The state of the Eastern Cape schools in a period almost the second decade of democracy

Thobeka Ncanywa

ERSA working paper 486

December 2014

Economic Research Southern Africa (ERSA) is a research programme funded by the National Treasury of South Africa.

The state of the Eastern Cape schools in a period almost the second decade of democracy

Thobeka Ncanywa*

December 11, 2014

Abstract

This paper explore the state of the Eastern Cape schools by employing an education production function approach using the Ordinary Least Squares (OLS) and quantile regression techniques in 2013, a period almost twenty years into democracy in South Africa. The study benefited from the availability of Annual National Assessment (ANA) results from the examination directorate as a measure of schooling outcomes. In the education production function, scores from ANA were estimated against educator characteristics, school characteristics and community characteristics. The results of this study indicated that in the Eastern Cape learner performance is strongly influenced by educator quality, school and community characteristics. Keywords: education production function, Quantile regressions, annual national assessments

'*The Eastern Cape was the hub of education for black people for a long time, a province with a wonderful history of production of educated heroes and heroines of struggle; individual teachers, learners, parents and all community based organisations should play a leading and pivotal role in rebuilding the Eastern Cape Education; no individual or an organisation should be allowed to act in a manner that is against the spirit of rebuilding the Eastern Cape Education Department,' Ms Makgate.*

Parliament of the Republic of South Africa: April 2012

1 Introduction

The South African government transformed from a previously segregated education system to a democratic education system in 1994. The Eastern Cape department of education had to address high levels of inequality within the province that fused education departments from Ciskei, Transkei and former

*Department of Economics, University of Fort Hare, Alice, South Africa and University of Cape Town, Cape Town, South Africa. E-mail: TNcanywa@ufh.ac.za or Ncntho004@myuct.ac.za

Republic of South Africa, taking over from the poorly managed homeland system. The homeland system had schools of poor infrastructural quality such as shack schools, mud schools and prefabricated schools. Some areas had no schools at all. In spite of the challenges inherited from the apartheid system, the democratic government had to acknowledge the right to education conferred on South Africans by Chapter 2 of the Constitution (RSA, 1996). The Bill of Rights gives all South Africans the right to basic education, including adult basic education.

Advances have been made in funding and equity in the schooling system since the advent of a democratic order in the country in 1994 (DoE, 2003). Initiatives have been put in place to improve the qualifications of educators, increase in operational expenditures such as provision of infrastructure, stationary and learner teacher support material. Considering this expenditure in education, it is essential that schools improve in quality, productivity and accountability (DoE, 2007). The extent of productivity in education institutions can be measured by education production functions. In educational microeconomics, the education production function or frontier is an education process that transforms input indicators into outcomes.

However, in spite of the concerns around the performance of learners and what inputs determines the schooling outcomes, little is known about the Eastern Cape education production function. The main focus of the paper is to find out what determines Eastern Cape schooling outcomes as measured by the Annual National Assessment (ANA) results. The ANA is an examination tool that is used to measure the level of performance of grade one to nine learners in Numeracy and Language. The Eastern Cape represents an ideal focus of attention as a poor and rural province getting a larger share of the national budget allocation in basic education but producing disappointing learner outcomes. For example, 95% of public schools in 2013 are 'no fee schools' and learners are in a nutrition programme.

Findings on the link between inputs into educational production and outputs yielded contradicting results. For instance, looking at the work of Hanushek (1989), there is little or unconvincing evidence of a link between inputs and the quality of schooling output. However, another set of studies does establish a statistical link between school quality and schooling outcome with Glewwe and Kremer (2002) being often cited. A number of competing reasons have been advanced for this divergence of findings, such as bias caused by unobservable variables and estimators. The study furthers the understanding of the link between inputs and outputs in Eastern Cape schools by examining what constitute schooling outcomes using Ordinary Least Squares (OLS) and quantile regressions.

Based on the above conflicting studies, the study will contribute to knowledge by finding the association of inputs with schooling outcomes in the Eastern Cape education system. The study attempts to provide estimates of the determinants of schooling outcomes in the period covering almost twenty years of democracy, through the availability of various schooling data sets. The study adopts an education production function approach. Section 2 below provides

institutional background, section 3 and 4 provide theoretical and empirical literature respectively. Section 5 specifies the model and section 6 provide data sources and descriptive analyses of the data. Section 7 provides an overview of the estimation techniques, followed by section 8 with results and discussion and then section 9 concludes.

2 Institutional background

The Eastern Cape Province is characterised by different types of schools such as village, farm, township and town schools with a large number of village schools in the former Transkei (map1). There are twenty three school districts demarcated across the province with a district director reporting to the provincial head office (PECE, 2012). The districts are grouped into three clusters that are managed by cluster chief directorates in the head office. The clusters are grouped in terms of geographic demarcation as, **Cluster A**: Libode, Lusikisiki, Maluti, Mbizana, Mt Fletcher, Mt Frere and Qumbu; **Cluster B:** Butterworth, Cofimvaba, Dutywa, Lady Frere, Mthatha, Engcobo and Sterkspruit; **Cluster C**: Cradock, East London, Fort Beaufort, Graaff-Reinet, Grahamstown, Queenstown, Uitenhage and Port Elizabeth (Map 1).

Cluster A and Cluster B are found at the eastern side of the province where a higher percentage of schools is located in villages and in small towns. These clusters are 99.9 % rural and consist of schools that belonged to the former Transkei homeland in the apartheid era. The region is underdeveloped in infrastructure and overpopulated with learners who belonged to single headed households or have no parents at all. Cluster C is a combination of rural, farm and urban schools, where East London and Port Elizabeth are developing towns and have seaports, railway stations and airports.

Cluster C has a number of former model C schools found in each district and three Universities while Cluster A and B have one University. Graaff-Reinet, Port Elizabeth, Uitenhage, Cradock and Grahamstown have no village schools. Schools situated in villages from the former Transkei and Ciskei has challenges of a large number of schools with mud, inadequate or prefabricated building structures. Furthermore, table 1 tabulate the performance of ANA district and cluster in grade 3, 6 and 9 for Mathematics, Home language and First Additional language. There is a drastic decrease in Mathematics performance across the grades.

3 Theoretical underpinnings: Conceptualised facts on education production functions

In an attempt to understand the causal relationship between educational inputs and academic achievements a useful assumption is that each household maximizes a utility function, subject to constraints (Glewwe & Kremer, 2006). Examples of constraints are the production function for learning, the impacts

of years of schooling and skills obtained for future earnings, budget constraints or an agricultural production function. This study adopts a production function approach and therefore focuses on the production functions for learning constraint. Due to lack of information on some characteristics of structural relationships of production functions such as learner and parent information the structural model is not easily measured (Boardman, Davis, & Sanday, 1977) Therefore, a causal relationship which is a reduced form equation in the following form can be used[1]:

$$A = f(S, Q, H) \tag{1}$$

A is a scalar representing achievement which can be captured by standardised test scores or grade attainment. S is a vector of school characteristics usually captured by learner educator ratios, location where the school operates and so on. Q represents educator characteristics usually captured by educator experience and educator qualifications. H is a vector of household characteristics which act as control for socio-economic characteristics and are captured by quintile category here. The study attempts to measure $A'(.)$; which provides the total derivative of A with respect to each of the individual variables defined within vectors.

According to Hanushek (1989), the most appropriate and useful approach is the production function approach which is also called input-output or cost-quality approach. The production function for schools focuses on the relationship between school outcomes and measurable (observable) inputs into the educational process. Learner outcomes can be measured at discrete points in time because the education process is cumulative. This means that the current levels of achievement are affected by the past inputs. Educational systems have no single defined production function, and no well defined indicators of input and output. In most studies of the education production function, the measure of input and output is limited by the availability of data. Therefore, various educational outcomes can result from a variety of different combinations of inputs.

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. Therefore, various school and nonschool inputs are used in frontiers to produce multiple outputs (Hanushek, 1989). Examples of output found in literature include academic performance; skills, attributes and values that favour workplace and social integration; communication and interpersonal skills, respect for the environment, physical fitness; and political, social and personal responsibility, grade repetition rates, or dropout rates. As one purpose of education is to develop the student's basic cognitive skills, output is often measured by the scores in tests, by the number of students passing grade twelve, by student's success in gaining admission to institutions of higher education, or by student's future earning potential. The majority of existing studies mainly concentrate on academic performance in terms of test scores as output (Ray, 1991)[2]. Hanushek explain that test scores are indicators of future success either

[1] (Glewwe & Kremer, 2005)

[2] (Giménez, Prior, & Thieme, 2006); (Schwartz & Stiefel, 2004); (Thieme, Giménez, &

in schooling or in the labour market (Hanushek, 1989).

On the other hand, inputs which are also referred to as instructional expenditures can be grouped as school, educator; student, household and peer characterised (Hanushek, 1989). School and teacher inputs can be controlled, monitored and evaluated by policy makers in a country. School inputs include school organisation such as class size typically measured by the student-teacher ratio, expenditure per pupil, facilities, buildings and administrative expenditures. Smaller classes meaning smaller student teacher ratios should improve student learning. Educator inputs such as educator quality is measured by educator education level, years of experience, race and sexual orientation (Borge & Naper, 2006). Higher levels of educator education and educator experience cost more and should be beneficial to student outcomes. Household or family inputs include parental education, income and family size. Student inputs involve the inner capabilities and individual motivation of students which can be the unobservable variable in education records. A number of researchers found that socioeconomic and environmental factors such as peer factors significantly affect achievement scores (Hanushek, 1989). Generally speaking, higher educator salaries, better facilities and better administration should lead to better student performance.

4 Empirical literature on education production function approach

The work on education production function studies builds on related literature that dates back from Coleman 1966 and Hanushek 1986. Literature on developed and developing countries has yielded mixed results on the relationship between instructional education inputs and student performance (Borge & Naper, 2006); (Addonizio, 2013). Researchers have used different estimation approaches for production frontiers and some researchers recommend the use of more than one approach to formulate strong conclusions (Chakraborty, Basudeband, & Lewis, 2001). There is abundant information of empirical literature on these different approaches (Hanushek, 1989)[3]. The challenge for both educators and economists is how to mix the inputs in the right proportions to achieve the most efficient outcome. The other challenge concerns the output that we should measure. Education economists need to find solutions to the controvential and complex education system and determine what works to achieve desired schooling outcomes, but are faced with challenges. To address the problem of missing counterfactual and addressing the problem of selection

Prior, 2011)

[3](Ray, 1991); (Deller & Halstead, 1994); (Case & Deaton, 1999); (Fedderke & Luiz, 2002); (Vignoles, Levcic, Walker, Machin, & Reynolds, 2000); (Chakraborty et al., 2001); (Koenker & Hallock, 2001); (Giménez et al., 2006); (Woodbury, Dollery, & Prasada, 2008); (Berg, 2013); (Gustafsson & Taylor, 2005); (Bhorat & Oosthuizen, 2006); (Crouch & Mabogoane, 1998); (Taylor, 2009); (Taylor, 2011); (Badri & Mourad, 2012); (Borge & Naper, 2006); (Jacob & Lefgren, 2004), (Ludwig, Miller, Malme, & Morse, 2007).

bias in programme targeting and participation, any of the following methods can be used. The estimation approaches include Randomised evaluation, matching methods, double differencing, Instrument Variables, Regression Discontinuity Design and Quantile regressions among others.

Various researchers examined education productions functions in developed and developing countries and yielded contradicting results. For example, Hanushek (1986) found an insignificant relationship between school expenditures and student performance in developed countries (Hanushek et al., 1990). The researcher revealed that the basic determinants of instructional expenditures used by the study in a district were teacher experience, teacher education, and class size. On the contrary, Fuller (1986) found a significant relationship between instructional inputs and student performance. Other studies like Hoxby (2000); Borge and Naper (2006) conducted studies on efficiency of schools and are in line with Fuller because they have found a link between education inputs such as school resources and outcomes such as pupil achievements.

Bhuchinsky (1994) examined the changes in wage structure and schooling return in the United States using quantile regression analysis. One group and 16 group models were employed, where it was found that the return of schooling differed across quantile of the wage distribution (Bhuchinsky, 1994). Eide and Showalter also used quantile regressions to investigate the relationship between school quality and performance. The pupil teacher ratio, district per pupil expenditure, the fraction of teachers with an advanced degree, enrolment at school and the length of the school year were used as input variables in the regression. The quantile regression results revealed different school quality effects at different points in the test score gain conditional distribution (Eide & Showalter, 1998). Angrist and Lavy (1999) used an instrumental Regression Discontinuity Design (RDD) to find out that a reduction in class size positively influenced schooling outcomes (Angrist & Pischke, 2008).

Aslam and Siddiqui (2003) analysed the determinants of pupil achievement of middle school students using purpose built data collected on 8th grade pupils in private and government schools in urban and rural Lahore (2002-2003), Pakistan. The OLS method supplemented with the Instrument Variable (IV) estimation to control for endogeneity bias and the Heckman two-step to control for the endogenous sample selection of students into private and government schools was used. The results indicated that measures of teacher quality such as teachers' education, training and pay were poor indicators of quality. Other school variables, such as pupil-teacher ratio, peer group variables and school resources, which are more amenable to policy were significant determinants of pupil achievement but differ across subject areas and by school-type. Finally, both home background and school-related factors were found to be important in explaining pupil achievement in Pakistan (Aslam & Siddiqui, 2003). McEwan and Carnoy (2000) used RDD in Chile to measure the impact of school feeding schemes and found that the calorie content meal has no effect on student achievement (McEwan & Carnoy, 2000).

A number of South African researchers have analysed the education production functions using various approaches. For example, Crouch and Mabogoane

(1998) conducted a bivariate regression analysis of factors that influence learner achievement in South African schools. Their study revealed the importance of teacher quality, computers and media centers as significant factors that contribute to student success (Crouch & Mabogoane, 1998). Case and Deaton (1999) found a link between pupil teacher ratios and South African schooling outcomes towards the end of the apartheid rule. The allocations showed marked disparities in average class sizes and significant effects on enrolment, education attainment and test scores for numeracy (Case & Deaton, 1999). Fedderke and Luiz (2003) used time series data from 1910 to 1993 and employed a Vector Error Correction Model to estimate the South African education production function. From their findings that education inputs matter in determining education output, they also found out that political instability and institutional dispensation significantly affect the education production function. Van der Berg and Burger (2003) employed an OLS regression analysis in a study to investigate what causes efficiency in the education system of the Western Cape. The production function used in the study indicated a significant need for managerial interventions, shortages in complementary teaching materials in improving efficiency of the education system in the province (Van der Berg & Burger, 2003).

Bhorat and Oosthuizen (2006) analysed what determines grade twelve pass rates in the South African schooling system. This study employed OLS regression and quantile regression techniques 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 pupil-teacher ratio was insignificant, physical resources were irrelevant, infrastructure was crucial and teacher characteristics were a key priority focus for any policy programme aimed at improving Grade 12 performance levels (Bhorat & Oosthuizen, 2006). One redeeming feature, from a policy perspective, of the results is that household vulnerability is a weak predictor of performance.

Van der Berg and Louw (2006) used an OLS and Hierarchical Linear modelling to analyze the performance of the South African schooling system. The findings were that the South African Mathematics performance was worse than expected relative to other African countries. This was a consequence of the poor ability of a large part of the schooling system to convert advantages in school resources, parent education and socioeconomic status into better school Mathematics performance (Van der Berg & Louw, 2006). Gustafson (2007) conducted a research on the South African school production function using 2000 Southern and Eastern African Consortium for Monitoring Educational Quality (SACMEQ) data. The study employed ordinary least squares (OLS) and hierarchical linear production function models (HLM). The findings of the study indicated the importance of physical infrastructure, textbooks and nutrition budgets. The study further highlighted and emphasised that correct allocation of teaching and management time in schools, less learner repetition, and better teaching methodologies stand out as important school and classroom

management imperatives (Gustafsson, 2007). Van der Berg (2007)[4] examined a South African education production function using regression analysis. The findings indicated the important role played by better teacher resources and school management strategies in school performance if certain efficiency conditions were met. Van der Berg (2006) argued that 'school resources do not necessarily make a difference but that the ability of schools to convert resources into outcomes is the crucial factor.' Taylor (2011) argued that 'South African school functionality or efficiency remains something of a black box'.

5 Model specification

The study adopts a reduced form of the production function approach and is written as (Glewwe, 2002)

$$A = \beta_0 + \beta_1 S + \beta_2 Q_i + \beta_3 R + u \tag{2}$$

Where A is observed output at establishment i represented by ANA pass rates; S educator characteristics, Q school characteristics and R community characteristics *are* vectors of inputs; and β *are* vectors of parameters which describe the transformation process of inputs to output. u reflects measurement error in A, S, Q and R; unobserved aspects of explanatory variables on cognitive skill acquisition.

ANA pass rates were used to measure the level of performance of learners in the General Education and Training (GET) band in Numeracy (Mathematics) and Language. ANA pass rates were provided as an average by grades covering grades {1, 2, 3, 4, 5, 6 & 9} and measures cognitive skills learnt by learners and therefore represented schooling outcomes. Educators' characteristics were represented by educator experience and qualifications. School characteristics were represented by educator learner ratio, language of instruction, former department of education, whether a school has a boarding school or not. Community characteristics were represented by the quintile category of schools.

Educator experience was the numbers of consecutive years' educators were employed. Educator qualification was represented by REQV, where REQV 10 (the referent variable), 11 and 12 represent unqualified or under qualified educators. REQV 13 represented educators who obtained grade 12 qualifications and 3 years training such as a teacher diploma; REQV 14 referred to grade 12 qualifications plus 4 years training such as a Bachelor of Pedagogic degree. REQV 15, 16 and 17 was an increase in training such as Honours and Masters[5] Degrees.

The ratio of educators to learners per school was calculated by dividing the number of learners by the number of educators and was used as proxy for class size in schools. A dummy for the quintile category referred to the poverty status

[4] (van der Berg, 2007); (Taylor, 2011)

[5] Further explanation of REQV can be obtained from the Department of Education website, as educators are placed in a REQV category according to the qualifications obtained by the educator.

of the school and the community in which the school was placed. Quintile 1 (the referent variable) referred to the poorest schools and quintile 5 referred to well resourced schools. Schools, especially at the foundation phase, choose a language to use for teaching all subjects. Government policy allows learners to be taught in their mother tongue in the foundation phase, but English and Afrikaans are the only languages of instruction in South Africa in grades higher than grade 3. The language of instruction dummy was created as Afrikaans (the referent variable), English, other South African languages and IsiXhosa.

ANA examinations assessed Mathematics and two languages per learner, where one language was a home language and the other was the additional language. The former department of education before 1994 in the apartheid regime were Cape of Education Department which is the referent, followed by Ciskei, Department of Education and Training, House of Assembly, House of Delegates, House of Representatives, KwaZulu Natal Education Department, New Education Department, to be updated and Transkei.

The Eastern Cape Education was grouped into districts namely, Butterworth the referent, Cofimvaba, Cradock, Dutywa, East London, Fort Beaufort, Graaff-Reinet, Grahamstown, King William's Town, Lady Frere, Libode, Lusikisiki, Maluti, Mbizana, Mt Fletcher, Mt Frere, Mthatha, Ngcobo, Port Elizabeth, Queenstown, Qumbu, Sterkspruit, To be updated, Uitenhage. Location represented whether the school was in a rural or urban location, where rural was the referent variable. The dummy for boarding gave information on whether the school had hostels for accommodation or not.

6 Data and descriptive analysis

The dataset used to determine the Eastern Cape education production function is combined from datasets obtained from different directorates in the department. The South African National Department of Basic Education (DBE) website provided the master list of all ordinary schools. The master list contains information on the school's characteristics which include the school's unique identifier (natemis), whether the school is public or independent, urban or rural; the provincial demarcation such as districts or municipalities or ex-departments, learner enrolments and number of educators. The Province of the Eastern Cape Department of Education (PECE) provided Education Management Information Systems (EMIS) in the EMIS directorate. The EMIS information provided information similar to the master list including information on educator quality. The Examination directorate provided annual national assessment (ANA) scores. The study employed a merged dataset from the mentioned sources.

Given the nature of data used in the study, the cross sectional data for 2013 and the reduced form production function, it is important to acknowledge some drawbacks suffered by the data within the production function context. For example, the unit of analysis was at school level not at individual level, and therefore constrained the validity of the estimates derived from the available variables as there was some bias. For instance, in educator characteristics there

was no direct link of educators to specific classrooms resulting in a measurement error bias. Lacking learner and parent information again suffered from omitted variable bias. However, it was assumed that parents had no choice but to enrol their kids in the available schools. There is some selection bias through the existence of learners who were not available at school when the ANA examinations were written, through absenteeism or dropping out from school. Therefore the estimates were regarded as causal effects in estimating conditional quantile functions with a large sample of Eastern Cape schools and are not necessarily indicative of their true impact on education performance.

ANA pass rate which represented the dependent variable had the mean percentage of 32.9, the median of 35.3 and standard deviation of 23.6 (table 2). This implied generally low performance in the Eastern Cape schools when measured by ANA examinations. The data further suggested a mean of 15.6 median of 16 years of educator experience and standard deviation of 10.4. This indicated that the province had a significant number of educators with more than ten years experience. The age of educators presented a mean of 31 and median of 32 with standard deviation of 9.7. The mean that is less than the median indicates that there are no outliers and no skew on the data (Gujarati, 1992). Therefore the pass rate, educator experience and age of educators assumed a normal distribution in the analyses. The mean value for learner enrolment of 381.6 with standard deviation of 311.4 and median of 287 indicates outliers and skew to the data as the mean is greater than the median. The mean for educator learner ratio is 1:32.3 with standard deviation of 12.6 and median of 31.88. The table indicates the mean qualification of educators as REQV 12 with standard deviation of 1.9 and median of REQV13. The mean quintile was 2.4$\sim$ 2 with standard deviation of 1.42 and median 2.

About 56.4% of the province was rural and 41% was urban. About 73% of schools were found in the former Transkei, Kwazulu Natal and Ciskei ex-departments, which were constituted with high volume of villages (map1). Former Transkei department displayed a dominant share of 64.1%, followed by the former Department of Education and Training with 11.17% and Ciskei with 9.4%. The apartheid legacy was seen in the language of instruction where English was used in a dominantly Xhosa speaking province. The data indicated a share of 43% for ANA results in Mathematics, a share of 37% for English, 17% for IsiXhosa and 3% for Afrikaans. About 97.6% of schools were public schools and less than 3% were independent schools. Of these public schools 95% were allocated quintile 1, 2 and 3 in 2013, implying no fee schools and benefiting from nutrition.

Some land is privately owned by farmers or churches but within the farms there are public schools. This has an impact in the school governing body, as the owner of land should form part of this legal entity even if he has no child attending the school. The Eastern Cape has a share of 2.6% of privately owned land against the 97.4% of publicly ownership of land. The share distribution of public schools and public ownership of land indicate that the government has huge responsibility in schooling in the Eastern Cape. The availability of telephone was used as proxy for infrastructure and represents the flow of communication

into the schools, where 99.6% indicated the availability of a telephone in the school.

The province has a large share of combined schools (53.6%) followed by 40.19 stand alone primary schools and 5.98% secondary schools. Data to represent the innate ability of learners was not available; therefore learner characteristics were unobservable in the analysis. There was no available data on detailed infrastructural development at school level, and even the data available at provincial level did not indicate any change in infrastructural development[6] of the province. The above descriptive evidence suggests that learner, educator, school and community characteristics display distinct statistical characteristics. Therefore, it was interesting to investigate how these variables simultaneously affected learner performance.

7 Estimation

The analysis began by looking at the Kernel density distribution of the dependent variable according to the ex-department, quintile and subject. The kernel density estimates of ANA are presented in figure 1 below and are estimated considering different categories. Kernel density is a non-parametric way of estimating the probability density function of a random variable. Through the Kernel density estimation, data is smoothed and inferences about the population are made based on the finite data sample (Parzen, 1962). The Kernel density distribution features of variables should reveal a normal distribution; otherwise the model may yield partial information about the estimates of the explanatory variables.

The kernel density graphs indicated a distribution around the zeros implying data with a huge number of schools with zeros. The fact that the dependent variable does not behave the same way in different percentiles across ex-departments, quintiles and subjects suggests supplementing the OLS with quantile regression analysis so as to obtain robust results. Furthermore, data used at school level was aggregated and suffered biases. To correct for these biases and add to robust of the results, an OLS should be supplemented by other methods (Hanushek et al., 1990), hence the quantile regression.

Quantile regression refers to the generalized case of least absolute deviations estimator. According to Maddala (2001) 'Quantile regression involves constructing a set of regression curves each for different quantiles of the conditional distribution of the dependent variable, where the dependency of y on x is taken from its tails.' This set of regression estimates will provide a more detailed analysis of the entire relationship between the dependent and independent variables than a standard regression model would (Madala, 2001). OLS is a sufficient regression method in the context that the dependent and independent variables followed a bivariate normal distribution but if data suffer measurement errors and data revisions, it will lead to biased estimates. The advantage of using quantile regression rather than simple OLS regression only is that the

[6] See chapter 2 of the main thesis for provincial infrastructure analysis

stochastic relationship between random variables can be portrayed much better and with much more accuracy (Buhai, 2004).

The quantile regression provides a sense of the kind of impact that explanatory variables have on schools that are at different points in the distribution of outcomes. It provides a guide to policymakers on how effective expenditure patterns are for schools with different pass rates. Estimators of quantile regressions are invariant to affine transformation and monotone transformation of the outcome and the explanatory variables. The equivariance property allows for the estimation of β without any distributional assumptions. There is no closed form of the estimator. Estimates are robust to changes of the outcome (Koenker & Bassett, 1978).

Quantile regression is a statistical procedure intended to estimate conditional quantile functions. The quartiles divide the population into four segments, the quintiles divide the population into five parts; the deciles into ten parts with equal proportions of the reference population in each segment. The quantile is a general case of these divisions (Koenker & Hallock, 2001). In quantile regression, quantile of the conditional distribution of the response variable are expressed as functions of the observed covariates. Quantile regression analysis shows the equivariance properties and the joint asymptotic distribution to permit a natural generalization to the linear model of certain well-known robust estimators of location. The estimators of the regression quantile minimize the sum of absolute residuals (Koenker & Bassett, 1978).

According to Koenker & Bassett (1978) the classical model of quantile regression is based on ordinal quintiles. The ordinal quantile is $F(y) = P_r(Y \leq y)$, then for $\tau \epsilon [0;1]$ the τ^{th} quantile of Y is $Q(\tau) = \inf\{y : F(y) \geqslant \tau\}$. The median is then $Q(\frac{1}{2})$, the first quartile $Q(\frac{1}{4})$, the first deciles $Q(\frac{1}{10})$. 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: $R(e) = \sum_{i=1}^{N} \rho_\tau(y_i - e)$. Considering a quantile simple model $y_i = \beta_0 + x_i\beta_1 + u_i$ the conditional quantile function of y is $Q_y(\tau/x) = \beta_0 + x\beta_1 + F_u^{-1}(\tau)$, taking the general approach at different quintiles in the distribution, the regression quantile estimated minimizes the equation:

$$\min_{\beta \epsilon R^k} \left[\sum_{i \epsilon \{i; y_i > X_{i\beta}\}} \theta |y_i - X_i\beta| + \sum_{i \epsilon \{i; y_i > X_{i\beta}\}} (1 - \theta |y_i - X_i\beta|) \right] \tag{3}$$

The dependent variable is y_i, X_i is $kx1$ vector of explanatory variables and β is the coefficient vector. The analyses further estimated the impact of the explanatory variables on the differences in pass rates by estimating the inter-quantile regressions set at $90^{th} - 10^{th}, 90^{th} - 50^{th}$ and $50^{th} - 10^{th}$. The inter-quantile approach took the form:

$$Q_\theta(y_i) - Q_{\theta'}(y_i) = (\alpha_\theta - \alpha_{\theta'}) + (\beta_\theta - \beta_{\theta'})X_i \tag{4}$$

Where Q_θ and $Q_{\theta'}$ referred to the specific quantiles for y_i the dependent variable.

8 Results and Discussions

The study ran the OLS estimates in conjunction with the quantile regressions. The quantile regressions allowed estimation at different points in the distribution of the dependent variable. The results were reported for the OLS and θ *was set* at 0.1; 0.25; 0.5; 0.75, 0.9. Most of the 90^{th} -10^{th} and 90^{th} -50^{th} percentile covariates were insignificant and are therefore not reported in this paper (table 3).

Educator experience was positive and significant in all specifications indicating a strong positive effect on pass rates. A strong relationship was also seen when educator experience was squared (table 3). This is consistent with the studies that have investigated the effect of educator experience on learner performance (Glewwe & Kremer, 2005); (Shepherd, 2011). The more experienced the educators are the more they can handle the changing curriculum and dynamic teaching and learning techniques.

The educator qualifications which were represented by the REQV category from REQV 10 to REQV 17 were reported in most specifications with mixed signs and significant covariates (table3). Significant covariates were generally negative in lower REQV's up to REQV 12 and the REQV 14, at ?=0.1 were positive and significant except for the negative REQV 12 and 16. At REQV 15 and 16 all significant estimates were generally positive. REQV 17, the least frequent category (0.85%) indicated negative and significant results except for the positive and significant results in 10^{th} percentile. The most observed outcome was that learner performance was positively affected by the educators in REQV 13, 15 and 16; and most REQV's in the 10^{th} percentile.

The educator qualification results were consistent with the findings of Shepherd (2011); and Crouch and Mabogoane (1998) that qualification was significant at higher REQV. However, Eide and Showalter (1998) found negative and insignificant results when they considered the effects of an advanced degree on pass rates, which is consistent with REQV 17 in this study. The need for higher REQV raises concerns as the province has about 5% of educators in REQV 15 and 16; and a high percentage (52%) of under-qualified educators in REQV 10, 11 and 12. This is consistent with what Glewwe (2005) points to when he refers to the low quality of education provided in schools in developing countries, as learners learn less in school than the curriculum states. The mismatch is due to the shortage in qualified educators or better trained educators which result in poor teaching and remedial education techniques.

The educator experience and educator qualifications provided for a powerful policy message as they have a direct influence on educator payment. Therefore, a system that strengthens the links between educators' pay and learners' performance could be beneficial for all stakeholders. In spite of the importance of these variables on performance, Occupation Specific Dispensation (OSD) which is the educator payment system in South Africa was not fully implemented in the Eastern Cape as items such as rural allowance, scarce skill allowance and performance related pay were not implemented and therefore OSD is a weaker educator incentive programme. Moreover, the department should strengthen the

existing programmes to improve educator qualifications such as the Advanced Certificate in Education (ACE), a mathematics, science and technology programme to focus in the FET band subjects and provision of remedial education programmes.

In this study, the educator learner ratio at levels yielded mixed estimates across specification. When non-linearity was controlled by squaring this variable, it turned out that the OLS and lower quantiles up to the 50^{th} percentile favoured large classes and high performing schools in the 75^{th} to the 90^{th}percentile favoured small classes (table 3). The educator learner ratio served to increase the gap in relative performance between the 50^{th} and the 10^{th} percentile. These results are in line with what is found by the South African scholars that the educator learner ratios differ by quantile and by racial group (Crouch & Mabogoane, 1998)[7]. However, the results contradicted with some studies that differed in foci and data coverage, such as Eide and Showalter (1998), Angrist and Lavy (1999). The educator learner ratio generally gave different results in literature as informed by factors such as location and race (Hanushek et al., 1990).

The quintile category of a school is a variable that is under the control of policy makers to alleviate the poverty status in schools, and therefore is a proxy for socioeconomic status or community characteristics. The policy context of this variable is viewed as the amount of money given to schools per learner, provision of nutrition programmes and non- payment of fees by parents. The results generally indicated large and significant magnitudes across specifications (table 3). Very poor schools (quintile 2) were negative and significant in low performing schools and yielded different insignificant results elsewhere. Quintiles 3 to 5 were significant determinants of pass rates for low and middle performing schools, and were significantly strong to increase the dispersion on pass rates. This positive impact on Eastern Cape schools implied that the democratic government's redress of poverty was a welcome development which needed to continue. The results contradicted McEwan (2012) who used a Regression Discontinuity Design and found no effect on Chile's caloric content feeding scheme. However, the change from caloric content feeding scheme to nutritious meals was recommended which in the South African context is the nutrition programme.

The language of instruction, which is usually the mother tongue at the foundation phase and English or sometimes Afrikaans in the former model C schools (DoE, 2010), was generally negative and significant with reference to Afrikaans (table 3). This indicated that Afrikaans learners, who were taught and examined in their mother tongue performed better than other learners who had no choice in being instructed in English. It should be noted that even former model C schools were populated by non-Whites as a large population of middle class households enrolled their children in these schools. The results of this study indicated that the Eastern Cape education system still embodied elements of the colonial system as English was used as a national language in a predomi-

[7] Case and Deaton (1999); Fedderke and Luiz (2002); Berg and Burger (2003); Bhorat and Westhuizen (2006) and Shepherd (2011)

nantly Xhosa speaking province. This challenge could be reduced by providing teaching and assessment in both English and the mother tongue as scholars such as Marshall (2011)[8] and Patrinos and Velez (2009) had used different estimation techniques and found a significant impact of the language of instruction in Guatemala. Durán, Roseth, & Hoffman (2010) also found a positive and significant impact of the provision of bilingual education in learner outcomes in Spain.

Assessing learners in Mathematics generally indicated negative and significant results indicating that assessing learners in Mathematics has a negative association with learner performance. The single centrally set curriculum resulted to distortions in the education system, as it accommodated elite learners while leaving many learners behind. Glewwe and Kramer (2005) referred to this as a mismatch between the curriculum and the typical learner, as the system does not cater for remedial interventions and leave many learners behind. The researchers further indicated large gaps in Mathematics performance between developed and developing countries, which reflect differences in family backgrounds and low school quality in developing countries (Glewwe & Kremer, 2005).

Another set of variables that were robust across the specified models were control variables such as location, boarding school, ex-departments and districts. A striking outcome on rural schools, which is a referent to location, was that they performed better than urban schools on the ANA. In fact, this is in line with the Minister of Basic Education's remark that the Free State is rural and yet performs well in grade 12 examinations (DBE, 2012). The presence of hostel facilities in boarding schools was generally positive and significant with high estimates in magnitude. This may be some evidence of the high quality of education provided in these schools and better quality schools. The Cape Education department (White learners), the referent variable, performed better than other ex-departments with the exception of the House of Delegates (Indian learners). This reflected the legacy of racially based schooling system which favoured White and Indian learners in the Eastern Cape, who constituted about 3.5% in our sample (see table 2).

Cluster dummy was omitted because of multicollinearity, and districts were used, where Butterworth district was used as a referent variable. Cradock, East London, Grahamstown, King William's Town, Lusikisiki, Maluti, Mt Frere, Port Elizabeth, Queenstown and Qumbu performed better than Butterworth. It was interesting to find out that Uitenhage and Fort Beaufort were performing worse than Butterworth while they were in cluster C which seemed to be performing better than other clusters. Lusikisiki, Maluti and Mt Frere are in cluster A, which favoured the conclusion that schools in cluster C were the best performing schools followed by cluster A schools.

[8] (Marshall, 2011); (Patrinos & Velez, 2009); (Durán, Roseth, & Hoffman, 2010)

9 Conclusion

The research determined the Eastern Cape education production function using Annual National Assessment data as a measure of schooling outcomes and input data that included educator, school and community characteristics. In order for the analysis to have more robust results, the OLS was supplemented with quantile regressions as estimation techniques to find what determined Eastern Cape schooling outcomes. The analysis revealed a number of interesting results. Firstly, the educator characteristics significantly explained that schooling outcomes based on more educator experience and higher educator qualifications had a positive influence on learner performance. These two characteristics were also determinants of educator payments. This outcome provides the powerful policy message 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. 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.

The educator learner ratio showed an association of low and middle performing schools with large classes and high performing schools with small classes. These results are indicative of issues of inequality between high and low performing schools in the Eastern Cape. From a policy perspective, the educator learner ratios would not be the recommended route as it is applied in other countries, but could be used as a reference point to address and eradicate inequality and poverty in the province. The results also revealed the positive impact of placing schools in quintiles to learner performance. While the government addressed the eradication of poverty by providing nutrition and declared 'no fee schools' to quintile 1 to 3 schools, budgets on learner teacher support material (ltsm) should include the purchasing of calculators and laptops or computers for e-learning. There should also be monitoring of the use of support material sent to schools in terms of proper textbook usage and maintenance.

Furthermore, the analysis showed the strong impact of the language of instruction on learner performance, and provided evidence that learners instructed in their mother tongue do better than others. The curriculum should address the language used for instruction by providing teaching and assessment in both English and the mother tongue language. The curriculum should cater for all types of learners so that there is no child left behind and, hereby improving pass rates, especially of the Mathematics performance in the province. As researchers such as Jacob and Lefgren (2004)[9] and Levin and Lockheed (2012) have found a positive impact on remedial programmes and provision of instructional material, it is advisable for the Eastern Cape to adopt these strategies so as to improve the learner performance.

Other variables such as location, boarding school, ex-departments and districts yielded interesting results in the study. Rural schools performed better than the urban schools on the ANA pass rates. The presence of hostel facilities

[9] (Jacob & Lefgren, 2004); (Levin and Lockheed, 2012)

in boarding schools showed a positive and significant impact on performance. There were reflections of the legacy of racially based schooling system which favoured a low proportion of White and Indian learners in the Eastern Cape. School districts from cluster A and C performed better than school districts in cluster B.

An 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 (Levin and Lockheed, 2012). Bearing in mind this ideal schooling system, it turns out that the Eastern Cape schooling outcomes can be strongly explained by educator characteristics such as educator experience and educator qualifications; community characteristics such as quintile categories; and school characteristics such as educator learner ratios, language of instruction, former education departments, location, boarding school and district demarcations.

Finally, within the context of the future research, it would be prudent to examine very specific schools in districts. For example, it would be interesting to find out why schools influence results differently and to get a more detailed analysis of what the key factors and parameters are, which make for high performing schools. This research was limited by the unavailability of some specific data such as school infrastructural data, the quality of school management teams, educational levels of school governing bodies, to name but a few.

References

[1] Addonizio, M. E. (2013). X-Efficiency and Effective Schools?: A New Look at Old Theories *. *Journal of Education Finance*, 35(1), 1–25.

[2] Angrist, J. D., & Pischke, J. (2008). Mostly Harmless Econometrics?: An Empiricist ' s Companion. *Massachusettts I Nstitute of Technology and The London School of Economics*, (March).

[3] Aslam, M., & Siddiqui, R. (2003). The Determinants of student Achievement in Government and Private schools in Pakistan. *The Pakistan Development Review*, 42(4), 841–876.

[4] Badri, M. A., & Mourad, T. El. (2012). Determinants of School Efficiencies in Abu Dhabi Using DEA. In *2012 International conference on management and education innovation* (Vol. 37, pp. 314–318).

[5] Berg, S. Van Der. (2013). South African perspective. *Research on Socio-Economic Policy*, (September), 1–35.

[6] Bhorat, H., & Oosthuizen, M. (2006). Determinants of Grade 12 Pass Rates in the Post-Apartheid South African Schooling System. *Development Policy Research Unit*, (October), 1–27.

[7] Bhuchinsky, M. (1994). Changes in the U.S. wage structure 1963-1987: Application of Quantile Regression. *Econometrica*, 62(2), 405–458.

[8] Boardman, A., Davis, O., & Sanday, P. (1977). A Simultaneous Equations Model of the Educational Process?: The Coleman Data Revisited with an Emphasis on Achievement. *Journal of Public Economics*, 7, 23–49.

[9] Borge, L.-E., & Naper, L. R. (2006). Efficiency Potential and Efficiency Variation in Norwegian Lower Secondary Schools. *Mohr Siebeck GmbH & Co. KG*, 62(2), 221–249. doi:10.1628/001522106X120677

[10] Buhai, I. S. (2004). *Quantile Regression?: Overview and Selected Applications* (pp. 1–20).

[11] Case, A., & Deaton, A. (1999). School inputs and educational outcomes in South Africa. *Quarterly Journal of Economics*, 114(3), 1047–84.

[12] Chakraborty, K., Basudeband, B., & Lewis, C. (2001). Measurement of Technical Efficiency in Public Education?: A Stochastic Production Function Nonstochastic Approach. *Southern Economic Journal*, 67(4), 889–905.

[13] Crouch, L., & 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. *Social Dynamics*, 27(1), 60–78.

[14] DBE. (2012). *National Senior Certificate School performance report* (pp. 1–238).

[15] Deller, S. C., & Halstead, J. M. (1994). Efficiency in the Production of Rural Road Services?: The Case of New England Towns. *Land Economics*, 70(2), 247–259.

[16] DoE. (2007). img-430121613-0001 provincial govt expenditure.pdf.

[17] DoE. (2010). *Education for All (EFA) 2009 Country Report?: South Africa*.

[18] Durán, L. K., Roseth, C. J., & Hoffman, P. (2010). An experimental study comparing English-only and Transitional Bilingual Education on Spanish-speaking preschoolers' early literacy development. *Early Childhood Research Quarterly*, 25(2), 207–217. doi:10.1016/j.ecresq.2009.10.002

[19] Eide, E., & Showalter, M. (1998). The effect of school quality on student performance: A quantile regression approach. *Economics Letters*, 58, 345–350.

[20] Fedderke, J., & Luiz, J. (2002). Production of Educational Output?: Time-Series Evidence from Socioeconomically Heterogeneous Populations — the Case of South Africa , 1927 – 1993 *. *Economic Develpoment and Cultural Change*, 51(1), 161–187.

[21] Giménez, V., Prior, D., & Thieme, C. (2006). Technical efficiency, managerial efficiency and objective-setting in the educational system: an international comparison. *Journal of the Operational Research Society*, 58(8), 996–1007. doi:10.1057/palgrave.jors.2602213

[22] Glewwe, P., & Kremer, M. (2005). Working Papers. *Center for International Development at Harvard University*, (122).

[23] Glewwe, P., & Kremer, M. (2006). Schools, teachers and education outcomes in developing countries. *Center for Global Development*, 2(06). doi:10.1016/S1574-0692(06)02016-2

[24] Gujarati, D. (1992). *Essentials of Econometrics*. McGraw-Hill.

[25] Gustafsson, M. (2007). *Using the hierarchical linear model to understand school production in South Africa. stellenbosch economic working paper* (pp. 1–13).

[26] Gustafsson, M., & Taylor, S. (2005). Treating schools to a new administration The impact of South Africa ' s 2005 provincial boundary changes on school performance. *Bureau for Economic Research*, 1–23.

[27] Hanushek. (1989). Impact of differential expenditures on school performance. *American Educatiion Research Association*, 18(4), 45–51.

[28] Hanushek, E. A., Humpage, S., Ravina, R., Behrman, J. R., Engle, P., Fernald, L., Alderman, H. (1990). Developing Countries. In P. Glewwe (Ed.), *Education Policy in Developing Countries*.

[29] Jacob, B. A., & Lefgren, L. (2004). Remedial education and student achievement: A Regression-Discontinuity analysis. *The Review of Economics and Statistics*, 86(1), 226–244.

[30] Koenker, R., & Bassett, G. (1978). Regression quantiles. *Enonometrica*, 46(1), 33–50.

[31] Koenker, R., & Hallock, K. F. (2001). Quantile Regression. *Journal of Economic Perspectives*, 15(4), 143–156.

[32] Levin and Lockheed. (2012). *Effective Schools in Developing Countries*. Routledge.

[33] Ludwig, J., Miller, L., Malme, R., & Morse, J. (2007). Does Head Start Improve Children ' s Life Chances?? Evidence from a Regression Discontinuity Design. *The Quarterly Journal of Economics*, 122(1), 159–208.

[34] Madala, G. S. (2001). *Introduction to Econometrics* (third.). England: John Wiley and sons.

[35] Marshall, J. H. (2011). School quality signals and attendance in rural Guatemala. *Economics of Education Review*, 30(6), 1445–1455. doi:10.1016/j.econedurev.2011.07.011

[36] McEwan, P. J., & Carnoy, M. (2000). The Effectiveness and Efficiency of Private Schools in Chile ' s Voucher System. *American Educatiion Research Association*, 22(3), 213–239.

[37] Parzen, E. (1962). On Estimation of a Probability Density Function and Mode. *The Annals of Mathematical Statistics*, 33(3), 1065–1076.

[38] Patrinos, H. A., & Velez, E. (2009). Costs and benefits of bilingual education in Guatemala: A partial analysis. *International Journal of Educational Development*, 29(6), 594–598. doi:10.1016/j.ijedudev.2009.02.001

[39] PECE. (2012). *Annual Report 2011/2012* (pp. 1–288). Zwelitsha.

[40] Ray, S. (1991). Resource-use efficiency in public schools: a study of Connecticut data. *Management Science*, 37(12), 1620–1628. Retrieved from http://pubsonline.informs.org/doi/abs/10.1287/mnsc.37.12.1620

[41] Schwartz, E. A., & Stiefel, L. (2004). Immigrants and the Distribution of Resources Within an Urban School District. *Educational Evaluation and Policy Analysis*, 26(4), 303–327.

[42] Shepherd, D. (2011). Constraints-to-school-effectiveness: what prevents poor schools from delivering results.

[43] Taylor. (2009). The state of South African schools Part 1?: Time and the regulation of consciousness. *Journal of Education*, (46), 1–24.

[44] Taylor, S. (2011). *Uncovering indicators of effective school management in South Africa using the National School Effectiveness Study* (pp. 1–51).

[45] Thieme, C., Giménez, V., & Prior, D. (2011). A comparative analysis of the efficiency of national education systems. *Asia Pacific Education Review*, 13(1), 1–15. doi:10.1007/s12564-011-9177-6

[46] Van der Berg, S. (2007). Apartheid's Enduring Legacy: Inequalities in Education. *Journal of African Economies*, 16(5), 849–880. doi:10.1093/jae/ejm017

[47] Van der Berg, S., & Burger, R. (2003). Education and socio-economic differentials: A study of school performance in the Western Cape. *The South African Journal of Economics*, 71(September), 496–522.

[48] Van der Berg, S., & Louw, M. (2006). Unravelling the mystery: understanding South African Schooling Outcome in regional context. *Centre for the Study of African Economies*, (467), 1–32.

[49] Vignoles, A., Levcic, R., Walker, J., Machin, S., & Reynolds, D. (2000). The relationship between resource allocation and pupil attainment: A review. *London School of Economics*, (September), 2–16.

[50] Woodbury, K., Dollery, B., & Prasada, R. (2008). Is local government efficiency measurement in Astralia adequate? An Analysis of the Evidence. *Public Performance and Mamgement Review*, 27(2), 77–91.

Table 1: Eastern Cape average % marks for grade 3, 6 and 9 in Mathematics, Home Language and First Additional Language per district, 2013

DISTRICT	GRADE 3		GRADE 6			GRADE 9		
CLUSTER A	MATHS	H L	MATHS	H L	FAL	MATHS	H L	FAL
LIBODE	53.7	47.9	33.4	39.8	42.6	19	31.8	33.4
LUSIKISIKI	48	43.6	31.8	40.9	40.5	17.3	33.5	31.2
MALUTI	47.2	44.2	28.1	55.8	38.1	13.4	38.6	34.2
MBIZANA	52.4	49.0	35.28	44.8	44.8	19.2	29.0	28.7
MT FLETCHER	48.8	45.3	27.4	40.1	38.2	13.3	28.6	33.5
MT FRERE	50.5	48.7	35.2	46.0	44.7	21.9	35.3	33.1
QUMBU	48	43.5	38.2	46.8	44.2	22.2	37.4	33.4
CLUSTER B								
BUTTERWORTH	53.4	51.5	32.6	43.6	45.9	27.4	36.6	33.4
COFIMVABA	54.2	54.7	34.3	44.3	45.2	22.6	33.3	37.1
DUTYWA	48.4	39.8	29.3	42.3	36.2	17.7	35.3	28.2
LADY FRERE	48.9	47.7	25.8	39.7	41.7	15.5	31	33.1
MTHATHA	48.2	45.0	33.3	39.3	45.6	18.0	31.7	36.5
NGCOBO	46.5	46.9	30.1	44.0	41.5	17.0	34.1	34.5
STERKSPRUIT	42.9	39.5	26.7	35.6	37.6	13.4	34.6	37.1
CLUSTER C								
CRADOCK	53.4	49.3	30.8	48.8	39.2	10.3	42.0	32.1
EAST LONDON	56.7	52.7	37.8	51	49.2	13.3	36	40.3
FORT BEAUFORT	50.5	50.8	33.5	33	43.6	11.8	33.3	16.1
GRAAFF-REINET	45.8	38.3	30.9	45.2	37.3	9.7	36.2	31.3
GRAHAMSTOWN	47.8	42.7	35.9	51.1	50.6	13.2	32.1	39.1
KING WILLIAMS TOWN	47.1	43.9	33.8	44.7	45.2	12.8	31.4	32.3
PORT ELIZABETH	53.8	48.6	44.3	50.5	47.2	14.2	42.0	25.8
QUEENSTOWN	57.7	55.6	35.4	55.2	47.4	13.3	45.1	32.3
UITENHAGE	51.8	47.5	35.5	45.7	51.0	15.7	44.7	32.1

Table 2: Descriptive statistics, 2013

Variable	Mean	Standard deviation	Median
Pass rates	32.907	(23.646)	35.3
Years experience of educators	15.642	(10.433)	16.0
Learner enrolment per school	381.6	(311.362)	287
Age of educators	31	(9.67)	32
Educator qualification	REQV 12	(1.9315)	REQV 13
Educator learner ratio	32.26	(12,558)	31.88
Quintile category	2.385	(1.42)	2

Source: South African Department of Basic Education and Province of the Eastern Cape Education, 2013

Notes:

1. Standard deviations are in parenthesis.
2. The educator qualification is represented by REQV, where REQV 13 is an educator with grade 12 and three years of professional training and REQV ranges from REQV 10 to REQV 17.
3. In case of binary variable, only one variable is reported

Variable	% Frequency	Variable	% Frequency	Variable	% Frequency
REQV		Location		Cluster	
Share REQV 10	42.08	Share rural	56.36	A	39.45
Share REQV 11	6.43	telecommunication	99.96	B	35.05
Share REQV 12	4.04	School phase		C	25.50
Share REQV 13	29.44	Combined	53.57	Language of	
Share REQV 14	12.65	Primary	40.19	instruction	
Share REQV 15	2.17	Secondary	5.98	Afrikaans	5.33
Share REQV 16	2.34	Poverty level;		English	61.41
Share REQV 17	0.85	Share quintile 1	37.23	IsiXhosa	31.83
Former department		Share quintile 2	29.43	Subjects assessed	
before 1994		Share quintile 3	26.80	by ANA	
CED	3.33	Share quintile 4	1.54	Afrikaans	2.96
Ciskei	9.43	Share quintile 5	2.42	English	36.54
DET	11.17			IsiXhosa	16.58
HOA	0.05			Mathematics	42.79
HOD	0.21			other	1.14
HOR	4.95				
KZN	3.39				
New Education	3.35				
Transkei	64.08				
Ownership of land					
Share public	97.40				

Source: South African Department of Basic Education and Province of the Eastern Cape Education, 2013

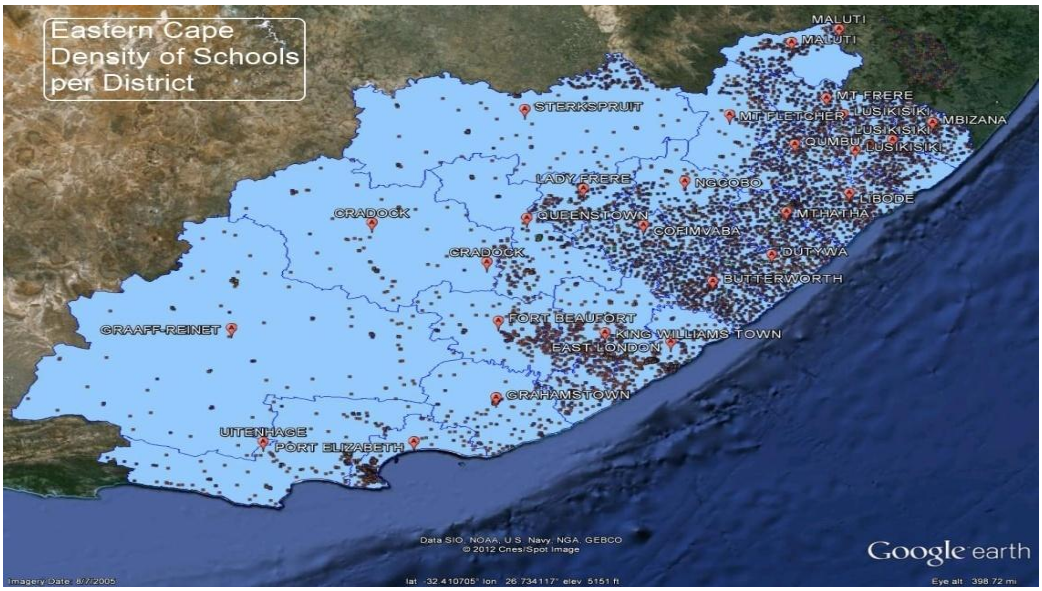
Table 3: Eastern Cape education production function estimates

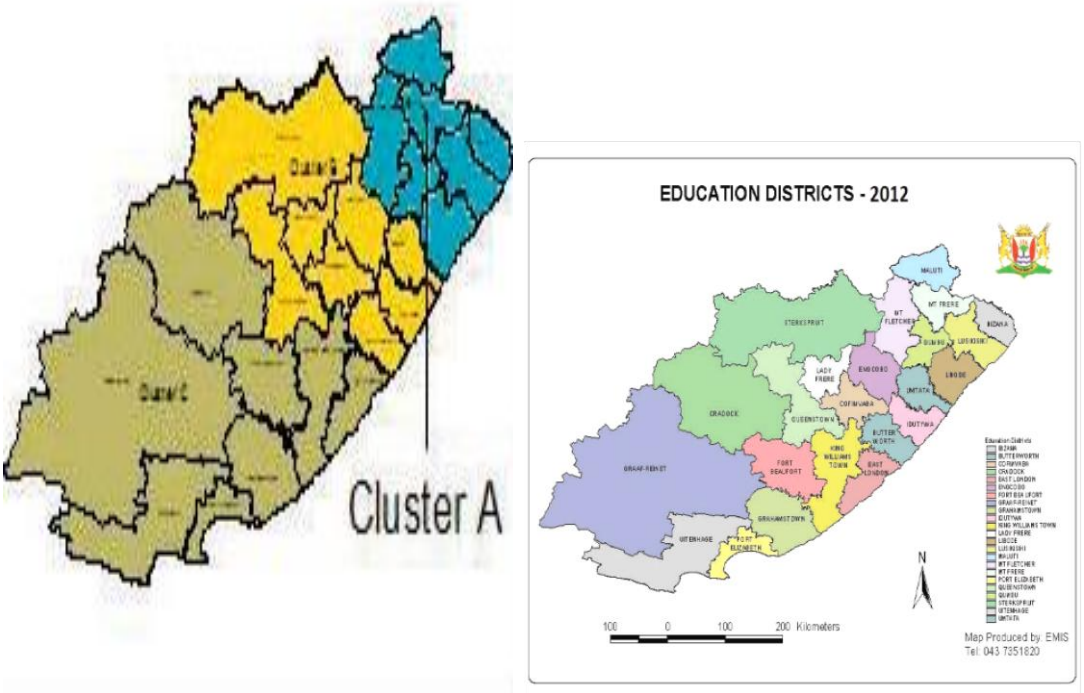
VARIABLES	OLS	Θ=0.1	Θ=0.25	Θ=.5	Θ=0.75	Θ=0.9	50th_10th
Dependent Variable: ANA pass rates							
Educator experience	0.325***	0.00958*	0.0993***	0.247***	0.362***	0.441***	0.238***
	(0.0239)	(0.00549)	(□.0218)	(0.0361)	(0.0386)	(0.0616)	(0.0408)
Ed experience^2	-0.00583***	-0.000504**	-0.00204***	0.00468***	-0.00536***	0.00605***	-0.00418***
	(0.000484)	(0.000200)	(0.000291)	(0.000367)	(0.000537)	(0.00136)	(0.000547)
Educator qualification REQV 11	-3.819***	0.897***	1.286	-5.395***	-3.649***	-6.241***	-6.291***
	(0.534)	(0.130)	(0.958)	(0.946)	(0.949)	(0.968)	(0.996)
REQV 12	0.745	-0.630***	1.030	5.720***	0.421	-5.404***	6.350***
	(0.528)	(0.0669)	(1.387)	(0.936)	(1.032)	(0.763)	(1.022)
REQV 13	2.479***	0.482***	2.021**	0.480	1.956***	2.106***	-0.00219
	(0.329)	(0.0752)	(0.947)	(0.648)	(0.707)	(0.613)	(0.754)
REQV 14	-1.936***	0.146**	0.633	-7.509***	-5.824***	-3.300***	-7.655***
	(0.347)	(0.0740)	(0.895)	(0.863)	(0.908)	(0.613)	(0.761)
REQV 15	19.63***	28.69***	21.67***	20.85***	25.96***	16.16***	-7.930***
	(0.662)	(0.416)	(1.643)	(1.831)	(0.802)	(0.812)	(1.994)
REQV 16	7.599***	-0.947	10.16***	-0.613	11.08***	4.240**	0.334
	(0.946)	(1.063)	(3.377)	(1.246)	(2.107)	(2.005)	(1.404)
REQV 17	-6.582***	6.926***	0.901	-3.730*	-21.38***	-26.84***	-10.66***
	(0.821)	(2.619)	(1.217)	(1.938)	(1.469)	(1.442)	(3.054)
Ed-learner ratio	0.0387	-0.549***	0.0172	-0.0571	0.459***	0.770***	0.492***
	(0.0364)	(0.0191)	(0.0654)	(0.0513)	(0.0931)	(0.0714)	(0.0536)
Ed l ratio^2	0.00157***	0.00775***	0.00132**	0.00144***	-0.00334**	-0.00683***	-0.00631***
	(0.000440)	(0.000229)	(0.000624)	(0.000516)	(0.00135)	(0.000951)	(0.000554)
quintile 2	0.0445	-0.337***	-2.004**	-1.472	1.025	0.0893	-1.134
	(0.424)	(0.0921)	(1.001)	(1.171)	(0.977)	(0.879)	(1.124)
quintile 3	3.728***	0.523***	9.545***	7.872***	-1.289	-2.475*	7.348***
	(0.492)	(0.0923)	(1.113)	(1.368)	(1.085)	(1.273)	(1.288)
quintile 4	7.766***	2.708***	25.23***	13.17***	1.090	-5.355***	10.46***
	(0.741)	(0.562)	(1.526)	(1.639)	(1.505)	(1.625)	(1.871)
quintile 5	14.99***	10.13***	21.37***	17.61***	12.39***	2.988	7.487**
	(0.821)	(2.696)	(1.776)	(1.784)	(1.970)	(1.923)	(3.096)
instructed in English	-9.107***	-3.253***	-11.07***	-8.177***	-7.057***	-11.47***	-4.925***
	(0.511)	(0.256)	(1.104)	(0.692)	(0.858)	(1.038)	(0.747)
instructed in other language	-17.71***	-1.788	-6.991***	-17.14***	-27.77***	-31.85***	-15.36***
	(1.123)	(5.349)	(1.946)	(1.980)	(1.570)	(2.635)	(5.285)
instructed in Xhosa	-15.87***	-3.286***	-16.69***	-14.78***	-17.69***	-21.86***	-11.49***
	(0.613)	(0.226)	(1.342)	(1.236)	(0.967)	(1.795)	(1.450)
assessed in Mathematics	-2.875***	-0.147**	-3.788***	-3.404***	-1.183**	-2.947***	-3.257***
	(0.245)	(0.0657)	(0.694)	(0.680)	(0.500)	(0.412)	(0.579)
location- To be updated	-2.524	3.738***	11.32**	-2.604	-20.20***	-14.83***	-6.343*
	(1.683)	(1.203)	(4.570)	(2.961)	(3.359)	(4.868)	(3.393)
location- urban	-2.072***	-0.250***	-1.852**	-0.411	-3.145***	-4.056***	-0.161
	(0.317)	(0.0710)	(0.841)	(0.683)	(1.002)	(0.696)	(0.663)
boarding school	9.592***	1.584	11.43***	6.961***	13.45***	8.468***	5.378**
	(0.554)	(2.632)	(1.173)	(1.112)	(0.692)	(0.957)	(2.549)

exdept Ciskei	-12.91***	-22.60***	-4.021**	-25.85***	-19.56***	-15.99***	-3.246
	(0.893)	(3.143)	(1.727)	(1.283)	(1.478)	(1.510)	(3.593)
exdept DET	-12.68***	-21.98***	-8.402***	-19.29***	-16.53***	-17.81***	2.688
	(0.694)	(2.530)	(1.578)	(0.869)	(1.450)	(1.266)	(2.944)
exdept HOD	24.62***	19.85***	26.41***	23.95***	36.91***	22.48***	4.097
	(1.939)	(3.174)	(5.108)	(3.041)	(3.679)	(5.711)	(4.197)
exdept HOR	-20.79***	-24.51***	-17.92***	-25.29***	-20.00***	-23.27***	-0.779
	(0.589)	(2.600)	(1.094)	(1.094)	(1.203)	(0.836)	(2.895)
exdept New department	-4.889***	-22.70***	4.055*	-8.275***	-2.050	-15.83***	14.42***
	(1.355)	(2.510)	(2.321)	(2.081)	(3.268)	(2.481)	(3.595)
exdept To be updated	-6.531***	-1.936	-10.33***	-11.33***	-5.620*	0.426	-9.390
	(1.609)	(6.694)	(3.635)	(2.831)	(3.281)	(2.268)	(6.109)
exdept Transkei	-2.564**	-21.71***	5.187**	-3.731**	-1.610	-5.974***	17.98***
	(1.049)	(2.497)	(2.128)	(1.643)	(1.844)	(2.210)	(3.267)
district Cofimvaba	-2.771*	1.560	-6.821*	-6.074**	8.477*	15.24***	-7.634
	(1.473)	(5.087)	(3.716)	(2.756)	(4.368)	(2.254)	(5.103)
district Cradock	7.845***	7.138***	16.31***	19.18***	13.61***	4.066*	12.04***
	(1.265)	(1.846)	(3.594)	(1.819)	(1.942)	(2.202)	(2.749)
district Dutywa	-13.42***	0.318	-11.79***	-19.19***	-11.60**	-7.525***	-19.51***
	(1.319)	(1.302)	(3.126)	(3.343)	(4.788)	(1.770)	(4.398)
district East London	8.625***	-6.621***	8.971**	10.40***	12.10***	21.87***	17.02***
	(1.480)	(1.429)	(3.542)	(1.610)	(4.489)	(2.164)	(2.350)
district Fort Beaufort	-7.546***	4.118	-2.067	-2.069	-12.06***	7.182**	-6.187*
	(1.706)	(2.917)	(3.410)	(1.675)	(2.424)	(3.526)	(3.538)
district Graaff Reinet	-4.926***	-17.48***	-7.662*	-6.140***	6.670**	6.758**	11.34**
	(1.627)	(4.516)	(4.648)	(2.363)	(3.317)	(2.833)	(4.886)
district Grahamstown	17.16***	2.009	17.30***	27.01***	22.97***	7.562***	25.00***
	(1.432)	(2.161)	(4.253)	(2.342)	(2.295)	(2.688)	(2.921)
district King William' town	12.32***	3.224	11.26***	27.99***	26.60***	12.97***	24.77***
	(1.388)	(2.880)	(4.066)	(1.733)	(1.955)	(2.237)	(3.424)
district Lady Frere	-4.491	2.186*	-3.107	-14.07*	0.134	8.938	-16.25*
	(3.130)	(1.278)	(3.824)	(7.928)	(7.754)	(5.920)	(8.859)
district Libode	-5.145***	3.090**	-3.315	-5.493***	-3.466**	-4.562***	-8.583***
	(1.030)	(1.249)	(3.242)	(1.786)	(1.517)	(1.769)	(1.961)
district Lusikisiki	1.806	3.306***	2.237	5.132**	7.093***	4.165**	1.825
	(1.281)	(1.256)	(3.407)	(2.235)	(1.994)	(2.078)	(2.382)
district Maluti	6.982***	4.111	7.019	9.951***	13.06***	7.332***	5.839
	(1.900)	(3.379)	(4.540)	(2.805)	(3.155)	(2.824)	(4.119)
district Mbizana	-1.978*	3.359***	-4.381	0.843	4.286**	3.200**	-2.516
	(1.016)	(1.262)	(3.113)	(2.361)	(2.132)	(1.579)	(2.411)
district Mt Fletcher	-0.0416	3.438***	5.244	0.230	5.273***	-2.618	-3.208
	(1.016)	(1.264)	(3.401)	(1.674)	(1.843)	(1.683)	(2.037)
district Mt Frere	1.838	2.722**	-3.810	5.256***	7.509***	11.89***	2.533
	(1.308)	(1.277)	(3.786)	(1.821)	(2.234)	(2.977)	(2.247)
district Mthatha	-7.512**	1.619	-7.373*	-11.40**	-2.573	2.857	-13.02**
	(3.068)	(1.521)	(4.203)	(5.051)	(7.027)	(5.967)	(5.370)
district Ngcobo	-7.307***	2.210	-10.76***	-5.622*	-4.902*	-5.278*	-7.840**
	(1.711)	(1.368)	(3.765)	(3.059)	(2.805)	(3.111)	(3.278)
district Port Elizabeth	7.216***	2.662**	8.929***	10.93***	18.06***	15.19***	8.271***
	(1.279)	(1.272)	(3.443)	(1.990)	(2.252)	(2.142)	(2.478)
district Queenstown	18.25***	27.68***	35.45***	22.45***	20.14***	17.16***	-5.229

	(1.648)	(3.041)	(3.521)	(1.713)	(2.235)	(1.899)	(4.301)
district Qumbu	0.303	3.034**	-10.16***	4.953***	7.377***	6.979***	1.919
	(1.019)	(1.261)	(3.426)	(1.787)	(1.557)	(1.667)	(2.037)
district Sterkspruit	-11.16***	1.593	-11.23***	-12.94***	-5.395**	-10.39***	-14.53***
	(1.377)	(1.245)	(3.835)	(2.198)	(2.484)	(2.226)	(2.579)
district Uitenhage	-6.136***	3.381***	-0.592	-3.689**	-3.575	-5.825**	-7.070***
	(1.252)	(1.261)	(3.344)	(1.707)	(2.300)	(2.318)	(2.126)
grade 2	-7.610***	-0.719***	-12.32***	-14.25***	-9.082***	0.887	-13.53***
	(0.500)	(0.0762)	(1.290)	(1.396)	(1.215)	(0.947)	(1.406)
grade 3	-3.249***	-0.277***	-2.323	-9.273***	-3.586***	-1.614*	-8.996***
	(0.506)	(0.0926)	(1.616)	(1.054)	(0.883)	(0.980)	(1.153)
grade 4	-9.301***	-0.532***	-3.135**	-9.465***	-13.98***	-12.45***	-8.933***
	(0.571)	(0.175)	(1.481)	(1.078)	(1.135)	(1.229)	(1.266)
grade 5	-7.973***	0.0174	-5.921***	-14.41***	-11.68***	-4.359***	-14.43***
	(0.548)	(0.170)	(1.359)	(1.352)	(1.070)	(1.126)	(1.301)
grade 6	-13.40***	-0.147	-9.709***	-14.81***	-16.77***	-12.45***	-14.67***
	(0.553)	(0.153)	(1.540)	(1.565)	(1.088)	(1.198)	(1.590)
grade 9	-18.72***	-0.130	-15.63***	-26.12***	-21.17***	-21.95***	-25.99***
	(0.556)	(0.156)	(1.586)	(1.057)	(0.834)	(1.088)	(1.072)
Constant	47.83***	31.42***	24.35***	55.10***	52.74***	66.14***	23.68***
	(1.657)	(2.826)	(4.390)	(3.405)	(3.313)	(3.215)	(4.243)
Observations	30,554	30,554	30,554	30,554	30,554	30,554	30,554
R-squared	0.365	0.1199			0.3077	0.292	

Map 1: Eastern Cape districts with population density of schools





Source: PECE emis website

Figure 1: The kernel density graphs per subject, ex-department and quintile