools of Ekit State in Nigeria by Olusola and Omotade (2014).These authors equally view students in rural areas as being more disadvantaged than those in urban areas. When assessing the impact of school location on academic achievement of students in science in senior secondary certificate examinations, Olusola and Omotade (2014) found a significant rural-urban difference in academic achievement. The study used a sample of 120 senior secondary students of science subjects in Ekit West Local Government area of Ekit State. The t-tests used to analyse the data indicated a significant mean difference in student scores between urban and rural areas. This finding is in support of the study which was done earlier in the same state of the same country by Olatude (2011) who conducted a similar study. The two studies from Nigeria only differed in the sample sizes and the study participants. Whereas Olusola and Omotade (2014) selected 120 students from 6 secondary schools, the study by Olatude (2011) selected 50 secondary schools, but the study does not show how many students were participants.

The study findings above are dissimilar to those from the study conducted in secondary schools in Ondo State of Nigeria by Bosede and Enuloju (2013) who view students in rural areas as equally advantaged as those in urban areas. Bosede and Enuloju (2013) used a final examination data set for senior students in public secondary schools of Ondo State. The study hypothesised that there was no significant difference in academic performance between students from rural environments and those from urban environments. Results of the t-tests led the authors to accept the null hypothesis. The authors concluded that there was no statistically significant mean difference in student performance between students in rural areas and those in urban areas. The authors justify this result by citing Government

efforts to improve rural areas by providing these areas with similar facilities like television and cinema halls for rural students to get updated.

Findings from the above studies, however, need to be understood only in the context of the methods applied, study environments and sample representativeness of the larger population. The 2008 survey in Zambia (MOE, 2008) for example, only considered performance of primary school students. These findings may not hold for older students in higher grades for which the present study intends to investigate. The conclusions deriving from the findings in the studies by Usaini and Abu Bakar (2015) were based on subjective views of respondents as they provided responses to the self-administered questionnaires. It is likely that results could be different if complete secondary data were utilised. Further, this study does not indicate the geographical location of the schools characterised by the described school environment. The present study separates the school environment into two areas; the rural and urban.

3.3 Significance of Gender in Student Academic Performance

Gender balance and equal participation of women in development programmes has been a serious policy concern not only in developing countries but, world over. Efforts have been directed towards closing the gap that exists between males and females in some countries by formulating and implementing policies that give women an equal share of participation at all levels of decision making (GRZ, 2000). Yet literature draws different conclusions about academic achievement of males and females. In Zambia females in primary grades perform better than males. This situation is however, reversed at secondary grades where males perform better than females in examinations (MOE, 2008).

Some studies conducted in other countries report female students performing better than their counterparts even in secondary schools. A longitudinal study conducted on university students in Jordan establish a significant difference in Grade Point Average (GPA) scores

between male and female undergraduate students (Khwaileh and Zaza, 2011). The study utilised a longitudinal dataset of undergraduate student records from the university database covering a period of 2002 to 2007 academic years. The study used a sample of 26,122 students. Results from one way Analysis of Variance (ANOVA) and t-tests found that females scored higher than males on GPA. This finding was similar to Farooq et al. (2011) who used similar analysis and found a statistically significant difference in marks with secondary school females obtaining a higher mean mark than males.

Two recent studies from Nigeria also found that secondary school females had better scores than males. Abdullahi, Mlozi and Nzalayaimisi (2015) used a cross-sectional study sample of 300 secondary school students from Katsima State to investigate the association of student achievement in agriculture science with student gender, age and school factors. Results from t-tests indicated a statistically significant mean difference between female and male students with females performing better than males. The other finding was that student age was positively related to student performance. Igbudu (2015) also found that in a sample of 412 secondary school students who sat for West African Examinations Council exam in Government subject between 2009 and 2011, females excelled more than males on average. Analysis of Grade 'A', 'B' and 'C' results shows that 51 percent of students with Grade 'A' were females and 49 percent males. For Grade 'B' category there was a reverse with males representing 51 percent. Category 'C' showed 52 percent females and 48 percent males leading the study to conclude that females performed better than males.

An earlier study in New Zealand also confirms these findings. Fergusson and Horwood (1997) conducted a longitudinal investigation of gender differences in educational achievement in a New Zealand birth cohort that was studied from birth to the age of 18years. Data on a sample of 982 students were analysed using a mixture of t-tests, and multiple logistic regression. The study found that at all grade levels male students had lower standardised scores than females in reading, mathematics, written expression and spelling before the age of 18. At the age of 18 the study still found that females performed better than males on a school certificate exam. These findings contrast those from a study by Olusola and Omotade (2014) who found no statistically significant difference in mean

scores between male and female secondary school students in both rural and urban schools.

Drawing inferences from some of these study findings needs caution. Some study conclusions were based on the assessment in one or two subjects while others were based on non-standardised assessments. The sample sizes used in some studies as well as the environment in which they were conducted and the methods applied had a bearing on the nature of results obtained. This means that the findings may not be generalised beyond the study settings.

Summary of the Literature

The literature above demonstrates application of survey research designs and data analyses using t-tests, ANOVA and OLS regressions in modeling student performance on various factors. Very few studies have quantitatively used instrumental variables to account for simultaneity and omitted variable biasness in studies on attendance. There is also no consensus on the role of various factors in measuring student performance on standardised assessments.

Generally, the literature supports the notion that attending class significantly influences student outcomes. Other important school and family characteristics are school resources and family level of education and socio-economic status. A higher economic status increases the probability of a child to attend and excel in school. There are however, mixed findings on the influence of age, parents' education, gender and school location on student scholastic attainment.

In framing appropriate attendance policies, empirical findings from other studies need to be contextualised as the environmental factors play important roles in shaping study results. It is noteworthy to state that the subject of attendance in Zambia has become more

important than ever before due to the current two-career pathway curriculum changes that have subjects that demand hands-on learning. The implication is that students who miss a day's practical demonstration will have a greater disorientation in skills progression.

Our review of the literature indicates that only a handful of empirical studies have taken into account and corrected for the endogenous nature of attendance. Most of the reviewed studies from African countries on this subject have been qualitative, making it difficult for policy guidance. Furthermore, the reported findings of attendance effect on performance do not explicitly indicate the level of significance of the parameter estimate. Suggestion on several variables including age, education, and gender and school location is inconclusive and numerous studies do not provide propositions as to how variables used in their studies influence student learning outcomes. In Zambia, and Kafue in particular there is scanty evidence of literature specific to quantitative assessment of student attendance in relation to performance in public day secondary schools.

The preceding arguments therefore, indicate that there is still a gap in our understanding of how and what exact factors affect student performance and to what extent mandatory class attendance can improve student academic welfare in Zambia and other developing countries.

CHAPTER FOUR

METHODOLOGY

4.0 Theoretical Framework of Class Attendance Behaviour and Academic Performance

The theoretical basis of this study is derived from Bratti and Staffolan (2002). The authors studied student time allocation and educational production function. The analysis considers an individual who decides to undertake higher education in order to obtain a qualification in a given course in a certain number of years taken as first period. After graduation in the second period, the individual starts working. In this model, Bratti and Staffolan (2002) estimate the relationship between student exam score and student time allocation on self-study and class attendance.

The authors define individual student's utility as:

$$U_j = u\,(c_{1j}, \ell_{1j}) + \beta u\,(c_{2j}, \ell_{2j}) \qquad (4.1)$$

where, after dropping the individual subscript j to simplify the notation:

1. $u(.)$ are period utility functions, whose first derivatives are positive and second derivatives negative in the arguments
2. β depends on the relative duration of working life with respect to the period devoted to schooling and the discount rate;
3. $\ell_1 = 1 - \sum_{i=1}^{N}(s_i + \lambda a_i)$ is first period leisure time, with s_i representing time committed to self-study and a_i being time committed to class attendance in subject i. The present study treats s_i as number of days during the school calendar a student commits to self-study and a_i being number of days committed to class attendance

of curriculum subjects. N is the total number of subjects to be passed and $\lambda > 1$ is a coefficient which takes account of the fact that λa_i hours are needed to produce a_i hours of class attendance and depend on the student's time taken to travel to class;

4. $\ell_2 = \bar{l}_2$ is leisure time in the second period which is assumed to be exogenous. Then, $(1 - \bar{l}_2)$ is exogenous second period labour supply;
5. $c_1 = \bar{c}_1$ is first period consumption equated to exogenous revenue from the family or institutions (such as bursary) net of the cost of education during the period of schooling.
6. c_2 is second period consumption which is assumed to solely depend on labour income. The individual's second period wage is also assumed to be an increasing concave function g of the scores earned in exams (G_i). It is assumed that $c_2 = wg(1 - \bar{l}_2)$, where w is a non-graduate wage and g is some markup due to academic performance.

Scores (G_i) obtained in the exams are taken to be an indicator of academic performance and define g as follows:

$$g = g(G_1, G_2 \ldots\ldots\ldots\ldots G_N) = \sum_{i=1}^{N} In(Gi) \qquad (4.2)$$

This formulation is coherent with the hypothesis of an increasing concave relationship between wage and scores earned in the single exams. The scores obtained in each exam is a function of three components; time committed to self-study; time spent attending classes and of unobserved individual ability (e_i) (Arulampalam, Naylor and Smith, 2007).

Thus, $G_i = G(s_i, a_i)\, e_i$. The assumption on e_i is that it is log-normally distributed with mean zero. Second period consumption is therefore given as:

$$c_2 = w \sum_{i=1}^{N} In\,(G\,(s_i, a_i)\, e_i)\,(1-\bar{l}_2). \tag{4.3}$$

Each student maximizes the expected utility of equation (4.1) subject to; period one time constraint in time allocation between school and leisure; non-negativity constraints for all time variables (self-study, attendance and leisure); the constraint $ai \leq \overline{ai}$ where $\overline{ai}$ is maximum attendance time in subject *i* and the budget constraint in the second period:

$$\max_{si,ai} E\,[u(\bar{c}_1, 1 - \sum_{i=1}^{N}(s_i + \lambda a_i) + \beta u\,(w\sum_{i=1}^{N} In(G_i\,(s_i, a_i)\, e_i)\,(1-\bar{l}_2), \bar{l}_2] \tag{4.4}$$

E[.] is the expectation operator at time 0, i.e. at enrollment. The student's choice variables are the self-study time and class attendance time denoted by s_i and a_i respectively in N subjects.

Applying the first period constraint, it turns out that $u'a_i = u'a_j = -\lambda\, u'_{\ell 1}$ and $u'_{si} = u'_{sj} = -u'_{\ell 1}$ for every subject *i, j* where u' indicates first derivative. The first order conditions of the problem are in the case of an interior optimum:

$$u'_{\ell 1} = \theta E\,(u'c_2)\frac{G'_{si}}{Gi} \quad \text{for } i = 1.....N \tag{4.5}$$

$$\lambda u'_{\ell 1} = \theta E\,(u'c_2)\frac{G'_{ai}}{Gi} \quad \text{for } i = 1.....N \tag{4.6}$$

where $\theta \equiv \beta w\,(1-\bar{l}_2)$.

$u'c_2$ is the first derivative of the second period utility with respect to consumption. G'_{si} and $G'a_i$ are the first derivatives of scores in exam *i* with respect to self-study and class attendance respectively.

It is next supposed that the score obtained in subject *i* is a constant elasticity function of self-study and class attendance with elasticities denoted by ε_{si} and ε_{ai} respectively. A

Cobb-Douglas relationship between score obtained in subject *i*, time devoted to self-study and time devoted to class attendance is assumed and specified as $G_i = z_i s i^{\alpha} a i^{\beta} e_i$; where $\alpha = \varepsilon_{si}$; $\beta = \varepsilon_{ai}$ and z_i is a subject specific scale parameter. The specification can be generalised to include also person specific attributes such as individual's ability.

The previous system can be represented as:

$$s_i u'_{\ell 1} = \theta E(u' c_2) \varepsilon_{si;} \qquad \text{for } i = 1.....N \tag{4.7}$$

$$\lambda a_i u'_{\ell 1} = \theta E(u' c_2) \varepsilon_{ai;} \qquad \text{for } i = 1.....N \tag{4.8}$$

hence: $s_i^* = \frac{\varepsilon_{s_i}}{\varepsilon_{s_j}} s_j^*$ and $a_i^* = \frac{\varepsilon_{a_i}}{\varepsilon_{a_j}} a_j^*$ so that

$$a_i^* = \frac{1}{\lambda} \frac{\beta}{\alpha} s_i^* \tag{4.9}$$

this shows that if $\beta > 0$ and $\alpha > 0$ then a_i^* and s_i^* are positively associated. When time taken by a student to travel to attend classes is accounted for, the optimal ratio between time committed to class attendance and time committed to self-study is given by the ratio between score elasticities with respect to the two time inputs.

The ratio between times for self-study or class attendance in two different subjects must be equal to the ratio between the elasticities of performance to self-study or class attendance in the two subjects.

With the optimal class attendance and self-study time we can then specify the Education Production Function (EPF) of each exam:

$$\ln(Gi^*) = \ln(z_i) + \varepsilon s_i \ln(s_i^*) + \varepsilon a_i \ln(a_i^*) + \ln(e_i)$$

$$= \ln(z_i) + \alpha \ln(s_i^*) + \beta \ln(a_i^*) + \ln(e_i) \tag{4.10}$$

Summing across equations we also obtain g^*, the index of overall academic performance evaluated at the optimum:

$$g^* = Z + \sum_{i=1}^{N} [\varepsilon s_i \ln(s_i^*) + \varepsilon a_i \ln(a_i^*)] + \sum_{i=1}^{N} \ln(e_i)$$

$$= Z + \sum_{i=1}^{N} [\alpha \ln(s_i^*) + \beta \ln(a_i^*)] + \sum_{i=1}^{N} \ln(e_i) \quad (4.11)$$

where $Z = \sum_{i=1}^{N} \ln(z_i)$. From these expressions, it follows that by estimating equation (4.11) the log coefficients of class attendance and self-study times in given subjects provide the elasticity of the scores obtained in the exam with respect to the same variables. In fact, equation 4.11 yields:

$$\frac{\partial g*}{\partial ln(si*)} = \frac{\partial ln(Gi*)}{\partial ln(si*)} = \varepsilon si = \alpha.$$

Similarly,

$$\frac{\partial g*}{\partial ln(ai*)} = \frac{\partial ln(Gi*)}{\partial ln(ai*)} = \varepsilon ai = \beta$$

Replacing the expression for s_i^* in equation (4.9) into the exam performance equation (4.10) we see that;

$$\ln(Gi^*) = \ln(z_i) + \varepsilon_{si} \ln\left(\lambda \frac{\varepsilon_{si}}{\varepsilon_{ai}}\right) + (\varepsilon_{si} + \varepsilon a_i) \ln(a_i^*) + \ln(e_i)$$

$$= \ln(z_i) + \alpha \ln\left(\lambda \frac{\alpha}{\beta}\right) + (\alpha + \beta) \ln(a_i^*) + \ln(e_i) \quad (4.12)$$

This specification is more commonly used in empirical literature in which inclusion of time for self-study is rarely done. A potential endogeneity pitfall could therefore emerge for time variables owing to the correlation between these variables and a student-subject specific unobserved ability $\ln(e_i)$. Equation (4.12) demonstrates that a regression omitting

other dimension of time use such as that devoted to self-study, generates an overestimation of performance elasticity with respect to time of class attendance. The size of overestimation is the elasticity of performance with respect to self-study. In this case the coefficient of a_i^* is the total elasticity of scale. In other words education production functions like (4.12) allow the researcher to estimate the aggregate effect of self-study and class attendance time only, but not the separate effects of the two time inputs. Provided that potential endogeneity issues are addressed, $(\alpha + \beta)$ is still a parameter of interest since it gives information about the change in student performance that one should expect in response to policies changing class attendance. This, nevertheless, serves as a motivation to apply Instrumental Variables strategy in the analysis of the influence of class attendance on student performance.

Some studies have tried to resolve the potential endogeneity of class attendance by instrumenting it using students' travel costs. In this model this is accomplished theoretically by imposing more structure on the problem. To solve the students' optimization problem, we sum across equations (4.7) and (4.8). We arrive at $\sum_{i=1}^{N}(s_i + \lambda a_i)$ on the left hand side, which is equal to $1-\ell_1$. Therefore:

$$u'_{\ell 1}(1-\ell_1)=\theta E(u'c_2)\sum_{i=1}^{N}(\varepsilon si+\varepsilon a_i) = \theta E(u'c_2)\sum_{i=1}^{N}(\alpha+\beta) \qquad (4.13)$$

To obtain an explicit solution for the maximization problem, we suppose that $u_t = kc_t + ln\,(\ell_1)$ for t=1, 2 so that $u'_{\ell 1} = \frac{1}{\ell 1}$ and $u'c_2$ = k. The optimal leisure time in t=1 period is driven thus;

$$\ell_1^* = \frac{1}{1+\theta k\sum_{i=i}^{N}(\varepsilon_{si} + \varepsilon_{ai})} = \frac{1}{1+\theta k\sum_{i=1}^{N}(\alpha+\beta)} \qquad (4.14)$$

the optimal time to be committed to self-study being

$$s_i^* = \frac{\theta k \varepsilon_{si}}{1+\theta k \sum_{i=i}^{N} (\varepsilon_{si} + \varepsilon_{ai})} = \frac{\theta k \alpha}{1+\theta k \sum_{i=1}^{N} (\alpha + \beta)} \quad (4.15)$$

and the optimal time devoted to class attendance is

$$a_i^* = \frac{\theta k \varepsilon_{a_i}}{\lambda(1+\theta k \sum_{i=i}^{N} (\varepsilon_{si} + \varepsilon_{ai}))} = \frac{\theta k \beta}{\lambda(1+\theta k \sum_{i=1}^{N} (\alpha + \beta))} \quad (4.16)$$

Thus this simple model suggests that travel time λ(for instance proxied by student residence) can be used as an instrument for class attendance time. Nevertheless there are reasons to question the validity of this "seemingly good" instrument. This is because more motivated students could possibly choose to rent a room closer to school in order to attend classes, thus creating a bias in estimating the effect of class attendance time on performance.

4.1 Data

The methodology and variables in this study were selected taking into account their relative importance on the theoretical and empirical basis. This was also consistent with data chosen by other researchers. The data used in this study were cross-sectional and were from three sources. Student data on Grade 9 final exam scores for 2014 were obtained from the Examinations Council of Zambia (ECZ). Data on student attendance, socio-economic status, family size, ratio of text books to students, ratio of students to classrooms, ratio of students to classroom furniture, presence of school library and ratio of students to grade-specific qualified teachers were obtained from school administrative records. To ascertain consistency in data from school records, a cross-check was done using 2014 data on schools from the District Education Office. The District Education Office consolidates annual data from individual schools that are submitted through the

school monthly returns and Annual School Census Questionnaire (ASCQ). School data captured in the ASCQ come from schools to the district office in three copies. The District Education Office retains one copy from each school while the remaining two copies per school are transmitted to the Provincial Education Office and Ministry of General Education headquarters. At ministry headquarters these school data are also entered and stored in the ED*ASSIST files.

The study unit of analysis was students and consisted of all 4,405 Grade 9 candidates who wrote examinations in 2014 in Kafue district. To arrive at the eligible number of study units for inclusion in the analysis, student data on scores were matched with data from school administrative records. Data from ECZ on student final examinations contain information on name of student, student exam score, year, student gender and name of school where an examination was written.

Each student's examination score was then linked to a particular school whose administrative record contained information on student personal, family and school characteristics. From this matching, students with missing data from school administrative records, those from boarding schools and those from the only private day secondary school with Grade 9 candidates in the district were dropped from the analysis. The justification for dropping students from boarding schools was that attendance patterns in these schools are often consistent and not affected by student distances since students lodge at the respective schools. For those in a private school the distribution of their scores was sharply skewed to the right with 100 percent pass rate at Grade 9, which was over and above the district mean score. In addition, school records from the private schools did not have complete information on the students as was the case for the GRZ schools. After dropping

some students using the above criteria, the number of eligible students included in the analysis was 2,579 Grade 9 students who sat for 2014 examinations in both rural and urban day GRZ secondary schools. Data ware analysed using STATA 13 software package.

4.2 Study Variables

4.2.1 The Dependent Variable

This study used for its dependent variable, examination composite scores of 2014 Grade 9 candidates that were obtained from the ECZ. At junior secondary school level, aggregate scores range from 0-600 in the best six subjects for junior secondary certificate award and subsequent transition to the 10^{th} Grade.

4.2.2. The Independent Variables

The independent variables used in the study were based on the considered theoretical and empirical literature (Nizamettin and Bekir, 2015; Mbugua et al., 2012).

Student Annual Attendance Days

These are total days that a student attended class as monitored by class teachers and reported in the class attendance register. In this study we refer to Grade 9 total attendance days in 2014. According to the school calendar, each school term has 13 weeks of learning or 65 days. This means that a consistent student would attend a maximum of 195 mandatory school calendar days in a year. A number of day secondary schools in Kafue district that were supplied with relief bicycles for students with a long mileage have strictly followed an agreed attendance policy which requires either extending the term by 5 days or opening the new term 5 days earlier.

Gender of a Student

Gender was categorised by applying a dummy variable with the code 1 to represent a female student and 0 for male. Conclusions on influence of student gender on student performance in most literature is mixed, with some literature finding female dominance and another finding no relationship (Khwaileh and Zaza, 2011; Olusola and Omotade, 2014).

Age of the Student

Age was also considered in the analysis. In Zambia the mandatory age of enrolling for the first grade is 7 years and by the 9^{th} Grade a student is expected to be 15 years old. The explanation in some literature that find student age to significantly influence performance is that as students grow up to a certain age their performance is high, ceteris paribus (Erdogan, Bayram and Levent, 2008).

Number of Students per Grade Specific Teacher

In most studies this variable is treated as student teacher-ratio computed by dividing the total number of students in school by the total number of teachers. However in this study this variable was computed by dividing the number of students in a given grade by only the number of teachers qualified and assigned to teach that grade. Literature suggests a positive impact of student-teacher ratio on student performance (Nizamettin and Bekir, 2015; Mbugua et al., 2012).

Number of Text Books per Student

This ratio was computed by dividing the total number of text books prescribed at junior secondary level by the total number of Grade 9 students. The text books include those

used in the 8th grade because the final examination material covers the 8th and 9th grade content. Empirical evidence suggests a positive correlation of this variable with student performance (Mlozi et al., 2013).

Number of Students per Classroom

The number of grade 9 students was associated with the allocated classrooms to find this ratio. It was calculated by dividing the total number of grade 9 students by the number of classrooms allocated and used by these students. Studies conducted on class size show existence of a negative correlation of this variable with student performance (Moshoe, 2015).

Number of Students per Desk

This is a ratio of students to available desks in a classroom, computed by dividing the number of Grade 9 students by number of desks allocated to grade 9 classes.

Income of Parent or Guardian

Income of either parent or guardian was proxied by parent employment status and entered in the analysis as a dummy coded 1 if parent or guardian was working and 0 if not. Literature provides strong evidence of the influence of parent's income and socio-economic factors on the performance of students (Farooq et al., 2011).

Size of the Family

In this study size of the family was taken to be the number of members of the household where the student came from. Some literature presents evidence on the significant

influence of family size on student academic performance (Tenibiaje, 2009; Ella, Odok and Ella, 2009).

Presence of School Library

School library was used as a dummy and coded 1 if the school had a library and 0 otherwise. Presence of school library most likely increases students' allocation of time for self-study (Mlozi et al., 2013; Melack, 2014).

4.3 Empirical Model and Estimation of the Class Attendance Parameter

The interest was in estimating the parameter characterizing the relationship between class attendance and student exam scores. It was assumed that a student score was the output of an educational production function that reflected the match between two types of factors: academic input and student input. This study adopted the modelling of an education production function by Andrietti, D'Addazio and Gómez (2008) in which student performance was a function of class attendance. Among others this approach has also been used by Gottfried (2009). The OLS basic learning model based on the assumption of linearity was specified as follows:

$$Y_{ik} = \beta_0 + \beta_1 A_{i\,k} + \beta_2 X_{i\,k} + u_{i\,k} \qquad i = 1, 2 \ldots\ldots\ldots\ldots\ldots n \tag{4.17}$$

Where Y_{ik} is the educational achievement for individual i in school k, measured by exam composite score. A_{ik} is class attendance days of individual i in school k. X_{ik} is a vector of selected inputs into the achievement process of individual i in school k and u_{ik} is an error term containing all the other factors influencing academic performance.

Input measures are those suggested by both theoretical considerations and results of previous studies. They include student level characteristics, school resources, family

background and socio-economic variables. In this study, these variables are subsumed in the following vector (X) of covariates for the i^{th} individual specified as:

$X_{i=}$	Student gender (entered as a dummy)	X_1
	Student Age	X_2
	Square of student age	X_3
	Number of students per grade specific qualified teachers	X_4
	Number of text books per student	X_5
	Number of students per classroom	X_6
	Income of parent or guardian (entered as a dummy)	X_7
	Family size	X_8
	Number of students per desk	X_9
	Presence of school library (entered as a dummy)	X_{10}

Among the several factors which matter for academic performance, there are also certain unobservable student characteristics, such as student ability (Cameron and Trivedi, 2005:183). Since these same variables are potentially correlated with students' class attendance, omitting them from the model would give rise to a problem of omitted variable bias.

From both empirical and theoretical literature the attendance variable suffers from the problem of endogeneity. This results in Ordinary Least Squares (OLS) parameter estimation to produce inconsistent and biased estimates. The endogeneity problem of attendance is attributed to the fact that student ability is correlated with both student attendance and performance. An Instrumental Variables (IV) approach was therefore, used to address the problem of reverse causality in order to estimate a causal relationship between class attendance and performance (Cameron and Trivedi, 2005:183).

More formally, in a multiple regression model with a typical observation

$$y = x'\beta + u \tag{4.18}$$

with K regressors , so that x and β are K x 1 vectors, an r x 1 vector of instruments Z with r ≥ K for the endogenous variable x satisfies the following conditions (Cameron and Trivedi, 2005):

a. $Cov(X,Z) \neq 0$ (4.19)

b. $Cov(Z,u) = 0$ (4.20)

c. Z is strongly correlated with the regressor vector x (4.21)

Instrumental variable (IV) estimation proceeded as follows:

Given a model

$$y = \alpha + \beta X + u \tag{4.22}$$

Multiplying through by the instrument Z to obtain

$$Zy = Z\alpha + \beta ZX + Zu \tag{4.23}$$

Implies that

$$Cov\,(Z, y) = Cov\,[Z\alpha + \beta ZX + Zu]$$

$$= Cov\,(Z\alpha) + Cov\,(\beta Z, X) + Cov\,(Zu) \tag{4.24}$$

Since $Cov\,(Z\alpha) = 0$ (using rules on covariance of a constant)

and $Cov(Zu) = 0$ (if assumption above about the properties of instruments is correct) then

$$Cov\,(Z,y) = 0 + \beta Cov\,(Z,X) + 0$$

Solving $Cov\,(Z, y) = 0 + \beta Cov\,(Z, X) + 0$ for β

leads to the calculation of the instrumental variable estimator

$$\hat{\beta}_{IV} = \frac{Cov(Z,y)}{Cov(Z,X)} \tag{4.25}$$

which, in a multiple regression, is equivalent to

$$\hat{\beta}_{IV} = \frac{(Z'Z)^{-1}Z'y}{(Z'Z)^{-1}Z'X} = \quad (Z'X)^{-1}Z'y \tag{4.26}$$

The study then ran a two-stage least squares regression using distance covered to reach schools (in kilometres) and student orphan hood (1=Orphan) as instruments for attendance by specifying the following model:

$$Y_{ik} = \beta_0 + \beta_1 A_{ik} + \beta_2 D_{ik} + \beta_3 O_{ik} + \beta_4 X_{ik} + \mu_{ik} \tag{4.27}$$

Where Y_{ik} is the educational achievement for individual *i* in school *k*, measured by exam composite score. A_{ik} is class attendance days of individual *i* in school *k*. D_{ik} is distance covered to school *k* by student *i*. O_{ik} is student orphan hood status. X_{ik} is a vector of selected inputs into the achievement process of individual *i* in school *k* and u_{ik} is an error term containing all the other factors influencing academic performance.

Information on distances covered and student orphan hood was extracted from school administrative records. These records contain lists of residential addresses of students and approximated distances. Schools in Kafue District compile distances covered by individual students as a requirement by the District Education Office in order to determine the number of students who are eligible for relief bicycles from World Bicycle Relief. School records also contain names of orphans and vulnerable children (OVC) by grade and gender.

Justification for Use of Distance and Orphan hood as Instruments

A number of essential features are vital for making an educational environment favorable for learning. To begin with, accessibility to school is cardinal. The longer the distances between students' homes and schools, the less likely students will be to attend school regularly. While the impact of school distance is intuitive, evidence quantifies the effect, indicating that each additional kilometre a student lives from school leads to attendance declining by 20 percent or more. Erratic attendance has a strong effect on learning. Evidence from Honduras shows that, a student's school attendance record is a major predictor of dropout and grade repetition. Moreover, distance from home to school can impact girls and boys differently, reinforcing the need to ensure reasonable access to schools in order to prevent gender-based learning disparities (USAID, 2008: 6). On the other hand, orphan hood often affects a child's development by increasing the risk of missing out on educational opportunities, living in a home which is food insecure and may cause suffering from anxiety or depression. Orphans most likely miss class in order to provide for their siblings (CSO, 2012).

4.3.1 The Instrumental Variables by the Two-Stage Least Squares

The 2SLS estimator gets its name from the result that it can be obtained by two consecutive OLS regressions. In this study the first part involved the first stage regression of the endogenous variable on instruments to obtain its predicted value $\hat{X}$ while the second stage involved regressing the dependent variable y on the predicted values of the endogenous variable to obtain the 2SLS estimate. The First stage regression was given as:

$$\hat{X} = Z(Z'Z)^{-1}Z'X \tag{4.28}$$

Where $\hat{X}$ was the predicted value of X from the OLS regression of X on a set of instruments Z.

In this study, regressing student composite scores on predicted attendance produced the 2SLS estimator given as:

$$\hat{\beta}_{2SLS} = [X'Z(Z'Z)^{-1}Z'X]^{-1}[X'Z(Z'Z)^{-1}Z'y] \quad (4.29)$$

Endogeneity of student attendance has been addressed by panel data methods that have employed fixed effect estimators (Gottfried 2009). These have mostly used lagged variables such as previous student scores. However, the present study used cross-sectional data and thus opted not to use such methods.

4.4 Determination of Differences in Student Performance by Location and Gender

The reviewed literature that considered differences in student performance by location of the school and gender of the student highly applied the methodology that employed use of t-tests. The t-tests were used to test the null hypotheses of no mean difference between two independent samples. Most of these studies found statistically significant differences (Khwaileh and Zaza, 2011; Olatude, 2011; Bosede and Enuloju, 2013; Abdullahi et al., 2015). This study applied similar t-tests to establish whether student performance varied by school location and student gender.

4.5 Diagnostic Tests

Diagnostic tests were run on the data to ensure precise and meaningful estimation of our models. Specifically, it was ensured that the models were correctly specified, there was

no heteroskedasticity, no endogeneity, no influential data points (outliers) and there was no serious multi-collinearity.

4.5.1 Model Specification

The study considered a possibility of specification error by either omission of relevant variables or inclusion of irrelevant variables in the OLS regression. In this study two tests were used to check for model correctness. Results of these tests indicate that the OLS model was not correctly specified as it had omitted variables. The null hypothesis tested was that the model was correctly specified and had no omitted variables. The following table presents the test statistics of the two tests side by side:

Table 2: Results of the link test and Ramsey specification test

Type of Test	**Variable**	**t-Statistic**	**Prob > |t|**
Link test	Hatsq	4.96	0.000
Type of test	**Variable**	**F-Statistic**	**Prob > F**
Ramsey specification	Powers of Fitted Values of score	F(3, 2564)= 27.21	0.000

Based on the results in table 2, we reject the null hypothesis and conclude that our model had omitted variables. We used Instrumental Variables to address the omitted variables problem.

4.5.2 Normality

Normality in the dependent variable distribution of residuals was tested in the study. STATA results of the "qnormal" plot presented in the appendix 1 show that the quantile distribution of student scores fairly followed the quantiles of a normal distribution. The Shapiro-Wilk W test for normal data however, indicated that the scores were not normally distributed. We rejected the null hypothesis of normality as shown by the following statistics:

(W=0.996; V= 6.718; Z=1 4.890), Prob>Z (0.000)

4.5.3 Heteroskedasticity

The learning model was first estimated by OLS regression in which the relationship between class attendance and student performance was modelled. This model was run without controlling for heteroskedasticity. However, when the diagnostic white test with Cameroon and Trivedi's IM-test decomposition was performed, it showed that the model suffered from heteroskedasticity. An alternative robust OLS regression that controlled for potential heteroskedasticity was then run with use of Huber-White standard errors. Wooldridge (2000) indicates that the advantage of using robust standard errors is that they are asymptotically valid in the presence of many forms of heteroskedasticity.

4.5.4 Endogeneity

Literature on class attendance shows that the attendance variable is endogenous. To test this endogeneity, a Durbin chi-square and Wu-Hausman F-tests of endogeneity on the model with two instruments were used. Results are reported in the findings section.

4.5.5 Test for Weak Instruments

In the implementation of IV regression, two instruments (distance and orphan hood) were used. To assess the strength of the used instruments, first stage regression partial correlations of the instruments with the endogenous variable together with the values of their *F*-statistics were computed in STATA. Results for this test are also reported in the findings section.

4.5.6 Test for Over Identifying Restrictions

Testing for correct model specification and instrument validity was accomplished by running an over identification test. The next chapter on findings presents the results.

4.5.7 Outlier Test

Cook's D and DFITS in STATA ware used to identify influential data points in the model. These tests seemed to suggest that the data used in this study had some outliers. It is not always clear how to handle outliers and it may be left to the consideration of the researcher (Musole, 2015; Gujarati, 2004). No data point on scores was dropped on account of being influential as the scores were distributed within range.

4.5.8 Multicollinearity

The study was concerned about the possibility of multicollinearity among regressors. Considering the nature of the data used, there was no prior motivation to suspect multicollinearity to be a serious problem. However, a variance-inflating factor (VIF) besides a pairwise correlation matrix was run to check for possibility of multicollinearity. The 1.82 mean value of the VIF in appendix 2 indicates low levels of multicollinearity. This is significantly lower than the benchmark of 10 beyond which multicollinearity

would be a problem (Gujarati, 2004:362). Tolerance values (1/VIF) of regressors also showed that they were greater than the 0.1 threshold, ruling out the presence of multicollinearity in regressors.

CHAPTER FIVE

PRESENTATION OF RESULTS

5.0 Results

5.0.1 Descriptive Statistics

The data show that 2,579 sampled students who participated in the 2014 Grade 9 examinations were comprised of 1,436 male and 1,143 female students, representing 55.70 percent and 44.30 percent respectively. The mean score for the full sample of students was 291.88 with a standard deviation of 83.84, the minimum score being 7 and maximum 544. It was observed that 47.38 percent of the sample obtained mean scores and below while 52.62 percent had scores above the mean.

The data also show the full sample mean value of the attendance variable of 160.98 days and standard deviation of 22.42. Student annual attendance days ranged from a minimum of 10 days to a maximum of 187 days. The percentage of students with mean attendance days and below was 40.75 percent and that with attendance days above the mean was 59.25 percent. The other observation from the data was that 58.90 percent of the sample covered a mean distance of 1.81 Km and below to school, while 41.10 percent covered distances above this average.

The summary of the descriptive statistics of all variables for individual, school and socio-economic characteristics for the sample are presented in Table 3.

Table 3: Descriptive statistics of variables used in analysis

Variable Name	Mean	Std. Dev.	Min.	Max.
Student composite score	291.876	83.842	7	544
Student annual attendance days	160.982	22.419	10	187
Gender (1=Female 0=Male)	0.443	0.497	0	1
Age	16.307	1.354	10	21
Number of students per grade-specific teacher	17.085	9.195	1.7	32.4
Number of text books per student	0.648	0 .712	0.02	3.3
Number of students per class room	64.308	15.135	29.5	90
Parental employment status (1 if parent works; 0 = otherwise)	0.487	0.50	0	1
Size of family a student comes from	5.592	1.976	2	16
Distance in km covered by student	1.812	1.579	0.01	10
Number of students per desk	2.585	0.762	1.2	4.3
School location (1 if school is in rural and 0 urban)	0.333	0.471	0	1
Student vulnerability (1 if an orphan; 0 other wise)	0.223	0.416	0	1
Presence of school library (1 if school has library, 0 otherwise)	0.694	0.461	0	1

5.1 Effect of Class Attendance on Student Performance

To examine the effect of class attendance on student performance as measured by the ninth grade examination composite scores, results from the estimated learning model by robust OLS regression show that attendance had a significant positive effect on student performance at 1 percent level. The results suggest that attending an extra day of class

increases average examination composite score by about 1.237. Table 4 presents the robust OLS results.

Table 4: Regression results for OLS regression

Variable	OLS
Class Attendance	1.237***
	(0.071)
Gender (1=Female 0=Male)	-20.672***
	(2.975)
Age	-57.043**
	(18.804)
Age squared	1.510**
	(0.567)
Number of students per grade-specific teacher	-0.063
	(0.241)
Number of text books per student	18.817***
	(2.765)
Number of students per class room	-0.565**
	(0.173)
Parental employment status (1if parent works; 0= otherwise)	6.509*
	(2.992)
Size of family a student comes from	0.810
	(0.723)
Number of students per desk	12.532***
	(3.544)
Presence of school library (1 if school has library, 0 otherwise)	15.502***
	(3.661)
Constant	602.073***
	(156.687)
R^2	0.26

Significance levels: * 10 percent level, ** 5 percent level; *** 1 percent level, robust standard errors are in parentheses

Given that the model had omitted variables, an Instrumental Variables approach was used in which a two-stage least squares (2SLS) regression with distance and orphan hood as

instruments for class attendance was run. The reason for instrumenting attendance with distance and orphan hood is that attendance is an endogenous variable. It is however, important to note that attendance is negatively correlated with these instruments and that the instruments have no direct linkage with the dependent variable except through the endogenous variable. Correlation coefficients of attendance with the instruments are presented in the two-way correlation table in appendix 3. The results show high significant negative correlations between class attendance and the two instruments; distance and orphan hood status.

Results of formal tests to justify the endogenous nature of the attendance variable are presented in Table 5.

Table 5: Test results for endogeneity

Ho: Variables are exogenous		
Test	**Value**	**Significance**
Durbin (score) chi2(1)	16.684	(p = 0.000)
Wu-Hausman F(1,2567)	16.708	(p = 0.000)

From the above table, it is noted that the reported Durbin chi-square and Wu-Hausman *F*-statistics have very small p-values. This makes the null hypothesis of variable exogeneity rejected in favour of variable endogeneity. This means treating attendance variable as endogenous was correct.

Results of another test on instruments for instrument strength indicate that the instruments chosen were fairly good and strong as shown by the STATA output below.

(Partial R-sq = 0.190; F (2, 2566) = 301.39; prob. > F =0.000)

Examination of the *F*- statistic showed that 301.39 was greater than any of the given critical values at the given percentage levels. This meant that the null hypothesis of weak instruments was rejected and that the instruments used were fairly good.

Results of the test for correct model specification with instruments and instrument validity are presented in Table 6. The null hypothesis tested was that instruments were valid and that the model was correctly specified.

Table 6: Test results of over identifying restrictions

Test	Value	Significance
Sargan (score) chi2(1)	0.184	(p = 0.668)
Basmann chi2(1)	0.183	(p = 0.669)

The p-values of the Sargan and Basmann tests in Table 6 indicate that the null hypothesis is not rejected, indicating correct model specification and instrument validity.

The first stage of the 2SLS regression was run where attendance was regressed on the two instruments and other control variables to obtain predicted attendance, controlling for heteroskedasticity and the results are shown in appendix 4. In the second stage scores were regressed on predicted attendance to obtain the Instrumental Variables estimate of

attendance. In this study an over identified model where two instruments were used for one endogenous regressor was run. The results in Table 7 show Instrumental Variables estimates.

Table 7: Regression results for IV regression

Variable	IV
Class Attendance	0.670***
	(0.167)
Gender (1=Female 0=Male)	-20.402***
	(2.986)
Age	-55.164**
	(20.013)
Age squared	1.478*
	(0.605)
Number of students per grade-specific teacher	-0.461
	(0.262)
Number of text books per student	18.701***
	(2.886)
Number of students per class room	-0.666***
	(0.175)
Parental employment status (1 if parent works; 0 = otherwise)	7.888**
	(3.043)
Size of family a student comes from	0.709
	(0.729)
Number of students per desk	15.863***
	(3.667)
Presence of school library (1 if school has library, 0 otherwise)	14.660***
	(3.642)
Constant	676.514***
	(167.886)
R^2	0.24

Significance levels: * 10 percent level, ** 5 percent level; *** 1 percent level, robust standard errors are in parentheses

The results in Table 7 indicate that the estimated effect of attendance on performance showed a more reduced effect under IV approach than under OLS. It declined from 1.237

to about 0.670, but was still positively related to student performance and significant at 1 percent level. This result suggests that under IV estimation, attending an extra day of class would instead, increase average student composite score by 0.670. Moreover, the findings show that estimates that do not take into account the endogenous nature of class attendance in the determination of student performance over estimate the effect.

Some control variables included in both the OLS and IV regression analysis had a significant effect on student performance, while others were insignificant. In both the OLS and IV specifications gender of a student had a statistically significant effect at 1 percent level. Female's mean composite score was 20.672 lower under OLS and 20.402 under IV estimations than males'. Student age was significant at 5 percent level in both OLS and IV specifications. The results indicate that as the age of a student increased by a year their average composite score declined by 57.043 under OLS and 55.164 under IV. Student-teacher ratio though negatively related to performance was not statistically significant in both OLS and IV specifications.

Findings on the number of text books per student indicated that the variable had a positive relationship with student performance. In both specifications, it was significant at 1 percent level. The parameter estimate suggested that increasing the textbook-student ratio by a unit would increase average composite score by 18.817 under OLS and 18.701 under IV. The finding on the ratio of students to classrooms was interesting. While the ratio had a negative relationship with student performance in both OLS and IV, its level of significance changed from 5 percent in OLS to 1 percent in IV estimation. The results suggested that a unit increase in student-classroom ratio would reduce composite scores by 0.565 under OLS and 0.670 under IV. Another interesting result was income of parent,

proxied by a dummy variable of the working status of a parent or guardian. While income had a positive relationship with student performance in both OLS and IV, its level of significance changed from 10 percent under OLS to 5 percent under IV estimation. The results suggested that increasing the income of a parent by a unit amount would increase average composite scores by 6.509 under OLS and 7.890 under IV.

Family size in both OLS and IV was not statistically significant and carried an unexpected positive sign. Student-desk ratio was statistically significant at 1 percent level in both OLS and IV estimations. However, though statistically significant, this ratio did not yield the expected sign. It presented the unexpected result that seemed to suggest that a positive change in the student-desk ratio by a unit would increase the average composite score by 12.532 and 15.863 under OLS and IV respectively. The finding on the estimated effect of school library on student performance indicates that the mean composite score for schools with libraries was 15.502 higher under OLS and 14.660 under IV than those without.

5.2 Student Performance by Geographical Location

One of the objectives was to establish whether a difference existed in performance between students in rural and urban day secondary schools as measured by the ninth grade examination composite scores. Applying a two independent sample mean t-test the study found a statistically significant mean difference (1 percent level) in academic performance between students in urban schools and students in rural schools as summarised in Table 8. Results from the table show that students in urban schools performed better on average than those in rural schools. The results are consistent with literature.

Table 8: Test of difference between mean composite scores, by school location

Variable	Mean Score	Std. Dev.
Urban (n=1720)	303.81	85.24
Rural (n=859)	268.00	75.55
Difference (Urban-rural)	35.82***	3.30

Significance levels: * 10 percent level, ** 5 percent level; *** 1 percent level

5.3 Student Performance by Gender

A similar comparison of mean composite scores by gender shows that males' average performance was higher than that for females at 1 percent level of statistical significance. The results are presented in Table 9.

Table 9: Test of difference between mean composite scores, by gender

Variable	Mean Score	Std. Dev.
Male (n=1436)	304.46	87.46
Female (n=1143)	276.07	76.21
Difference (Male-Female)	28.38***	3.23

Significance levels: * 10 percent level, ** 5 percent level; *** 1 percent level

Further investigation of this gender difference in performance was carried out by filtering the full sample into two distinct sub-samples; the urban and rural sub-samples. Analysis of differences in gender performance in these sub-samples demonstrated that males still

performed better than females in urban day schools. However, this difference was only significant at 10 percent level in the case of the rural sub-sample. Table 10 presents the results.

Table 10: Test of difference between mean composite scores, by gender in urban and rural sub-samples

Variable	Mean Score	Std. Dev.
Urban area		
Male (n=949)	321.07	89.78
Female (n=771)	282.55	74.01
Difference (Male-Female)	38.52***	3.95
Rural area		
Male (n=487)	272.07	72.63
Female (n=372)	262.64	79.00
Difference (Male-Female)	9.43*	5.25

Significance levels: * 10 percent level, ** 5 percent level; *** 1 percent level

To establish whether these performance differentials were linked to class attendance behaviors of students by school location and gender, further two-sample t-tests were performed on student attendance by school location and gender. Results of these t-tests revealed that students in urban schools had higher mean attendance days than those in rural schools (t = 2.093; two tailed p-value = 0.037). Similarly, males in the full sample had attended more classes on average than females (t = 1.656; two tailed p-value = 0.098).

These mean differences in attendance rates were, however, only statistically significant at 10 percent level in both cases.

The study also examined the marginal effect of class attendance on scores with respect to gender and school location. Computation of marginal effects is shown in appendix 5A. Results in appendix 5B and appendix 5C show that the marginal effect of an additional day of class attendance on composite scores is 0.832 for females and 0.490 for males. Results in appendix 5D and appendix 5E indicate that the marginal effect of an additional day of class attendance on composite scores is 1.01 for rural students and 0.392 for urban students.

CHAPTER SIX

DISCUSSION OF FINDINGS

6.0. Discussion

This chapter discusses the results of the study. Determination of the effect of class attendance on student performance as the main focus of the study involved use of Instrumental Variables (IV) regression strategy. Attendance was noted to be an endogenous variable and instrumented it with distance and student orphan hood status. Similar studies that have instrumented attendance on distance are those by Gottfried (2009) and Andrietti et al. (2008). The study observed a decrease in the value of the attendance parameter estimate from its OLS value of 1.237 to its IV value of 0.670. This finding however, contradicts these earlier studies that reported higher IV coefficients of the attendance variable than that based on simple OLS. This finding seems to suggest that there exists a direction and magnitude of the influence of class attendance on student scores. In this study the influence is found to be unidirectional. The use of an over identified specification in the present study however, differs from Gottfried (2009) who used a single instrument in a just-identified model, but was similar to the Spanish study by Andrietti et al. (2008) who also used two instruments. This study's point of departure from the Spanish study is on the sampled subjects and the second instrument used. Whereas this study focused on junior secondary school students' performance in six subjects and used orphan hood as the additional instrument, the Spanish study focused on university student performance in microeconomics and used student's working status as an additional instrument to distance.

The other aspect of convergence of this study with Gottfried (2009) and Andrietti et al. (2008) is on the correlations of instruments with attendance. In both this study and those by Gottfried (2009) and Andrietti et al. (2008), the attendance variable was negatively correlated with the instruments.

The observed reduction in the coefficient of the attendance parameter through the two instruments signifies the critical nature of school distances and student vulnerability in relation to class attendance. In Kafue district where some schools are far from student homes and where a considerable number of students are orphaned, class attendance of some of these students in day secondary schools could easily be affected by these long mileages and their vulnerability status.

The significance of increasing class attendance rates is noted by Aucejo and Romano (2014) who report that reduction in student absences produces a 5.8 percent gain in maths and 3 percent gain in reading scores. The study strongly prefers a policy that finances a programme aimed at reducing absenteeism to the one in support of extending the school calendar. They argue that extending a school calendar by 10 days increased mathematics and reading scores only by 0.8 percent. The plausible explanation of low attainment with the school calendar extension is that the study dealt with students in lower grades whose attention and concentration may be limited to a prescribed length of the school term.

Another collaborating study on the impact of student attendance is that by Oghuvbu (2010) who reports a parallel positive relationship between student attendance and performance in secondary schools of Delta State, Nigeria. The Nigerian study finds that increasing attendance by a percentage produces an increase in average performance by 0.192. The only weakness of this study, though, is that the author did not take into account

the possibility of attendance having correlations with other factors. The method used in the Nigerian study differs substantively from the current study and those of Gottfried (2009), Andrietti et al. (2008) and Aucejo and Romano (2014). These studies including the current study made use of instrumental variables in recognition of the endogenous nature of attendance. To the contrary, earlier Instrumental Variables findings by Caviglia-Harris (2006) do not support the findings of this study, neither do they agree with those by Gottfried (2009), Andrietti et al. (2008) and Aucejo and Romano (2014). Caviglia-Harris (2006) did not find the number of days a student attended class to have a significant impact on their academic performance in an economics course. The study however, acknowledges the fact that absences in a course could adversely impact on the course grades. The author's further argument is that the claim of positive effect of attendance on performance which is often associated with the OLS regression disappears after controlling for simultaneity bias and endogeneity. The study instead finds student's previous knowledge of the subject to be the most significant predictor of student performance.

The other finding from this study shows that geographical location of the school has an influence on the performance of students and this result supports findings of the National Assessment Survey (MOE, 2008) that reported performance of students in urban schools to be higher than those in rural areas of Zambia. Findings by Ramous et al. (2012) also indicate that students from rural areas performed lower than those in urban areas. While Ramous et al. (2012) attributes poor academic performance of students in rural areas to low socio-economic conditions in these areas, this study could not make the same claim of socio-economic conditions to be explanatory of this rural-urban difference. This is so

because the present study did not extend its coverage to assessing the disparities in socio-economic conditions between rural areas and urban areas of Zambia. Yet another study by Olusola and Omotade (2014) confirms these geographical differences with urban secondary schools in Ekit State of Nigeria performing better in science examinations than rural schools. This difference in performance by location was also noted by Usaini Abu Bakar (2015) in the study of secondary schools in Kuala Terrengganu in Malaysia. The author claims that the environment of the school determined by availability of learning facilities had a positive relationship with student performance. However, since the current study had not taken into consideration the differences or similarities in learning facilities between rural and urban schools through the analytical framework, this attribute may not have influenced the finding in the current study. From the findings of this study and others that students' performance in rural schools was lower than that for students in urban schools there are a number of factors that could have been at play. For rural schools in Kafue district and other similar schools in Zambia, socio-cultural factors that make students combine school with household chores such as cattle herding, farming and other chores could have significantly affected class attendance and subsequently impact on performance. Contrary to findings of this study and others (MOE, 2008; Ramous et al., 2012; Olusola and Omotade, 2014; Usaini Abu Bakar, 2015), results from Ondo State of Nigeria by Bosede and Enuloju (2013) find students from rural schools as equally advantaged as those in urban schools. The study did not find a significant difference in student performance between rural and urban secondary schools in the final examinations.

Gender difference in scores was statistically significant in this study with mean scores for males being higher than females. These results conform to those by the National

Assessment Survey that reports that better performance of females in primary schools is reversed at secondary level with male students performing better than females in examinations (MOE, 2008). In Kafue district for instance, primary school female students persistently perform better than males in their Grade 7 final composite examinations. This pattern is nevertheless, reversed at Grade 9 where, besides poor performance, a number of girls dropout of school and others choose to be absent even in examinations (ECZ, 2013). The study findings on males performing better than females are, however, contradictory to those studies that report females as performing better than males instead. Findings from a study in Jordan by Khwaileh and Zaza (2011) indicate that university undergraduate female students scored higher on their pre-university Grade Point Average (GPA) than males. Farooq et al. (2011) also report similar results that the mean mark of secondary school females was, in their study, statistically significantly higher than that for males. Two comparable results on females exhibiting better performance are reported in Nigeria by Abdullahi et al. (2015), who report that secondary school female students in Katsima State had performed better in Agricultural Science than males. The results by Igbudu (2015) equally indicate that females had better overall grades in the West African Examinations Council exams in the Government subject than males. These results support earlier findings in New Zealand where females performed better than males on school certificate exam (Fergusson and Horwood, 1997). Findings from the current study and others are however, not supported by those from the study by Olusola and Omotade (2014) who report that there was no statistically significant difference in mean performance between male and female secondary school students regardless of school geographical location. The low performance of female students compared to males could possibly be attributed to the different domestic activities at home such as cooking, cleaning, fetching

water and general family care which consume more of their time for school. Poor sanitation facilities in some schools with no washrooms could also have affected attendance of females.

Class attendance in this study was not the only variable that had a significant effect on student performance. There were some variables included in the study as controls for class attendance. Some of these control variables indicated a statistically significant relationship with student performance under both OLS and IV analysis, while others that were not significant under OLS were also not significant under IV. Study results show that the ratio of students to teachers, for example, was not statistically significantly related to student performance under both IV estimation and OLS. This result was contrary to what was expected. Student-teacher ratio is an important policy variable for Ministry of General Education in the measuring of quality of school resources. The finding on non-significance of student-teacher ratio is contrary to that by Idowu and Oluwole (2014) who report a significant correlation between student performance in mathematics and student-teacher ratio for senior secondary schools of River State, Nigeria. Nizamettin and Bekir (2015) extend this significance by adding that a city in Turkey with lowest student-teacher ratio ranked higher in performance in examinations. On the influence of income on student achievement, this study finds income of parent or guardian under IV estimation to be more significantly related to student performance at 5 percent level than the 10 percent level under the OLS estimation. This result is supported by Farooq et al. (2011), Mbugua et al. (2012) and Melack (2014) who report income of parent to be a significant determinant of student academic success.

Teaching and learning materials are also a significant factor in the performance of students. The study finds the ratio of textbooks to students to be significantly related to student examination composite scores. The study results affirm Mbugua et al. (2012) who attribute poor performance of secondary school students in mathematics in school certificate examinations to inadequate availability of teaching and learning materials. Similarly, teaching and learning materials are reported to have a significant influence on the performance of students in Sumbawanga and Mbeya districts of Tanzania (Melack, 2014; Mlozi et al., 2013). This study further collaborates with Melack (2014) on the relationship of presence of school library with student performance. The study finds that presence of school library was positively and significantly associated with student scores. The other controls that were significantly related to student performance were ratio of students to classrooms (1 percent level) and ratio of students to desks. However, the student-desk ratio, though statistically significant at 1 percent level, did not carry the expected sign. The size of family a student came from not only was statistically insignificant, but also carried an opposite sign from what was expected.

The conflicting findings with previous studies on some of the study variables show a need for caution in extending results from the analysis of attendance, school location and student gender. Different study environments may lead to different findings and conclusions about the interaction of these study variables. The unrelated results could partly be a reflection of the different policy options on education. Reporting of some parallel results on the other hand, show that some of the findings from this study on school location, gender and attendance may be replicated in similar settings.

CHAPTER SEVEN

CONCLUSION AND STUDY RECOMMENDATIONS

7.0 Conclusion and Policy Recommendations

The study was based on three objectives. In the first objective, the study aimed at determining the magnitude of class attendance effect on student performance in public day secondary schools. In the second objective, the study sought to ascertain if there was a significant difference in the performance of students between students in rural and urban day secondary schools. The third objective of the study was to ascertain if there was a difference in performance between male and female students. The study utilised cross-sectional data that were sourced from three sources. Student data on final exam scores were obtained from the Examinations Council of Zambia. Data on student attendance, socio-economic status, family size, ratio of textbooks to students, ratio of students to classrooms, ratio of students to classroom furniture, presence of school library and ratio of students to grade-specific qualified teachers were obtained from school administrative records and was cross-checked with records from DEBS office.

Instrumental Variables (IV) regression was run on the same data to determine the magnitude, direction and significance of the influence of student class attendance on student performance. T-tests were also run on the same data to analyse the differences in student performance by school location and gender. Results from Instrumental Variables (IV) regression analysis show the significance of attending class. The findings indicate that increasing student class attendance by a day translates into a student's average

composite score increasing by 0.670. This finding is suggestive of both the magnitude and the unidirectional relationship of class attendance effect on student scores.

Data analysis also revealed that there was a statistically significant mean difference of 35.82 in student scores between students in urban schools and those in rural schools with t and p-value of 10.86 and 0.000 respectively indicating that students in urban schools performed better than those in rural schools. Similarly, a mean difference of 28.38 between male and females with the t and p-values of 8.80 and 0.000 respectively indicates that males outperformed females in the exams. Compared to males, females in the rural-sub sample performed poorly possibly on account of other school factors other than class attendance. This argument is based on the observed attendance patterns between males and females. In rural schools the mean attendance difference between males and females was statistically insignificant compared to their urban counterparts.

From the study findings, the following recommendations are made:

i. Government and communities need to work together to build more secondary schools to reduce inter school distances to facilitate daily class attendance.
ii. Policy makers should consider providing all schools with enough resource allocations for educational materials and libraries
iii. Government and school authorities should sensitise communities to support the education of girls and invest more resources to make schools more “girl friendly” to sustain their class attendance and daily endurance.

7.1 Recommendation for Future Research

For further proof of class attendance effect on student performance in day secondary schools, this study can be replicated on day secondary schools in other districts in order to provide further informed evidence that will be relevant for policy actions in Zambia.

REFERENCES

Abdullahi, H.A., Mlozi, M.R.S. and Nzalayaimisi, G.K., 2015. Determinants of students' academic achievement in agricultural sciences : A case study of secondary schools in Katsina State , Nigeria. *African Educational Research Journal*, 3(1),pp.80-88.

Andrietti, V., Addazio, R.D. and Velasco, C., 2008. Class Attendance and Academic Performance among Spanish Economics Students. *Working Paper 09-61*, Economic S (October 2008).

Arulampalam, W., Naylor, R.A. and Smith, J., 2012. Am I missing something? The effects of absence from class on student performance. *Economics of Education Review*, 31(4), pp.363–375.

Aucejo, E.M. and Romano, T.F., 2014. Assessing the effect of school days and absences on test score performance. *CEE Discussion Paper*, (1302), pp.1–33.

Bosede, A.F. and Enuloju, A.A., 2013. Rural and Urban Differential in Students' Academic Performance amoung Secondary School Students in Ondo State, Nigeria. *Journal of Educational and Social Research.*

Bratti, M. and Staffolani, S., 2002. Student Time Allocation and Educational Production Functions. *SSRN Electronic Journal*, (July), p.27. Available at: http://papers.ssrn

Cameron,C.A. and Trivedi, K.P.,2005.*Microeconometrics:Methods and Applications*. Cambridge, UK: Cambridge University Press.

Caviglia-harris, J.L., 2006. Attendance and Achievement in Economics: Investigating the Impact of Attendance Policies and Absentee Rates on Student Performance . *Journal of Economics and Finance Education*, 4(2),pp.1-15.

Central Statistical Office, 2012. *Living Conditions Monitoring Survey Report 2006 and 2010*. CSO: Republic of Zambia.

Dolton, P., Marcenaro, O.D. and Navarro, L., 2001. The effective use of student time : a stochastic frontier production function case study.*Centre for the Economics of Education*, pp.547–560.

Examinations Council of Zambia, 2013. *Research Study on Learner Absenteeism from Public Examinations: An Inquiry into the Extent and Causes of Absenteeism at the Primary and Junior Secondary School Level Examinations*.Lusaka: Examinations Council of Zambia.

ED*ASSIST, 2010. *Education Automated Statistical Software Information Toolkit*. Lusaka: Ministry of Education.

ED*ASSIST, 2014. *Education Automated Statistical Software Information Toolkit*. Lusaka: Ministry of Education.

Educating our Future, 1996. *National Policy on Education*. Lusaka: Zambia Educational Publishing House.

Education Reform, 1977. *Proposals and Recommendations*. Lusaka: Government Printer.

Education for All, 2008. *Education International's Response to the Global Monitoring Report*. Education International.

Ella, E.R., Odok,O.A. and Ell, E.G., 2015. Influence of Family Size and Family Type on Academic Performance of Students in Government in Calabar Municipality, Cross River State , Nigeria. *International Journal of Humanities Social Sciences and Education*, 2(11), pp.108–114.

Erdogan, Y., Bayram, S. and Levent, D., 2008. Factors That Influence Academic Achievement And Attitudes in Web Based Education. *International Journal of Instruction*, 1(1), pp.31–47.

Farooq, M.S.,Chaudhry,H.A.,Shafiq,M.and Berhanu,G., 2011. Factors Affecting Students' Quality of Academic Performance: A Case of Secondary School Level. *Journal of Quality and Technology Management*, 7(2), pp.1–14.

Fergusson, D.M. and Horwood, L.J., 1997. Gender Differences in Educational Achievement in a New Zealand Birth Cohort. *New Zealand Journal of Educational Studies*, 32(1), pp.83–96.

Focus on Learning, 1992. *Strategies for Development of School Education in Zambia: Report of the Team Appointed to Review Investment Strategies in Education*. Lusaka: Ministry of Education.

Gottfried, M.A., 2009. Evaluating the Relationship between Student Attendance and Achievement in Urban Elementary and Middle Schools: An Instrumental Variables Approach. *American Educational Research Journal*, 20(10), pp.1–32.

Government of the Republic of Zambia, 2000. *National Gender Policy*: *Gender in Development, Division Office of the President* . Lusaka: Cabinet Office.

Gujarati, D.N., 2004. *Basic Econometrics*. 4 th ed. Mac Graw-Hill Companies.

International Bank for Reconstruction and Development. and World Bank, 2014. *World Development Indicators*. Worshington DC.

Idowu, A.R. and Oluwole, A.B., 2014. Effects of Student-Teacher Ratio on Academic Achievement of Selected Secondary School Students in Port Harcourt Metropolis, Nigeria. *Journal of Education and Practice.* 5(24), pp.100–107.

Igbudu, U., 2015. Influence of Gender on Students' Academic Achievement in Government Subject in Public Secondary Schools in Oredo Local Government Area of Edo State, Nigeria. *Journal of Educational and Social Research*, 5(2), pp.101–106. Available at: http://www.mcser.org/journal/index.php/jesr/article/view/6563.

Khwaileh, M.F. and Zaza, I.H., 2011. Gender Difference in Academic Performance among Undergraduates at the University of Jordan: Are they real or Stereo typing?. *College Student journal.* 45(3).

Mbugua, Z.K., Kibet, K., Muthaa,G.M. and Nkonke, G.R., 2012. Factors Contributing To Students' Poor Performance in Mathematics at Kenya Certificate of Secondary Education in Kenya: A Case of Baringo County, Kenya. *Aijcrnet.Com*, 2(6), pp.87–91. Available at: http://aijcrnet.com/journals/Vol_2_No_6_June_2012/11.pdf.

Melack, D.N., 2014. *Determinants of Poor Academic Performance of Secondary School Students in Sumbawanga District, Tanzania* (Doctoral Dissertation, Sokoine University of Agriculture).

Ministry of Education, 2005a. *Annual School Census 2005*.Lusaka: Ministry of Education.

Ministry of Education, 2005b. *Educational Statistical Bulletin 2005*. Lusaka: Ministry of Education.

Ministry of Education, 2006. *Educational Statistical Bulletin 2006*. Lusaka: Ministry of Education.

Ministry of Education, 2008a. *Annual School Census 2008*. Lusaka: Ministry of Education.

Ministry of Education, 2008b. *Zambia National Assessment Survey Report 2008: Learning Achievement at the Middle Basic Level*. Lusaka, Zambia : Printech (Z) Ltd.

Ministry of Education, 2010. *Infrastructure Operational Development Plan*. Lusaka: Ministry of Education.

Ministry of Finance, 2006. A prosperous Middle-Income Nation by 2030: Lusaka: Ministry of Finance.

Ministry of Finance, 2008. *Zambia Millennium Development Goals Progress Report*. Lusaka: Ministry of Finance.

Minisry of Finance, 2014. *The Revised Sixth National Development Plan 2013-2016: People Centred Economic Growth and Development*: Lusaka: Minisry of Finance.

Ministry of General Education, 2014. *Zambia Educational Statistical Bulletin 2014*. Lusaka: Ministry of General Education.

Ministry of General Education, 2015. *Kafue 2015-2018 Learner Performance District Strategic Plan*. Kafue : Kafue District Education Office.

Ministry of General Education, 2016. *Annual School Census 2016*. Lusaka: Ministry of General Education.

Mlozi, S.R., Kaguo, E.F. and Nyamba, Y.S., 2013. Factors Influencing Academic Performance in Community and Government Built Secondary Schools in Tanzania: A Case of Mbeya Manucipality. *International Journal of Science and Technology*,2(2), pp.174–186.

Moshoeshoe, R., 2015. Average and Heterogeneous Effects of Class Size on Educational Achievement in Lesotho. *ERSA working paper*, 496(February). Available at: http://www.econrsa.org/system/files/publications/working_papers/working_paper_496.pdf.

Musau, L.M. and Abere, M.J., 2015. Teacher qualification and students ' academic performance in science mathematics and technology subjects in Kenya. *International Journal of Education Administration and Policy Studies*, 7(3), pp.83–89.

Musole, K.K., 2015. *Determinants of Healthcare Seeking Patterns and Out-of- Pocket Expenditure for Malaria Treatment in Zambia* (Masters Dissertation, University of Zambia. Lusaka).

Nizamettin, K. and Bekir, C., 2015. The Impact of Number of Students per Teacher on Student Achievement. *Procedia - Social and Behavioral Sciences*, 177(July 2014), pp.65–70.Available at: http://www.sciencedirect.com/science/article/pii/S1877042815016894.

Oghuvbu, E.P., 2010. Attendance and Academic Performance of Students in Secondary Schools : A Correlational Approach. *Stud Home Comm*, 4(1), pp.21–25.

Olatude, Y.P., 2011. School Location and Academic Achievement of Secondary Schools in Ekiti State, Nigeria. *Asian Social Science*, 7(5), pp.170–175.

Olusola, R. and Omotade, A., 2014. Impact of School Location on Academic Achievement of Science Students in Senior Secondary School Certificate Examination. *International Journal of Scientific and Research Publications*, 4(9), pp.4–7.

Ramos, R., Duque, J. and Nieto, S., 2012. Decomposing the rural-urban differential in student achievement in Colombia using PISA microdata. *IZA Discussion Paper*, (6515), pp.1–25. Available at: http://www.econstor.eu/handle/10419/58718.

Tenibiaje, D.J., 2009. Influence of Family Size and Family Birth Order on Academic Performance of Adolescents in Higher Institutions. *Pakistan Journal of Social Sciences*, 6(3), pp.110–114.

United Nations Economic Social Council, 2011. *Imperative for Quality Education for All in Africa:Ensuring Equity and Enhancing Teaching Quality*. Lome, Togo:UNICEF and UNESCO.

United States Aid for International Development, 2008. *Opportunity to Learn : A high Impact Strategy for Improving Educational Outcomes in Developing Countries*. Washington, D.C.

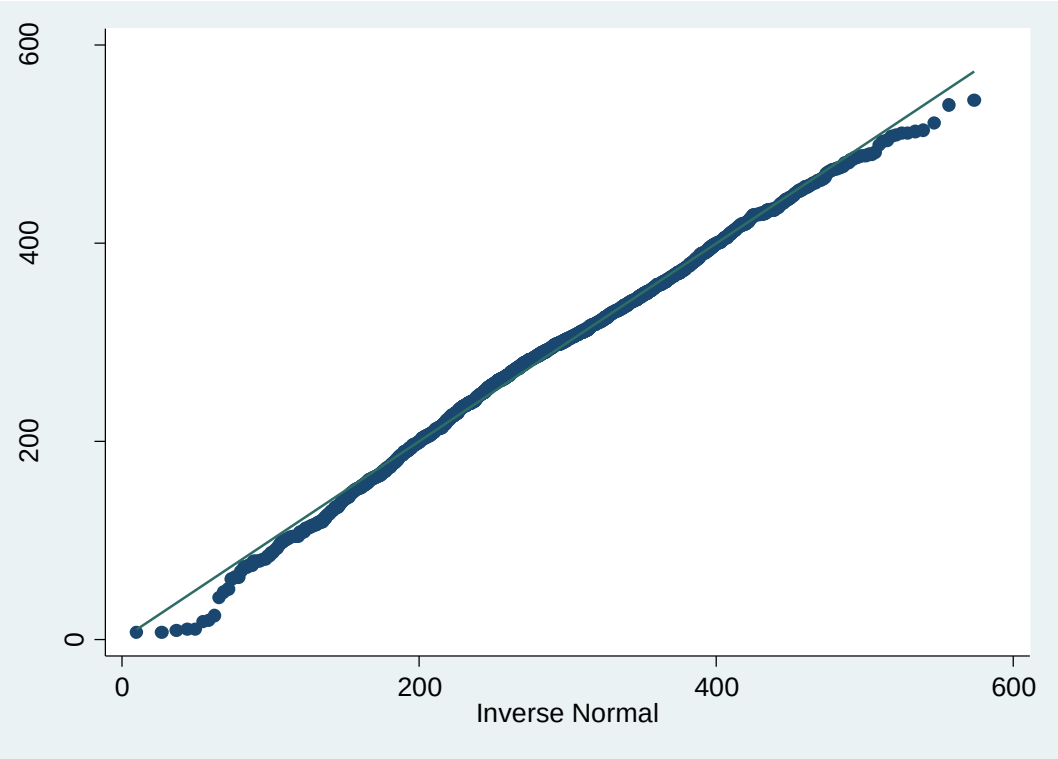
Usain, M.I.and Abubakar,N., 2015. The Influence of School Environment on Academic Performance of Secondary school Students in Kuala Terrenganu, Malaysia . *Proceedings of ICI International Conference on Empowering Islamic Civilisation in the 21st Centuary*.

Wooldridge, J.M., 2002. *Econometric Analysis of Cross Section and Panel Data*. London,England: The MIT Press.

World Bicycle Relief, 2011. *Mobility Equals Education: Bicycles for Educational Empowerment Programme, 2011 Report*. www.worldbicyclerelief.org. Accessed 6, May,2016.

APPENDICES

Appendix 1: Quantile distribution of scores against quantile normal distribution



Appendix 2: Variance-inflating factors

Variable	VIF	1/VIF
Student-classroom ratio	3.76	0.266
Student-desk ratio	3.37	0.297
Student-teacher ratio	2.29	0.436
Book-student ratio	1.91	0.525
Presence of school library	1.33	0.752
Age	1.26	0.792
Attendance	1.13	0.883
Income dummy	1.08	0.923
Gender	1.06	0.947
Family size	1.05	0.955
Mean VIF	1.82	

Appendix 3: Correlation of attendance with instruments

N= 2579	Attendance	Distance	Orphan hood
Attendance	1.0000		
Distance	-0.4629***	1.0000	
Orphan hood	-0.0554***	0.0259	1.0000

Significance levels: * 10 percent level, ** 5 percent level; *** 1 percent level,

Appendix 4: First-stage regression results for IV regression

Variable	OLS
Distance (Instrument)	-5.950***
	(0.367)
Orphan (Instrument)	-0.294
	(0.956)
Gender (1=Female;0=Male)	1.902*
	(0.795)
Age	6.767
	(5.077)
Age squared	-0.167
	(0.152)
Number of students per grade-specific teacher	-0.589***
	(0.068)
Number of text books per student	-0.707
	(0.837)
Number of students per class room	-0.143***
	(0.043)
Parental employment status (1 if parent works; 0 = otherwise)	2.262**
	(0.790)
Size of family a student comes from	-0.207
	(0.196)
Number of students per desk	4.262***
	(0.857)
Presence of school library (1 if school has library, 0 otherwise)	-1.354
	(1.063)
Constant	115.082**
	(42.810)
R^2	0.29

Significance levels: * 10 percent level, ** 5 percent level; *** 1 percent level, robust standard errors are in parentheses

Appendix 5A: Marginal effects of attendance between male and female students and between urban and rural students

Score= β_0 +(λ_1 x gender) + $\beta_1$attendance +($\lambda_1$attendance x gender) + $\beta_2$stud-tr ratio + ($\lambda_2$stud-tr ratio x gender) + $\beta_3$stud-class ratio + ($\lambda_3$stud-class ratio x gender) + $\beta_4$stud-desk ratio + ($\lambda_4$stud-desk ratio x gender) + $\beta_5$age + ($\lambda_5$age x gender) + $\beta_6$agesq + ($\lambda_6$agesq x gender) + $\beta_7$bk-stud ratio + ($\lambda_7$bk-stud ratio x gender)+ $\beta_8$income + ($\lambda_8$income x gender) + $\beta_9$family + ($\lambda_9$family x gender) + β_{10}library + (λ_{10}library x gender)

Taking derivatives, we find that the marginal effect of class attendance on scores with respect to gender is given by:

$$\frac{\partial Score}{\partial Attendance} = \beta 1 + \lambda 1 gender$$

Similarly, replacing gender with school location in the above regression and taking partial derivatives, we find that the marginal effect of class attendance on scores with respect to geographical location is given by:

$$\frac{\partial Score}{\partial Attendance} = \beta 1 + \lambda 1 location$$

STATA outputs for four separate regressions for each category are displayed in the next appendices.

Appendix 5B: IV results of marginal effects of attendance on scores for females

Variable	IV
Class Attendance	0.832***
	(0.218)
Age	-131.407***
	(32.186)
Age squared	3.842***
	(0.986)
Number of students per grade-specific teacher	-1.553***
	(0.458)
Number of text books per student	-4.765
	(5.116)
Number of students per class room	0.822**
	(0.342)
Parental employment status (1 if parent works; 0 = otherwise)	-4.371
	(4.361)
Size of family a student comes from	0.352
	(1.048)
Number of students per desk	-11.759*
	(6.617)
Presence of school library (1 if school has library, 0 otherwise)	-0.977
	(6.094)
Constant	1264.468***
	(267.084)
R^2	0.16

Significance levels: * 10 percent level, ** 5 percent level; *** 1 percent level, robust standard errors are in parentheses

Appendix 5C: IV results of marginal effects of attendance on scores for males

Variable	IV
Class Attendance	0.490*
	(0.251)
Age	-22.590
	(23.354)
Age squared	0.493
	(0.701)
Number of students per grade-specific teacher	-0.262
	(0.352)
Number of text books per student	26.275***
	(3.903)
Number of students per class room	-0.958***
	(0.241)
Parental employment status (1 if parent works; 0 = otherwise)	13.120***
	(4.338)
Size of family a student comes from	1.339
	(1.019)
Number of students per desk	22.513***
	(5.066)
Presence of school library (1 if school has library, 0 otherwise)	20.508***
	(4.988)
Constant	420.317**
	(195.591)
R^2	0.27

Significance levels: * 10 percent level, ** 5 percent level; *** 1 percent level, robust standard errors are in parentheses

Appendix 5D: IV results of marginal effects of attendance on scores for rural students

Variable	IV
Class Attendance	1.005***
	(0.212)
Age	-32.421
	(22.076)
Age squared	1.058
	(0.655)
Number of students per grade-specific teacher	-1.161*
	(0.601)
Number of text books per student	-4.560
	(3.561)
Number of students per class room	0.970**
	(0.419)
Parental employment status (1 if parent works; 0 = otherwise)	6.137
	(5.767)
Size of family a student comes from	1.318
	(1.195)
Number of students per desk	-30.117***
	(6.401)
Presence of school library (1 if school has library, 0 otherwise)	-2.170
	(6.937)
Gender (1=female;0 otherwise)	-9.947**
	(5.059)
Constant	380.945**
	(194.215)
R^2	0.14

Significance levels: * 10 percent level, ** 5 percent level; *** 1 percent level, robust standard errors are in parentheses

Appendix 5E: IV results of marginal effects of attendance on scores for urban students

Variable	IV
Class Attendance	0.392
	(0.304)
Age	-36.038
	(31.858)
Age squared	0.865
	(0.995)
Number of students per grade-specific teacher	-0.071*
	(0.424)
Number of text books per student	51.247***
	(6.443)
Number of students per class room	-0.273
	(0.310)
Parental employment status (1 if parent works; 0 = otherwise)	-8.337**
	(3.785)
Size of family a student comes from	0 .920
	(0.889)
Number of students per desk	-2.144
	(7.189)
Presence of school library (1 if school has library, 0 otherwise)	-13.969
	(10.609)
Gender (1=female;0 otherwise)	-17.516***
	(3.714)
Constant	606.606**
	(258.835)
R^2	0.32

Significance levels: * 10 percent level, ** 5 percent level; *** 1 percent level, robust standard errors are in parentheses