DETERMINANTS OF LEARNER PERFOMANCE IN A COMBINED SCHOOL IN MPUMALANGA PROVINCE: EDUCATION PRODUCTION FUNCTION APPROACH

by

ZAKHELE CEDRICK SIBIYA

Dissertation submitted in fulfilment of the requirements for the degree of

MASTER OF COMMERCE IN ECONOMICS

in

FACULTY OF MANAGEMENT & LAW

(School of Economics and Management)

at the

UNIVERSITY OF LIMPOPO

SUPERVISOR: PROF T. NCANYWA

2019

DECLARATION

I declare that **determinants of learner performance in a combined school in Mpumalanga province: Education production function approach** is my own work and that all the sources that I have used or quoted have been indicated and acknowledged by means of complete references and that this work has not been submitted before for any other degree at any other institution.

……

…….

Signature Date

ACKNOWLEDGEMENTS

I wish to express my sincere gratitude to:

- My supervisor, Dr Ncanywa, for her wonderful support, encouragement and wisdom in guiding me in writing and completing this dissertation.
- Prof. Mongale for his words of advice and guidance.
- My Heavenly Father, for being with me all the way, and for granting me wisdom and courage, strength and perseverance to complete this study.
- My dearest parents, who loved education so much that they always empowered me with the words of wisdom that education is the key to success.
- My family at large, particularly my wife Itumeleng Mashabela, who has been my pillar of strength; and my children, for being patient with me.

ABSTRACT

This study examined the determinants of learner performance by employing an education production function approach using the descriptive statistics, ordinary least squares (OLS) and quantile regression techniques in 2016. The study utilised the data obtained from SA-SAMS of Bankfontein combined school at Mpumalanga province. In the education production function, learner performance was estimated against variables such as age, gender, days absent and socio-economic status.

The results of this study indicated that in the rural combined school, learner performance is strongly influenced by age, absenteeism and socio economic status. For instance, results revealed that absenteeism had a negative effect on learners' educational performance. An increase in absenteeism by 1 day led to a reduction in learner's examination score by approximately 0.1 percentage points during the chosen period. The "socioeconomic status" variable revealed a statistically significant and negative impact on learners' educational performance.

The results demonstrate that poverty leads to poor educational performance as measured by examination scores. It is recommended that schools should manage learner diversity (age, gender and socio-economic factors), introduce learner motivation programmes, teacher performance improvement interventions, and improve organisational planning and development, parental involvement among others to retain learners at school. Furthermore, schools should enforce education policies that stipulate entry and exit age at different levels of schooling.

KEY CONCEPTS: Quantile Regression, Learner performance, SA-SAMS, Production Function, Ordinary least Squares, Socio Economic Status

JEL CODES: I21, I31

TABLE OF CONTENTS

DECLARATION ii

ACKNOWLEDGEMENTS iii

ABSTRACT iv

LIST OF FIGURES ix

LIST OF TABLES x

LIST OF APPENDICES xi

ACRONYMS xii

CHAPTER 1 1

ORIENTATION TO THE STUDY 1

1.1 Introduction and background 1

1.1.1 South African Educational System 2

1.1.2 Educator Quality 5

1.1.3 Overview of Bankfontein combined school 8

1.2 Statement of the problem 9

1.3 Research aim and objectives 10

1.3.1 Aim of the study 10

1.3.2 Objectives of the study 10

1.4 Research questions 10

1.5 Definition of concepts 10

1.6 Ethical considerations 11

1.7 Significance of the study 11

1.8 Structure of the of the study 11

CHAPTER 2 14

LITERATURE REVIEW 14

2.1 Introduction 14
2.2 Theoretical Framework 14
2.2.1 The education production function approach 14
2.2.2 The background of the production function 16
2.2.3 The integration of education into the production function 18
2.3 Empirical literature 20
2.3.1 Review of input factors trend of learner performance 20
2.3.2 Input-output relationship of learner performance 31
2.4 Learner perfomance in South African schools 36
2.5 Summary 39
CHAPTER 3 41
RESEARCH METHODOLOGY 41
3.1 Introduction 41
3.2 Data 41
3.3 Model specification 42
3.4 Estimation techniques 43
3.4.1 Descriptive Statistics 44
3.4.2 Regression estimates 48
3.4.2.1 Ordinary Least Squares (OLS) 48
3.4.2.2 Quantile Regressions 49
3.5 Summary 51
CHAPTER 4 52
DISCUSSION / PRESENTATION / INTERPRETATION OF FINDINGS 52
4.1 Introduction 52
4.2 Descriptive statistics results 52
4.2.1 General Education and Training (GET) Descriptive Statistics Results 52

4.2.2. Further Education and Training (FET) Descriptive Statistics Results 54
4.3 Regression Estimates Results.....55
4.3.1 GET learners' Ordinary Least Squares (OLS) Regression Results.....55
4.3.2 FET learners' Ordinary Least Squares (OLS) Regression Results.....57
4.4. Quantile regression estimates results.....59
4.4.1 GET Learners' Quantile Regression Results.....59
4.4.2 FET Learners' Quantile Regression Results.....63
4.5 Summary.....68
CHAPTER 5.....70
SUMMARY, CONCLUSION, RECOMMENDATIONS.....70
5.1 Summary and Interpretation of Findings.....70
5.2 Conclusions.....72
5.3 Recommendations of the study 73
5.4 Limitations of the study.....76
LIST OF REFERENCES.....80
APPENDICES.....88
Appendix A.....88
Appendix B.....103
Appendix C.....104
Appendix D.....106

LIST OF FIGURES

Figure 2.1: Distribution of the logarithmic pupil-teacher ratio and the logarithmic expenditure per pupil for Africans who are enrolled……..…23

LIST OF TABLES

Table 2.1: Learner: Educator Ratio…….……21

Table 2.2: Pass rates between 2007- 2017…….……..38

Table 2.3: South African education statistics according to school wealth…….……39

Table 3.1: Correlation coefficient…..…….……..……45

Table 4.1: Descriptive statistics for GET learners, 2016……..53

Table 4.2: Pairwise correlations of GET leaners...……53

Table 4.3: Descriptive statistics of FET learners……...……54

Table 4.4: Pairwise correlations of FET leaners……55

Table 4.5: GET learners' OLS educational performance model..……...…56

Table 4.6: FET learners' OLS educational performance model……..57

Table 4.7: GET learners' Quantile educational performance model……..…59

Table 4.7.1: GET learners' 10% or 0.1 quantile educational performance model……..61

Table 4.7.2: GET learners' 25% or 0.25 quantile educational performance model……61

Table 4.7.3: GET learners' 75% or 0.75 quantile educational performance model……62

Table 4.7.4: GET learners' 95% or 0.95 quantile educational performance model……63

Table 4.8: FET quantile regression educational performance model……..…64

Table 4.8.1: FET learners' 10% or 0.1 quantile educational performance model……...65

Table 4.8.2: FET learners' 25% or 0.25 quantile educational performance model……..66

Table 4.8.3: FET learners' 75% or 0.75 quantile educational performance model……..67

Table 4.8.4: FET learners' 95% or 0.95 quantile educational performance model……68

LIST OF APPENDICES

APPENDIX A: Spread Sheet of Data

APPENDIX B: Data Coding

APPENDIX C: Results

APPENDIX D: School Picture

ACRONYMS

ACE: Advanced Certificate in Education

ANA: Annual National Assessments

ANOVA: Analysis of Variance

B Ed: Bachelor of Education

CDE: Centre for Development and Enterprise

DBE: Department of Basic Education

DEPS: District of Colombia Public Schools

DHET: Department of Higher Education and Training

DoE: Department of Education

ECD: Early Childhood Development

EFA: Education For All

EMIS: Education Management Information systems

FDE: Further Diploma in Education

FET: Further Education and Training

GDP: Gross Domestic Product

GET: General Education AND Training

GPA: Grade Point Average

HDE: Higher Diploma in Education

HLM: Hierarchical Linear Model

LTSM: Learner-Teacher Support Material

MEC: Member of the Executive Council

NAEP: National Assessment of Educational Progress

NNSSF: National Norms and Standards for School Funding

NQF: National Qualification Framework

OLS: Ordinary least squares

PGCE: Post Graduate Certificate in Education

PPM: Post Provisional Model

REQV: Relative Education Qualification Value

SACE: South African Council of Educators

SACMEQ: Southern and Eastern African Consortium for Monitoring Education Quality

SAQA: South African Qualification Authority

SA-SAMS: South African School Administration Management System

SASS: Schools and Staffing Surveys

SAT: Scholastic Aptitude Test

SEM: Structural Equation Model

SES: Socioeconomic Status

SGB: School Governing Body

SRN: School Register of Needs

STAR: Student Teacher Achievement Ratio

TIMSS: Trends in International Mathematics and Science Study

TVAAS: Tennessee Value-Added Assessment System

CHAPTER 1

ORIENTATION TO THE STUDY

1.1 Introduction and background

In the post-apartheid era South African government has increased efforts to improve the state of education in the country. Considering all sorts of spending by the government, matriculation results still seem to be a disappointing component among all government spectrums. According to the report released by the Department of Education, there are priority areas outlined to be improved over the period up to 2014 (DoE, 2010). This plan was part of the school realization of year 2025 government objective. The report further outlined the aim of improving learner performances by overcoming the weaknesses in the educational system. This is to be done particularly by improving access to education, providing safe buildings, providing learning and teaching materials promptly and by improving educator quality (DoE, 2010).

The challenges of the education system were suggested to be addressed by medium-term expenditure framework. For example, according to National Treasury (2013), South Africa spent R227 billion which was about 19.7 per cent of total government expenditure on education, equivalent to 6.5 per cent of GDP. Also, National Treasury (2010) illustrates the components of South African education and their allocated shares of funding for the period 2010/11 to 2016/17. Primary and secondary school education, administered by provinces, received the largest share of the total education budget about 57.7 per cent in 2013/14. The ration of spending visa versa schooling outcomes does not compare favourably with other developing countries. According to DoE (2015), Mpumalanga Education MEC Reginah Mhaule says that performance in lower grades needs to improve in order to ensure better matriculation results and had ensured that Mpumalanga reached its target of a 70 percent matriculation pass rate in 2012. Thereafter, the province had set a target of 80 percent for 2013 which it has never achieved but falling to 77 percent (DoE, 2015).

Mpumalanga's matriculation results have improved steadily over the years but not in accordance to the allocated budget. According to DoE (2012), the province obtained a 47.9 percent matriculation pass rate in 2009, 56.8 percent in 2010 and 64.8 percent pass rate in 2011 where of the 47 889 matriculates who wrote the 2012 examinations, 33 504 passed of these, 9 495 qualified for bachelor programmes, 14 277 for diplomas, 9 633 for higher certificates and 99 achieved the National Senior Certificate. The leading education district was Ehlanzeni District, which produced a pass rate of 74 percent, followed by Nkangala at 73 percent, Gert-Sibande at 69 percent and Bohlabela at 62.5 percent. The Bohlabela education district has recorded the highest improvement even though is not proportionately to the budget and attention it has received over the period, from a dismal 28.8 percent in 2009 yet the budget and skilled educators were already existing in the district.

1.1.1 South African Educational System

Education in South Africa is governed by two national departments, namely the department of Basic Education (DBE), which is responsible for primary and secondary schools, and the department of Higher Education and Training (DHET), which is responsible for tertiary education and vocational training. Prior to 2009, these two departments were represented in a single Department of Education. The Department of Basic Education deals with public schools, private schools (also referred to by the department as independent schools), early childhood development (ECD) centres, and special needs schools. The public schools and private schools are collectively known as ordinary schools, and comprise large percentage of schools in South Africa. The nine provinces in South Africa also have their own education departments that are responsible for implementing the policies of the national department, as well as dealing with local issues (DoE, 2015).

The Department of basic Education in South Africa has grouped the schooling system into primary, secondary and combined schools (DoE, 2015). Combined schools include both primary and secondary schools. Schools with grades R to 9 are under the General Education and Training (GET) band and those with grades 10 to 12 are under the Further Education and Training (FET) band. Policies, rules and regulation of governing schools are contained in the South African Schools Act,

which guides the school governing bodies. The ideal schooling system should take place in a well-equipped building, have a well-thought curriculum in terms of scope and sequence, set platform for the culture of teaching and learning, provide adequate material inputs and provide highly qualified and well-paid educators (DoE, 2015).

According to DoE (2014) the South African basic education system comprised 12 644 208 learners, 30 586 schools, and 439 394 teachers in 2014. The department of Basic Education is headed by the director-general Bobby Soobrayan, and its policy is made by the minister Angie Motshekga and the deputy minister Enver Surty by 2018. The South African Schools Act (1996) recognises two categories of school: public and independent. Public schools are state controlled and independent schools are privately governed. Section 29(3) of the Bill of Rights contained in Chapter 2 of the Constitution of the Republic of South Africa, Act 108 of 1997 states clearly the intention of the state to make it possible for independent educational institutions to exist. Furthermore, to provide for their possible subsidisation by the state but those subsidies are not meant to be made available to private education institutions for profit (Khumalo, 2014).

According to the statistics from DoE (2013) there were 1 584 independent schools with 513 804 learners being taught by 33 194 educators. 37.4 percent of these schools are in Gauteng which also has the highest number of learners (229 984) and educators (15 466). This statistic shows that the learner: educator ratios at independent schools are considerably lower than at public schools ranging from 11:1 in the Western Cape to 22:1 in Limpopo. The smaller classes are made possible by the relatively high school fees from which teacher salaries at independent schools that typically do not receive a government subsidy are paid. In these schools, all school related costs, such as learner and teacher support materials and maintenance of infrastructure, are paid by parents and sometimes private sponsors.

The Minister determines a school as a no fee school by publishing its name in the Government Gazette, currently all schools ranked in quintiles 1, 2 and 3 should be no fee schools. According to DoE (2000), School fees are determined by the parents

at a general meeting of parents at schools, currently quintile 4 and 5 schools are allowed to charge school fees.

An independent study by Masondo (2016) found that undue union influence and critical educational factors, including weak institutional functionality, uneducated teachers, and insufficient learning time, were responsible for the poor educational state of South Africa. This means that there are many shortfalls in the South African educational system and it prevents the country from achieving its goals, and need to be addressed urgently. Some of the areas within education that require special attention include; educator training and development; timely provision of learning and teaching support materials (LTSM); and school infrastructure at all levels (DoE, 2014).

In the Annual National Assessment (2013) the matric examinations in South Africa are coordinated and monitored by Umalusi, to solve problems with external examinations, proposals have been accepted to extend external examinations into the GET band. Given that in schooling there is likely to be a very limited extent to which learners want to exit the system after grade 9, on the focus for quality assurance should be of teacher assessment and an inspectorate model, conducted by provincial departments, which Umalusi could monitor. However, standardized testing, if thought of as part of systemic evaluation of the system, and not in relation to the achievements of individual students, it could be used as a tool with which to ascertain levels of learning at specific grades (DoE, 2012).

Viewing the current educational funding system, Fataar (2010) argues that in view of this analysis, an understanding of education policy requires that the multidimensional process and dynamics that shape education must be analysed. In this case the quintile school funding system reinforced by equity and redress needs to be analysed. Christie (2008) suggests that in order to address the inequalities in education, a provision was made in the National Norms and Standards for School Funding of 1999 (NNSSF) to pay special attention to poor schools. According to the NNSSF policy of 1999, schools were to be categorised according to quintiles using the target list of all the schools in the province. Schools were categorised from

National Quintile 1 for the poorest, to National Quintile 5 for the least poor. Quintile 3 was categorised as medium and quintiles 4 and 5 were categorised as least poor. Motala (2006) noted the following as the criteria according to which the quintile system categorised schools:

- Poverty levels of the community in which the school is situated.
- Literacy levels of the community
- Geographical location of the school

1.1.2 Educator Quality

It has been reported that out dated teaching practices and lack of basic content knowledge have resulted in poor teaching standards (Makgato and Mji, 2006). The poor standards have also been exacerbated by a large number of under-qualified or unqualified teachers who teach in overcrowded and non-equipped classrooms. The Education For All (EFA) 2000 assessment (2005) also reported that, in spite of approximately 85 percent of mathematics educators, being professionally qualified only 50 percent have specialized in mathematics in their training. Similarly, with 84 percent of science educators professionally qualified, only 42 percent are qualified in science.

The DoE (2000) introduced a new policy which presents the new Norm and standards for educators, the cornerstone of the new policy is the seven roles for educators and their associated applied competence that should be integrated into the purpose and exit level outcomes of the qualification but it is difficult to practice those roles in South Africa because of the nature of the public rural school which makes it to be almost impossible.

Educators are the most valuable asset in the schooling system and the variables associated with educator quality that is often found in literature as follows (Hanushek, 2008). Content knowledge: This is associated with the major subject of the educator, the educator that teaches his or her major subject is likely to be effective on student achievements and it is a consistently strong predictor of student performance even though some studies differ on how strong is the effect. Goldhaber

and Brewer (1996) found that the presence of teachers with at least a major in their subject area was the most reliable predictor of student achievement scores in math and science. Rivkin et al. (2005) and Ferguson and Ladd (1996) found out that teachers with graduate-level training in a content area performed better than did teachers with no degree in their content area. Educators, on the other hand, admitted to shortcomings they had with respect to certain sections of the content they were teaching.

Teaching experience: Harris and Sass (2007) found out generally positive, but mixed, evidence on the effects of experience and little or no evidence of the efficiency of advanced degrees for teachers and that the first few years of experience substantially increase the productivity of elementary and middle school teachers but have little impact on the effectiveness of teachers at the high school level. Harris and Sass (2007) further argued that experience enhances teacher productivity at all grade levels in reading and in both elementary and middle-school math, though experience effects decline as the study progress from elementary to middle and high school.

Teacher training and credentials: This measure is in a form of formal qualification obtained by the educator. According to DoE (2010), educator's qualifications are classified in terms of M+3 or M+4 known as REQV13 or REQV14 respectively. However, there are educators who are not professionally qualified in the educational system who have studied other qualifications besides teaching such as BSc, BComm, BA and those who are under qualified meaning REQV 10 to 12. Harris and Sass (2007) argued that at the middle and high school levels there is evidence that prior professional development training has positive effects on the productivity of Math educators. These positive effects are primarily due to increase exposure to content-focused training but the other types of in-service coursework, such as pedagogical training, are not found to enhance educator productivity.

According to DoE (2000) the Certificate in Education (120 credits) and the Diploma in Education (240 credits) are entry and exit points on the Bachelor of Education (480 credits) path, a teacher with either or both of these two qualifications (placing them on REQV 11 or 12) will not be regarded as professionally qualified. To be

registered with SACE as a professionally qualified educator a minimum of REQV 13 is required or 360 SAQA credits at level 5 or above. DoE (2000) states that the Diploma in Education comprises a total of 240 credits which includes the 120 credits of the Certificate in Education.

The DoE (2000) further argued that first Bachelor's degrees should include sufficient credits in appropriate subjects so that the teacher will be competent in his/her chosen specialisation. After a Bachelor of Education degree or a first Bachelor's degree and a Post-Graduate Certificate in Education, the teacher could further his/her studies on either a horizontal level by enrolling for an Advanced Certificate in Education or vertically by completing a Bachelor of Education (Honours) degree (DoE 2000). The Advanced Certificate in Education is a newly introduced qualification on level 6 that has replaced the FDEs (which are in many cases offered on level 5). The ACE is intended to replace the FDE and the HDE and may be used for up-grading or further training in a specialisation or for re-training.

Entry into the Advanced Certificate in Education may be vertically from a three-year diploma in education (REQV 13) and equivalent qualifications or horizontally from a PGCE or B Ed or from an NQF level 7 or 8 qualification. In other words, the ACE will follow either a general formative appropriate degree (B A, B Comm or B Sc) together with by a PGCE (which will replace the current HDE post-graduate) or it will follow a new 480 credit B Ed (DoE, 2000). Existing educators who are in possession of a three-year college diploma may also be admitted to the new level 6 ACE. A student may not advance from a 480 credit Bachelor of Education degree to a 240 credit Master's degree. He/she could however enrol for a 360 credit Master's degree programme or a 120 B Ed (Hons) programme followed by a 240 credit Master's degree programme.

The new B Ed (Honours) will replace the existing post-graduate B Ed. The new B Ed (Honours) may be designed with a particular focus on the advanced development of either academic or professional or occupational competence. DoE (2000) states that, it is not necessary to award the new B Ed (Hons) retrospectively to past B Ed graduates. The new 480 credit B Ed in the Norms and Standards will be endorsed to the effect to indicate that it is a first or initial professional teaching qualification, such

as Foundation Phase, or Further Education and Training Phase: Mathematics. That will also help to distinguish between the old and the new B Ed qualification (DoE, 2000).

In order to achieve educational progress South Africa needs an institutional structure that promotes good teaching and that attracts and retains the best teachers (Van der Berg, Taylor, Gustafsson, Spaull, Armstrong, 2011). The institutions should encompass teacher pay, bursary programmes and other interventions targeting existing teachers. Effective schools require well selected individuals as principals together with management teams that understand and fulfil their roles as leaders of the curriculum, ensuring that an organised environment conducive to learning is present.

1.1.3 Overview of Bankfontein combined school

Bankfontein combined school is situated at Nkangala District in Mpumalanga province. It is a rural based school which was built by a farmer around 1960s however; it is classified as a public school and falling into quantile one category. The school absorbs the capacity of about 1000 learners per year but employed only 27 educators. The national poor performance of the South African schools is also true in Bankfontein combined school as they achieved only 58 percent 2014, 60 percent in 2015 and 73 percent in 2016. The school is relatively poorly resourced and the building structures are well presented in annexure D.

Crouch and Mabogoane (2001) and Spaull (2013) report that out of all the factors in South Africa that have an impact on learner's achievement, socioeconomic ones, especially poverty, may be the most important. Contextual constraints affecting learner performance in Bankfontein are complex, intertwined and often structural. These include a lack of facilities and resources, large class sizes, inadequate teacher education, poor learner commitment and discipline, inadequate parental involvement, to name but a few. Most of these derive from the country's apartheid history and the very high levels of inequality in society.

It is common cause that the quality of the South African schooling system as a whole is poor and that levels of literacy and numeracy are dismally low. Learner achievement in mathematics and science, while slowly improving, is still at an unacceptably low level. In 2016 the Bankfontein teacher learner ration was 1:55 which is inconsistent with the policies of the department of basic education of 1:30. The small results improvement of the school is solely helped by the extra effort which the matric educators are putting in the form of extra classes on weekends and holidays.

1.2 Statement of the problem

Since 1994, there has been a disappointing schooling outcomes (matric pass rates) in South Africa, this problem has been witnessed at national level and also true in the Mpumalanga province (DoE, 2015). However, there are several programs and propositions that have been documented and deployed in the past which were intended to improve the devastating schooling outcomes. Also, even though there has been a massive budget of R204 billion allocated to the Department of Basic Education no positive results were yielded up to this far especially in subjects such as Mathematics, Accounting and Physical Sciences compared to small budget of other developing countries. Van Der Berg and Low (2006) argued that even the steep improvement in South African test scores observed for the upper most quintiles is not particularly impressive in regional context. Van Der Berg and Low (2006) concluded that learners in Tanzania and the Seychelles achieve better results compared to South Africans. According to National Treasury (2016) 5, 94 percentage of GDP was allocated to the department of education in 2016.

In the 2003 Trends in International Mathematics and Science Study (TIMMS), the grade 8 Mathematics scores for both former black and former white schools were below the international average. In fact, South Africa's score was the lowest of all 53 participating countries; even the other five much poorer African countries in the study outscored South Africa (Taylor, Muller and Vinjevold, 2003). The TIMSS results of 2015 show that South Africa was ranked second last out of 48 countries for Grade 4 mathematics, second last for Grade 8 mathematics and stone last for Grade 8 science out of 38 countries. According to Gustaffson (2007) policy attention in South

Africa has been increasingly focusing on what is by now clear evidence of unacceptably low levels of learner performance across the bulk of historically disadvantaged schools, but even, by international standards, poor quality in historically advantaged schools. Hence, it was interesting to determine factors that can positively influence learner results and ultimately influence schooling outcomes.

1.3 Research aim and objectives

1.3.1 Aim of the study

The aim of the study is to analyse the determinants of the learner performance at Bankfontein Combined School in Mpumalanga province in the period 2016.

1.3.2 Objectives of the study

The objectives of the study are as follows:

- To find out which input factors can influence learner performance in Bankfontein Combined School.
- To examine the relationship between education input factors and learner performance.
- To formalize policy recommendations based on the research findings.

1.4 Research questions

- Which input factors can influence learner performance in Bankfontein combined school?
- What is the relationship between education input factors and learner Performance?

1.5 Definition of concepts

SA-SAMS: according to South African school administration and management system is a computer application database meant to assist in administration, management and governance in schools (DoE, 2012).

EDUCATION PRODUCTION FUNCION: is the input-output relationship between school and student inputs and a measure of school output (Hanushek, 2007).

QUANTILE REGRESSIONS: is an estimation technique which aims at estimating either the conditional median or other quantiles of the response variable (Koenker, and Hallock, 2001).

REQV: stands for Relative Education Qualification Value where a person must have a qualification of at least three years after the senior certificate (M+3), but also the qualification must include professional training as an educator (DoE, 2000).

LEARNER PERFOMANCE: is an output or an outcome in terms of learner's assessment results (Van Der Berg and Low 2006).

1.6 Ethical considerations

The study uses secondary data from SA-SAMS of the Department of Education in a school in Mpumalanga province. The data will be utilized with integrity and be kept confidential, and information of other authors will be acknowledged properly. The information utilized will be for academic purposes only.

1.7 Significance of the study

The study is aimed at focusing on all secondary grades (8 – 12) to evaluate significance of factors contributing to the pass rate at Bankfontein Combined School in Mpumalanga Province under the Nkangala district. The Mpumalanga Education MEC believes that performance in lower grades needs to improve in order to ensure better matric results (DoE, 2015). The level of concentration on lower grades by the government is very minimal as compared to grade 12 or matric. This great focus on matric is resulted by the fact that in South Africa, matric pass rate is a standardized and reliable measure of schooling outcomes. However, the position of this study is to reveal that matric is only one of the components of the schooling outcomes, in fact; lower grades are also an important factor in the schooling outcomes. Bankfontein was chosen to represent the rural based schools with the lack of resources because

it has been underperforming over the last years and it is surrounded by the low income community, this is aimed at providing a clear picture on the factors affecting the performance of poor schools.

As it is mentioned above that the focus is much on matric also the previous studies have never been including much lower grades and be school specific in analyzing the factors contributing in school outcomes as it has been emphasized that all grades are equally important in the analyses of schooling output. The results of this study will inform policy makers on areas that can be improved so as to obtain better schooling outcomes.

The study adds to the existing body of knowledge concerning the causes of and ways of dealing with poor performance among learners in South Africa. Schools experiencing similar problems are thus able to craft appropriate intervention programmes basing on this study's research finding. The research also contributes to provide insights into; learner performance improvement strategies at GET and FET levels; the reasons why some schools fall short of achieving acceptable performance targets and; ways of ensuring that learning outcomes improve. Another contribution of the study is that it provides ways by which the quality of teaching can be improved through teacher intervention strategies. With better quality teaching, the process of learning and its outcomes are bound to improve.

According to Crouch and Mabogoane (1998) over the past few years, South African Department of Education (DoE) attempts to transform its educational system. This is due to the fact that perhaps most of the intellectual attention and bureaucratic effort in the systems area has gone to the redistribution or allocation of quantitative resources. Resources such as educator re-deployment, based on a quantitative norms relating to learner: teacher ratios, emphasis on redistribution of material and financial resources, building schools in areas which previously have not been well-provided, and focuses on the establishment of the education budgets in dialogues between national parties, inside provincial administrations, and between provincial and national line and fiscal departments.

A large amount of time and emotional energy is devoted to this above mentioned issues but results sometimes are better in poor schools than in schools with resources. This suggests that there may be a poor management to those resources

which opens a room for investigating the causes of that outcome; however, this study is aimed at focusing on the poor schools and investigates its determinants of learner performance. The topic is considered as a very pertinent topic given the low learner performance in the education sector, although this study is limited to one school, the results of this study provide insights not only about the reasons of poor performance but also about the problems faced by teachers and learners. The results of this study suggest possible solutions to make teaching and learning more meaningful. The study further aims at suggesting ways to eliminate problems which continue to detract from the performance of Matric.

1.8 Structure of the dissertation

The dissertation is prearranged as follows: chapter 2 reviews the theoretical framework and empirical literature review.

Chapter 3 presents the methodology of the study where ordinary least squares and quantile regressions are employed.

Chapter 4 discusses the research findings from the econometrics tests performed in the study.

Chapter 5 presents the conclusion and policy recommendation of the study.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

This chapter presents a review of relevant literature and factors affecting academic Performance. In South Africa, there are several factors that have contributed to the learner's performance over the years. Given the imbalances of the past coursed by skewed distribution of resources, location and background of the learners, a range of factors can be identified. While these factors are identified and discussed, it is imperative to review the theoretical and empirical literature with regards to education production functions. Therefore, the first section of this chapter focuses on the theoretical underpinnings of the study. The second section focuses on the empirical literature pertaining to education production function studies. Since this study is school specific it is also essential to review provincial education departments which are responsible for implementation to clearly outline the present educational landscape, budget and expenditure trends, service delivery achievements and the medium-term outlook.

2.2 Theoretical Framework

This section discusses the production function approach and its efficiency as applied by different authors. The section also considers how educational output could be best examined efficiently using production function.

2.2.1 The education production function approach

A production function is a process where there is a combination of inputs that are transformed into output. According to Gustaffson (2007) there is considerable and probably growing interest in the education production function approach. This may be caused by the fact that education authorities are often interested in ranking

schools on the basis of performance or school outcome. It has been alluded that educational systems have no single defined production function, and no well-defined indicators of input and output, in fact most studies of the education production function the measure of input and output are limited by the availability of data (Ncanywa, 2014). Therefore, various educational outcomes can result from a variety of different combinations of inputs. Ncanywa (2014) further argued that the measurement of education production functions assumes that the output of the educational process, which is achievement of individual students, is related to a series of inputs.

This function can be represented as follows

Y_i= f (X_{1i}…… …..X_{ni} ,A, Z) (1)

Y, represents learner performance at the end of the year examination; X represents educator quality such as teacher level of satisfaction, teacher salary, teacher qualification (REQV), teacher experience and teacher's age; A represent school characteristics that can be measured by Learner: teacher Ratios, quintiles and school condition. Lastly, the function must view the effects of the education resources as conditional on the level of socio-economic status (SES) of learners (Z). This is because of the clearly proven positive correlation between SES and performance, independent of the education resources available (Hanushek, 2003). Typically, Z will include measures of both material welfare and the level of education of the learner's parents. The level of the analysis, or i, may be the individual learner or the school, and in this study it refers to the learner (Gustaffson, 2007). Most analyses of education production functions have directed their attention at a relatively small set of resource measures, and this makes it easy to summarize the results (Hanushek, 2003).

2.2.2 The background of the production function

Tan (2008) explains that Cobb-Douglas functional form of production function is widely used to represent the relationship of an input-output relationship. It was proposed by Knut Wicksell (1851 - 1926), and tested against statistical evidence by Charles Cobb and Paul Douglas in 1928. In 1928 Charles Cobb and Paul Douglas published a joint paper in which they modelled the growth of the American economy during the period 1899 – 1922. They considered a simplified view of the economy in which production output is determined by the amount of labour involved and the amount of capital invested (Humphrey, 1997).

Dewy (2000) argued that an education production function is an application of the economic concept of a production function to the field of education. It relates various inputs affecting a learning process such as school resources, families, peers, neighbour hoods, expenditures among others. Measured outputs including subsequent labour market success, college attendance, graduation rates, and, most frequently, standardized test scores. Coleman, Campbell, Hobson, McPartland, Mood, Weinfeld, and York (1966) draw a lot of interest by applying production function and published a report which concluded that the marginal effect of various school inputs on student achievement was small compared to the impact of families and friends. During that period the use of the education production function was not popular, and then soon after Coleman Report many economists (Hanushek, 1979; Krueger, 1999; Angrist, and Lavy, 1999) chose to apply production functions to answer educational question in such a way that the 90 individual publications that appeared before 1995 contain 377 separate production function estimates (Hanushek, 2008). However, Hanushek (2008) further reported a very high correlation between adjusted growth rate and adjusted test scores after introducing the structure of production to the consideration of student learning outcomes.

Ncanywa (2014); Kimani and Bhorat (2014); Bhorat and Oosthuizen (2006) are among a large number of successive studies published in South Africa today on educational production functions, as increasingly involving economists, produced consistent results about the impact of school resources on student performance, leading to significant policy discussions. The interpretation of the various studies has

been very controversial, in part because the findings have a direct influence on policy debates and decision making in the educational systems. One separate line of study have been conducted by Ncanywa (2014) using production function approach and found that, the educator characteristics significantly explained that schooling outcomes based on more educator experience and higher educator qualifications had a positive influence on learner performance and are also determinants of educator payments. Ncanywa (2014) further submitted to policy makers that educators, as an important human resource, need to be properly trained and well paid to be attracted to the teaching profession and to be retained in the education system, and these could also help in attracting educators to teach in remote rural villages of the Eastern Cape and could entice students to choose teaching as a career.

Kimani and Bhorat (2014) observed that the overall policy question of whether added funds to schools are likely to produce higher achievement has entered into legislative debates and court consideration of school finance systems in many countries. However, this study by Kimani and Bhorat (2014) further provides answers after examining the relationship between school inputs such as pupil: teacher ratio, expenditure per pupil and educational attainment. The study employed a reduced form production function, applying a partial generalized ordered probity which allows identifying heterogeneous effects of the controls at different levels of education. The government has increased its spending in basic education mainly through an increase in government employed teachers in an effort to reduce the pupil: teacher ratio, and through an increase in expenditure per pupil allocated by school quintile in a bid to attain equality of resources at school. It is found that both pupil-teacher ratio and expenditure per pupil have strong and significant effects on educational attainment of Africans in South Africa. These inputs have higher effects on attainment of lower education levels. Kimani and Bhorat (2014) is a case study of South Africa which means these findings can only be relevant in countries with similar conditions to South Africa.

2.2.3 The integration of education into the production function

Shephard (1970) claimed that if Cobb-Dougla's production function suggests that, for example, production of goods and services involves transforming resources (inputs) such as labour power, raw materials, and the services provided by facilities and machines into finished products. This implies the production function is a mathematical representation of the various technological recipes from which an organisation can choose to configure its production process. In particular, the production function tells us the maximum quantity of output to be produced given the quantities of the inputs that it might employ (Caves and Barton, 1990).

The production function could be written this way:

$$Q = f(L, K) \quad (2)$$

Where Q is the quantity of output, L is the quantity of labour used, and K is the quantity of capital employed. This expression tells us that the maximum quantity of output the organisation can get depends on the quantities of labour and capital it employs.

However, this study is examining the impact of different variables in the schooling outcome using the very same production function approach; it is obvious that the inputs in this case are going to be different from the above mentioned ones to measure the schooling output. As Hanushek (2008) argued most applied inputs in the past that affect schooling outcome are firstly; family background which usually characterized by such socio-demographic characteristics as parental education, income, and family size. Secondly; peer inputs, when included, are typically aggregates of student socio-demographic characteristics or achievement for a school or classroom. Thirdly; school inputs typically include teacher background (education level, experience, sex, race, and so forth), school organization (class sizes, facilities, administrative expenditures, and so forth), and district or community factors (for example, average expenditure levels).

Then education production function will be presented as follows as explained in Equation 1 with the educational variables as inputs in this case.

Y_i= f (X_{1i}…… …..X_{ni} ,A, Z) (3)

Hansen (1970) defined an educational production function as the relationship between school student inputs and a measure of school output. The study emphasized that this method of representing the educational production process is likely to be of particular value both in descriptive studies of human capital formation and in normative investigations to determine optimal educational resource allocation. According to Hansen (1970) a simple production model is mostly applied in the analysis of the education economics. The common inputs are variables like school resources, teacher quality, and family attributes, and the outcome is usually student achievement. For example, Coates (1998) applied production function contributing to the literature on the effectiveness of schools. Coates (1998) argued that, the data employed include the minutes of instruction per five-day week in each of four subjects for all the public schools in Illinois, while very few education production function papers have any information on the amount of instruction students receive in a given subject. Therefore, this variable was unique but still applicable and relevant to the education production function.

Coates (1998) also included variables such a class size where it is argued that it should interact with learner, teacher, and school characteristics as well as with the instructional times. The Coates (1998) empirical analysis tests of hypothesis pointed out that the small marginal effects of class size or of teacher qualifications may result because the harm of larger class sizes is undone by better trained personnel, and the benefits of better trained teachers is undone by large class sizes. In Coates (1998) the data span three years enabling the use of panel data techniques. Individual school specific effects are estimated and found to be strongly significant. Moreover, in the presence of these effects, class size always has the theoretically predicted sign and is highly significant and the time series aspect of the data allows for tests of the comparability of test scores over time. The results reveal that year to year differences can be substantial and statistically significant.

Nye et al. (2004) conducted the study through the Tennessee Department of Education's Student Teacher Achievement Ratio (STAR) project which was an experiment aimed at evaluating the effects of smaller classes on student achievement over the period of four years. The experiment randomly assigned students from various racial and socioeconomic backgrounds to small and regular-size classes in 79 schools across the United States of America. STAR's reliance on randomized samples, combined with the data-tracking capacity of the Tennessee Value Added Assessment System (TVAAS), offered an important and unique opportunity to examine variations in student achievement where the only difference between classes was the teacher. With the application of the production function the analyses shows that teachers had a substantial effect on student achievement as it clearly outlines achievement gains associated with smaller class sizes, a stronger achievement gain is associated with educator quality.

2.3 Empirical literature

In this section of the study an empirical literature review relating to the different variables and their significant on student performances will be presented. This review is intended to show previous work in the area of educational efficiency in both developed and developing countries including South Africa.

2.3.1 Review of input factors trend of learner performance

2.3.1.1 School characteristics

DBE's 2010 statistics report (published in 2012), the departmental objective is to achieve a national average of 30 learners per teacher, for example if there are 480 learners per school, there should be 16 teachers per school. The ratio of learners per teacher is almost the same in all provinces, but the ratio of learners per school varies per province. For example, in Gauteng there are 800 learners per school and 28 teachers per school, whereas in the Eastern Cape there are 350 learners per school and 12 teachers per school.

DoE (2013) released a report that public school enrolments have been steadily rising over the past years. In 2013 there were 11 975 844 learners enrolled in 24 136 public ordinary schools, being taught by 391 829 educators. There was an average of 496 learners per school. Gauteng and the Western Cape, which are primarily urban provinces, had the highest learner: school ratios with 924 and 690 respectively. The learner: school ratios for mostly rural provinces such as Eastern Cape (338) and Limpopo (424) were lower, mainly due to the higher number of smaller schools in these provinces

Minister of Basic education Engy Musthega during parliamentary questioning argued that the Department has, as a strategic objective, to reduce the class size at schools. Funding was secured in 2009 and each Provincial Department of Education would have received their equitable share of this funding. This funding provides support to targeted schools in the form of additional posts. Through the reduction of class size, the learner-educator ratio will decrease. It indicates that some provinces have already achieved the 30:1 ratio. Nationally, the country is at 30.4:1 on average. Table 2.1 below shows that some provinces are even better off at a ratio below 30:1.

Table 2.1: Learner: Educator Ratio

Province	learner: teacher ratio
Eastern Cape	29.1:1
Free State	27.1:1
Gauteng	31.4:1
KwaZulu-Natal	31.2:1
Limpopo	30.1:1
Mpumalanga	31.1:1
Northern Cape	31.8:1
North West	30.6:1
Western Cape	30.6:1
South Africa	30.4:1

Source: Department of Education (2012)

South African average number of learners to educators shown in table 2.1 above was 30.4:1 includes educators employed by the state as well as those employed by school governing bodies (SGBs). If educators employed by SGBs are excluded from the calculation, the average learner: educator ratio increases to 37:1 in the Western Cape and 36:1 in Gauteng. This indicates that there is significant number of educators at schools employed by SGBs and their inclusion in the system has a noteworthy impact on the learner: educator ratio.

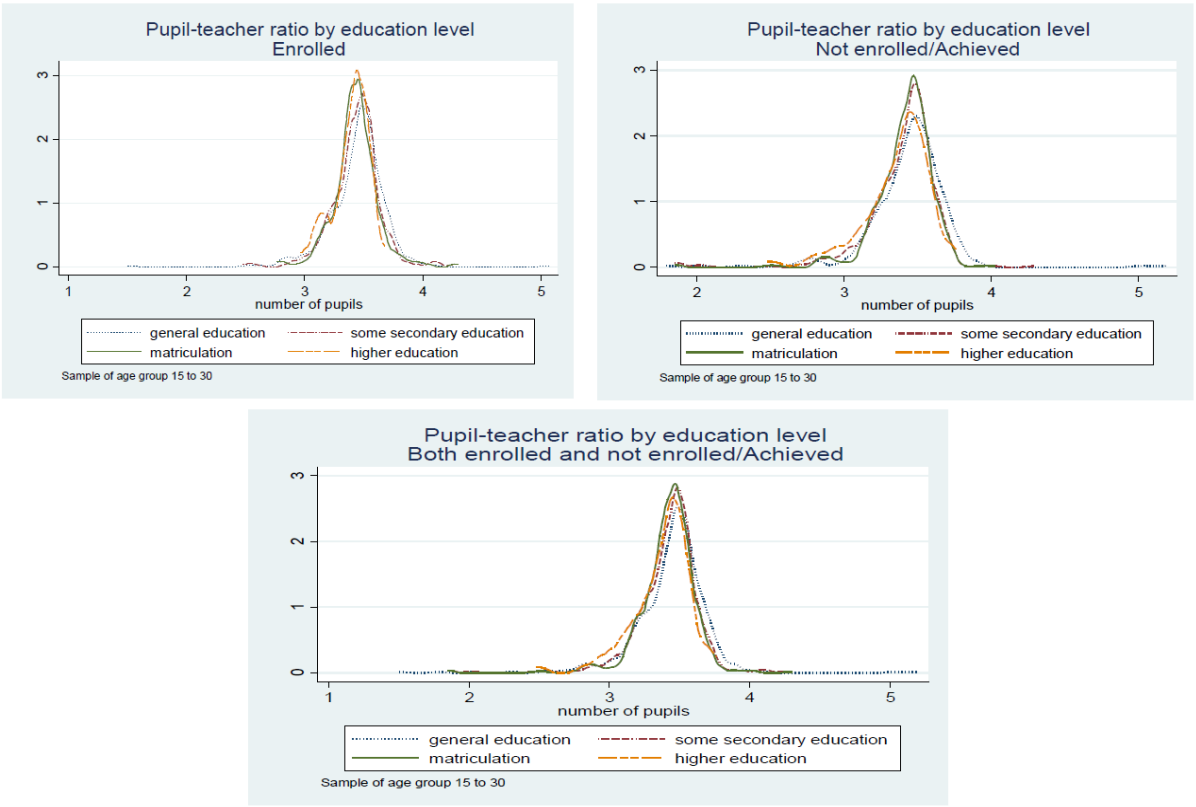
In 1998 post provisioning norms (Government Notice 1676 of 1998) were developed, and became fully implemented in 2000. This represented the country's first attempt at applying an equitable policy to the distribution of publicly funded educator posts across public schools and technical colleges. Regulations made in terms of this Act determine that an MEC must create a pool of posts in accordance with funds available for this purpose, after which the relevant head of department must distribute these posts among schools in accordance with the post-provisioning model (PPM).

Parliament (2012) noted that distribution of posts among schools should be based on each school's relative needs in this regard and in order to determine the school's relative need for posts, in relation to that of other schools, the PPM attaches weightings to all learners based on their relative needs for teachers and, in doing so, determines a weighted learner enrolment for each school, irrespective whether it is primary or secondary. The current norms overrode the 1995 collective agreement. Furthermore, the post provisioning norms as a policy currently implemented do not specify timeframes to achieve a 30:1 learner-educator ratio. However, as part of the department's commitment to strive towards reducing class size, the department has already reached a national average of 30:1 (Parliament, 2012). The STAR data have been examined extensively by an internal team of researchers; this analysis has found that students in small classes tended to perform better than students in larger classes (Kruger, 1999).

Kimani and Bhorati (2014) argued that, in South Africa the apartheid era government was discriminatory in its provision of educational services and resources; it prioritized

White schools at the expense of Black schools. This meant that there were fewer and less qualified teachers and less (if any) expenditure allocated to schools meant for Blacks, the majority of the population. At the end of apartheid the teacher-learner ratio stood at between 1:20 and 1:30 for Whites and between 1:40 and 1:70 for African pupils. Given the legacy of apartheid that differentiated education provision by race, a pooled analysis of all population groups is likely to mask effects of school inputs on African's educational attainment, which has triggered Kimani and Bhorati, (2014) to conduct a study and presented empirical evidence with the estimation sample that made up of 3023 African individuals who are either not enrolled (1352) or enrolled in institutions of learning (1671).

Figure 2.1: Distribution of the logarithmic pupil-teacher ratio and the logarithmic expenditure per pupil for Africans who are enrolled



Source: Kimani and Bhorat, (2014)

Kimani and Bhorat, (2014) applied descriptive statistics and interpreted figure 2.1 as follows: in both enrolled and achieved samples, on the one hand primary education is more skewed to the right, has a lower peak, and has longer tails than the other levels. On the other hand, higher education is skewed to the left, has the highest

peak for those enrolled, and has the shortest tail. A naive interpretation of these distributions is that a lower pupil-teacher ratio is associated with enrolling or attaining a higher education while a higher ratio is associated with enrolling or attaining a primary education. Kimani and Bhorat, (2014) also conduct a Kolmogorov-Smirnov test to find out if individuals who are enrolled at institutions of learning, and those who have achieved a given level of education are from a population with identical distribution functions. The results for pupil-teacher ratio show that the individuals who are enrolled and individuals who have achieved a primary, some secondary, a matriculation, or a higher education are from a population with identical distribution functions. This means that at all levels of education there is no significant difference between the ratio for those enrolled and for those who have achieved.

2.3.1.2 Gender

Billy-Jackman et al. (2011) report that gender variable is among factors affecting schooling outcomes into the educational system. The analysis around this variable aims to provide data to inform the development of a comprehensive gender policy intended to ensure gender equity in access to education and the attenuation of gender gaps in educational outcomes. The study used the different statistical analyses such as quantile regression, frequency distributions, scale reliability, One-way analysis of variance (ANOVAs), chi square tests of independence, and Pearson product-moment correlation coefficients to test the gender and school achievement.

Billy-Jackman et al. (2011) revealed that more females were enrolled in secondary schools than males. However, males still dominated classes in mathematics, computer technology, and technical vocational studies while the test scores data revealed that males generally underachieve in Mathematics and English which are basic entry level requirements for post-secondary and tertiary level enrolment. Males appear to be underachieving in traditional male-dominated vocational subjects compared to females. It was concluded that even though male under-perform in the important subject but the male matric pass rates relatively do not differ significantly from those of females.

Ghazvini and Khajehpour's (2011) study examines gender differences existing in various cognitive motivational variables (locus of control, academic self-concept and use of learning strategies) and in performance attained in school subjects of Language and Mathematics. For this purpose, a sample of 363 students was selected from the high school. Results show the existence of gender difference in variables under consideration, with girls showing internal locus of control, using attitude, motivation, time management, anxiety, and self-testing strategies more extensively, and getting better marks in Language. With boys using concentration, information processing and selecting main ideas strategies more, and getting better marks in Mathematics. Ghazvini and Khajehpour (2011) concluded that differences exist in the cognitive-motivational functioning of boys and girls in the academic environment, with the girls have a more adaptive approach to learning tasks hence achieve better results on average. However, the influence of contextual variables that may differently affect boys' and girls' motivation was not taken into account. But, Walla (2015) is presenting a different view in his study where the results showed a significant effect of age in reading performance, with the older children having better scores than younger ones for reading fluency, reading comprehension, and the total reading performance while gender was not found to play an important role in reading performance.

Lavy (2004) study the importance of gender stereotypes in the evaluation of student outcomes at upper-secondary level in Israel. Using a natural experiment which allows for comparisons between a gender-blind test score and a non-gender- blind score, he finds, contrary to his expectations, that the gender bias is in favor of girls. Girls have systematically higher scores on the non-blind test compared to the blind test, but there seems to be a small systematic difference based on teacher's gender in the discrimination against boys. The gender bias in favour of girls is larger among male math teachers than female math teachers, but the opposite is true for physics, biology and chemistry. Holmlund and Sund (2005) investigate whether the gender performance gap can be attributed to the fact that the teacher profession is female dominated, that is, is there a causal effect on student outcomes from having a same-sex teacher. An OLS regression supports the initial hypothesis that a same-sex teacher is positively associated with student grades and Holmlund and Sund (2005) concludes that there are a number of possible hypotheses that can explain such a

finding. First, teachers might have preferences over students of their own sex, and hence female (male) teachers will, given student performance, reward girls (boys) more highly in terms of grades and support. Eddy (2011) in his research found out that females have also been shown to have better language ability than males.

Although it is fair to argue that males and females are equal human beings, they demonstrate different features not just physically; which in most cases is rather obvious, but also mentally. They are said to perform differently in everyday activities, to think in different ways or sometimes even to transform a surface structure to different deep structures and to misunderstand each other in this way. As the way of thinking is closely related to the ability to comprehend; if the thinking of the two sexes differs, it is quite predictable that the ways they learn and outcomes will be different.

2.3.1.3 Age

Eddy (2011) found that there is a critical period for learning. This is supported by proponents of Critical Period Hypothesis, which states that human beings are optimally suited to learn certain types of behavior (including foreign/second language abilities) during a certain age span, and that after this period has passed, learning such behaviour is difficult or impossible.

Lightbown and Spada (1993) come with the following findings,

- Older learners are superior to children in rate of acquisition.
- Older children learn more rapidly than younger children
- With regards to morphology and syntax, the adolescents do best, followed by the adults and then the children
- Grammar differences diminish over time, and children begin to catch up, but older learners outperform children in the short term

Lightbown and Spada (1993) further suggest that older learners learn faster than children, and this is more applicable to grammar than pronunciation, although in the case of formal learning situations old learners seem to do better even in the pronunciation area. It is not clear when children start to catch up.

Concerning demographic characteristics, age and gender were found to have an effect on academic performance (Krause, 2005). Similarly, Clarke and Ramsey (1991) found age to correlate with academic performance and the general finding was that older learners perform better than younger learners. However, in some subjects such as Mathematics and Science, mature-age students were adversely affected because their learning speed tended to decrease with age, while the depth of learning increased.

2.3.1.4 Days Absent

Aucejo and Romano (2014) jointly estimate the effect of absences and length of the school calendar on test score performance using Simple Ordinal Least square, Quantile regression and descriptive statistics with administrative data from North Carolina public schools. The study exploits a state policy that provides variation in the number of days prior to standardized testing and find substantial differences between these effects. Aucejo and Romano (2014) revealed that extending the school calendar by ten days' increases math and reading test scores by only 0.8 percent and 0.2 percent of a standard deviation, respectively; a similar reduction in absences would lead to gains of 5.8 percent and 3 percent in math and reading. Performing a number of robustness checks including utilizing u data to instrument for absences, family-year fixed effects, separating excused and unexcused absences, and controlling for a contemporaneous measure of student disengagement. The results are robust to these alternative specifications. In addition, the findings indicate considerable heterogeneity across student ability, suggesting that targeting absenteeism among low performing students could aid in narrowing the gaps in performance.

Gaete-Romeoy (2017) regress students' academic outcomes on students' school absenteeism rate during secondary school. The study was based on the 2011 Chilean student strikes, led by university students but promptly joined by hundreds of thousands of secondary school students, triggered a major drop in public secondary school class attendance during that year (a decline of nearly 20 percent in all four grades). Attendance then returned to normal levels in 2012. The evidence shows

that the strikes led to persistent negative effects for public secondary school students' results in a high-stakes math exam (taken after completing secondary education) and university enrollment rates. Gaete-Romeoy (2017) estimates that, for each of the four cohorts of public secondary school students in 2011, scores in their math exam fell between 3.2 - 4.0 percent of a standard deviation and their associated university enrollment rates fell between 9.8 - 15.3 percent. In contrast, there is no significant effect on their performance in the high-stakes language exam. The results are neither driven by the sorting of students across schools or cohorts, nor by other factors such as disruptiveness at the time of the high-stakes exams, school environment, class size or teacher effects. Gaete-Romeoy (2017) used the type of school that students attended during the strikes as an instrument for school attendance. OLS and instrument variable (IV) point estimates suggest that a 10 percentage point decrease in attendance during secondary school is related to a 9.5 percent of a standard deviation reduction in the math exam score, and a 3.2 percentage point reduction in the associated probability of university enrollment. The IV point estimate suggests that a 10 percentage points decrease in the attendance rate during secondary school reduces the language exam score by 2.9 percent This effect is only significant at the 10 percent level.

Balkis, Erdinc and Arslan (2016) examine the direct and indirect relationship between student school absenteeism, personal factors (academic self- perception, attitudes towards teacher and school, goal valuation and motivation/ self-regulation), family factors (parents' educational level and income), and academic achievement in structural equation model (SEM) analyses. Four hundred and twenty-three high school students participated in the study. The findings revealed that student absenteeism was negatively related to academic self-perception, attitudes towards teacher and school, goal valuation, motivation/ self-regulation, and academic performance. Results also revealed that student absenteeism differed in respect to parents' educational level and income. Results from correlation and SEM analyses noticed that personal and family factors significantly predict previous and current student absenteeism. SEM analyses also revealed that previous student absenteeism significantly predicts previous academic achievement. Finally, SEM analyses noticed that previous student absenteeism and previous academic achievement can predict current student absenteeism. While Balkis et al. (2016)

major findings showed that students' absenteeism is negatively related with academic achievement. Klem and Connell, 2004 come to a same conclusion by suggesting that students who attend school regularly have higher academic achievement than students with high absences. This finding of this study is consistent with those studies indicating that absentees have low level of academic achievement.

2.3.1.5 Socio-economic Status (SES)

Spaull (2012) analysed the South African data from the most comprehensive measurement of educational performance across a number of African countries, the Southern and Eastern African Consortium for Monitoring Educational Quality (SACMEQ). This data was collected during the last quarter of 2007 from 9,083 grade six students and 1,488 grade six teachers in nearly 400 schools across South Africa. Schools were split into four wealth groups, with group 1 being the poorest 25 percent of schools. To determine the socioeconomic status of schools, students were asked possession questions whether items such as a daily newspaper, bed or piped water were present in the place they stayed at during the school week. The study found that pupils in South Africa's wealthiest 25 percent of schools outperform students in the remaining 75 percent of schools. Spaull (2012) wrote that there are two types of school systems in South Africa, largely split along historical-school-system and socioeconomic lines.

Van der Berg (2007) concludes that poor South African children are performing worse than equally poor children in the other African countries despite favourable characteristics in South Africa in terms of learner: teacher ratio. Van der Berg (2007) further demonstrates that SES has a strong influence on learner achievement in South Africa. The entire history of family background variables and school resources may contribute to student's achievement in a given year.

Spaull's (2012) conducted a research on grade 9 pupils in South Africa and presented similar results to the primary school data. Trends in International Mathematics and Science Study (TIMSS) are a cross-national study that measures mathematics and science achievement. The study tested 11,969 pupils in 285 South

African schools in 2010 and 2011. Of the 48 countries that participated in TIMSS, South Africa came 47th for mathematics and 48th for science. South Africa's Human Sciences Research Council separated participating schools into 5 groups, ranging from 1 (poorest) to 5 (least poor). Their analysis showed that the least poor 20 percent of schools significantly outperformed the remaining 80 percent of schools in both Science and Mathematics. There is support by researchers such as Van der Walt, Maree and Ellis (2008) and Ndlovu (2011) in their claim that some of the reasons for the poor performance of South African learners in mathematics include the poor socioeconomic background of learners which makes impossible to study at home, lack of appropriate learner support materials, general poverty of school environment, general poor quality of teachers and teaching including poor subject knowledge and poor motivation, language of instruction often not the same as learner's mother tongue and an inadequate study orientation.

Khumalo (2014) reports that the size of the grant paid by government to the school is determined largely by the poverty level of the neighbourhood in which the school is situated, as well as unemployment rate and general education rate of the population in that neighbourhood. Consequently, schools in more rich areas have to raise more money from other sources to maintain the high standard of education, and those very same schools often have so much additional income that their standard of education is much higher than that of poor schools anyway.

DoE, (2006) argued that poor parents are exempted from paying the school fees exemption depends on the income of the parents in relation to school fees. The DoE (2006) further argued that, if the school fees equal 10 percent of the joint income, parents qualify for full exemption, there is only partial exemption if fees are between 2.5 percent and 10 percent of income, also depending on the number of children the parents have attending a public school. The aim of the exemption program is to improve the chances for every South African child to access a quality education therefore It is a duty of the department to ensuring that no child is prevented from accessing a public school based on poverty or is in anyway discriminated against or has to face punitive measures imposed by a school in the case of non-payment of fees. If learners meet certain criteria, they should not have to pay school fees in any public school and the Department has a number of initiatives and programmes in

place to make it easier for learners to access quality education, these include No Fee Schools; Partial Fee Exemptions and Automatic Fee Exemptions (DoE, 2006).

Hedges and Greenwald (1996) suggested that family structures have deteriorated, more children are in single parent households, more are living in poverty, and more have less connection to books or well educated people. Consequently, the task of educating children has grown more difficult over time.

Baker and Jones (2005) conclude that there is a correlation between low socioeconomic status and learner's poor performance schools. However, some evidence has suggested that it is not the socioeconomic status per se but factors associated with home resources and background experiences that affect the learner's performance. According to Saiduddin (2003), factors such as unstable homes, drug abuse and teenage pregnancy contribute to poor performance among learners. Teenage pregnancy in South African schools is increasing with an increase rate; consequently, the learner tends to drop out of school, resulting in an on-going cycle of poverty in the home, community, province and country hence persistent low socioeconomic status.

2.3.2 Input-output relationship of learner performance

Kruger (1999) analyzed data on 11600 students using econometric analysis and Null Hypotheses of the only large-scale randomized experiment on class size ever conducted in the United States, the Tennessee Student/Teacher achievement Ratio experiment, known as Project STAR. This analysis has found that students in small classes tended to perform better than students in larger classes. Kruger (1999) applied the following statistical models to see the advantage of a randomized experiment in estimating the effect of school resources on student achievement,

$$Yij = aSij + bFij + \mathrm{e}\ ij \qquad (3)$$

Kruger (1999) analyzed the model assuming that *Yij* is the achievement level of student *i* in school *j*, *Sij* is a vector of school characteristics, *Fij* is a vector

representing the family background of the student, and e *ij* is a stochastic error component. In principle, *Sij* and *Fij* include information cumulated over the student's life; with the example of classroom size and teacher qualifications for each year the student attended school. A simple comparison of mean achievement between children in small and large classes provided an unbiased estimate of the effect of class size on achievement. Then the following regression equation for students in each grade level was used

$$Yics = \beta_0 + \beta_1 Small_{cs} + \beta_2 Reg/A_{cs} + \beta_3 X_{ics} + a_s + \varepsilon_{ics} \quad (4)$$

where *Yics* is the average percentile score on the SAT test of student *i* in class *c* at school *s*, *Smallcs* is a dummy variable indicating whether the student was assigned to a small class that year, *Reg*/*Acs* is a dummy variable indicating whether the student was assigned to a regular-size class with an aide that year, and *Xics* is a vector of observed student and teacher covariates a separate dummy variable is included for each school to absorb the school effects, a_s, and finally ε_{ics} is assumed to consist of two components: a class-specific random component that is common to all members of the same class, and an idiosyncratic error term. The equation was therefore estimated by ordinary least squares (OLS).

Kruger (1999) presented the regression results with the following conclusions (1) on average, performance on standardized tests increases by four percentile points the first year students attend small classes; (2) the test score advantage of students in small classes expands by about one percentile point per year in subsequent years; (3) teacher aides and measured teacher characteristics have little effect; (4) class size has a larger effect for minority students and those on free lunch; (5) Hawthorne effects were unlikely.

Van Der Berg and Low (2006) conducted a study to determine factors affecting South African student's performance. The study used regressions and hierarchical linear models and revealed that the relative SES advantage not translating into good test scores is echoed with respect to South African parent education levels. It has been found that children of parents with higher level of education perform better at school. Indeed, having better school resources (including higher teacher education)

do not appear to make the difference one might expect it would in South Africa. Finally, the study revealed widespread inadequacy in resource management in the dysfunctional (formerly black) part of South Africa's schooling system which suggests the strong need for policy attention at this end of the spectrum. Crouch and Mabogoane (1998) argued that of all the factors in South Africa that impact on mathematics and science achievement, socioeconomic ones, especially poverty, may be the most important.

Bhorat and Oosthuizen (2006) investigated the determinants of the grade twelve pass rates in the South African schooling system. In the study ordinary least squares and quantile regression techniques were applied to estimate an educational production function. The data set contained 5612 schools for which there was a mean grade 12 pass rate, a series of physical infrastructure, services, school and classroom-type characteristics, and household and community characteristics and to a very limited extent, indirect pupil and teacher characteristics. The results from the analysis indicated that the Learner: teacher ratio is insignificant, physical resources are irrelevant, infrastructure is crucial and teacher characteristics are a key priority.

Couch and Mabogoane (1998) carried out the research on the role of learning resources, social advantage, and education management in improving the performance of South African school's results. Correlation and regression analyses were used for analyses which required the data to be put up from matric examinations from different schools. Data was compared to the data from the Education Management Information System (EMIS) in order to get a good idea of the quality and quantity of personnel resources each school uses, and also to compare it with data from the School Register of Needs (SRN) to get a good idea of the physical condition of the school, and finally to compare it with a socio-economic database Commissioned by the National Department of Education. The results confirmed that one more year of REQV (years of training) amongst the educators is associated with a 16 point increase in the pass rate. This is really quite extraordinary and was considered most significant to the pass rate.

Spaull, (2012); Stols et a.l (2007); Taylor and Moyana, (2005) and Carnoy, 2008) have drawn attention to weak teacher content knowledge. Although the evidence is

accumulating, it is not clear what can be done about teacher content knowledge. The DBE's existing strategy of short in-service training courses does not seem to be particularly effective. Taylor (2011) concludes that short courses of the order of 3-5 days have little impact on a schooling outcome, therefore it is suggested that intensive in-service training, in the order of weeks per year, is required to equip existing teachers with the knowledge they need to teach effectively.

Gustaffson (2007) used ordinary least squares (OLS) and hierarchical linear production function models (HLM) and analysing the data from 2000 SACMEQ, for the country and for a sub-set of historically disadvantaged schools. In this study it is concluded allocation of time in school and school management, certain teaching methodologies, grade repetition approaches and infrastructure improvements are very important factors in schools' performance.

Hanushek (2007) determines achievement and the importance of the various inputs into student performance. The results show that, for classroom resources, only 9 percent of estimates for teacher education and 14 percent for teacher-pupil ratios yielded a positive and statistically significant relationship between these factors and student performance. However, the conclusion was drawn that the existing research suggests inefficiency in the provision of schooling. It does not indicate that schools do not matter nor does it indicate that money and resources never impact achievement. The accumulated research surrounding estimation of education production functions simply says there currently is no clear, systematic relationship between resources and student outcomes.

Hanushek (2008) further investigated the impact of expenditures on school performance used education production function and the experimental effects from STAR project. The study indicated that commonly purchased inputs to schools such as class size, teacher experience, and teacher education bear little systematic relationship to student outcomes, implying that conventional input policies are unlikely to improve achievement. In contrary to the findings of Hanushek (2008) Darling-Hammond (2000) used ordinary least square regression analyses to analyse the data from a 50 state survey of policies, state case study analyses, the 1993-94 Schools and Staffing Surveys (SASS), and the National Assessment of Educational

Progress (NAEP). This study examines the ways in which teacher qualifications and other school inputs are related to student achievement across the states. The findings pointed out that, there is a large number of variables associated with teacher quality appear to bear a significant relationship to student achievement and measures of teacher preparation and certification are by far the strongest correlates of student achievement in reading and mathematics.

Adnot et al. (2016) was holding a view that teacher turnover appears to have negative effects on school quality as measured by student performance. He, therefore, conducted a study evaluating the effects of teacher turnover on student achievement, the unique performance-assessment and incentive system in the District of Columbia Public Schools (DCPS). Employing a quasi-experimental design based on data from the first year, and finds that, on average DCPS replaced teachers who left with teachers who increased student achievement by 0.08 SD in math. When isolate the effects of lower-performing teachers who were induced to leave DCPS for poor performance, it was found that student achievement improves by larger and statistically significant amounts (that is 0.14 SD in reading and 0.21 SD in math) and finally the study concluded that the effect of exits by teachers not sanctioned is typically negative but not statistically significant.

Kane et al. (2006) applied the production function approach at USA to analyse the relationship between undergraduate grade point average (GPA) and teacher productivity in elementary and middle school. As with the other measures of undergraduate education, it was discovered that there is no significant relationship between GPA and subsequent teacher performance.

Harris and Sass (2007) carried out a study using ordinary least Square model (OLS) with unique panel data set of school administrative records from Florida (USA) that helped to overcome many of the challenges associated with measuring the impact of education and training on teacher productivity. The data covered all public school students throughout the state and included student-level achievement test data for both math and reading in each of grades 3-10 for the years 1999-2000 through 2004-2005. The results showed that experience enhances teacher productivity at all grade levels in reading and in both elementary and middle-school math, though

experience effects decline as the study progress from elementary to middle and high school. The positive effects of early career experience (1-2 years) in elementary and middle school found in the model without teacher effects remain largely unchanged; however, the accuracy is substantially reduced for experience effects above 5 years in elementary mathematics and middle school reading.

2.4 Learner performance in South African schools

Centre for Development and Enterprise (CDE) (2014) reported that the South African overall matriculation pass rate has been on the rise since 2011 reaching 78.2 percent in 2013. There has been an increase in the number of passes, from 109 697 in 2009 to 171 755 in 2013 after a decline in 2010, However, the percentage of candidates who achieve a university entrance pass remains low at 30.6 percent in 2013, up from only 20.1 percent in 2008. DoE (2015) the level of pass achieved by learners in mathematics has also increased between 2010 and 2013. The percentage of those candidates who passed mathematics at 40 percent has increased from 29.4 percent to 40.5 percent. This trend is encouraging, but the declining number of learners passing mathematics remains a great concern a decrease by more than half from just above 260 000 in 2010 to just above 100 000 in 2013 (CDE, 2014).This rise concern on a skills shortage in the country which tempers with economic growth.Van der Berg at al. (2011) added that South African learners have performed very poorly in international benchmark assessments such as the Progress in International Reading Literacy Study, Trends in International Mathematics and Science Study, and Southern and Eastern Africa Consortium for Monitoring Education Quality (SACMEQ).

Spaull (2012) argued that a conducted SACMEQ III survey of Grade 6 mathematics and reading, South Africa performed below most African countries that participated in the study. An alarmingly high proportion of Grade 6 learners have clearly not mastered even the most basic reading and numeracy skills. Using a categorisation of competency levels provided by SACMEQ as a benchmark, learners who have not reached Level 3 in the reading and mathematics tests can be regarded as functionally illiterate and functionally innumerate in the sense that they have not

acquired the basic reading and numeracy skills necessary to function meaningfully in society (Shabalala, 2005). Spaull (2012) further argued that of the 15 education systems that participated in the study, South Africa has the third highest proportion of functionally illiterate learners at 27 percent, and the fifth highest proportion of functionally innumerate learners 40 percent.

DoE (2013) the annual national assessments (ANA) was used to test grade 1 to 6 and grade 9 learners in literacy/languages and numeracy/mathematics in 2013. According to the results of ANA, (2013) the average score for language 51 percent in grade 3 to 59 percent in grade 6 home language, for grade 6 first additional language the average score is 46 percent, the average score for grade 9 home language is 43 percent, and for first additional language 33 percent and the very low score of 14 percent for grade 9 mathematics. These results suggest that there is a need to improving literacy and numeracy outcomes. Another area of concern is the trend of more learners choosing to take the less demanding option of mathematical literacy instead of mathematics (CDE, 2014).

The Department of Education announced the matric pass rate for 2017, at national level 75.1 percent which is 2.6 percentage points higher than in 2016. According to DoE (2018) 629,155 full-time grade 12 learners sat in for the 2017 NSC examinations, while an additional 173,276 part-time learners also wrote. Of these candidates, 534,484 full-time candidates, and 117,223 part-time candidates wrote the exams. The education department explains this discrepancy is caused by the number of progressed learners (learners which have failed Grade 11 twice) which will write their remaining exams later the following year, if these students are excluded, the pass rate rises to 76.5 percent. The DoE (2018) revealed that as many as 58 subjects were written in 2017, of which marks for 16 were increased, and marks for four were decreased as part of Umalusi’s standardisation process.

Table 2.2 Pass rates between 2007- 2017

Year	Pass rate	Change
2007	65.2%	-1.3
2008	62.5%	-2.7
2009	60.6%	-1.9
2010	67.8%	+7.2
2011	70.2%	+2.4
2012	73.9%	+3.7
2013	78.2%	+4.3
2014	75.8%	-2.4
2015	70.7%	-5.1
2016	72.5%	+1.8
2017	75.1%	+2.6

Source: National Department of Education (2018)

Table 2.2 above shows the matric pass rate from 2007 to 2017, in 2014 the matric pass rate declined by 2.4 percentage points, this was followed by a 5.1 percentage point decline in the 2015. However, this has been followed by increases in both 2016 and 2017.The matric pass rate has shown major improvement in 2017-2018 financial years approaching an all-time high of 78.2 percent in 2013 which triggered an argument between politicians claimed to celebrate the results of massive hard work and academics who claimed to have not seen any real improvement in the quality of South African education over time.

South African education statistics according to school wealth is presented in table 2.3 blow where schools have been divided into four wealth groups, ranging from the poorest schools categorized by 1 to the wealthiest schools at category 4. The figures are presented in percentages as follows.

Table.2.3: South African education statistics according to school wealth

Variables	1	2	3	4
At least one parent has a degree	4.70	7.80	10.70	28.70
At least one parent has matric	29.90	40.6	49.30	77.20
Gets homework (most days of the week)	49.90	52.10	46.10	75.80
Has own mathematics textbook	27.60	35.80	32.30	50.90
Has own reading textbook	34.40	42.30	38.2	66.10
More than 10 books at home	17.30	23.00	30.80	67.2
Proportion functionally illiterate	43.30	33.30	25.60	4.10
Proportion functionally innumerate	56.90	48.60	44.80	8.40
Pupil has used a PC before	11.80	39.90	51.40	94.90
Pupil very old for grade 6 (14y+)	23.70	20.10	14.00	2.00
Pupil-teacher ratio	36.30	34.80	35.50	30.50
Repeated at least 2 grades	10.90	9.30	10.30	1.80
School in urban area	5.50	21.40	31.20	73.30
Self-reported teacher absenteeism (days)	24.20	22.70	20.10	11.60

Source: Spaull, 2012

Table 2.3 shows that the wealthier schools are advantaged in most variables for instance wealthier schools have 77.20 parents who has matric compered to poorer schools with only 29.90. And pupil-teacher ratio is lower in rich schools that in the poor school. Pupil very old for grade 6 the percentage is very high at 23.70 percent while is at 2 percent in rich school. These statistics clearly outline that wealthier school are very much resourced and are likely to produce good results compared to the rural poor schools.

2.5 Summary

The information reviewed in this chapter points out that the education sector needs to invest significant resources in order to improve literacy and numeracy outcomes in the country. The annual national assessments should be utilised as an opportunity to ensure that interventions are directed at the appropriate areas. Interventions can be targeted to support specific schools or classes; the cornerstone of such interventions is the ability of education district offices to convey the intervention message to school managers and teachers, and to help them with implementation. Government should be committed to improve the quality of basic education; one element of this is to provide an environment that is conducive for learning and teaching by addressing

the school infrastructure backlogs. Finally, the most important factor that can never be ignored is the fact that educators have a significant impact on education quality since they develop learner skills and knowledge thus contributing to the economy in the future. The training and development of educators should be a government priority and departments of education should encourage more students to enter the teaching profession to address learner: educator ratio.

From the above discussion with the consideration of all this sort of Government spending on education it is clear that teachers, principals and officials in the department of Education lack competence and commitment in the workplace. Moloi and Chetty (2010) substantiate that limited teacher knowledge, coupled with low levels of accountability is one of the challenges that impact learner's performance in a South African context. Teacher's competencies may be improved by means of advancing their knowledge by further studies such as Bachelor of Education (Honours) and above. This is where teachers can be advanced by means of subject content knowledge and pedagogical content knowledge which can introduce different approaches in the teaching sector. Teachers should also be encouraged to participate in local and national conferences to share teaching strategies and techniques with their colleagues. The government should devise appropriate means to bring back the value of teaching in order for teachers to maximize their potential.

CHAPTER 3

RESEARCH METHODOLOGY

3.1 Introduction

Research methodology comprises a variety of techniques used for collecting data and analysing it within a specific framework of research (Waghid, 2002). Silverman (2001) posits that research methods are specific research techniques that include quantitative techniques, such as statistical correlations, as well as techniques like observation, interviewing and audio recording. Silverman (2001) further argued that these techniques themselves are not true or false, rather prove to be more or less useful, depending on their fit with the theories and methodologies being used and the hypothesis being tested or the research topic that is selected. This study is quantitative in nature, as informed by literature survey studies of school effectiveness and other input-output studies have generally used ordinary least-squares regression strategies and some regressions such as quantile regressions. In the light of the above cited articles the focus of this chapter is to discuss the research techniques that were employed in this study.

3.2 Data

According to Sapsford (2014) data collection is the process of gathering and measuring information on targeted variables in an established systematic fashion, which then enables one to answer relevant questions and evaluate outcomes. This study use secondary data obtained from Bankfontein Combined School Nkangala district in South African Department of Basic Education in the Mpumalanga province. South African School Administration and Management System (SA-SAMS) is utilized to obtain the learner and educator information.

The study utilized the data on the learners from grade 8 to grade 12 and the teachers who are teaching those grades at Bankfontein combined schools. The aim

of selecting grade 8 to 12 learners, teachers and quintile category of the school was to better understand both the degree and the nature of challenges they are faced within the learning and teaching. The data that was taken into account include the observations in the period 2016. This period is significantly chosen due to availability of input and output data. For example, the South African School Administration and Management System (SA-SAMS) was introduced only in 2014 at Bankfontein Combined School.

Except for the educator quality variables (teacher experience and qualifications), data observed for all variables was in the student-level form. The advantage of using student-level data over aggregated data is to address issues of endogeneity and omitted variable bias. However, the data used in this research suffered some drawbacks, as there is an element of incomplete information supplied by SA-SAMS. For example, there are intra-classroom variations in the FET band as learners are placed in different subject groups and some parent information such as parent qualifications is missing. Furthermore, the exclusion of dropout and grade repetition rates lead to selection bias as this exclusion result to overestimation of outcomes (Bhorat & Oosthuizen, 2006). However, the study is undertaken due to available large sample of student-level data of Bankfontein school. Also, its contribution to education economics literature as it is the first school representative study of its kind post the apartheid era. This education production function is also benefiting from the newly developed data bases like SA-SAMS which controls for the above biases and add to the robustness of the results.

3.3 Model specification

Zellner (2001) argued that analysts often think that complex problems require complicated regression equations. However, studies reveal that simplification usually produces more precise models. The specification of this study is not based only on statistical measures. In fact, the foundation of the model selection process of this research is dependant largely on theoretical concerns that are already raised in the literature review chapter.

According to Zellner (2001) model specification is the process of determining which independent variables to include and exclude from a regression equation. The critical motive of this study is to understand the determinants of learner performance [output] as a function of the following vector inputs. The education production function can be presented by the following model:

$$Y = f(E, A, Q, S) \quad (1)$$

Equation 1 shall be transformed into a linear model defined as follows:

$$Y_i = \mu_O + \mu_1 E_i + \mu_2 A_i + \mu_3 Q_i + \mu_4 S_i + u \quad (2)$$

Where Y is a dependent variable representing learner performance which is determined by learner achievement, and are end of the year examination scores or grade attainment in this study. E represents a vector of educator quality such as educator experience and qualification, A school characteristics such as educator learner ratio, public or private, rural/ urban, Q parent quality or socio economic issues such as quintile category, S represent learner quality like learner age, attendance and gender, u represent the error term to cater for measurement error. Finally, μs are regression coefficients.

3.4 Estimation techniques

With refernce to Zellner (2001), estimation is the process of finding an estimate, or approximation, which is a value that can be used for some purpose even if input data may be incomplete, uncertain, or unstable. Estimation determines how much money, effort, resources, and time it will take to achieve a specific outcome.

Estimation is based on

- Past Data/Past Experience
- Available Documents/Knowledge

- Assumptions
- Identified Risks

This study used descriptive statistics and regression analysis of both ordinary least-squares (OLS) and quantile regression to test the relationship between inputs variables and outputs (learner performance).

3.4.1 Descriptive Statistics

The study used descriptive statistics to analyse the educational output on collected variables. According to Dodge (2003) some descriptive statistics is summary statistics that quantitatively describes or summarizes features of a collection of information, it includes a process of using and analysing those statistics. Dodge (2003) further explains that descriptive statistics aims to summarize a sample, rather than use the data to learn about the population that the sample of data is thought to represent. Mann (1995) elaborate that some measures that are commonly used to describe a data set are measures of central tendency and measures of variability or dispersion. Measures of central tendency include the mean, median and mode, while measures of variability include the standard deviation (or variance), the minimum and maximum values of the variables, kurtosis and skewness (Mann 1995). However, this study used the bellow explained measures with regards to the relevance of the study.

3.4.1.1 Pairwise Correlations

Nikolic, Muresan, Feng and Singer (2012) describes correlation as a broad class of statistical relationships involving dependence, though in common usage it most often refers to the extent to which two variables have a linear relationship with each other. Nikolic et al. (2012) further emphasised that correlations are useful because they can indicate a predictive relationship that can be exploited in practice. The purpose of using correlations in this study is to determine if the relationships between variables are negative or positive and the strength of relationships as outline in the following table.

According to table 3.1 below the main result of a correlation is called the correlation coefficient or r. It can range from 0.0 to 1.00. or 0.00 to -1.00. Zero correlation means there is no relation between two variables and a correlation of 1.00 (either + or -) means perfect correlation. The sign of the correlation (+ or -) depends only on how the two variables are set. Infect the sign (+ or -) of a correlation tells only which direction a relation goes; not how strong it is. If r is positive, it means that as one variable gets larger the other gets larger. If r is negative it means that as one gets larger, the other gets smaller often called an inverse correlation.

Table 3.1: correlation coefficient

Value of r	Strength of relationship
-1.0 to -0.5 or 1.0 to 0.5	Strong
-0.5 to -0.3 or 0.3 to 0.5	Moderate
-0.5 to -0.3 or 0.3 to 0.5	Weak

Source: Nikolic et al. 2012

According to Croxton, Cowden and Klein (1968) like all statistical techniques, correlation is only appropriate for certain kinds of data. Correlation works for quantifiable data in which numbers are meaningful, usually quantities of some sort. Correlations are useful because they can indicate a predictive relationship that can be exploited in practice, especially for cross sectional data.

3.4.1.2 R squared

All the coefficient or values of r have been squared to obtain r squared in order to make the analyses simpler. As Croxton et al. (1968) suggests that while correlation coefficients are normally reported as r (a value between -1 and +1), squaring them makes then easier to understand. The square of the coefficient or r squared is equal to the percent of the variation in one variable that is related to the variation in the other. After squaring r, the decimal point is ignored for example an r of .5 means 25

percent of the variation is related (.5 squared =.25) and r value of .7 means 49 percent of the variance is related (.7 squared = .49).

Finally, a correlation report also shows a second result of each test - statistical significance. In this case, the significance level explains how likely it is that the correlations reported may be due to chance in the form of random sampling error.

3.4.1.3 Frequency Distribution

The study uses tables that display the frequency of various outcomes in a sample of data. According to Jaynes (2007) each entry in the table contains the count of the occurrences of values within a particular group or interval. The tables used in the study shows us a summarized grouping of data divided into mutually exclusive classes and the number of occurrences in a class and this is a way of showing unorganized data.

3.4.1.4 Central Tendencies

The mean, median and mode are used to analyse the data in the study. Dodge (2003) defined central tendency saying that it refers to the idea that there is one number that best summarizes the entire set of measurements, a number that is in some way central to the set.

The Mean: is the most common measure of central tendency used by researchers and people in all kinds of professions because of its mathematical qualities. Brown and Mood (1951) defined mean as just another name for average, where all the values are added together and divide by the number of values in the data set. Brown and Mood (1951) further argued that, since it is the measure of central tendency which is also referred to as the average a researcher can use the mean to describe

the data distribution of variables measured as intervals or ratios, these are variables that include numerically corresponding categories or ranges like race, class, experience, gender, or level of education, as well as variables measured numerically from a scale that begins with zero like household income or the number of children within a family.
The formula for a mean can be presented as:

$$\overline{x} = \frac{1}{N}\sum_{i=1}^{N} x_i = \frac{x_1 + x_2 + \cdots + x_N}{N} \quad (3)$$

The Median: Brown and Mood (1951) distinguished median from mean as the value separating the higher half of a data sample, a population or a probability distribution, from the lower half therefore in simple terms, it may be thought of as the middle value of a data set when those data are organized from the lowest to the highest value. This measure of central tendency can be calculated for variables that are measured with ordinal, interval or ratio scales. And Dodge (2003) argued that the median is actually a better measure of centrality than the mean if your data are skewed, meaning lopsided.

The Mode: According to Brown and Mood (1951) the mode is the measure of central tendency that identifies the category or score that occurs the most frequently within the distribution of data. In other words, it is the most common score or the score that appears the highest number of times in a distribution. The mode can be calculated for any type of data, including those measured as nominal variables, or by name.

3.4.1.5 The standard deviation

The standard deviation is the degree to which scores deviate from the mean, to be more precisely, it measures how far all the measurements are from the mean, square each one, and add them all up, the result is called the variance and the square root of the variance is the standard deviation. Like the mean, it is the expected value of how far the scores deviate from the mean (Mann, 1995).

The formula for standard deviation is presented like this:

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \overline{x})^2} \quad (4)$$

3.4.2 Regression estimates

Regression estimates has included Ordinary Least Squares and Quantile regression separately for the Further Education and Training as well as General Education and Training.

3.4.2.1 Ordinary Least Squares (OLS)

The OLS regression model has used multiple explanatory variables by simply adding additional variables to the equation. This form of the model ensures that learner performance which is output (Y) is predicted by multiple explanatory variables E, A, Q and S as depicted in equation 1 (Hutcheson, 2011). According to Hutcheson (2011) OLS regression is a generalized linear modelling technique that may be used to model a single response variable which has been recorded on at least an interval scale. The technique maybe applied to single or multiple explanatory variables and also categorical explanatory variables that have been appropriately coded.

According to Montenegro (2001) Ordinary Least Square is the best linear and unbiased estimator, if following four assumptions are satisfied:

1. The explanatory variable is non-stochastic

2. The expectations of the error term are zero

3. Homoscedasticity - the variance of the error terms is constant

4. No autocorrelation

This study chose to complement OLS with the quantile regression to ensure credible results should one of the above assumptions is violated. Because frequently one or more of these assumptions are violated, it is wise to also apply quantile regression because it can tackle following issues:

1. Frequently the error terms are not necessarily constant across a distribution thereby violating the axiom of homoscedasticity.
2. By focusing on the mean as a measure of location, information about the tails of a distribution are lost.
3. OLS is sensitive to extreme outliers that can distort the results significantly (Montenegro 2001).

3.4.2.2 Quantile Regressions

To correct for some biases suffered by the data and add to robust of the results, an OLS should be supplemented by other methods (Hanushek, 2003), hence the quantile regression. Quantile Regression as introduced by Koenker and Bassett (1978) seeks to complement classical linear regression analysis. It is the extension of ordinary quantiles from a location model to a more general class of linear models in which the conditional quantiles have a linear form (Buchinsky, 1998). Koenker (2001) describes a quantile as a value that corresponds to a specified proportion of an ordered sample of a population. For instance, a very commonly used quantile is the median, which is equal to a proportion of 0.5 of the ordered data, this corresponds to a quantile with a probability of 0.5 of occurrence. However, this study has also estimated upper and lower quantiles. Quantiles help to mark the boundaries of equally sized, consecutive subsets (Koenker, 2001).

According to Cook and Minning (2013) quantile regression complements OLS regression and related methods, which typically assume that associations between independent and dependent variables are the same at all levels. They further argued that, quantile methods allow the analyst to relax the common regression slope assumption. The reason of using both methods is that, in OLS regression, the goal is to minimize the distances between the values predicted by the regression line and the observed values. In contrast, quantile regression differentially weights the

distances between the values predicted by the regression line and the observed values, and then tries to minimize the weighted distances.

The main advantage of quantile regression over least-squares regression is its flexibility for modelling data with heterogeneous conditional distributions. Quantile regression provides a complete picture of the covariate effect when a set of percentiles is modelled, and it makes no distributional assumption about the error term in the model. Quintile regressions with different percentiles will be estimated. Furthermore, inter quantile regressions will be estimated. The quantile regression classical model is based on ordinal quantiles.
The ordinal quantile as stipulated in Koenker and Bassett (1978):

$$F(y) = P_r(Y \leq y) \quad (5)$$

Then for $\tau \in [0; 1]$

The τ^{th} quantile of Y is

$$Q(\tau) = inf\{y: F(y) \geq \tau\} \quad (6)$$

The median is then$Q\left(\frac{1}{2}\right)$, the first quartile$Q\left(\frac{1}{4}\right)$, the first deciles $Q\left(\frac{1}{10}\right)$,

The sample median can be derived through minimizing the sum of absolute residuals. The empirical objective function that should be minimised is non-differentiable at N points and can be flat at the optimum:

$$R(e) = \sum_{i=1}^{N} \rho_\tau (y_i - e) \quad (7)$$

Considering a quantile simple model

$$y_i = \beta_0 + x_i\beta_1 + u_i \quad (8)$$

The conditional quantile function of y is

$$Q_y(\tau/x) = \beta_0 + x\beta_1 + F_u^{-1}(\tau) \quad (9)$$

Lee (2005) differentiates quantile regression from other regression with the following factors:

1. The entire conditional distribution of the dependent variable can be characterized through different values.
2. Heteroscedasticity can be detected. If the data is heteroscedastic, median regression estimators can be more efficient than mean regression estimators.
3. The minimization problem be solved efficiently by linear programming methods, making estimation easy.
4. Quantile functions are also equivariant to monotone transformations. That is, for any function quantiles are robust in regards to outliers.

3.5 Summary

This chapter has outlined the research methods, which are consistent with the purpose and the theoretical framework of the study. Elements of the descriptive statistics such as correlations, r squared, frequency distribution, central tendencies including (mean, mode and median), standard deviation are employed. Data collection procedure was also important to give an overview of where, how and when the data was collected to substantiate its reliability, integrity and validity. The production function that comprises of school, socioeconomics, educator and learner characteristics is estimated with OLS supplemented with quantile regressions.

CHAPTER 4

DISCUSSION / PRESENTATION / INTERPRETATION OF FINDINGS

4.1 Introduction

This chapter provides results and discussions on the analysis of the determinants of learner performance for General Education & Training (GET) and Further Education & Training (FET) in a combined school. To examine the relationship between education input factors and learner performance, the Ordinary Least Squares (OLS) and Quantile regression methods were applied in estimation of the results using data for 2016. The analysis began by computing descriptive statistics and pairwise correlations.

4.2 Descriptive statistics results

The descriptive statistics discussed in this section include the arithmetic mean, standard deviation, and the minimum and maximum values observed on each variable series of the dataset used in the study. The sample was divided into two categories of leaners, namely the GET which comprised of grades 8 and 9, and FET comprised of grade 10, 11 and 12.

4.2.1 General Education and Training (GET) Descriptive Statistics Results

This sub-section provides descriptive statistics and pairwise correlations for the GET learners. From the total 310 GET learners, 161 (52%) learners were females and the remaining 149 (48%) learners were males.

Table 4.1: Descriptive statistics for GET learners, 2016

Variable	Obs	Mean	Std. Dev.	Min	Max
Age	310	16.283	1.872	13	22
Days absent	310	2.622	5.181	0	36
Examination score	310	43.803	6.463	26	67

Source: Author's calculations using Stata version 14

The descriptive statistics results in table 4.1 reveal that from a sample size (n) of 310 learners, the least age was 13 years while the greatest age was 22 years, and average (arithmetic mean) age of 16 years. While some learners were never absent from school at all, the maximum number of days one of the learners was absent from school was 36 days. On average, the average number of days the learners were absent was approximately 3 days. In terms of educational performance based on examination scores, the least mark obtained was 26 percent while the highest mark obtained by the learner was 67 Percent.

The average examination score of approximately 44 percent shows that academic achievement by GET learners was generally poor. The computed standard deviation statistics show that variability in observations was relatively more pronounced for the variables examination score (SD = 6.4) and days absent (SD = 5.1).

Table 4.2: Pairwise correlations of GET leaners

	Age	Days absent	Examination score
Age	1.0000		
Days absent	0.2503	1.0000	
Examination score	-0.2919	-0.2533	1.0000

Source: Author's calculations using Stata version 14

Table 4.2 results reveal existence of a statistically low correlation among the variables used in the analysis. Positive but low correlation is found between age versus days absent (correlation = 0.250), Negative correlation between days absent

and Examination Score (correlation= -0253) and Age versus Examination Score (correlation = -0.291)

4.2.2. Further Education and Training (FET) Descriptive Statistics Results

This sub-section provides descriptive statistics and pairwise correlations for the FET learners. From the total 407 GET learners, 248 (61%) learners were females and the remaining 159 (39%) learners were males.

Table 4.3: Descriptive statistics of FET learners

Variable	Obs	Mean	Std. Dev.	Min	Max
Age	407	18.668	1.831	13	26
Days absent	407	3.518	8.590	0	87
Examination score	407	39.356	8.031	20	70

Source: Author's calculations using Stata version 14

Descriptive statistics results provided in Table 4.3 show that from a sample size (n) of 407 learners, the least age among FET learners was 13 years while the highest age was 26 years, and yielding an average age of approximately 19 years. Though some learners were never absent from school, the maximum number of days some of the learners were absent from school was 87 days in 2016. On average, the average number of days the learners were absent was approximately 4 days. With regards to educational performance based on examination scores, the least mark obtained was 20 percent while the highest mark obtained by the learner was 70 Percent.

The average examination score of about 44 percent demonstrates that academic achievement by the FET learners in 2016 was generally poor, similar to GET learners. The calculated standard deviation statistics show that variability in observations was relatively more substantial for the variables days absent (SD = 8.5) and examination score (SD = 8.0).

Table 4.4: Pairwise correlations of FET leaners

	Age	Days absent	Examination score
Age	1.0000		
Days absent	-0.0006	1.0000	
Examination score	-0.2396	-0.0926	1.0000

Source: Author's calculations using Stata version 14

Table 4.4 results divulge existence of a statistically low correlation among the variables used in the research study. Positive but low correlation was found between examination score versus days absent (correlation = 0.005). Negative and low correlations were found between age versus days absent (correlation = -0.001), age versus examination score (correlation = -0.239) and days absent versus examination score (correlation = -0.092).

4.3 Regression Estimates Results

This section presents and explains the results of the ordinary least squares (OLS) regression and median quantile regression for both GET and FET learners.

4.3.1 GET learners' Ordinary Least Squares (OLS) Regression Results

Table 4.5 below provides OLS regression estimates for the GET learner's education production function made up of examination score as the dependent variable, while gender, age, days absent and socioeconomic status are independent variables.

Table 4.5: GET learners' OLS educational performance model

Source	SS	df	MS		Number of obs	= 310
					F (4, 403)	= 1829.00
Model	583317	4	145829		Prob > F	= 0.000
Residual	24397.8	306	79.7316		R-squared	= 0.9599
					Adj R-squared	= 0.9593
Total	607715	310	1960.37		Root MSE	= 8.9293

Examination score	Coeff.	Std. Err.	T	P > I t I	[95% Conf.	Interval]
Gender	-0.582	1.050	-0.55	0.580	-2.650	1.485
Age	2.767	0.103	26.65	0.000	2.563	2.971
Days absent	-0.490	0.100	-4.90	0.000	-0.687	-0.293
Socioeconomic status	1.355	1.588	0.85	0.394	-1.769	4.481

Source: Author's calculations using Stata version 14

Table 4.5 results of the Analysis of Variance (ANOVA) section show that from the total sum of square of 607715, about 583317 was accounted for by the model, whereas 24397.8 remained unexplained. From the 310 total degrees of freedom, 4 were consumed by the model while 306 were left for the residual. The OLS regression estimates show that age had a statistically significant and positive impact on learners' educational performance. Computed results (coefficient = 2.76; t-statistic = 26.65) show that an increase in age by year led to an increase in examination score by about 2.7 percentage points. These results are consistent with the finding by Lightbown and Spada (1993) that old learners are superior to children in rate of acquisition, hence old students learn faster than younger children.

Statistically insignificant variable "socioeconomic status" with (coefficient = 1.355; t-statistic = 0.85). These results are inconsistent with the finding reported by Van der Berg (2007) that further demonstrates that SES has a positive influence on learner achievement in South Africa. Baker and Jones (2005) further emphasise existence of a statistically significant correlation between low socioeconomic status and learner's poor performance schools.

Absenteeism measured by the variable "days absent" demonstrated a statistically significant and negative effect on learners' educational performance. Results in this

research study (coefficient= -0.49; t-statistic = -4.9) reveal that an increase in absenteeism by 1 day led to a reduction in learner's examination score by approximately 0.5 percentage points. This result is consistent with the finding reported by Balkis et al. (2016) that students' absenteeism negatively affects learners' academic achievement. Klem and Connell (2004) express the similar conclusion by suggesting that students who attend school regularly have higher academic achievement than students with high absences. The computed F(4, 403) statistic (= 1829; Prob > F = 0.000) rejected the null hypothesis that all model coefficients excluding a constant are equal to zero. With no constant, the model's computed Adjusted R-squared shows that about 95.9 Percent overall variation in the GET learners' examination scores was accounted for by the covariates gender, age, days absent, and socioeconomic status.

4.3.2 FET learners' Ordinary Least Squares (OLS) Regression Results

Table 4.6 below provides OLS regression estimates for the FET learner's education production function made up of examination score as the dependent variable, while gender, age, days absent and socioeconomic status are independent variables.

Table 4.6: FET learners' OLS educational performance model

Source	SS	df	MS		Number of obs	=	407
					F (4, 403)	=	1665.88
Model	619154	4	154788		Prob > F	=	0.000
Residual	37445	403	92.917		R-squared	=	0.943
					Adj R-squared	=	0.942
Total	656600	407	1613.26		Root MSE	=	9.639

Examination score	Coeff.	Std. Err.	T	P > I t I	[95% Conf.	Interval]
Gender	0.849	0.957	0.89	0.375	-1.032	2.731
Age	2.059	0.077	26.55	0.000	1.906	2.211
Days absent	-0.057	0.055	-1.03	0.302	-0.166	0.051
Socioeconomic status	-2.701	1.140	-2.37	0.018	-4.942	-4.459

Source: Author's calculations using Stata version 14

Table 4.6 results of the Analysis of Variance (ANOVA) segment show that from the total sum of square of 656600, about 619154 was accounted for by the model, while merely 37445 remained unexplained. From the 407 total degrees of freedom, 4 were consumed by the model while 403 were left for the residual. The MSE equals to 9.63 shows that the model was statistically stable.

The OLS regression estimates show that age had a statistically significant and positive impact on learners' educational performance at 1 percent. Computed results (coeff = 2.05; t-statistic = 26.55) show that an increase in age by year led to an increase in examination score by approximately 2.1 percentage points. This result remains consistent with the finding reported by Lightbown and Spada (1993) that old learners are superior to children in rate of acquisition, therefore old students are deemed to learn relatively faster than younger children.

The covariate "socioeconomic status" demonstrated a statistically significant and negative impact on learners' educational performance measured by the examination score at 5 percent. Based on the arithmetic mean statistic which shows that the majority of learners come from families with household incomes below the poverty line, the estimated coefficient (coefficient = -2.70; t-statistic = -2.37) shows that poverty leads to poor educational performance as measured by examination scores. This result is consistent with the finding reported by Van der Berg (2007) that further demonstrates that SES has a strong influence on learner achievement in South Africa. Baker and Jones (2005) further emphasise existence of a statistically significant correlation between low socioeconomic status and learner's poor performance schools.

Absenteeism measured by the variable "days absent" demonstrated a statistically insignificant and negative effect on learners' educational performance. Though statistically insignificant, this result is consistent with the finding reported by Balkis et al. (2016) that students' absenteeism negatively affects learners' academic achievement. Klem and Connell (2004) again shares a similar conclusion that students who attend school regularly have higher academic achievement than students with high absences.

The computed F(4, 403) statistic (= 1665; Prob > F = 0.000) rejected the null hypothesis that all model coefficients excluding a constant are equal to zero. With no constant, the model's computed Adjusted R-squared shows that about 94.2 Percent overall variation in the FET learners' examination scores was accounted for by the covariates gender, age, days absent, and socioeconomic status.

4.4 Quantile regression estimates results

A median quantile estimates for both GET and FET are presented and it was important to include the lower and upper quantile estimates for both phases using a production function.

4.4.1 GET Learners' quantile regression results

Table 4.7 below provides median quantile regression estimates for the GET learner's education production function made up of examination score as the dependent variable, while gender, age, days absent and socioeconomic status are independent variables.

Table 4.7: GET learners' quantile educational performance model

Median regression		Number of obs	= 310
Raw sum of deviations	772.5 (about 44)		
Min sum of deviations	723.285	Pseudo R-square	= 0.063

Examination score	Coeff.	Robust Std. Err.	T	P > I t I	[95% Conf.	Interval]
Gender	-1.642	0.757	-2.17	0.031	-3.132	-0.153
Age	-0.642	0.179	-3.58	0.000	-0.995	-0.289
Days absent	-0.214	0.104	-2.06	0.041	-0.419	-0.009
Socioeconomic status	1.000	0.825	1.21	0.227	-0.623	2.623
_cons	56.928	2.963	19.21	0.000	51.098	62.759

Source: Author's calculations using Stata version 14

The median estimates of quantile regression presented in Table 4.7 show the covariates gender, age and days absent demonstrated statistically significant and negative impacts on learners' examination scores. Computed results (coefficient = -0.642; t-statistic = -3.58) show that an increase in age by year led to a reduction in examination score by about 0.64 percentage points. This result is contradictory to the finding reported by Lightbown and Spada (1993) that old learners are superior to children in rate of acquisition, hence old students learn faster than younger children.

Absenteeism measured by the variable "days absent" exhibited a statistically significant and negative effect on learners' educational performance. Results in this research study (coefficient = -0.214; t-statistic = -2.06) reveal that an increase in absenteeism by 1 day led to a reduction in learner's examination score by approximately 0.2 percentage points. This result is consistent with the finding reported by Balkis et al. (2016) that students' absenteeism negatively affects learners' academic achievement. In addition, Klem and Connell (2004) expresses the similar conclusion by suggesting that students who attend school regularly have higher academic achievement than students with high absences.

Socioeconomic status demonstrated statistically insignificant results, the variable exhibited (coefficient = 1.00; t-statistic = 1.21). This result is inconsistent with the finding reported by Van der Berg (2007) that further demonstrates that SES has a strong influence on learner achievement in South Africa. Baker and Jones (2005) further emphasise existence of a statistically significant correlation between low socioeconomic status and learner's poor performance schools.

4.4.1.1 GET Learners' lower quantile and upper quantile regression results

This sub-section presents quantile regression estimates for both lower (10% and 25%) and upper (75% and 95%) quantiles to explore how GET students behave at different quantile performance levels. The results for lower quantiles are presented in Tables 4.7.1 and 4.7.2, while results for upper quantiles are presented in Tables 4.7.3 and 4.7.4.

Table 4.7.1: GET learners' 10% or 0.1 quantile educational performance model

0.1 Quantile regression					Number of obs	= 310
Raw sum of deviations	325.9 (about 36)					
Min sum of deviations	308.6453				Pseudo R-square	= 0.0529

Examination score	Coeff.	Robust Std. Err.	T	P > I t I	[95% Conf. Interval]	
Gender	-0.459	0.999	-0.46	0.646	-2.426	1.507
Age	0.040	0.270	0.15	0.881	-0.491	0.573
Days absent	-0.377	0.100	-3.75	0.000	-0.576	-0.179
Socioeconomic status	1.5	1.831	0.82	0.413	-2.103	5.103
_cons	36.686	4.030	9.10	0.000	28.755	44.616

Source: Author's calculations using Stata version 14

The 10% or 0.1 quantile regression estimates presented in Table 4.7.1 indicate that only days absent (coefficient = -0.377; t-statistic = -3.75) had a significant and negative impact on examination performance. The result show that an increase in days absent by 1 day led to a drop in examination score by about 0.38 percentage points. Age and socioeconomic status had statistically insignificant impacts on examination scores.

Table 4.7.2: GET learners' 25% or 0.25 quantile educational performance model

0.25 Quantile regression					Number of obs	= 310
Raw sum of deviations	605.75 (about 40)					
Min sum of deviations	574.8071				Pseudo R-square	= 0.0511

Examination score	Coeff.	Robust Std. Err.	T	P > I t I	[95% Conf. Interval]	
Gender	-1.054	1.087	-0.97	0.333	-3.194	1.085
Age	-0.315	0.222	-1.42	0.158	-0.753	0.122
Days absent	-0.260	0.100	-2.59	0.010	-0.459	-0.062
Socioeconomic status	2.315	1.502	1.59	0.124	-0.641	5.272
_cons	47.097	3.878	12.14	0.000	39.466	54.729

Source: Author's calculations using Stata version 14

Table 4.7.2 presents 25% or 0.25 quantile regression estimates which show that only days absent (coefficient = -0.260; t-statistic = -2.59) had a statistically significant and negative impact on examination score. The results show that an increase in days absent by 1 day led to a reduction in examination score by about 0.26 percentage points. Gender and age had statistically insignificant impacts on examination scores, and socioeconomic status had a statistically insignificant impact on examination score.

Table 4.7.3: GET learners' 75% or 0.75 quantile educational performance model

0.75 Quantile regression				Number of obs	= 310
Raw sum of deviations	325.9 (about 36)				
Min sum of deviations	308.6453			Pseudo R-square	= 0.0529

Examination score	Coeff.	Robust Std. Err.	T	P > | t |	[95% Conf. Interval]	
Gender	-0.40	1.071	-0.37	0.709	-2.509	1.709
Age	-1.00	0.266	-3.76	0.000	-1.523	-0.476
Days absent	-0.20	0.103	-1.92	0.055	-0.404	0.004
Socioeconomic status	3.00	2.198	1.36	0.173	-1.326	7.326
_cons	64.6	4.409	14.65	0.000	55.923	73.276

Source: Author's calculations using Stata version 14

The estimates in Table 4.7.3 for the 75% or 0.75 quantile regression reveal that age (coefficient = -1.00; t-statistic = -3.76) had a statistically significant and negative impact on examination score at 1 percent. The results imply that an increase in age by 1 year led to a reduction in examination score by about 1 percentage point. Days absent had negative statistically significant impacts on examination scores at 5 percent. Gender and socioeconomic status had statistically insignificant impacts on examination scores.

Table 4.7.4: GET learners' 95% or 0.95 quantile educational performance model

0.95 Quantile regression					Number of obs	= 310
Raw sum of deviations	236.55 (about 55)					
Min sum of deviations	194.2537				Pseudo R-square	= 0.1788

Examination score	Coeff.	Robust Std. Err.	T	P > l t l	[95% Conf. Interval]	
Gender	-1.888	1.337	-1.41	0.159	-4.520	0.742
Age	-1.703	0.381	-4.47	0.000	-0.453	-0.953
Days absent	-0.185	0.131	-1.41	0.160	-0.443	0.073
Socioeconomic status	-0.111	1.355	-0.08	0.935	-2.777	2.555
_cons	84.444	7.893	10.70	0.000	68.912	99.976

Source: Author's calculations using Stata version 14

Table 4.7.4 95% or 0.95 quantile regression estimates show that only age (coefficient = -1.703; t-statistic = -4.47) had a statistically significant and negative impact on examination score. The results imply that an increase in age by 1 year led to a reduction in examination score by about 1.7 percentage points. Gender, days absent and socioeconomic status all had statistically insignificant impacts on examination scores at the 95^{th} quantile.

4.4.2 FET Learners' Quantile Regression Results

Table 4.8 below provides median quantile regression estimates for the FET learner's education production function made up of examination score as the dependent variable, while gender, age, days absent and socioeconomic status are independent variables.

Table 4.8: FET quantile regression educational performance model

Median regression		Number of obs	= 407
Raw sum of deviations	1238.5 (about 39)		
Min sum of deviations	1194.366	Pseudo R-square	= 0.035

Examination score	Coeff.	Robust Std. Err.	T	P > | t |	[95% Conf. Interval]	
Gender	-1.892	1.020	-1.86	0.064	-3.898	0.113
Age	-0.580	0.231	-2.50	0.013	-1.036	-0.124
Days absent	-0.053	0.065	-0.83	0.410	-0.181	0.074
Socioeconomic status	-2.634	1.038	-2.54	0.012	-4.675	-0.593
_cons	52.612	4.626	11.37	0.000	43.517	61.708

Source: Author's calculations using Stata version 14

The quantile regression results presented in Table 4.10 show the covariates gender, age and socioeconomic status demonstrated statistically significant and negative impacts on learners' examination scores. Computed results (coeff = -0.580; t-statistic = -2.50) show that an increase in age by year led to a reduction in examination score by about 0.58 percentage points. This result is contradictory to the finding reported by Lightbown and Spada (1993) that old learners are superior to children in rate of acquisition, hence old students learn faster than younger children.

Absenteeism measured by the variable "days absent" exhibited a statistically insignificant on learners' educational performance. This result is inconsistent with the finding reported by Balkis et al. (2016) that students' absenteeism in relative terms has statistically negative affects learners' academic achievement. In addition, Klem and Connell (2004) express the similar conclusion by suggesting that students who attend school regularly have higher academic achievement than students with high absences.

The exogenous variable "socioeconomic status" confirmed a statistically significant and negative impact on learners' educational performance measured by the examination score. Based on the arithmetic mean statistic which shows that the majority of learners come from families with household incomes below the poverty

line, the estimated coefficient (coeff = -2.634; and a t-statistic = -2.54) demonstrates that poverty leads to poor educational performance as measured by examination scores. This result is consistent with the finding reported by Van der Berg (2007) that further demonstrates that SES has a strong influence on learner achievement in South Africa. Baker and Jones (2005) further emphasise existence of a statistically significant correlation between low socioeconomic status and learner's poor performance schools.

4.4.2.1 FET Learners Lower and Upper Quantile Regressions Results

This sub-section presents quantile regression estimates for both lower (10% and 25%) and upper (75% and 95%) quantiles to explore how FET students behave at different performance levels. The results for lower quantiles are presented in Tables 4.8.1 and 4.8.2, while results for upper quantiles are presented in Tables 4.8.3 and 4.8.4.

Table 4.8.1: FET learners' 10% or 0.1 quantile educational performance model

0.1 Quantile regression				Number of obs	= 407
Raw sum of deviations	539.5 (about 29)				
Min sum of deviations	467.2976			Pseudo R-square	= 0.1338

Examination score	Coeff.	Robust Std. Err.	t	P > | t |	[95% Conf. Interval]	
Gender	-5.000	0.876	-5.70	0.000	-6.723	-3.276
Age	-0.982	0.205	-4.77	0.000	-1.386	-0.577
Days absent	-0.035	0.045	-0.78	0.435	-0.124	0.053
Socioeconomic status	-1.964	0.917	-2.14	0.033	-3.768	-0.160
_cons	56.715	4.056	13.98	0.000	48.741	64.690

Source: Author's calculations using Stata version 14

Table 4.8.1 presents 10% or 0.1 quantile regression estimates which reveal that age (coefficient = -0.982; t-statistic = -4.77) and socioeconomic status (coefficient = -1.964; t-statistic = -2.14) had statistically significant and negative impacts on examination score. The results demonstrate that an increase in age by 1 year led to a reduction in examination score by about 0.98 percentage points. The significant

negative impact of socioeconomic status confirms the calculated mean statistic for the variable which shows that the majority of learners came from families whose household incomes levels were below poverty line, hence poor socioeconomic status led to a decrease in examination score by about 1.96 percentage points.

Table 4.8.2: FET learners' 25% or 0.25 quantile educational performance model

0.25 Quantile regression				Number of obs	= 407
Raw sum of deviations	962.25 (about 35)				
Min sum of deviations	893.8183			Pseudo R-square	= 0.0711

Examination score	Coeff.	Robust Std. Err.	t	P > l t l	[95% Conf. Interval]	
Gender	-2.972	0.891	-3.34	0.001	-4.724	-1.220
Age	-0.743	0.225	-3.30	0.001	-1.185	-0.300
Days absent	-0.060	0.010	-5.65	0.000	-0.081	-0.039
Socioeconomic status	-2.513	0.821	-3.06	0.002	-4.129	-0.897
_cons	52.836	4.194	12.60	0.000	44.590	61.081

Source: Author's calculations using Stata version 14

The estimates presented in Table 4.8.2 for the 25% or 0.25 quantile regression show that all the independent variables, namely gender (coefficient = -2.972; t-statistic = -3.34), age (coefficient = -0.743; t-statistic = -3.30), days absent (coefficient = -0.06; t-statistic = -5.65) and socioeconomic status (coefficient = -2.513; t-statistic = -3.06) had a statistically significant and negative impacts on examination score. Results indicate that an increase in age by 1 year led to a reduction in examination score by about 0.74 percentage points, while an increase in days absent from school led to a drop in examination score by 2.5 percentage points.

Table 4.8.3: FET learners' 75% or 0.75 quantile educational performance model

0.75 Quantile regression					Number of obs	= 407
Raw sum of deviations	1065.5 (about 44)					
Min sum of deviations	1017.3				Pseudo R-square	= 0.0452

Examination score	Coeff.	Robust Std. Err.	t	P > I t I	[95% Conf.	Interval]
Gender	-0.254	1.151	-0.22	0.825	-2.519	2.009
Age	-0.945	0.252	-3.75	0.000	-1.441	-0.449
Days absent	-0.090	0.078	-1.16	0.246	-0.244	0.062
Socioeconomic status	-3.818	1.179	-3.24	0.001	-6.136	-1.499
_cons	62.872	5.070	12.40	0.000	52.904	72.840

Source: Author's calculations using Stata version 14

Table 4.8.3 presents 75% or 0.75 quantile regression estimates which demonstrate that age (coefficient = -0.945; t-statistic = -3.75) and socioeconomic status (coefficient = -3.818; t-statistic = -3.24) had statistically significant and negative impacts on examination score. The results show that an increase in age by 1 year led to a reduction in examination score by about 0.94 percentage points. The significant negative impact of socioeconomic status confirms the calculated mean statistic for the variable which shows that the majority learners came from families whose household incomes levels were below poverty line, hence poor socioeconomic status led to a decline in examination score by 3.82 percentages points.

Table 4.8.4: FET learners' 95% or 0.95 quantile educational performance model

0.95 Quantile regression		Number of obs	= 407
Raw sum of deviations	398 (about 54)		
Min sum of deviations	362.294	Pseudo R-square	= 0.0897

Examination score	Coeff.	Robust Std. Err.	t	P > I t I	[95% Conf.	Interval]
Gender	2.746	2.604	1.05	0.292	-2.374	7.867
Age	-2.084	0.629	-3.31	0.001	-3.321	-0.848
Days absent	-0.169	0.080	-2.11	0.036	-0.327	-0.114
Socioeconomic status	-6.746	3.650	-1.85	0.065	-13.922	0.430
_cons	90.776	12.382	7.33	0.000	66.434	115.119

Source: Author's calculations using Stata version 14

The estimates presented in Table 4.8.4 for the 95% or 0.95 quantile regression show that age (coefficient = -2.084; t-statistic = -3.31), days absent (coefficient = -0.169; t-statistic = -2.11) and socioeconomic status had a statistically significant and negative impacts on examination score. Results indicate that an increase in age by 1 year led to a reduction in examination score by about 2.1 percentage points, while an increase in days absent from school led to a decrease in examination score by about 0.17 percentage points.

4.5 Summary

This chapter provided an analysis of statistical results on the educational performance functions for General Education & Training (GET) and Further Education & Training (FET) bands using data for 2016. The Ordinary Least Squares (OLS) and Quantile regression methods were applied in estimation of the results using Stata statistical software version 14 for windows. Statistical estimates computed for analysis include descriptive statistics and pairwise correlations.

Some key common differences were noted in results estimated for the two bands. The OLS regression estimates variable age consistently demonstrated a statistically significant and positive impact on learners' examination scores in both GET and FET groups. It turns out that the quantile regression estimates for the same variable age

consistently demonstrated a statistically significant and negative impact on learners' examinations cores in both GET and FET bands. For the GET learners, socioeconomic status had a statistically insignificant and positive impact on learners' examination score based on estimates of both OLS and quantile regression methods. In contrary, the socioeconomic status exhibited a statistically significant and negative effect on FET learners' academic performance based on estimates of both OLS and quantile regression methods. These results indicate that learner performance can be explained by the chosen inputs even at different level of learner performance as seen from different regression quantiles.

CHAPTER 5

SUMMARY, CONCLUSION, RECOMMENDATIONS

5.1 Summary and interpretation of findings

This chapter provides the summary, conclusion and recommendations of the study which aimed to analyse the determinants of the learner performance at Bankfontein Combined School in Mpumalanga province. In order to achieve this aim, the study was guided by the objectives such as; to find out which input factors to influence learner performance in Bankfontein Combined School; to examine the relationship between education input factors and learner performance and; to make appropriate recommendations based on the findings from the study. Also to be contained in the chapter are the study limitations.

In order to achieve the set objectives, descriptive statistics, ordinary least squares and quantile regression techniques were applied to analyse data. The study used data for 310 GET and 407 FET learner' information in 2016 obtained from the school database called SA-SAMS. The study found out that the average performance from the GET learners during the chosen period was poor as the computed average score was only approximately 44%. This is despite the fact that there was a low level of absenteeism during this period among GET learners.

For FET learners during the same period the average academic achievement was almost similar to the GET learners with an average of nearly 44% as well which meant that academic performance was generally poor. The results revealed that most FET learners came from poor socioeconomic background. Similar to the GET learners, the FET learners had a low rate of absenteeism with only an average of 4 days absent in that period. The pairwise correlations of FET learners were generally low though positive.

The Ordinary least squares for the GET learners revealed that age had a statistically significant and positive impact on learners' educational performance. Though statistically insignificant, the variable "socioeconomic status" exhibited a positive

impact on learners' educational performance measured by the examination score. Learner absenteeism demonstrated a statistically significant and negative effect on learners' educational performance. Results in this research study reveal that an increase in absenteeism by 1 day led to a reduction in learner's examination score by approximately 0.5 percentage points.

The quantile regression estimates for GET learners revealed that the covariates gender, age and days absent demonstrated statistically significant and negative impacts on learners' examination scores. The computed statistics showed that an increase in age by year led to a reduction in examination score by about 0.64 percentage points. The results showed that absenteeism had a negative effect on learners' educational performance. In addition to that the findings indicated that the 'socioeconomic status' had a positive impact on learners' educational performance. The GET learners' lower and upper quantile regression estimates showed that at lower levels of performance only the absenteeism had a significant and negative impact on examination performance. Whilst at higher levels of performance only age had a statistically significant and negative impact on examination score.

The ordinary least squares for FET learners revealed that that only age had a statistically significant and positive impact on learners' educational performance. Calculated results showed that an increase in age by year led to an increase in examination score by approximately 2.1 percentage points. The "socioeconomic status" demonstrated a statistically significant and negative impact on learners' educational performance measured by the examination score. This showed that poverty leads to poor educational performance as measured by examination scores. The OLS results revealed that absenteeism had a negative effect on learners' educational performance. An increase in absenteeism by 1 day led to a reduction in learner's examination score by approximately 0.1 percentage points.

The quantile regression results for FET learners showed that age and socioeconomic status demonstrated statistically significant and negative impacts on learners' examination scores. The calculated results showed that an increase in age by year led to a reduction in examination score by about 0.58 percentage points. The results also showed that absenteeism had a negative effect on learners' educational performance. An increase in absenteeism by 1 day led to a reduction in learner's

examination score by approximately 0.1 percentage points. The “socioeconomic status” variable revealed a statistically significant and negative impact on learners’ educational performance measured by the examination score. This demonstrates that poverty leads to poor educational performance as measured by examination scores.

The FET learners’ lower and upper quantile regression estimates revealed that at lower levels of performance gender and socio- economic status had statistically significant and negative impacts on examination score. The results demonstrated that an increase in age by 1 year led to a reduction in examination score by about 0.98 percentage points. They also demonstrated that a poor socioeconomic status led to a decrease in examination score by about 1.96 percentage points. However, at higher levels of performance the quantile regression results showed that age and absenteeism had statistically significant and negative impacts on examination score. Results indicate that an increase in age by 1 year led to a reduction in examination score by about 2.1 percentage points, while an increase in days absent from school led to a decrease in examination score by about 0.17 percentage points.

5.2 Conclusions

The aim of the study was to analyse the determinants of the learner performance at Bankfontein Combined School in Mpumalanga province in 2016. Based on the research findings the following conclusions can be made;

Objective 1: To find out which input factors to influence learner performance in Bankfontein Combined School.

The input factors and their influence on learner performance at Bankfontein Combined School have been found to be consistent and a general consensus can be mentioned. The chosen input factors have been key determinants of learner performance at the school. The results showed that there was consistently poor performance of both GET and FET learners. Academic performance was not in any way helped by the low rate of absenteeism which was another common factor within the same period. Another factor which was evident from the results was that an increase in age for both GET and FET learners translated to better performance

though the margin of improvement was very little. Absenteeism for the both groups though very low is also a contributing factor in poor performance.

Objective 2: To examine the relationship between education input factors and learner performance.

There is a strong relationship between education input factors and learner performance for both GET and FET learners. The education input factors influence the overall learner performance. Determinants of schooling outcome such as age, socio-economic status and absenteeism to a larger extent impact on learner performance. An increase in age means that the rate of acquisition gets better with older learners abler to grasp concepts faster than their younger peers. In the same vein the rate of absenteeism has a negative effect on the learner's academic performance. Regular school attendance of learners must therefore be prioritized as it contributes to higher academic performance. The socio- economic status also has a bearing on a learner's performance. Low economic status is in most cases related to poor performance. This may imply that rather than being motivated to learn in the hope of taking their families out of poverty, learners are in fact too demotivated by their circumstances to see the value of education. In view of these determinants of schooling outcomes, appropriate recommendations must therefore be tabled to improve learner performance.

5.3 Recommendations of the study

In light of the above conclusions, the following recommendations are therefore mentioned;

- Managing diversity

The challenge of diversity, as has been revealed in the study, is more pronounced in the FET level. Learners come with different ages, skills, knowledge and from different socio- economic backgrounds. The teachers must therefore be able to differentiate their learning methods to ensure that while the older students grasp things faster, the younger ones are not left behind. The teachers must be responsive to learner diverse needs so that they provide them with the vital support they need.

- Introduce a learner motivation programme

As it has been concluded that the socio- economic status might be demotivating learners to learn, focus must be on motivating learners. The learning process will be more successful if the learners are motivated. The school can acquire the services of a professional motivational speaker to come to the school at specified times to motivate the learners to work harder. The education districts may also introduce motivational programmes for learners through its officials who can move around schools motivating learners.

- Introduce holiday and extra classes

As poor performance of learners has actually been a trend that has been identified over years in the school, holiday classes and extra classes might be useful strategies that can be used to break the trend. They must not necessarily replace the regular classes but rather they must be used to complement them and enhance the performance of learners. Everyone must take part in these classes as the government only support grade 12 learners only presently; the poor performers, the average performers and the good performers in all grades. The classes must separate these groups of performers so that for each group differentiated methods of teaching can be used to assist them.

- All concerned stakeholders should be mobilised to support the school

The community and the development practitioners should be united and be involved in the planning processes for school development. At the centre of development activities should be the school's performance needs, from then on there should be an articulation of the support that can be offered by the stakeholders to improve the performance of the schools. Thus stakeholder involvement should take an inside out approach (starting from the school's needs going to the stakeholder input). A steering committee should be formed with representation of all stakeholders to coordinate all stakeholder involvement activities. The steering committee may have to decide on the frequency of its meetings to discuss educational challenges and other school project. The ultimate objective of the steering committee should be to find ways to improve the performance of learners.

- Improve on planning and organisation

The objective of planning and organizing is to improve the way through which the school functions. In dysfunctional schools as may be the case at Bankfontein Combined School, there is little room for meaningful learning even if effective teachers are available. The school management team (SMT) at Bankfontein Combined School should ensure that curriculum delivery initiatives are properly planned and organised. Planning and organizing should target the SMT because of the crucial role it plays in managing the school's technical operations which encompass strategic planning for the school, financial management and curriculum management.

- Teacher performance improvement interventions

Special attempts must be made to improve teacher performance at Bankfontein Combined School. Teacher performance interventions must be aimed at supporting teachers adequately with delivery and curriculum planning materials. These materials would include work schedules, learning programmes, lesson plans and learner assessment tools. The successful implementation of teacher performance interventions would mean that they will be able to successfully deliver the curriculum in classrooms which in turn leads to improved learner performance. The interventions would also ensure that the teacher is able to assess and monitor learner progress effectively as per requirement of the curriculum.

- Teacher competence enhancement interventions

Linked to teacher performance is the level of teacher competence. There is need to target the knowledge and skills that the current teachers possess in order to find ways to improve them so that they can effectively facilitate deliver in classrooms. Learners taught by the effective teachers are likely to progress faster than those taught by a less effective teacher. Teacher development programmes need to be designed to determine the level of competence, to assess teachers' strengths and areas where development is needed. With improved knowledge, professionalism, leadership, reasoning and expectations, learner performance will most likely improve as well.

- Parental involvement

There is a need to break the perception that parents cannot be effective contributors to their children's education simply because some of them are illiterate. The school needs to incorporate parents into its school improvement programmes. The involvement of parents would ensure that they also do their part to assist teachers through monitoring their children's progress at home. It would also be useful for the school to increase the number of visitation days by parents so that the parents feel part of their children's learning process. Parents need to be educated on the importance of assisting their children to manage time after school, study and reading for pleasure. This would go a long way in cultivating a culture of wanting to learn within the learners.

- District support

The role of the district education department needs not to be understated. As the authority closest to the school, the district education department should assist the school through providing all necessary resources, systems and monitor the schools. The district education department should also provide technical support to the school in terms of its strategic planning and organisation. The potential results of education districts' support include improved communication between the school and its stakeholders, assistance in project implementation and assistance in the mobilisation of financial resources for the school from its partners.

5.4 Limitations of the study

The study had the following limitations:

- The study was conducted at one school thereby limiting it in terms of generalizability to different contexts.
- The study is only a snapshot of the events of that occurred in 2016.

The study is limited to the quantitative methods that were used with focus on statistical analyses of records for 2016, for an in-depth analysis of the determinants of education performance qualitative methods could be used.

LIST OF REFERENCES

Adnot, M., Dee, T., Katz, V. and Wyckoff, J. 2016. *Teacher Turnover, Teacher Quality, and Student Achievement in DCPS.* CEPA Working Paper No. 16-03.

Angrist, J.D. and Lavy, V. 1999. *Using Maimondide's rule to estimate the effect of class size on scholastic achievement.* Quarterly Journal of Economics 114, 533–75.

Armstrong, P. 2009. *Teacher pay in South Africa: How attractive is the teaching profession?* Stellenbosch Economic Working Paper. No. 04/09.

Aucejo, E. M. and Romano, T.F. 2014. *Assessing the Effect of School Days and Absences on Test Score Performance.* CEP Discussion Paper No 1302.

Baker, R. and Jones, A. 2005. *How can international studies such as the International Mathematics and Science Study and the Programme for International Student Assessment be used to inform practice, policy and future research in science education in New Zealand?* International Journal of Science Education, 27: 145-157.

Balkis, M., Erdinc, D. and Arslan, G. 2016. *The School Absenteeism among High School Students: Contributing Factors.* Research-gate 307088341

Bhorat, H. and Oosthuizen M. 2006. *Determinants of grade 12 pass rates in the post- apartheid South African schooling system.* ASGISA 18-20 October 2006 Conference.

Billy-Jackman, A., Moore, Z., Jackman, W. M., Joseph, S. and Morris, J. 2011. *Gender Study on the Factors Affecting Male Re-Entry into the Post-secondary and Tertiary Education System.* Ministry of Science, Technology and Tertiary Education.

Botev, Z.I., Grotowski, J.F. and Kroese, D.P. 2010. *Kernel density estimation via diffusion*. Annals of Statistics. 38(5): 2916-2957.

Brown, G. W. and Mood, A. M. 1951. *On Median Tests for Linear Hypotheses. Proc Second Berkeley Symposium on Mathematical Statistics and Probability*. Berkeley, CA: University of California Press.

Buchinsky, M. 1998*. Recent Advances in Quantile Regression Models: A Practical Guidline for Empirical Research*. Journal of Human Resources. 33(1): 88–126.

Carnoy, M. L. 2008. *Towards Understanding Student Academic Performance in South Africa: A Pilot Study of Grade 6 Mathematics Lessons in South Africa*. Report prepared for the Spencer Foundation. Pretoria: HSRC.

Caves, R. and Barton, D. 1990. *Efficiency in U.S. Manufacturing Industries*. Cambridge, MA: MIT Press.

Centre for Development and Enterprise (CDE). 2014. *Teacher, Teaching and Learner Performance in Mathematics. Johannesburg:* Centre for Development and Enterprise.

Christie, P. 2008. *Opening the doors of learning*. Johannesburg: Heinemann.

Clark, D. and Ramsay, A.L. 1990. *Problems of Retention in Tertiary Education.* Education Research and perspectives, 17(2), 47-59.

Coates, D.1998. *Education Production Functions Using Instructional Time as an Input*. Baltimore. UMBC.

Coleman, J.S., Campbell, E.Q., Hobson, C.J., McPartland, J., Mood, A.M., Weinfeld, F.D. and York, R.L. 1966*. Equality of Educational Opportunity*. Washington, DC: US Government Printing Office.

Cook, L.B. and Manning, W.G. 2013. *Think beyond the mean: A Practical Guide for Using Quantile Regression Methods for Health Services Research*. Shanghai Arch Psychiatry. 25(1): 55-59.

Crouch, L. and Mabogoane, T. 1998. *No Magic Bullets, Just Tracer Bullets: The role of learning resources, social advantage and education management in improving the*

performance of South African schools. Research triangle park: Research triangle institute.

Croxton, F.E., Cowden, D.J. and Klein, S. 1968. *Applied General Statistics*, Pitman. ISBN 9780273403159

Darling-Hammond, L. 2000. *Teacher Quality and Student Achievement: A Review of State Policy Evidence. Education policy analysis archives:* Arizona State University. 8(1): 1-44.

Davies, H. 2004. *Some aspects of equality in the funding of public schools in South Africa*. Perspectives in Education, 22(3): 109-124.

Department of Education. 1996. *South African Schools Act No 84 of 1996*. Pretoria: Government Printers.

Department of Education, 2000. *National Norms and Standards for School Funding policy of 1999.* Pretoria: Government Printers.

Department of Education. 2000. Recognition and Evaluation of Qualifications for Employment in Education. DoE. Pretoria. Government printers.

Department of Education 2002. *National mathematics standard grades 10-12 (schools)*. Mathematics draft. DoE. Pretoria: Government Printers.

Department of Education. 2006. *NO fees Schools in South Africa - Policy Brief*. DoE. Pretoria. Government Printers.

Department of Education annual report. 2010. DoE. Pretoria. Government printers.

Department of Education. 2012. *Education Statistics Report in South Africa.* DoE. Pretoria. Government printers.

Department of Education. 2012. EMIS: School realities. DoE. Pretoria. Government printers.

Department of Education. 2012. Every Child is a National Asset: SA-SAMS. DoE. Pretoria. Government printers.

Department of Education. 2012. *Guide for Parents: School Fees in Public Schools*. DoE. Pretoria. Government printers.

Department of Education. 2013. *Report on the Annual National Assessments of 2013,* DoE. Pretoria. Government printers.

Department of Education. 2013. *School Realities 2013: Special Needs Education.* SNAP Surveys. DoE. Pretoria. Government printers.

Department of Education. 2015. *P4 Circuits Performances 2009 -2015*. DoE. Pretoria. Government printers.

Dewy, J., Husted, T. and Kenny, L. 2000. *The Ineffectiveness of School Inputs: A Product of Misspecification*. Economics of Education Review.

Dodge, Y. 2003. *The Oxford Dictionary of Statistical Terms.* OUP. ISBN 0-19-850994-4.

Eddy, E. 2011. *The Involvement of Cognitive Process in the Acquisition of English Grammar by Slovak Learners. University of Slovakia*

Education For All (EFA). 2000. Assessment 2005. Accessed Date 20-01-2018. http://education.pwv.gov.za/content/document

Fagerlind, I. and Saha, L.J. 1983. *Education and National Development: A Comparative Perspective.* Pergamon Press Ltd. UK.

Fataar, A. 2010. *Education policy development in South Africa's democratic transition, 1994-1997*. Stellenbosch. Sun Media.

Ferguson, R. F. and Ladd, H. F. 1996. *How and Why Money Matters: An Analysis of Alabama Schools, in Holding Schools Accountable: Performance Based Reform in Education*. Washington: Brookings Institution.

Gaete-Romeoy, G. 2017. *Follow the Leader: Student Strikes, School Absenteeism and Long Term Implications for Education Outcomes.* Job Market Paper

Ghazvinia, S. D. and Khajehpoura, M. 2011. *Gender differences in factors affecting academic performance of high school students.* Procedia Social and Behavioral Sciences 15 (2011) 1040–1045

Goldhaber, D. D. and Brewer, J. 1997. *Why Don't Schools and Teachers Seem to Matter? Assessing the Impact of Unobservable on Educational Productivity.* Journal of Human Resources. 32(3):505-23.

Gustaffson, M. 2007. *Using the Hierarchical Linear Model to understand school production in South Africa.* Working paper of the department of economics and the bureau for economic research at the University of Stellenbosch 01/07.

Gustafsson, M. and Patel, F. 2008. *Managing the teacher pay system: what the local and international data are telling us*? Cape Town: DPRU.

Hansen, J.B. 1970. *Test and Formulas Concerning Secondary Consolidation*. Geoteknisk Institute.

Hanushek, E.A. 1979. Conceptual and empirical issues in the estimation of educational production functions*. Journal of Human Resources* 14, 351–88.

Hanushek, E.A. 1999. Some findings from an independent investigation of the Tennessee STAR experiment and from other investigations of class size effects. *Educational Evaluation and Policy Analysis* 21, 143–63.

Hanushek, E. A. 2003. The failure of input based schooling policies*. The economic journal* 113 (485)

Hanushek, E. A. 2007. *Education Production Functions*. Hoover Institution, Stanford University Press.

Hanushek, E. A. 2008. *Education production functions.* The New Palgrave Dictionary of Economics. Second Edition. Palgrave Macmillan.

Harris, D. N. and Sass, T. R. 2007. *Teacher Preparation and Student Achievement. Centre for policy research*. University of Albany.

Hedges, L.V. and Greenwald, R. 1996. *The Relation Between School Resources and Student Performance: The Effect of School Resources on Student Achievement and Adult Success*. Washington, D.C: Brookings Institution Press.

Holmlund, H. and Sund, K. 2005. *Is the Gender Gap in School Performance Affected by the Sex of the Teacher?* Swedish Institute for Social Research (SOFI)

Humphrey, T.M. 1997. *Algebraic Production Functions and their Uses before Cobb-Douglas.* Federal Reserve Bank of Richmond Economic Quarterly bulletin.

Hutcheson, G. D. 2011. *Ordinary Least-Squares Regression*, The SAGE Dictionary of Quantitative Management Research. 224-228.

Jaynes, E.T. 2007. *Probability theory: the logic of science*. Cambridge University Press.

Kane, T.J., Jonah, E., Rockoff J.E. and Staiger, D.O. 2006. *What Does Certification Tell Us About Teacher Effectiveness? Evidence from New York City*. Working Paper 12155.Cambridge. MA: National Bureau of Economic Research.

Khoza, S. 2007. *Socio-economic rights in South Africa*. Bellville: Community Law Centre.

Khumalo, N. 2014. *Some Implications of the Quintile School Funding in South Africa*. University of Stellenbosch.

Kimani, E.M. and Bhorat, H. 2014. *The effects of pupil-teacher ratio and expenditure per pupil on educational attainment in South Africa*. University of Cape town

Klem, A. M. and Connell, J. P. 2004. *Relationships matter: Linking teacher support to student engagement and achievement*. Journal of School Health, 74, 262–273.

Koenker, R. 2001. *Quantile Regression: Economics Perspectives*. Econometrica.

Koenker, R. and Bassett, G.W. 1978. *Regression Quantiles*. Econometrica. 46.33.50

Koenker, R. and Hallock, K. F. 2001. *Quantile regression*. The Journal of Economic Perspectives.

Krause, D. 2005. *The First Year Experience in Australian Universities: Findings from a Decade of National Studies. Centre for the Stud of Higher Education*. University of Melbourne.

Krueger, A.B. 1999. *Experimental estimates of education production functions.* Quarterly Journal of Economics 114, 497–532.

Lavy, V. 2004. Do *Gender Stereotypes Reduce Girls' Human Capital Outcomes? Evidence From a Natural Experiment*. NBER Working Paper 10678.

Lee, S. 2005. *Lecture Notes for MECT1 Quantile Regression*. Accessed date 14-02-2018. http://www.homepages.ucl.ac.uk/~uctplso/Teaching/MECT/lecture8.pdf

Lightbown, P. M. and Spada, N. 1993. *How Languages are learned*. Oxford University press.

Makgato, M. and Mji, A. 2006. *Factors associated with high school learner's poor performance: a spotlight on mathematics and physical science.* South African Journal of Education, 26(2)253-266.

Mann, P. S. 1995. *Introductory Statistics (2nd ed.).* Wiley. ISBN 0-471-31009-3.

Masondo, S. 2016. *Education in South Africa: A system in crisis*. City Press.

Moloi, M.Q. and Chetty, M. 2010. *The SACMEQ IIIProject in South Africa: A Study of the Conditions of Schooling and the Quality of Education*. Pretoria. Government printers.

Montenegro, C. E. 2001. *Wage Distribution in Chile: Does Gender Matter? A Quantile Regression Approach.* Working Paper Series 20, The World Bank, Development Research Group.

Motala, S. 2009. *Privatizing public schooling in post-apartheid South Africa: Equity considerations*. Journal of Comparative & International Education, 39(2): 185-202.

National Department of Education. 2018. *South Africa's matric pass rate: 2007 – 2017*. Accessed Date 18-01-2018. https://businesstech.co.za.

National Treasury. 2010. National treasury annual report of 2010/2011. South Africa. Pretoria. Government printers.

National Treasury. 2013. *Estimates of National Expenditure*. Republic of South Africa. Pretoria. Government printers.

National Treasury. 2014*. Provincial Budget and Expenditure Review: 2010/11-2016/17.* Pretoria. Government printers.

Ncanywa, T. 2014. *The state of the Eastern Cape schools in a period almost the second decade of democracy.* ERSA working paper 486.

Ndlovu, M.C. 2011. *University-school partnerships for social justice in mathematics and science education: the case of the SMILES project at IMSTUS*. South African Journal of Education, 31:419-433.

Nikolic, D., Muresan, R.C., Feng, W. and Singer, W. 2012. *Scaled correlation analysis: a better way to compute a cross-correlogram*. European Journal of Neuroscience.

Nye, B., Konstantopoulos, S. and Hedges, L.V. 2004. *How Large are Teacher Effects?* Educational Evaluation and Policy Analysis 26(3):237-257.

Parliament. 2012. *School Realities*. Parliament South Africa. preliminary report. Pretoria. Government printers.

Rivkin, S.G., Hanushek, E.A. and Kain, J.F. 2005. *Teachers, Schools and Academic achievement*, Econometrica publishers.

Roithamyr, D. 2003. *Access, adequacy and equality: The constitutionality of school fee financing in public education.* South African Journal on Human Rights, 19(3): 382-429.

Saidudddin, J. 2003. *Factors Affecting Achievement at Junior High School on the Pine Ridge Reservation*. Spain: Ohio State University.

Sapsford, R. 2014. *Data Collection and Analysis*. Victor Jupp ISBN 0-7619-5046-X

Shabalala, J. 2005. The SACMEQ II Project in Swaziland: *A Study of the Conditions of Schooling and the Quality of Education*. Accessed Date 20-01-2018. http://www.sacmeq.org/education-swaziland.htm

Shephard, R.W. 1970. *Theory of Cost and Production Functions.* Princeton, NJ. Princeton University Press.

SILVERMAN, D. 2001. *Interpreting Qualitative Data, Methods for Analyzing Talk, Text and Interaction (2nd ed.).* London. Sage Publication Inc

Spaull, N. 2012. *Equity and Efficiency in South African Primary Schools: A preliminary analysis of SACMEQ III South Africa*. Stellenbosch Economic Working Papers No. 11/11.

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

Tan, B.H. 2008. *Cobb-Douglas Production Function*. Thomson Brooks.

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

Taylor, N., Muller, J. and Vinjevold, P. 2003. *Getting schools working.* Cape Town: Pearson Education.

Taylor, S. 2011. *Uncovering indicators of effective school management in South Africa using the National School Effectiveness Study*. Stellenbosch Economic Working Papers No. 10/11

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

Van der Berg, S. and Low, M. 2006. *Unravelling the mystery: Understanding the South African outcomes in the regional context*. University of Stellenbosch.

Van der Berg S., Taylor S., Gustafsson M., Spaull N. and Armstrong P. (2011). *Report for the National Planning Commission, Department of Economics,* University of Stellenbosch

Van der Walt, M., Maree, K. and Ellis, S. 2008. *A mathematics vocabulary questionnaire for use in the intermediate phase.* South African Journal of Education. 28:489-504.

Waghid, Y. 2002. *Democratic education: Policy and praxis*. Stellenbosch: Stellenbosch University Press.

Walla, P. 2015. *Effect of age and gender on children's reading performance: The possible neural underpinnings. University of Thessaly.* Argonafton and Filellinon. 38221 Volos. Greece

Wand, M.P. and Jones, M.C. 1995. *Kernel Smoothing*. London: Chapman & Hall/CRC. ISBN 0-412-55270-1.

Yamauchi, F. 2011. *School quality, clustering and government subsidy in post-apartheid South Africa*. Economics of Education Review, 30: 146-56.

Yen-Chi, C. 2017. *A Tutorial on Kernel Density Estimation and Recent Advances*. University of Washington

Zellner, A. 2001. *Simplicity, Inference, and Modelling: Keeping it Sophisticatedly Simple*. Cambridge University Press, Cambridge.

APPENDICES

Appendix A

FET [GRADE 10, 11 & 12] BAND DATASET

Gender	Age	Days_Absent	Examination_Score	SES
1	17	18	39	0
1	16	4	45	0
2	16	0	45	0
2	21	38	29	0
2	19	2	36	0
1	17	4	53	0
1	17	83	39	0
1	18	65	45	0
1	19	4	27	0
1	17	6	46	0
1	17	14	27	0
2	16	0	45	0
1	18	13	44	0
1	16	6	40	0
1	17	3	40	0
2	17	22	39	1
2	17	55	36	0
1	17	15	34	0
1	17	4	39	0
2	18	6	40	0
1	19	87	28	1
1	17	3	45	0
1	17	7	40	0
2	14	8	50	0
1	17	2	52	0
1	16	8	46	0
2	20	44	26	0
2	16	1	43	0
2	18	1	38	1
2	17	3	42	0
2	18	8	33	0
2	18	1	47	0
1	18	9	43	0
1	19	1	39	0
1	18	9	39	0
2	17	5	51	0
2	18	0	55	0
1	17	0	45	0
1	18	13	40	0
1	19	4	39	1
2	18	0	43	0
1	18	7	45	0

1	20	5	38	0
1	19	13	49	0
1	19	9	38	0
1	17	0	41	0
1	19	37	45	0
1	18	2	54	0
1	18	5	40	0
2	13	1	49	1
1	18	2	60	0
2	21	26	44	0
1	18	5	42	0
1	17	3	56	0
2	20	5	43	0
2	20	7	39	1
1	19	14	54	0
1	19	6	40	0
1	20	6	44	0
1	20	10	43	0
1	19	0	43	0
1	19	7	42	0
2	21	2	41	0
1	21	5	44	0
1	18	2	45	0
1	19	8	44	0
1	18	11	38	0
1	17	16	47	1
2	17	8	42	0
1	20	2	47	0
1	19	0	44	0
1	22	6	48	0
2	18	1	56	0
1	22	3	43	0
1	16	2	70	0
1	21	0	42	0
2	17	0	51	0
1	20	0	54	0
2	19	1	38	1
1	19	16	42	0
2	21	0	43	0
2	20	9	33	0
1	19	3	34	0
1	19	1	46	0
1	20	5	32	0
2	20	9	29	0
1	19	22	31	1
1	18	2	35	0
2	21	4	28	0
2	16	0	46	0
1	21	4	27	0
1	15	3	42	0

1	16	12	41	1
2	17	0	40	0
2	17	5	34	0
1	16	1	39	0
1	18	14	47	0
1	18	4	38	0
1	17	3	35	0
2	19	2	29	0
1	19	2	33	0
1	17	0	37	0
2	20	0	41	1
2	18	1	30	0
1	16	0	38	0
1	16	5	37	0
1	15	1	40	0
2	20	4	26	0
1	16	0	50	0
2	19	6	42	0
1	17	0	30	0
1	19	0	28	0
2	16	1	38	0
2	18	2	29	0
2	20	3	40	0
1	18	0	36	0
1	16	0	36	0
1	17	1	37	0
1	18	1	40	0
2	21	15	25	0
1	17	1	36	0
1	16	2	37	0
1	17	4	33	1
2	21	8	28	0
2	21	7	27	0
2	19	24	34	0
1	19	1	37	1
2	18	14	40	1
1	22	7	32	1
2	18	2	39	1
1	17	2	34	1
1	21	4	36	0
1	18	4	35	0
1	17	5	42	0
2	17	0	50	0
1	20	2	38	0
1	18	0	56	0
1	18	1	44	0
2	17	1	49	0
1	17	3	43	0
1	17	2	41	1
1	18	1	37	1

2	18	6	37	0
2	19	6	35	0
1	18	2	43	0
1	18	1	37	0
2	15	2	56	1
1	18	7	51	0
1	17	4	46	0
2	17	5	42	0
2	19	1	32	0
1	19	13	34	1
2	18	1	48	0
2	19	1	46	0
1	19	3	42	0
1	20	2	34	0
1	19	3	36	0
2	18	0	49	0
1	18	0	44	0
2	19	0	57	0
1	18	0	38	1
1	19	4	36	0
1	20	1	33	0
1	19	4	38	0
1	18	2	49	0
2	19	1	41	1
1	21	4	40	0
1	19	3	41	0
1	20	0	42	0
1	19	0	52	0
1	19	0	40	0
2	14	0	48	0
1	19	0	54	0
2	22	1	36	0
1	19	0	38	0
1	18	0	57	0
2	21	0	41	0
2	21	0	40	0
1	20	0	47	0
1	20	0	37	0
1	21	0	43	1
1	21	0	37	0
1	20	0	41	0
1	20	0	40	0
2	22	1	35	0
1	22	0	43	0
1	19	0	41	0
2	18	0	46	0
1	15	3	36	0
1	20	0	42	0
1	21	1	38	0
2	21	3	32	0

2	19	4	41	0
1	16	0	37	1
1	18	2	30	0
2	18	0	37	0
1	19	0	40	0
2	18	2	31	1
2	16	0	35	0
1	16	4	35	1
1	16	0	22	0
1	22	3	30	0
2	17	2	43	0
2	18	0	35	0
1	18	3	33	0
1	18	0	40	0
2	17	3	24	0
1	16	0	37	1
2	18	1	32	0
1	16	0	34	0
2	21	8	21	1
1	19	0	36	0
2	20	1	36	0
2	19	1	37	0
1	20	1	38	0
2	19	0	27	1
1	16	0	36	0
1	18	0	41	0
2	18	0	32	0
2	19	1	21	0
2	23	1	25	1
2	19	1	32	0
1	17	0	40	0
1	17	0	45	0
2	20	2	21	0
2	21	0	33	0
2	16	0	35	0
2	19	0	29	0
2	20	4	23	1
1	18	0	36	1
1	18	1	35	0
1	15	0	46	1
1	20	0	28	1
2	18	0	31	1
2	17	0	37	0
1	16	0	35	1
1	19	3	34	1
2	19	5	27	0
1	16	0	32	1
1	19	5	32	0
2	17	0	39	0
2	20	6	31	0

1	16	1	34	1
2	18	0	36	0
2	19	3	24	1
1	20	0	36	0
1	19	0	46	0
2	16	0	33	1
1	21	2	35	1
2	19	2	35	1
1	16	0	41	0
2	19	2	32	1
2	20	1	29	0
1	16	0	62	0
1	19	0	39	1
2	20	2	27	0
2	19	0	22	0
1	15	0	46	0
1	16	0	36	0
2	16	0	34	0
1	16	0	41	0
1	17	0	38	1
2	17	0	38	0
1	17	1	46	0
2	21	0	23	0
1	19	2	38	0
2	15	0	42	0
2	22	3	31	0
1	18	3	34	0
1	19	2	36	0
1	18	0	36	1
1	21	2	31	1
2	22	0	30	0
1	16	0	38	0
2	19	0	35	0
2	18	1	37	0
2	20	0	33	1
2	21	1	26	0
1	19	2	35	1
1	20	0	35	0
2	19	2	20	0
1	21	1	29	1
1	23	2	31	0
1	17	0	36	0
1	16	1	33	0
2	15	0	31	0
2	17	0	35	0
2	16	0	35	0
1	18	2	51	0
1	18	1	45	0
2	20	4	46	0
2	21	3	37	0

1	19	1	43	1
1	20	6	37	0
2	20	5	36	0
2	17	1	40	1
1	20	3	36	1
1	19	1	40	1
1	21	4	30	0
2	22	1	23	1
2	17	1	30	0
1	16	1	45	1
1	17	1	44	1
2	20	4	30	1
2	18	1	41	0
1	18	1	47	1
1	19	1	43	1
2	18	0	37	0
1	17	1	35	0
1	17	1	34	0
1	20	0	27	0
1	19	3	29	1
1	20	8	44	0
1	19	4	45	0
1	22	1	37	1
1	21	6	32	1
2	21	0	34	1
2	19	5	35	1
1	19	2	39	1
1	18	3	38	1
2	21	0	42	1
2	19	1	31	0
1	16	0	36	0
1	20	8	32	1
1	22	0	40	1
1	21	3	40	0
1	17	2	51	1
2	20	3	46	1
2	20	2	43	1
1	23	2	28	1
2	17	1	35	0
1	22	7	29	1
2	22	2	34	1
2	20	2	31	1
2	21	3	39	0
1	18	0	36	0
1	22	1	25	0
1	17	1	37	1
1	18	0	33	1
1	19	1	37	1
1	17	3	35	0
1	22	0	38	1

2	22	3	35	0
1	19	3	37	0
1	18	1	31	0
1	21	3	37	1
2	19	2	40	0
1	17	3	41	1
2	19	2	37	0
1	18	4	32	1
2	19	1	35	0
1	22	3	38	0
2	21	1	41	0
2	20	0	36	0
2	18	0	56	0
1	21	4	43	0
1	19	0	56	0
1	18	0	50	0
2	19	0	49	0
2	18	0	61	1
1	21	0	44	0
1	21	0	43	0
2	18	0	52	1
1	18	0	52	0
2	18	0	64	0
2	19	0	38	0
1	19	2	39	0
2	20	0	40	0
2	19	0	60	0
1	18	2	64	0
1	19	0	49	0
1	19	1	40	0
1	22	1	40	0
1	19	1	41	0
2	16	0	66	0
1	19	0	50	0
2	21	0	45	1
1	18	0	47	0
2	18	1	44	1
2	20	0	42	0
1	20	0	43	0
2	19	0	46	0
2	20	0	46	0
1	20	1	42	0
1	18	0	54	0
1	21	1	38	0
1	20	1	38	0
2	19	0	53	0
1	19	2	46	0
1	17	3	38	0
2	20	0	60	0
2	21	0	41	0

1	19	0	39	0
1	23	1	40	0
2	21	0	52	1
1	18	0	51	0
1	19	0	34	0
1	20	3	38	0
1	26	0	37	0
1	21	0	38	1
1	20	1	39	0
1	19	5	48	0
2	20	0	45	0
2	19	0	53	0
1	21	1	38	0
1	22	0	48	1
1	20	3	45	0

GET [GRADE 8&9 DATASET]

Gender	Age	Days_Absent	Examination_Score	SES
1	13	5	45	0
2	15	4	35	0
2	18	15	36	0
2	17	13	44	0
1	14	0	43	0
1	16	5	36	0
1	16	8	40	0
2	16	4	49	0
2	19	23	33	0
2	16	6	38	0
2	14	1	37	0
2	15	1	41	0
2	16	5	40	0
1	15	1	44	0
2	17	5	47	0
2	16	5	40	0
2	17	0	41	0
2	17	27	33	0
2	18	7	37	0
1	18	36	32	0
1	13	0	53	0
1	18	20	42	0
1	18	1	38	0
1	16	23	43	0
2	15	2	43	0
2	15	16	28	0
2	17	0	37	0
1	14	2	37	0
1	18	3	35	0
1	14	2	46	0

1	14	3	39	0
1	19	9	40	0
2	17	1	45	0
2	18	5	39	0
2	18	9	35	0
1	15	2	33	0
1	13	1	53	0
1	14	1	45	0
2	15	3	48	1
1	15	2	48	0
2	19	5	32	0
1	17	15	44	0
2	13	1	51	0
1	14	3	35	0
2	19	7	36	0
1	17	5	39	0
1	19	10	35	0
1	15	0	38	0
2	15	2	41	0
2	14	1	37	0
2	19	7	42	0
1	18	26	45	0
2	18	0	39	0
1	17	7	44	0
2	20	1	45	0
2	15	0	54	0
1	20	0	37	0
1	14	2	46	0
1	15	2	47	0
2	16	1	52	0
2	16	1	43	0
1	15	0	52	0
2	16	0	57	0
1	17	0	48	0
1	16	34	49	0
2	18	5	47	0
1	18	0	49	0
2	19	24	41	0
1	16	2	47	0
2	19	1	48	1
2	17	0	44	0
1	15	1	49	0
1	15	0	53	0
1	14	1	54	0
2	19	0	40	0
1	15	2	54	0
1	16	1	48	0
1	18	2	43	0
2	15	0	45	1

2	17	0	39	0
2	19	4	44	0
1	17	5	47	0
1	15	3	46	0
1	16	1	47	0
1	17	6	46	0
2	20	14	45	0
1	16	0	45	0
1	15	0	49	0
1	16	5	41	0
2	20	4	45	0
2	18	0	38	1
1	18	1	50	0
1	21	1	38	0
1	16	0	48	0
1	14	1	44	0
2	15	1	35	0
2	16	2	41	0
2	15	1	36	0
1	14	1	40	0
2	15	0	47	0
2	14	1	35	1
2	17	8	41	1
2	19	0	39	0
2	16	2	42	0
2	15	0	32	0
2	21	0	41	0
1	15	0	44	1
1	13	1	41	0
2	16	2	34	0
1	16	2	48	0
2	20	7	43	0
2	17	2	48	0
1	15	0	33	0
1	14	0	54	0
1	15	0	44	0
2	16	5	41	0
1	18	2	32	0
1	16	1	34	1
1	14	1	39	0
2	16	1	30	0
1	15	1	53	0
2	14	0	42	0
1	18	1	40	0
2	17	1	44	0
2	16	4	38	1
1	15	1	42	0
1	19	10	39	0

1	15	1	48	0
2	17	0	39	0
2	15	7	36	0
1	14	1	46	0
2	19	1	40	0
1	14	0	36	0
2	14	1	42	0
1	14	0	32	0
1	16	4	35	0
1	16	2	36	0
1	15	3	45	0
2	17	5	34	0
1	17	0	34	0
1	14	0	40	0
1	18	3	34	0
1	15	1	39	0
1	14	0	45	0
2	15	0	35	0
1	15	1	39	1
2	16	1	45	0
2	16	2	41	0
1	15	0	43	0
2	15	0	46	0
1	14	9	45	0
2	18	2	52	0
1	15	1	45	0
1	17	1	41	0
2	17	0	51	0
2	17	2	41	0
2	15	0	46	0
2	16	2	49	0
2	17	2	50	1
2	18	0	50	0
2	17	6	47	0
2	18	2	39	0
2	19	3	42	0
1	14	1	60	0
1	19	5	42	0
2	16	1	45	1
1	18	5	44	0
2	18	1	43	0
1	15	2	47	0
2	19	6	42	0
2	16	12	42	1
1	20	18	45	0
2	18	2	43	0
2	19	4	44	0
2	19	4	42	0
1	14	0	56	0

1	15	0	51	0
2	16	0	48	0
1	16	0	53	0
2	20	3	40	0
1	18	6	50	0
2	14	0	56	0
1	17	2	45	0
1	20	8	47	0
2	20	3	39	0
1	18	3	44	1
1	20	3	42	0
1	16	2	45	0
2	16	1	45	0
1	16	2	44	1
2	17	1	44	0
1	13	7	39	0
2	15	4	37	0
1	16	2	45	0
1	16	1	43	0
2	14	0	46	0
1	13	2	36	0
2	16	3	41	0
2	15	3	46	0
2	17	3	34	0
1	16	0	52	0
2	18	2	40	0
2	18	3	42	0
1	14	0	67	0
2	13	0	56	0
2	15	0	36	0
2	16	0	44	0
2	14	0	46	0
2	15	0	40	0
2	15	1	40	0
1	13	0	51	0
2	19	1	31	0
2	16	1	41	1
2	19	8	39	1
2	15	0	44	0
1	15	5	48	0
1	16	1	55	0
2	16	0	40	0
1	14	0	43	0
2	20	4	41	1
1	15	0	45	0
1	14	0	55	0
2	16	0	42	0
1	19	4	47	1

2	18	0	38	0
2	16	0	53	1
1	14	1	59	0
2	15	2	48	0
2	14	0	62	0
1	14	1	49	0
1	17	1	38	0
1	14	1	63	0
1	15	2	56	1
2	16	1	52	1
2	14	0	49	0
1	15	2	45	0
1	15	3	48	0
2	15	0	34	0
1	14	0	45	1
1	16	0	42	0
1	18	29	26	0
1	13	0	46	0
2	19	1	44	0
1	15	0	66	0
1	14	1	49	0
1	13	0	62	0
1	15	0	49	0
1	14	0	51	0
1	13	0	54	1
1	15	0	49	0
2	16	0	40	0
2	17	0	43	1
2	16	0	43	0
1	16	0	41	0
2	16	0	52	0
2	18	1	44	1
2	20	0	46	0
1	16	0	54	1
2	17	0	47	0
2	22	0	48	0
1	16	0	47	0
1	14	0	47	0
1	15	0	37	0
2	15	0	45	0
2	17	0	38	0
1	17	1	50	1
1	16	0	42	0
2	21	0	41	1
2	18	0	52	0
1	16	1	43	0
1	15	0	58	1
1	16	0	48	1

2	17	0	42	0
1	19	0	36	1
1	17	0	37	0
1	17	0	44	0
1	15	0	46	1
2	17	0	36	0
1	16	1	55	1
2	15	0	50	0
1	19	0	42	0
1	21	1	43	0
2	18	0	45	0
2	17	0	41	0
1	16	0	40	0
1	20	1	41	0
1	16	0	47	0
2	18	0	43	0
1	17	0	43	0
2	16	0	41	0
1	15	1	47	0
2	20	0	42	0
2	15	0	48	0
1	17	0	43	1
1	17	0	38	0
2	18	0	41	0
2	15	0	52	0
1	18	0	38	0
1	15	0	47	1
1	15	0	44	0
1	19	0	37	0
1	16	0	43	0
1	15	0	50	0
2	16	0	44	0
1	16	0	43	0
2	17	0	50	0
1	15	0	47	0
2	17	0	47	0
1	16	0	51	0
2	16	0	52	0

Appendix B

Data Coding

Grade

8 = Grade 8

9 =Grade 9

10 = Grade 10

11 = Grade 11

12 = Grade 12

Gender

1 = Female

2 = Male

School Type

0 = Public

1 = Private

SES

0=below poverty line

1=above poverty line

Appendix C

GET learners' frequency table

```
. tab Gender
```

Gender	Freq.	Percent	Cum.
Female	161	51.94	51.94
Male	149	48.06	100.00
Total	310	100.00	

FET learners' frequency table

```
. tab Gender
```

Gender	Freq.	Percent	Cum.
Female	248	60.93	60.93
Male	159	39.07	100.00
Total	407	100.00	

GET learner's descriptive statistics

Table 4.1: Descriptive statistics for GET learners, 2016

Variable	Obs	Mean	Std. Dev.	Min	Max
Age	310	16.283	1.872	13	22
Days absent	310	2.622	5.181	0	36
Examination score	310	43.803	6.463	26	67
SES	310	0.116	0.320	0	1

Source: Author's calculations using Stata version 14

GET learner's pairwise correlations

Table 4.2

	Gender	Age	Days absent	Examination score	SES
Gender	1.0000				
Age	0.2546	1.0000			
Days absent	-0.0147	0.2503	1.0000		
Examination score	-0.1698	-0.2919	-0.2533	1.0000	
SES	-0.0140	0.0365	0.0631	0.0797	1.0000

Source: Author's calculations using Stata version 14

FET learner's descriptive statistics

Table 4.3

Variable	Obs	Mean	Std. Dev.	Min	Max
Age	407	18.668	1.831	13	26
Days absent	407	3.518	8.590	0	87
Examination score	407	39.356	8.031	20	70
SES	407	0.228	0.420	0	1

Source: Author's calculations using Stata version 14

FET learner's pairwise correlations

Table 4.4

	Gender	Age	Days absent	Examination score	SES
Gender	1.0000				
Age	0.0543	1.0000			
Days absent	-0.0361	-0.0006	1.0000		
Examination score	-0.1165	-0.2396	-0.0926	1.0000	
SES	-0.0040	0.0603	0.0053	-0.1511	1.0000

Source: Author's calculations using Stata version 14

Appendix D

School picture