The effects of pupil-teacher ratio and expenditure per pupil on educational attainment in South Africa

Esther Mumbi Kimani*and Haroon Bhorat

1st September 2014

Abstract

In this study we examine the relationship between school inputs -pupil-teacher ratio and expenditure per pupil- and educational attainment in South Africa. Following a reduced form production function we apply a partial generalized ordered probit which allows us to identify 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 in resources at school. Controlling for individual, socio-economic and neighborhood characteristics we find 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.

1 Introduction

In accordance with the UN Millennium Development (MDGs) Goal 2 (United Nation; 2010, p15) quality education and increasing educational attainment are key parts in achieving universal primary education by 2015. There have been substantial gains in achieving universal education with net enrollment reaching 90% in developing countries, and 77% in Sub-Saharan Africa in 2011; a 7 percentage point increase from 2000 (United Nation; 2013, p15). Sadly this increase in enrollment has not been accompanied by an increase in school quality where by school quality we refers to effects of school inputs. Studying the role of education in economic growth in developing countries, Hanushek (2013) finds most students who have completed 9 years of schooling in developing countries have skills below the international levels and cannot compete with their peers. Hanushek argues that for the type of education offered to be of value it should at least impart some basic skills. However, he adds that the provision of a quality education

*for comments please email: esther.kimani@myuct.ac.za

requires both infrastructure and access (Hanushek; 2013). Given that public provision of basic education has mainly been a prerogative of governments a continued increase in public expenditure on education is largely inevitable.

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[1] 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-pupil ratio[2] stood at between 1:20 and 1:30 for Whites and between 1:40 and 1:70 for African pupils (Case and Deaton; 1999). Equally, in this era, there was racial divergence in allocation of expenditure per pupil; for instance, for every R1 spent on a White pupil only 19 cents per capita expenditure was spent on an African pupil (Moulder 1992 as cited in Bhorat and Oosthuizen 2008, p2). Resources allocated for Indians and Coloureds fell between these two population groups. This divergence in resource allocation had implications for the quality of education offered in the different schools. White schools had the highest quality while African schools had the lowest quality. To date, these formerly White schools (commonly known as 'Model C' schools) are better resourced and have had the best educational outcomes; the high matriculation[3] pass rate in these schools is a case in point. Although resource provision has changed a great deal, to redress past equity issues, inequality in school quality still persists (Van der Berg; 2007 and Yamauchi; 2011). This persistence has been in spite of the efforts by the various education stakeholders to transform the education system.

In accordance with the Constitution and the School Act, South African's are free to apply for admission into any school regardless of their race or religion. Heterogeneity in student population by race is, however, yet to be achieved in most schools. The population structure in schools is still homogeneous with 96% of Black pupils still attending formerly Black schools (Van der Berg; 2007). This lack of racial mixing can be associated with two main factors. First, the school admission policy gives priority to students within a 5km radius of the school. This spatial rule has meant that, in a racially homogeneous community, schools have mainly remained racially homogeneous. Spatial boundaries and the apartheid reinforcing past of populations being clustered by race have been critical barriers to entry in schools of choice particularly entry into 'model C' schools. In some cases, particularly in private schools, former White's Schools have been found to inflate school fees in an effort to discourage Black students from entry (Selod and Zenou; 2003). Second, financial constraints have been partly to blame for this lack of racial mixing in schools. These have been in the form of labour market constraints, and high land and housing prices which make it impossible for low income households to move to particularly formerly White areas (Yamauchi; 2011). As it stands, barriers to equal opportunities for entry into a school of

[1]Black in this case refers to Africans, Indians/Asians, and Coloured

[2]The pupil-teacher ratio indicates the number of educator provided in a school, it includes staff that might not be teaching such as the school management staff. The class size is the average number of learners(learner and pupil are used interchangeably throughout the text) in a class (DBE 2010b).

[3]Matriculation examination is undertaken at the end of grade 12 and is a prerequisite for university admission.

choice for some South African students still exist in spite of the various efforts by education stakeholders to erode them.

The heterogeneous provision of- and access to- education in the apartheid-era led to disparity in educational attainment by race and type (or level) of skill acquired by the different population groups. Van der Berg and Moses (2011) attributes the persistent income inequality to this heterogeneity in education attainment while Bhorat et al. (2009) finds wage income to be the main driver of the income inequality in the country. Income inequality is highly skewed by skill and by race; on the one hand Bhorat et al. (2009) show that in a decade, that is, between 1995 and 2005 wage inequality increased by 14%. They associate this increase with a skill premium. Gradin (2014, p75) on the other hand show that in 2005/06 the median income of Africans in South Africa was 9% that of Whites. The need to close the gap in skills particularly between the different population groups from a position where Whites were highly skilled, and Blacks mainly had low to no skills at all has been a priority in the country and the provision of a high quality education to all has been a necessary requirement. Education policies such as allocation of expenditure per pupil by school poverty quintile and increase in employment of teachers to lower the pupil-teacher ratio in all schools have been put in place by the government in an effort to improve quality and offer equal educational opportunities to all South Africans. Provision of opportunity to access equal education to all is perceived to enhance social mobility (CHE, 2004); a necessary constituent in post-apartheid South Africa overall transformation making a study of school quality in South Africa imperative.

In the recent years, government efforts to bring down the barriers, and to deal with inequality in provision of education have been through subsidies (Republic of South Africa; 1998b). Specifically, this has been through school-fee exemption for students from poor backgrounds attending fee paying schools (quintile 4 and 5 schools) as stipulated in the School Act of 1996, and no-fees[4] school as per the National Norms and Standards for School Funding of 2006. The objectives of these policies have been to try and ease the financial burden of education on poor families and to promote equitable access to quality education for all. The government implements this through a school quintile[5] system that is used to determine allocation of expenditure per pupil to schools. All ordinary public schools are assigned to a quintile based on the poverty level of the surrounding community[6], and the physical condition, facilities and population of the school [7](South African School Amendment Act, no 84 of 1996, as amended: 2006

[4]The no fee status of a school is determined by the school quintile which is arrived at following the national poverty ranking. It initially applied to 40% of the poorest schools (bottom two quintiles); both primary and secondary schools. These are supposed to receive 60% of the available resources. In 2011, no-fee policy was extended to 3rd quintile (South African School Act of 1996, Amended National Norms and Standards for School Funding Notice 33723, as amended: 2010 p2)

[5]The school quintile determines how much money is allocated per learner in a school. Schools in quintile 1 receive the highest while those in quintile 5 receive the least amount.

[6]In determining school poverty score the geographical area of the school and three indicators of poverty: income; dependency ratio (or unemployment rate); and level of education (or literacy rate) of the community are weighted (South African School Amendment Act, no 84 of 1996, as amended: 2006 p29).

[7]The provincial education department is mandated to create indices based on the range of physical facilities at the school,

p24-29). Schools in quintile 1 are the poorest schools and receive the highest per pupil allocation, and those in quintile 5 are the least poor schools and receive the lowest allocation per pupil. This allocation of expenditure per pupil ensures that more funds are allocated to pupils in the largest and neediest schools. The National Norms and Standards for School Funding stipulates that these funds are for recurrent expenditure that may comprise of: purchase of textbooks, and educational material or equipment for the school; school buildings improvement and maintenance; extra-mural curriculum; provision of services to the school. They, however, are not to be used for hiring teachers or personnel and/or for capital expenditure (South African School Act of 1996, National Norms and Standards for School Funding Notice 2362 of 1998 p30). Assessment of the impact of these funds has been minimal. In a qualitative analysis of the effect of school fees and funding on school quality, Hall and Giese (2008) finds that school fee exemption and no-fee schools have undoubtedly increased funding to poor schools; however, this increase has not culminated in school quality improvement (Hall and Giese, 2008). According to Hall and Giese these policies have failed to deal with unequal allocation of teaching capacity since there is no pro-poor allocation of funding to teachers, an important element in improving school quality. Although not many studies have been done on the effect of increased education funding in developing countries, existing studies on the effect of expenditure per pupil on educational outcome in developing countries have conflicting findings. In the case of South Africa, there are hardly any studies examining the relationship (if any) between this expenditure per pupil and educational attainment particularly for Africans. Using a nationally representative sample this is one of the objectives of this study. Given the existing debate on the effect of the increased funding on educational outcomes, there is a need for empirical evidence.

There is consensus that the quality of education in South Africa is low relative to what is spent; a view that even the government agrees with (Department of Education; 2003, p107). Discontent with the quality of education provided particularly in public schools among the general public has been on the increase. In addition to these quality concerns, inefficiency from the education department at the national level to the school level has been a concern (Taylor et al.; 2008). There has hardly been any growth in the quality of education for 80% of children in South Africa (Van der Berg and Moses; 2011) and the increase in resources to poor schools has not reduced the quality differentials between schools (Van der Berg; 2007). These concerns cannot be overemphasized particularly with education viewed as an equalizer of income inequality through better job prospects and as an engine for economic growth. In addition to the human capital needs of the economy, the argument for public provision of education has been based on the need to distribute incomes and build social cohesion. The issues of how government policies on schooling affects educational outcome of students are not just essential to policy makers but are important research questions in their own right.

Empirical research on the effects of school inputs (commonly referred to as school quality in literature) on educational attainment for South Africa is limited. Existing research has mainly focused on test

learner:classroom ratio, the overall condition and need for repairs, availability of basic services (South African School Act of 1996, National Norms and Standards for School Funding Notice 2362 of 1998 p28).

scores as the outcome variable and pupil-teacher ratio as the quality measure using survey data sets that are not nationally representative. The findings from these studies are, however, conflicting. In a study of school quality in the apartheid-era, Case and Deaton (1998) find that the pupil-teacher ratio had a marked effect on three educational outcomes -enrollment, education attainment and test score- of Black children. Two studies in post-apartheid South Africa find conflicting results despite the fact that they use similar data. Using 1999 and 2000 matriculation results Van der Berg (2007) on the one hand finds the teacher-pupil ratio has a negative effect on the matriculation score overall except in Black schools while Bhorat and Oosthuizen (2008) on the other hand find that the teacher-pupil ratio had no effect on the 2000 matriculation pass rate. This study extends the existing literature on effects of school inputs on educational attainment by examining the relationship between the pupil-teacher ratio and expenditure per pupil, and highest education level attained.

In literature the most commonly used measures in educational outcome studies are class size, pupil-teacher ratio and expenditure per learner. About a third of studies on effects of school quality on educational outcome consider test score as the outcome variable (Wilson; 2001). In this study we consider the highest level of education attained as the outcome measure, expenditure per pupil; a non-capital and non-personnel allocation by the South African government, and the pupil-teacher[8] ratio as our school quality measures (which in this study we refer to them as school inputs). In addition to data availability, there is a motivation to this choice: One, the pupil-teacher ratio is the most common quality measure in South African school quality literature. Second, there is hardly any consensus on the relationship between the pupil-teacher ratio and educational outcome in South Africa. Third, this specific type of expenditure per pupil is yet to be considered in an analysis involving a household survey data in South Africa -at least to the best of our knowledge. Lastly, most education school quality studies in South Africa consider the test score as the outcome variable. An analysis of the effect of school quality on education level attained would be an addition to the literature. The study's hypotheses are therefore: one a reduction in pupil-teacher ratio leads to an improvement in classroom instruction quality resulting from improved interaction between pupils and their teacher which in turn leads to an increase in the educational attainment. Second, an increase in expenditure per learner leads to an increase in the resources available to students resulting in an improved learning experience which in turn leads to an increase in educational attainment.

Our analysis educational attainment of Africans in Wave 1 (administered in 2008) of the National Income Dynamic Study (NIDS) a national representative panel using a partial generalised ordered probit Indicate that... Our least square analysis support this results

The remainder of the study is organized as follows: section 2 reviews the relevant literature, section 3 outlines the approach and method of analysis, and discussion on the data used. Descriptive statistics follow in section 4 followed by empirical results where a discussion on results from education level achieved and education level index. A conclusion is given in the last section.

[8]Teachers in the majority of schools are employed by the government.

2 Econometric estimation approach

Our outcome variable is the individual's highest education level attained. In the sample, this variable is discrete and non-normally distributed, that is, it is spiked at no schooling and at matriculation level. Some of the observations on highest level of education attained are right censored for respondents who were currently enrolled in 2008. Further, given our variable of interest are pupil-teacher ratio and expenditure per pupil we do not consider individuals with no schooling which makes education level attained left censored. In the presence of non-normality, discreteness, and censoring, using an ordinary least square estimation would potentially yield biased and inconsistent estimates (Tansel; 1997, Glick and Sahn; 2000 and Mani et al.; 2009, p12). Glick and Sahn (2000, p69) argues that the censoring problem can be overcome by restricting the sample to individuals old enough to have completed at a risk of considering older individuals whose schooling determinants could have changed at the time of the survey. We therefore restrict our sample to younger respondents (15 to 30 year old's) and to address the problem of censoring, we restrict respondents who have completed education into one sample and use an ordered probit specification to obtain unbiased and consistent estimates.

We employ an ordered probit to estimate the effect of school inputs on educational attainment for Africans not currently enrolled, that is, those who have achieved a given level. Education level attained has a natural order relation to it. However, the difference between having no schooling and having a primary(general) education is not the same as the difference between having a matriculation and having a higher education. The benefits of schooling are, therefore, unlikely to be the same thus ruling out the use of an ordinary least square that assumes a linear expected value locus. An ordered probit gives us the advantage of not treating the dependent variable as normally distributed since education level attained is discrete and bimodal, that is, it is spiked at no schooling and at matriculation level. We assume that we can explain all educational outcomes using a single index model. The underlying latent variable model in consideration is:

$$l_i^* = x_i\alpha + \varepsilon_i \quad with \quad \varepsilon \sim N(0,\sigma^2) \tag{1}$$

where l_i^* linearly dependent on x_i and $x_i = [m_i, f_i, c_i, h_i]$ represents all the observable variables. We assume the disturbance term ε_i is independently and normally distributed with a mean 0 and variance σ^2. We also assume a standard normal distribution hence $\varepsilon \sim N(0,1)$. Although l_i^* is unobserved, we observe where an individuals index lies given categories $j = 1,...,J$. The index is defined by its unknown lower bound μ_{j-1} and upper bound μ_J. We assume that $\mu_{j-1} = -\infty$ such that $F(-\infty) = 0$ and $\mu_J = \infty$ such that $F(\infty) = 1$. Education level attained is divided into four ordinal categories namely: general level (Grade 1 to grade 9), some secondary education (Grade 10, 11), matriculation(grade 12), and higher level (Grade 9 plus any further education and training certificate, and grade 12 plus any additional higher education level attained including certificate, diploma, or degree).

In the ordered probit model, the cumulative probability of the discrete outcome is related to the single

index as follows:

$$Pr[l \leq j|x] = F(\mu_j - x'\alpha) \quad j = 1,...,J \tag{2}$$

where μ_j and $\beta_{(k\times 1)}$denote unknown parameters in the model, and F(.) is a standard normal distribution function. For well defined probabilities we restrict μ_j>μ_{j-1} ,$\forall j$, and the bounds of $F(.)$ are as defined above. To ensure identification of these parameters we assume that $\alpha_0 = 0$, that is, x does not contain a constant, and we normalize the variance of the distribution of function F.

A common practice is to employ a standard probit in this kind of analysis, however, in analyzing marginal product effects, the standard ordered probit is not without limitations which mainly stems from its assumptions. These assumptions are the single crossing property, the constant threshold, and the distributional assumption that does not allow for heterogeneity between individual realization (Boes and Winkelmann; 2006). We first run a standard ordered probit and test for the assumption of parallel lines. A likelihood ratio test (using omodel in Stata) on the assumption of the parallel lines assumption in the standard ordered probit gives us a $\chi^2_{32} = 131.72$ with a $p - value = 0$ an indication that this assumption is violated at 1%-level. Due to this violations and to deal with the limitations of the standard ordered probit we follow a generalization of this model as proposed by Boes and Winkelmann (2006). This model relaxes the three assumptions of the standard ordered probit, and a partial generalization of this model by Williams (2006) gives more flexibility. This partial generalized ordered probit allows for heterogeneous controls that vary to vary in each category, and those that are constant to remain constant in each category. A test of the parallel lines for each variable in the partial generalized model indicate that the quadratic logarithmic pupil-teacher ratio, both the logarithmic expenditure per pupil and its quadratic, quadratic age, gender, attending independent homeland or a self-governing territory school, and distance to school violate the parallel line assumption and are therefore allowed to vary. A Wald test of the partial generalized model gives us a $\chi^2_{26} = 30.64$ with a $p - value = 0.242$ indicating that this model does not violate the parallel lines regression. This means that the effects of some of our controls vary with the education level attained while others remain constant at each level. This model generalizes the threshold parameters by making them dependent on covariates as follows:

$$\mu_j = \tilde{\mu}_j + x'\omega_j \tag{3}$$

where ω_j is a $k \times 1$-dimensional vector of response specific parameters. We replace μ_j in equation 2 to obtain the cumulative probabilities in the generalized threshold model as:

$$Pr[l \leq j|x] = F(\tilde{\mu}_j + x'\omega_j - x'\alpha) = F(\mu_j - x'\alpha_j) \quad j = 1,...,J$$

where similarly $\mu_j = -\infty$ and $\mu_j = \infty$ as before, and $\alpha_j \equiv \omega_j - \alpha$ since ω_j and α cannot to separately identified using the same x entering the index and generalized threshold models. This allows us to estimate the different parameters of x for each category.

2.0.1 Data

The National Income Dynamic Study (NIDS) is a nationally representative household and individual level panel that commenced in 2008 and is undertaken every two years. It focuses on household composition and structure, education, labour market participation and economic activity, health, agriculture among other issues (Leibbrandt et al.; 2009). These issues are covered in four questionnaires namely: household, individual adult, individual proxy, and child. In arriving at the households sample, it employed a stratified two-stage cluster sample design. Wave 1 response rate was 69% that is 7305 households and 28225 individuals responded across South Africa (Leibbrandt et al.; 2009). To take care of sample non-response and representativity, NIDS has both design weights and post-stratification weights (Wittenberg; 2002). We use Wave 1 to study the effects of pupil-teacher ratio and expenditure per pupil on education attainment. To ensure that we can make inferences from the African sample to the African population in South Africa we use post-stratification weights in the analysis.

In post-*apartheid* South Africa there remain differences between the different population groups. Our sample of interest in NIDS is Africans[9] respondents. 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, hence our focus on this population group. We consider the sample of African youth aged between 15 and 30 in the NIDS 2008 wave 1. Our choice of this age group is motivated by the following policies: in South Africa, schooling is compulsory until age 15 or grade 9 whichever comes first, and the South African National Youth policy defines the youth as individuals aged between 14 to 35 years. These two conditions imply that our sample focuses on individual who are past the legal age of minimum schooling and are considered as the youth of South Africa. An additional condition is that NIDS adult questionnaire has a focus on 15 to 30 year old on education responses. Further, in our view, it is more probable that the family background and neighborhood characteristics in NIDS are more likely to be accurate for this age group than for older respondents. This is because this group is more likely to be residing at their parents home. Ideally, a study of determinants of educational attainment ought to account for final level of education level attained and relate it to characteristics of the environment the individual grew up in. To ensure that our sample is that of respondents who are most likely to have been raised in a given household and are dependents of the household head, we further restrict our sample to respondents who are not the household head (resident or absent), and whose relationship with household head is not: husband or wife or partner, father or mother, father or mother-in-law, grandparent, great grandparent, uncle or aunt, household help, lodger or relative of lodger, other family member, and other non-family member. Our estimation sample is therefore made up of 3023 African individuals who are either not enrolled (1352) or enrolled in institutions of learning (1671).

[9] As of census 2011, Africans were 79.2% of the South African population. In NIDS our sample of individuals who are 15-30 year old and are dependents, Africans are over represented at 90%.

3 Descriptive statistics

Summary statistics of education attainment -the outcome variable- are shown in table 1. Table 1 shows the frequency of Africans who: *(i)* are not enrolled, that is, they have achieved a primary/general, some secondary, completed secondary(matriculated), or a higher education (have a further education and training certificate or any other certificate, diploma, or degree) or *(ii)* are enrolled at either a primary/general, is in grade 10 or 11, or is in grade 12, or is in an FET college or a higher education institution. From the table a significant number of individuals in our sample are enrolled in a secondary school, or have attained some secondary education, that is, 50.03% and 68.64% respectively if we take the second and the third categories. The rest, at most, are currently in primary school or have attained a primary school education, while a significantly small proportion are enrolled in a higher education institution or have achieved a higher education. The number of those currently enrolled at primary and secondary level are higher than those who have attained from which one could infer more individuals are likely to at least matriculate in the future.

Table 1: Education attainment by enrollment

Highest education level attained	Frequency(Percentage)					
	Enrolled		Not enrolled/Achieved		Total	
Primary	809	(48.41)	312	(23.08)	1121	(37.08)
Some secondary	643	(38.48)	453	(33.51)	1096	(36.26)
Matriculation	193	(11.55)	475	(35.13)	668	(22.10)
Higher	26	(1.56)	112	(8.28)	138	(4.57)
Sample Size	1671	(100)	1352	(100)	3023	(100)

Source: Own calculation from NIDS Wave 1 conducted in 2008. Note: Higher education includes those with a certificate, a diploma or a degree.

Table 2 shows the mean characteristics of individual, household, school, and neighborhood characteristics of our estimation sample. In the table, for a better picture of the distributions of these statistics we break them down by the highest education level attained or enrolled at. The pupil-teacher ratio averages at between 30 and 33 pupils per teacher across the different levels. Notably, in both enrolled and attained samples, primary level has the highest average at 33 while higher level has the lowest average at 30. It declined to 32.3 and 30.4 respectively for the state-paid teachers and the state-paid and School Governing Board-paid teachers in 2012 (DoE School Realities 2012 p4). The ratio in the sample is much more dispersed for those who have attained a primary education than for those who have attained a secondary education. A clearer picture of the ratio's distribution by level of education attained for those who are not enrolled/achieved and those who are still enrolled is presented in figure 1.

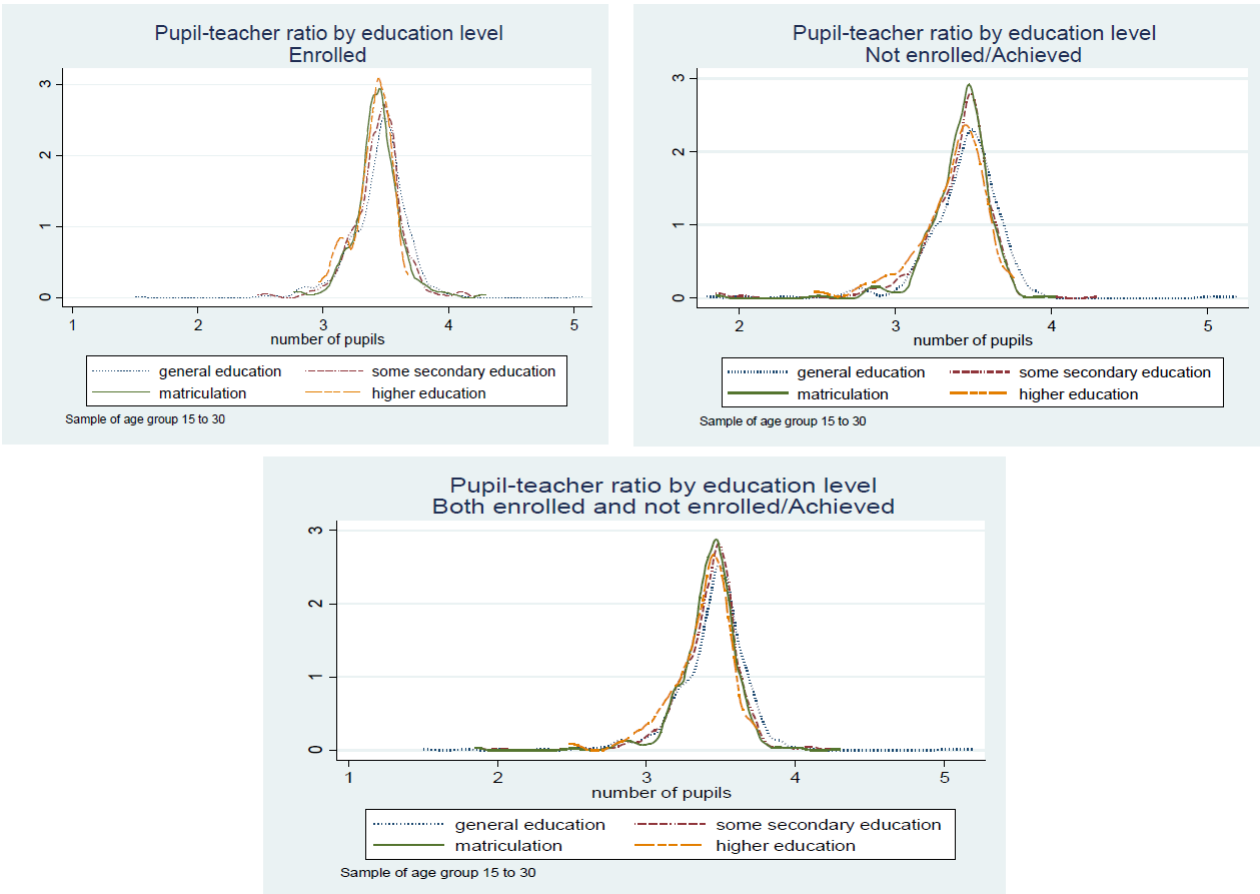
Table 2: Mean characteristics of the sample by education outcome

Variable name	Enrolled				Not enrolled/Achieved				Both enrolled and achieved			
	Primary	Some secondary	Matriculation	Higher	Primary	Some secondary	Matriculation	Higher	Primary	Some secondary	Matriculation	Higher
School inputs: Log(pupil-teacher ratio)	32.50	32.09	31.03	29.84	32.55	30.96	30.76	30.02	32.52	31.61	30.85	30.15
	(0.508)	(0.434)	(0.566)	(0.993)	(0.728)	(0.578)	(0.406)	(0.889)	(0.483)	(0.386)	(0.332)	(0.718)
Log(expenditure per pupil)	524.6	498.9	449.0	220.3	276.0	286.3	296.4	151.2	449.9	403.6	341.4	164.9
	(13.83)	(12.15)	(17.95)	(31.06)	(16.19)	(14.45)	(12.51)	(18.74)	(13.21)	(11.13)	(11.66)	(18.61)
Age	16.77	18.40	19.27	22.24	21.61	23.34	22.57	24.36	18.24	20.64	21.59	23.96
	(0.102)	(0.122)	(0.244)	(0.554)	(0.296)	(0.191)	(0.212)	(0.409)	(0.159)	(0.144)	(0.168)	(0.391)
Gender (Female=1)	0.439	0.540	0.506	0.597	0.438	0.684	0.580	0.646	0.440	0.604	0.556	0.644
	(0.022)	(0.024)	(0.045)	(0.107)	(0.040)	(0.032)	(0.032)	(0.067)	(0.021)	(0.021)	(0.025)	(0.057)
Parents education (%): No schooling	0.537	0.552	0.481	0.597	0.663	0.555	0.505	0.524	0.576	0.558	0.499	0.528
	(0.024)	(0.026)	(0.058)	(0.107)	(0.033)	(0.040)	(0.034)	(0.068)	(0.022)	(0.023)	(0.029)	(0.058)
Primary	0.372	0.340	0.301	0.370	0.317	0.377	0.387	0.294	0.354	0.353	0.361	0.318
	(0.025)	(0.025)	(0.041)	(0.106)	(0.032)	(0.044)	(0.030)	(0.054)	(0.021)	(0.025)	(0.024)	(0.049)
At least some secondary	0.0910	0.107	0.218	0.0329	0.0199	0.0675	0.108	0.183	0.0693	0.0886	0.140	0.154
	(0.014)	(0.017)	(0.047)	(0.023)	(0.009)	(0.020)	(0.019)	(0.058)	(0.010)	(0.014)	(0.021)	(0.045)
Number of siblings	2.750	2.725	2.073	2.950	2.776	2.625	2.697	2.468	2.763	2.670	2.505	2.543
	(0.113)	(0.122)	(0.221)	(0.466)	(0.177)	(0.187)	(0.173)	(0.233)	(0.111)	(0.123)	(0.151)	(0.209)
Household income per capita(monthly)	635.4	650.0	1392.2	1800.6	484.2	645.1	769.7	1129.7	589.2	645.1	953.8	1273.1
	(56.94)	(56.69)	(299.30)	(427.93)	(75.44)	(92.60)	(80.05)	(164.2)	(46.2)	(53.32)	(108.7)	(158.4)
Home location (%)n: Rural/Traditional	0.619	0.569	0.357	0.336	0.569	0.504	0.495	0.370	0.604	0.543	0.455	0.363
	(0.049)	(0.045)	(0.055)	(0.120)	(0.057)	(0.056)	(0.053)	(0.082)	(0.047)	(0.045)	(0.047)	(0.075)
Urban	0.340	0.377	0.623	0.610	0.376	0.393	0.449	0.597	0.350	0.382	0.501	0.608
	(0.048)	(0.044)	(0.056)	(0.125)	(0.056)	(0.054)	(0.052)	(0.083)	(0.046)	(0.043)	(0.047)	(0.075)
Farm	0.0413	0.0539	0.0195	0.0544	0.0557	0.104	0.0554	0.0331	0.0456	0.0755	0.0441	0.0296
	(0.017)	(0.018)	(0.016)	(0.040)	(0.018)	(0.035)	(0.022)	(0.019)	(0.015)	(0.022)	(0.017)	(0.015)
School Type (%): Independent homelands	0.237	0.118	0.175	0.127	0.267	0.153	0.173	0.104	0.247	0.133	0.176	0.111
	(0.034)	(0.021)	(0.046)	(0.115)	(0.043)	(0.028)	(0.032)	(0.044)	(0.031)	(0.019)	(0.029)	(0.043)
Self-governing territory	0.397	0.401	0.239	0.122	0.343	0.394	0.381	0.411	0.380	0.397	0.337	0.343
	(0.036)	(0.036)	(0.041)	(0.084)	(0.048)	(0.039)	(0.043)	(0.079)	(0.034)	(0.030)	(0.037)	(0.065)
Department of education	0.257	0.311	0.420	0.386	0.310	0.345	0.327	0.364	0.272	0.328	0.355	0.373
	(0.032)	(0.038)	(0.061)	(0.111)	(0.044)	(0.039)	(0.041)	(0.072)	(0.032)	(0.033)	(0.039)	(0.062)
House of assembly, representative and delegates	0.0457	0.0612	0.0845	0.281	0.0232	0.0324	0.0401	0.0477	0.0388	0.0478	0.0539	0.0972
	(0.011)	(0.015)	(0.028)	(0.100)	(0.010)	(0.011)	(0.013)	(0.022)	(0.009)	(0.010)	(0.012)	(0.028)
New schools	0.0641	0.109	0.0810	0.0835	0.0571	0.0749	0.0781	0.0726	0.0617	0.0939	0.0782	0.0759
	(0.014)	(0.023)	(0.027)	(0.052)	(0.017)	(0.019)	(0.019)	(0.028)	(0.012)	(0.019)	(0.017)	(0.028)
Distance to school	20.43	22.31	34.21	40.98	86.51	41.82	64.11	56.97	40.28	48.21	54.56	54.32
	(4.507)	(5.763)	(8.935)	(12.585)	(19.81)	(6.93)	(22.79)	(20.56)	(7.27)	(17.82)	(17.15)	(16.94)
Neighbourhood characteristics:												
Number of rooms per person	4.572	4.712	4.323	4.653	4.239	4.540	4.601	4.579	4.471	4.638	4.512	4.579
	(0.110)	(0.130)	(0.149)	(0.297)	(0.160)	(0.166)	(0.163)	(0.214)	(0.112)	(0.132)	(0.137)	(0.188)
% with electricity	0.724	0.801	0.881	0.950	0.719	0.793	0.812	0.850	0.722	0.797	0.834	0.871
	(0.035)	(0.026)	(0.022)	(0.019)	(0.044)	(0.029)	(0.029)	(0.032)	(0.035)	(0.024)	(0.023)	(0.026)
% shared toilet	0.140	0.170	0.208	0.237	0.155	0.206	0.200	0.249	0.145	0.184	0.201	0.250
	(0.024)	(0.023)	(0.029)	(0.054)	(0.026)	(0.032)	(0.026)	(0.043)	(0.022)	(0.025)	(0.024)	(0.038)
% refuse collected	0.301	0.349	0.591	0.614	0.314	0.383	0.413	0.540	0.305	0.362	0.466	0.563
	(0.045)	(0.043)	(0.055)	(0.114)	(0.050)	(0.050)	(0.047)	(0.068)	(0.043)	(0.042)	(0.044)	(0.063)
% street light	0.272	0.327	0.524	0.565	0.307	0.342	0.397	0.492	0.283	0.332	0.435	0.513
	(0.034)	(0.038)	(0.050)	(0.093)	(0.045)	(0.042)	(0.045)	(0.065)	(0.035)	(0.036)	(0.041)	(0.060)
Average household members	6.150	6.047	5.467	6.045	6.353	6.192	6.331	5.970	6.210	6.109	6.071	5.990
	(0.161)	(0.236)	(0.176)	(0.492)	(0.279)	(0.329)	(0.287)	(0.441)	(0.184)	(0.257)	(0.234)	(0.362)
% employed	0.141	0.172	0.224	0.237	0.171	0.199	0.193	0.244	0.150	0.183	0.202	0.244
	(0.009)	(0.010)	(0.019)	(0.029)	(0.012)	(0.016)	(0.011)	(0.022)	(0.009)	(0.011)	(0.011)	(0.019)
% with piped water	0.788	0.824	0.888	0.943	0.848	0.866	0.858	0.907	0.806	0.842	0.868	0.917
	(0.034)	(0.028)	(0.028)	(0.030)	(0.032)	(0.023)	(0.025)	(0.026)	(0.031)	(0.025)	(0.021)	(0.022)
% member own house	0.902	0.885	0.839	0.787	0.879	0.854	0.877	0.854	0.895	0.872	0.866	0.838
	(0.010)	(0.014)	(0.020)	(0.054)	(0.015)	(0.026)	(0.013)	(0.027)	(0.010)	(0.017)	(0.012)	(0.025)
% have landline telephone	0.0767	0.112	0.179	0.212	0.0779	0.0814	0.124	0.161	0.0769	0.0979	0.140	0.173
	(0.009)	(0.015)	(0.031)	(0.047)	(0.013)	(0.011)	(0.019)	(0.027)	(0.008)	(0.012)	(0.018)	(0.024)
Sample size	**1671**	**1671**	**1671**	**1671**	**1352**	**1352**	**1352**	**1352**	**3023**	**3023**	**3023**	**3023**

Note: *(1)* Standard errors in parenthesis. *(2)*Pupil-teacher ratio is as provided by the Department of Basic Education in the NIDS administration data. *(3)* Expenditure per is learner derived from matching respondent's school quintile variable in the NIDS administration data to the amount of expenditure allocated by the government to each student using school quintile every year. *(4)*For all the neighbourhood characteristics the variables are calculated by (a) calculating the average value in each cluster and (b) assigning each individual a value corresponding to his/her cluster. *(5)* The sample is of African respondents aged 15-30 years. *(6)* See Appendix for definitions of all variables used. *(7)* Design weights are used

From figure 1, in both enrolled and achieved samples, on the one hand primary education is more skewed to the right, has a lower peaks, 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. We conduct a Kolmogorov-Smirnov test to test whether 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.

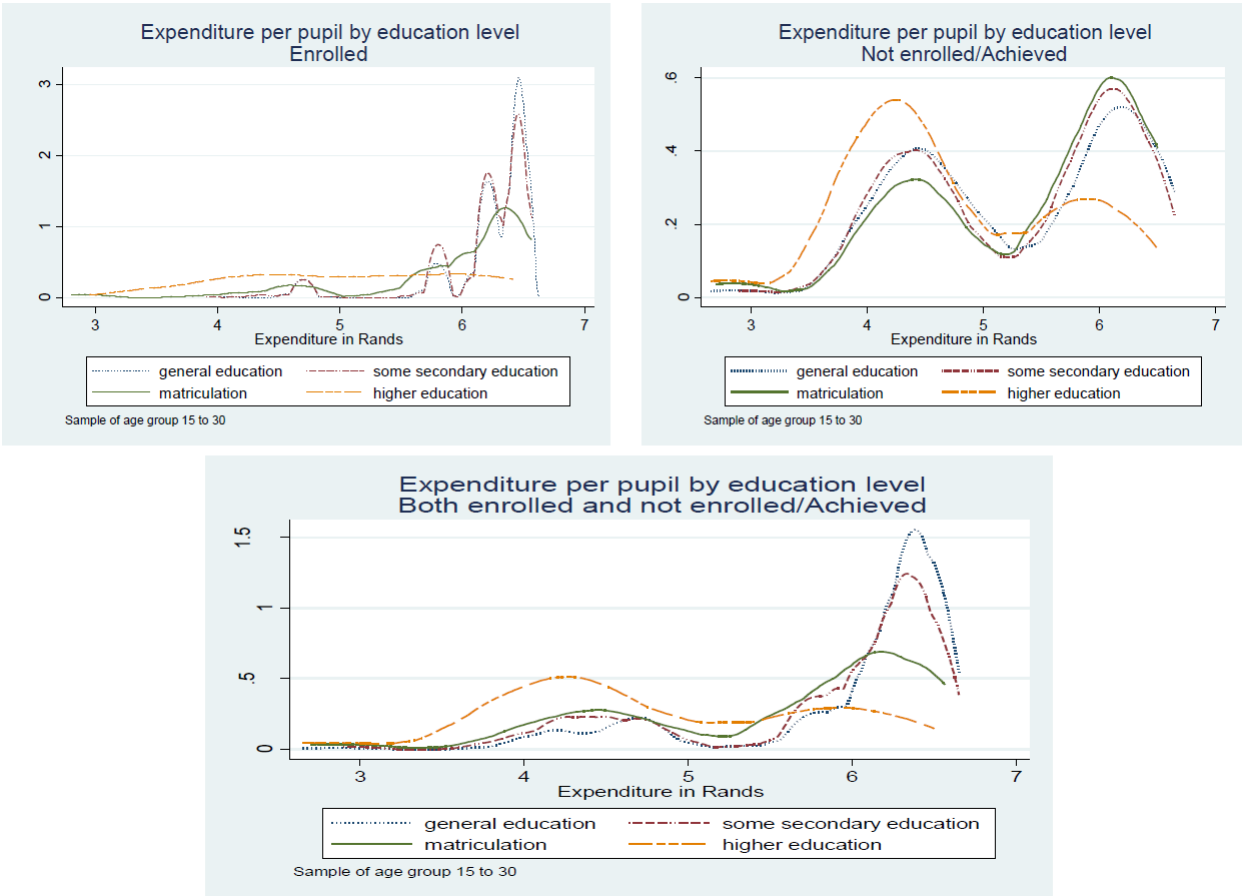
Figure 1: *Distribution of the logarithmic pupil-teacher ratio and the logarithmic expenditure per pupil for Africans who are enrolled*



From table 2 the average expenditure per pupil[10] is significantly different for those enrolled as compared with those not enrolled, that is, the average expenditure per pupil is R497 and R274 for those who are enrolled and those who have attained (not enrolled) respectively. The difference is evident in the densities presented in figure 2. From the figure, the distribution of expenditure per pupil for those enrolled in primary and secondary schools has multiple spikes and is skewed to the right. However, the distribution for those in higher education is rather flat, that is, platykurtic which could be an indication insensitivity to- or independence of- expenditure per pupil in individual's enrollment in higher education. The figure also indicates a bimodal distribution for those who have attained a given level of education. This distribution structure is dominant in the distribution of the full sample and is consistent with the 'advantaged' and 'disadvantaged' schools where the 'advantaged' schools which are mainly formerly White schools receive less funding but have higher matriculation rate and progression through higher education, and the 'disadvantaged' schools are mainly formerly Black schools which receive more funding per pupil but have lower matriculation rate and progression to higher education. A notable feature of the distributions of those who have attained a given level is the change in peaks between the different education levels as expenditure per pupil increases. At lower expenditure per pupil, higher education has the highest peak while matriculation has the lowest peak which implies that at low expenditure more individuals attained a higher education than at a matriculation. At higher expenditure per pupil more individuals have attained a matriculation, primary or some secondary education while fewer have attained a higher education.

[10] We use CPI to deflate the values taking 2008 is the base year.

Figure 2: *Distribution of the logarithmic pupil-teacher ratio and the logarithmic expenditure per pupil for Africans not enrolled/Achieved*



The full sample indicates all distribution are bimodal, but, both primary and secondary are more skewed to the right and higher education is skewed to the left. The skewness is, however, not surprising because the average age of those who have attained a given level of education is higher by about 3 years relative to those who are still enrolled and those who have attained a higher education are 4 percentage more than those still enrolled in higher education institutions. These differences in the distributions between enrollment and attainment are therefore more likely driven by time. The bimodal distribution is likely to be as a result of allocations pattern where quintile 1, 2, and 3 schools have progressively been allocated higher expenditure per pupil than quintile 4 and 5. Schools in quintile 1 and 2 were categorized as non-fee paying schools in 2007/08 and in 2012 schools in quintile 3 were also included. A significant proportion of Africans attend schools in quintile 1, 2, and 3. Schools in quintile 4 and 5 are mainly 'model C' schools, they receive substantially less allocation per pupil and have had a higher success rate at matriculation and enrollment into higher education. A naive comparison of these distributions tell us that, one there is no significant difference in expenditure per pupil between those who are enrolled or have attained a primary or some secondary education, second respondents who have attained a higher ed-

ucation have a lower expenditure per pupil than those who are currently enrolled, and third respondents who are enrolled in grade 12 or have achieved a matriculation received a higher expenditure per pupil which could be an indication that the increase in expenditure per pupil in formerly disadvantaged schools is increasing educational attainment. A Kolmogorov-Smirnov test to test show the distributions of expenditure per pupil for individuals who are enrolled, and for individuals who have achieved a primary, some secondary, and a matriculation are at 1%-level, and for higher education at 10% from a population with different distribution functions. This means that there is a significant difference between the expenditure per pupil for those enrolled and those who have achieved across the different levels of education. At all levels of education those who have attained are consistently from a lower expenditure per pupil distribution than those who are still enrolled in institutions of learning.

From table 2 the sample is of Africans on average in their 20's. The average age difference between those enrolled and those who have attained is 5 years. Those enrolled in primary school and at grade 12 are older for their age by 2 and 1 year respectively (This is based on grade 1 enrollment age of 7.). The sample is predominantly female with the exception of those enrolled or have attained a primary level at 44%. The majority of the individual's parents have on average no formal education while a relatively small proportion have at least some secondary education. Notably, 21.8% of individuals enrolled in grade 12(matriculation) have parents with at least some secondary education, and 10.8% and 18.3% of individuals who have attained a matriculation and higher education respectively have parents with at least some secondary education. One can naively infer from this that parental education does matter in their children's educational attainment. On average respondents have 3 siblings, which is a relatively small number. The average household monthly income per capita in the sample increases with education level the individual is enrolled at or has attained. The average income is lowest for individuals enrolled or have attained a primary education, and is highest for individuals enrolled or have attained a higher education. There is almost a threefold difference in the average household income for respondents enrolled or have attained a primary education and those enrolled or have attained a higher education.

In South Africa students in urban areas perform better than those in rural areas (van der Berg and Moses 2011). In the sample as indicated in table 2 the majority of individuals who are enrolled in primary, or in grade 10 or 11 live in traditional areas (which are defined as communally-owned land area and are mainly rural villages) while majority of those enrolled in grade12 or in higher institution of learning live in urban areas (cities, towns, 'townships', small towns, and hamlets). A similar pattern is repeated for those who have attained a given level except for those with a matriculation where about 50% live in rural areas. The majority (more than 80% except those enrolled in higher education at 63.5%) of individuals are enrolled or attended a formerly African only school, that is, independent homeland, self-governing territory, and Department of Education schools. On average 40% of individuals enrolled in primary school, or in secondary school at grade 10 or grade 11 attend formerly self-governing territory schools, and 42% and 38.6% of individuals enrolled in grade 12 and higher education respectively attend(ed) formerly Department of Education schools. It is interesting to note that a significantly higher propor-

tion of individuals enrolled in higher education attended formerly Department of Education schools as compared with those who have attained a higher education, a 24.3 percentage point difference. On average the highest proportion of individuals that have attained a given level of education attended formerly self-governing school at 38% followed by formerly Department of Education at 33.3%.

Most of these 10 neighborhood characteristics are related, we therefore consider it optimal to consolidate these characteristics into fewer categories for the regression analysis.We opt to use principal component analysis to identify and compute composite scores for these neighborhood characteristics. We name these components as: (i) serviced by local municipality, (ii) Community economic status , and (iii)Population density and housing quality In table 4 we present the mean characteristics of these components.

Table 3: Mean components

Components:	Enrolled				Not enrolled/achieved				Both enrolled and achieved			
	Primary	Some secondary	Matriculation	Higher	Primary	Some secondary	Matriculation	Higher	Primary	Some secondary	Matriculation	Higher
Serviced by municipality	-0.568	-0.260	0.546	0.768	-0.410	-0.177	-0.0341	0.393	-0.521	-0.229	0.141	0.490
	(0.160)	(0.137)	(0.166)	(0.279)	(0.172)	(0.142)	(0.152)	(0.200)	(0.154)	(0.126)	(0.135)	(0.180)
Community economic status	0.805	0.577	-0.0786	-0.147	0.662	0.429	0.473	0.0253	0.761	0.518	0.310	-0.0283
	(0.096)	(0.126)	(0.162)	(0.310)	(0.139)	(0.190)	(0.139)	(0.262)	(0.095)	(0.139)	(0.126)	(0.219)
Population density and housing quality	-0.137	-0.0262	-0.0834	0.126	-0.328	-0.262	-0.0693	-0.0554	-0.195	-0.128	-0.0763	-0.0265
	(0.072)	(0.090)	(0.177)	(0.273)	(0.092)	(0.119)	(0.116)	(0.166)	(0.068)	(0.094)	(0.111)	(0.156)
Sample size	**1671**	**1671**	**1671**	**1671**	**1352**	**1352**	**1352**	**1352**	**3023**	**3023**	**3023**	**3023**

Note: *(1)* Standard errors in parenthesis. *(2)* The sample is of African respondents aged 15-30 years. *(3)* See Appendix for definitions of all neighbourhood characteristics used to derive the components. *(4)* Design weights are used.

From table 4 fewer individuals who are enrolled or have attained a primary or grade 10 or 11 live in neighborhoods serviced by municipality while more individuals enrolled or have attained a higher education live in neighborhoods serviced by municipality. The majority of individuals enrolled or holding a primary education live in communities with better economic status as compared with other education level. More individuals enrolled or with a higher education live in neighborhoods of high density and lower quality of housing.

4 Empirical results

In this section we estimate the relationship between school inputs (pupil-teacher ratio and expenditure per pupil) and education level attained. We consider this relationship when other determinants of educational attainment are included. We present the estimated parameters under several model specifications. In the analysis of educational attainment we use a partial generalized ordered probit (Goprobit) to analyze actual education level attained on a sample of Africans who have achieved a given level of education, that is, those who are not currently enrolled in any educational institution.

4.1 Educational attainment

4.1.1 School inputs effects

Table 4 displays the estimated coefficients on logarithmic pupil-teacher ratio and logarithmic expenditure per pupil from the partial generalized ordered model. In the model, three different parameter vectors θ_1, θ_2, and θ_3 are estimated. The vectors give a series of estimated probit analyses, that is, θ_1 presents a probit analysis of having a primary education versus having some secondary, matriculation, and higher education, θ_2 presents a probit analysis of having primary and some secondary education versus having matriculation and higher education, and θ_3 presents a probit analysis of having a primary, some secondary, and matriculation versus having a higher education. The results are from our full specification which in addition to the school inputs controls for age, age squared, gender, parent's education, number of siblings, household monthly income per capita, home location, classification of school attended, distance to school, and neighborhood characteristics.

Table 4: Generalized ordered probit (Goprobit) education attainment models with school inputs as explanatory variables

	Goprobit		
Dependent variable: Education level attained	**θ_1**	**θ_2**	**θ_3**
Log(pupil-teacher ratio)	-0.513**	-0.513**	-0.513**
	(0.233)	(0.233)	(0.233)
Log(pupil-teacher ratio^2)	-0.258	-0.595***	-0.460
	(0.230)	(0.196)	(0.280)
Log(expenditure per pupil)	0.370***	0.0899	-0.313***
	(0.057)	(0.067)	(0.112)
Log(expenditure per pupil^2)	0.125*	0.0630	-0.0837
	(0.065)	(0.057)	(0.071)
Sample size	1352	1352	1352

Note: (1) Standard errors in parenthesis. (2) The * indicates significance difference at different levels: * p<0.1, ** p<0.05, and *** p<0.01. (3) control variables in the full specification include age, age squared, gender, parents education dummies, number of siblings, log of household monthly income per capita, school type per apartheid classification dummies, distance to school, home location, and neighbourhood characteristics components.(4) The sample is of African respondents aged 15-30 years. (5) Both expenditure per pupil and household income are in real Rands, deflated by CPI with 2008 as base year. (6) See Appendix for definitions of all variables used. (7) Post-stratification weights are used.

The coefficient on the logarithmic pupil-teacher ratio are negative and statistically significant at 5%-level[11], but they are nonlinear for θ_2 only. The pupil-teacher ratio has an identical effect on the likelihood of attaining any of the levels of education considered, that is, the ratio does not violate the parallel lines regression assumption. This can be interpreted to mean that a larger ratio increases the likelihood of attaining a primary education level or lower and decreases the likelihood of attaining a higher education holding all other characteristics constant. The quadratic ratio indicates the ratio is concave downward in all the categories, albeit only significantly so for θ_2. The coefficient on logarithmic expenditure per pupil have mixed signs an indication that it has a heterogeneous effect on educational attainment. A higher expenditure per pupil makes it more likely to attain more than a primary education as indicated by the positive and significant θ_1, has no effect on the probability of attaining a primary and some secondary education level as indicated by an insignificant θ_2, and has a negative and significant effect on the probability of attaining a matriculation or a lower education level as indicated by the coefficient for θ_3. Overall, these results imply that a higher expenditure per pupil is more likely to increase the likelihood of primary education attainment, but it is less likely to increase the likelihood of higher education attainment. The effect of an increase in expenditure per pupil decreases with the level of education. This makes sense to us since this expenditure per pupil represents current expenditure per pupil and is more likely to influence primary education attainment. Besides, expenditure per pupil has non-linear effects only on the likelihood of attaining a primary education or lower. See table 8 for estimates of the rest of the controls.

[11]This coefficient is larger both when we include no additional controls other than school inputs, and when we control for exogeneous variables, that is, age and gender and school inputs.

Table 5: Marginal probabilities of pupil-teacher ratio at representative values

Dependent variable: Education level attained	Primary	Some secondary	Matriculation	Higher
Pupil-teacher ratio by sample distribution				
Predicted probability at means	0.181***	0.371***	0.381***	0.0668***
	(0.0141)	(0.0199)	(0.0190)	(0.0102)
Log(18) (5 percentile)	-0.0740***	0.100*	-0.108***	-0.0955*
	(0.0175)	(0.0572)	(0.0341)	(0.0570)
Log(28) (25 percentile)	-0.0931***	0.0750**	-0.133**	-0.0715**
	(0.0157)	(0.0379)	(0.0597)	(0.0347)
Log(33) (50 percentile)	-0.100***	0.0629**	-0.139**	-0.0633**
	(0.0161)	(0.0273)	(0.0652)	(0.0274)
Log(36) (75 percentile)	-0.105***	0.0540***	-0.142**	-0.0581**
	(0.0167)	(0.0195)	(0.0675)	(0.0229)
Log(42) (95 percentile)	-0.112***	0.0402***	-0.145**	-0.0508***
	(0.0179)	(0.00930)	(0.0688)	(0.0170)

Note: (1) Standard errors in parenthesis. (2) The * indicates significance difference at different levels: * p<0.1, ** p<0.05, and *** p<0.01. (3) control variables in the full specification include age, age squared, gender, parents education dummies, number of siblings, log of household monthly income per capita, school type per apartheid classification dummies, distance to school, home location, and neighbourhood characteristics components.(4) The sample is of African respondents aged 15-30 years. (5) Both expenditure per pupil and household income are in real Rands, deflated by CPI with 2008 as base year. (6) See Appendix for definitions of all variables used. (7) Post-stratification weights are used.

We calculate marginal probability effects at representative values conditional on the mean in a bid to give a better interpretation of the effects of logarithmic pupil-teacher pupil ratio and logarithmic expenditure per pupil on education level attained, that is, primary/general, some secondary, matriculation, or a higher level.

From table 5 the marginal probability effects at the 5 presented ratios of the logarithmic pupil-teacher ratio are all positive and significant at 1% level. The magnitude of these effects, however, differ at the different values of the ratio within each level of education an indication of non-linearity, and at different levels of education attained within each value of the ratio. At the different values of the ratio the effects are increasing with the increase in the ratio values for both primary and some secondary education attained, that is, the effects are larger at higher ratios and smaller at a lower ratios, with the largest increase between the ratios being between 18 and 28 pupils per teacher. These marginal probability effects are, however, decreasing as the values of the ratio increases for attainment of matriculation and higher education, that is, larger effects at a lower ratio and smaller effects at a higher ratio, with the largest decrease being between 18 and 28 pupils per teacher. Overall, the differences in the effects of the ratio are highest at lower end of the distribution. This is evident in figure 3 below where the curves are steeper at the lower end of the distribution, they are steeper before 18 and steady out after 42.

From the table by the education level attained, the pupil-teacher ratio increases the probability of attaining some secondary education, but it decreases the probability of attaining any other level of educationholding other characteristics constant. For instance, for students attending a school with a pupil-teacher ratio of 18 their likelihood of attaining some secondary education is increased by 10 percentage

points. These effects are substantial compared to the average predicted probability of attaining some secondary education of about 37%. This increase is matched by a decrease of 9.6 percentage points of the likelihood of attaining a higher education. However, this effect is minimal given the average predicted probability of attaining a higher education of about 6.7%. The results of different ratios show the effects are increasing in magnitude as the ratio increase for primary and matriculation, but are decreasing for some seconday and higher education. For example, a doubling of the ratio from 18 to 36 leads to a 42% decrease in the likelihood of attaining a primary level as compared to 39% decrease in the likelihood of attaining a higher education. A higher ratio, therefore, has a larger effect on the likelihood of attaining a primary education than that of a higher education, and these effect is even larger the higher the ratio. From these results we can infer that the pupil-teacher ratio has varied effects on different education level, but a higher ratio is more detrimental at primary level.

Figure 3: Distributions of marginal probability of pupil-teacher ratio effects at representative values

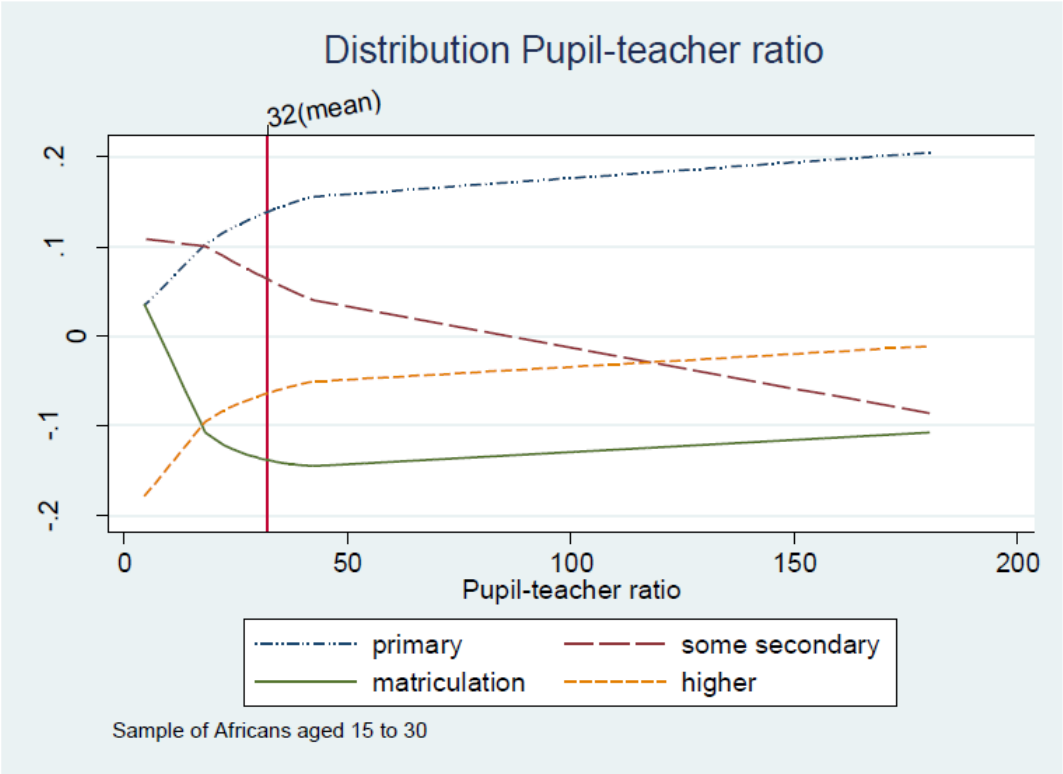


Figure 3 displays the marginal probability effects at different points of the sample distribution of the logarithmic pupil-teacher ratio from the partial generalized ordered model plotted against values of the pupil-teacher. From the figure an increase in the pupil-teacher ratio decreases the likelihood of attaining some secondary education and matriculation but increases the likelihood of attaining a primary education and a higher education but at a decreasing rate. A possible explanation for this is that: on the one hand younger children are more reliant on teachers than older children hence a decline in attainment with a

higher ratio for primary, and on the other hand a minimum class size is important in increasing the probability of attaining some secondary education or a matriculation a likely benefit of peer learning which is likely to occur in older children. Peer relations have been argued to influence education aspiration and achievement and attitude towards schooling (Johnson; 1981). From the figure and the results in table 5 we can conclude that a large pupil-teacher ratio has the least effect on higher education.

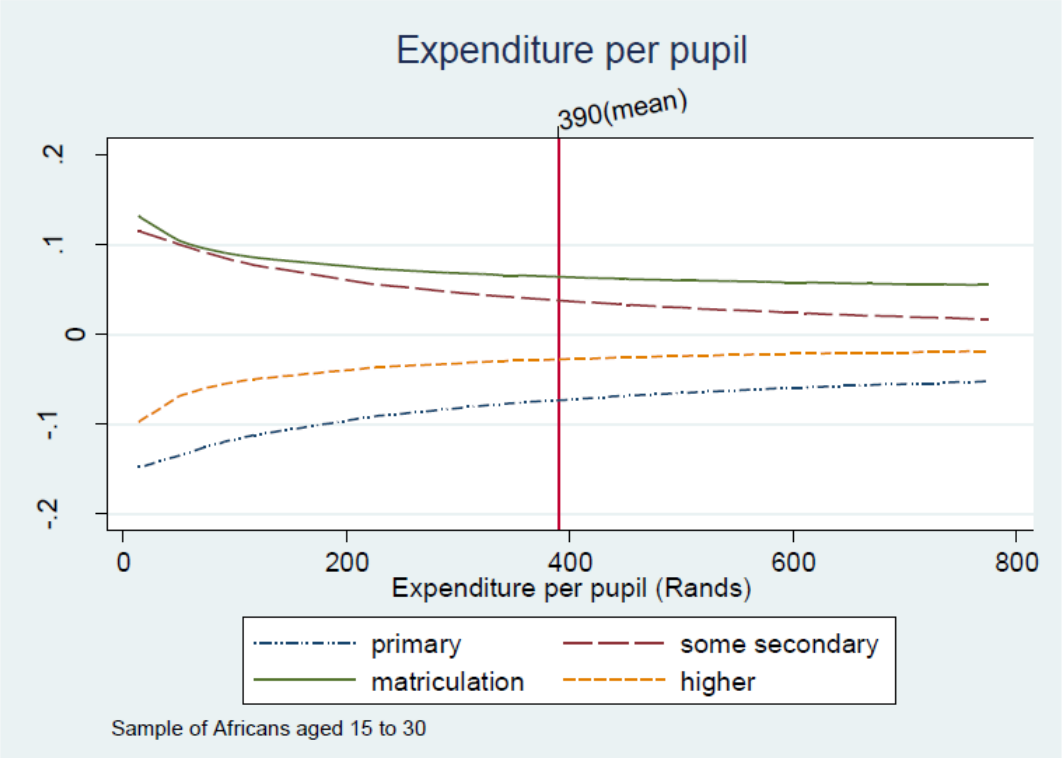
Table 6: Marginal probabilities of expenditure per pupil at representative values

Dependent variable: Education level attained	Primary	Some secondary	Matriculation	Higher
Expenditure per pupil by sample quintile average				
Predicted probability at means	0.181***	0.371***	0.381***	0.0668***
	(0.0141)	(0.0199)	(0.0190)	(0.0102)
Log(55) (Quintile 5)	-0.133***	0.0982***	0.102**	-0.0670*
	(0.0249)	(0.0227)	(0.0411)	(0.0350)
Log(165) (Quintile 4)	-0.101***	0.0659***	0.0783***	-0.0428**
	(0.0169)	(0.0213)	(0.0301)	(0.0171)
Log(257) (Quintile 3)	-0.0868***	0.0511**	0.0703**	-0.0346***
	(0.0127)	(0.0218)	(0.0280)	(0.0114)
Log(339) (Quintile 2)	-0.0777***	0.0420*	0.0658**	-0.0301***
	(0.0102)	(0.0225)	(0.0274)	(0.00849)
Log(391) (Quintile 1)	-0.0731***	0.0373	0.0636**	-0.0278***
	(0.00899)	(0.0229)	(0.0271)	(0.00717)
Expenditure per pupil by 2008 quintile allocation				
Log(129) (Quintile 5)	-0.109***	0.0739***	0.0832***	-0.0478**
	(0.0192)	(0.0214)	(0.0318)	(0.0207)
Log(388) (Quintile 4)	-0.0733***	0.0375	0.0637**	-0.0280***
	(0.00905)	(0.0229)	(0.0272)	(0.00724)
Log(581) (Quintile 3)	-0.0608***	0.0249	0.0581**	-0.0223***
	(0.00605)	(0.0243)	(0.0269)	(0.00430)
Log(711) (Quintile 2)	-0.0548***	0.0190	0.0556**	-0.0198***
	(0.00492)	(0.0251)	(0.0268)	(0.00333)
Log(775) (Quintile 1)	-0.0524***	0.0166	0.0546**	-0.0188***
	(0.00453)	(0.0253)	(0.0268)	(0.00304)

Note: (1) Standard errors in parenthesis. (2) The * indicates significance difference at different levels: * p<0.1, ** p<0.05, and *** p<0.01. (3) control variables in the full specification include age, age squared, gender, parents education dummies, number of siblings, log of household monthly income per capita, school type per apartheid classification dummies, distance to school, home location, and neighbourhood characteristics components.(4) The sample is of African respondents aged 15-30 years. (5) Both expenditure per pupil and household income are in real Rands, deflated by CPI with 2008 as base year. (6) See Appendix for definitions of all variables used. (7) Post-stratification weights are used.

In table 6, we display these marginal probability effects of logarithmic expenditure per pupil from the partial generalized ordered probit model. For the logarithmic expenditure per pupil we present marginal probability effects at averages of each quintile in the sample, and at amounts allocated to schools by quintile in 2008. From table 6, the marginal effects of logarithmic expenditure per pupil by both sample quintile averages and 2008 quintile allocation show that the allocation increases the likelihood of attaining some secondary education and matriculation, but decreases the likelihood of attaining a primary education and a higher education. For instance an expenditure of R129 increases the likelihood of

Figure 4: Distributions of marginal probability of expenditure per pupil effects at representative values



attaining a matriculation by 8.3 percentage points, but it decreases the likelihood of attaining a primary education by 10.9 percentage points. The effect on matriculation is substantial given the average predicted probability of attaining this level of about 38%. Higher allocations, however, have no effect on the likelihood of attaining some secondary education. The results also show the magnitude of the effect to be decreasing with the increase in the allocation. From both table 6 and figure 4 expenditure per pupil has the least effect on attainment of higher education, at least a twofold difference between the effect on primary and on higher education.

Figure 4 displays the marginal probability effects at different points of the sample distribution of the the logarithmic expenditure per pupil from the partial generalized ordered model plotted against values of the expenditure per pupil. From figure 4, the results show two distinct patterns: the likelihood of attaining a primary or a higher education increases as expenditure increases, but the likelihood of attaining some secondary education or matriculation decreases as expenditure increases. Although not much research has been done, the existing research shows that an increase in expenditure per pupil increases educational attainment. These findings agree with some studies (Hanushek; 1995, Glewwe et al.; 2011, and Gustafsson; 2007) which find expenditure per pupil has positive and significant effects on attainment particularly when used for items such as textbooks, workbooks, school equipment among others. This are essential items particularly for matriculation. In South Africa Van der Berg (2007)

and Bhorat and Oosthuizen (2008) also find resources to positively affect matriculation pass rate. Our results are also in agreement Yamauchi (2011) who finds increase in expenditure per pupil improves school quality. This results therefore indicate that increased expenditure per pupil increases education attainment particularly matriculation which is contrary to the public sentiment.

4.1.2 Individual, family and neighborhood effects

In table 7 we present our full specification results from the partial generalized ordered probit model, and two least square estimation of the relative education index and years of schooling. In all models we control for individual, family and neighborhood characteristics which have empirically been shown to be important determinants of educational attainment. In this section we discuss the results for the partial generalized ordered probit analysis as presented in the first three columns. We find age to be an insignificant determinant of the education level attained, however, it has the 'wrong' sign and has significant nonlinear effects. This is a likely to be as a result of small variation in age likely to be driven by the sub-sampling of 15-30 year old. We find being female increases the probability of attaining a higher education implying women are more likely to attain a primary education than men. Both age and gender have varying effects on the different education levels attained.

Family characteristics are important determinants of educational attainment. We consider parental education both as an indicator of parental taste in education and as a variable in its own right. The results shows that having a parent with at least some secondary education increases the likelihood of attaining a higher level of education relative to having a parent with no formal education. We can infer from these results that more educated parents are likely to give more educational opportunities to their children hence offering more social mobility for their generations. Gustafsson (2007) finds similar results for South Africa. An increase in household per capita income increases the likelihood of attaining a higher education. The positive correlation between household income and education attainment implies that education is a normal good. The effects of family characteristics on the likelihood of educational attainment do not vary with education level. We, however, find number of siblings and location to be of no importance.

Educational attainment is dependent on the type of school attended, in the case of South Africa schools differ by historical (*apartheid*) classification. From table 8 attending a formerly self-governing school significantly (at 10%-level) increases the likelihood of attaining a primary education or lower while attending any other school category has no significant effect relative to attending a formerly independent homeland school. The importance classification of school attended is in line with the findings byBhorat and Oosthuizen (2008) who found the matriculation pass was positively determined by type of school attended as defined by *apartheid* classification. A long distance to school increases the likelihood of attaining a primary education or lower. School classification and distance to school have heterogeneous effects at different levels of education.

Neighborhood characteristics are important in the determination of educational attainment particularly due to associated externalities. We find that living in a neighborhood that is serviced by the municipality has no effect on the likelihood of educational attainment. However living in a neighborhood of high economic status increases the likelihood of attaining a primary education or lower while, living in a less densely populated area with better quality housing increases the likelihood of attaining a higher education. Although we do not discuss the marginal effects of these characteristics at their mean, table 8 in the appendix presents these results.

Table 7: Determinants of education attainment for Africans

	GOprobit			Least square	
Variables:	θ_1	θ_2	θ_3	Relative education index	Years of schooling
School Inputs: Log(pupil-teacher ratio)	-0.513**	-0.513**	-0.513**	-0.0748***	-0.703***
	(0.233)	(0.233)	(0.233)	(0.027)	(0.255)
Log(pupil-teacher ratio^2)	-0.258	-0.595***	-0.460	-0.0785**	-0.722**
	(0.230)	(0.196)	(0.280)	(0.036)	(0.339)
Log(expenditure per pupil)	0.370***	0.0899	-0.313***	0.0283***	0.274***
	(0.057)	(0.067)	(0.112)	(0.007)	(0.071)
Log(expenditure per pupil^2)	0.125*	0.0630	-0.0837	0.0108*	0.116**
	(0.065)	(0.057)	(0.071)	(0.006)	(0.059)
Individual characteristics: Age	-0.0403	-0.0403	-0.0403	0.0067*	-0.109***
	(0.030)	(0.0230)	(0.030)	(0.003)	(0.034)
Age^2	-0.0207***	-0.0092**	-0.0129**	-0.00007	-0.030***
	(0.004)	(0.004)	(0.005)	(0.0002)	(0.002)
Gender (female=1)	0.530***	0.0612	0.209	-0.0223***	0.528***
	(0.110)	(0.107)	(0.162)	(0.008)	(0.080)
Parent's Education(No schooling=0):					
Primary	0.0939	0.0939	0.0939	0.0061	0.0744
	(0.081)	(0.081)	(0.081)	(0.009)	(0.084)
At least some secondary	0.722***	0.722***	0.722***	0.0672***	0.659***
	(0.156)	(0.156)	(0.156)	(0.013)	(0.118)
Number of siblings	0.0367	0.0367	0.0367	0.0021	0.0206
	(0.023)	(0.023)	(0.023)	(0.003)	(0.025)
Household income per capita(monthly Rands)	0.211***	0.211***	0.211***	0.0241***	0.233***
	(0.060)	(0.060)	(0.060)	(0.006)	(0.058)
Home location(Rural=1):					
Urban	-0.0098	-0.0098	-0.0098	0.0458***	0.430***
	(0.153)	(0.153)	(0.153)	(0.014)	(0.136)
Farm	-0.239	-0.239	-0.239	0.0303**	0.284**
	(0.216)	(0.216)	(0.216)	(0.015)	(0.144)
Ex-department classification(independent homelands=1)					
Self-governing territory	-0.367*	0.0081	-0.222	0.0485***	0.472***
	(0.214)	(0.207)	(0.301)	(0.019)	(0.179)
Department of education	-0.0931	0.112	0.271	0.0630***	0.592***
	(0.182)	(0.193)	(0.238)	(0.021)	(0.204)
House of assembly, representative and delegates	-0.130	-0.130	-0.130	-0.00001	-0.0001
	(0.150)	(0.150)	(0.150)	(0.00001)	(0.0001)
New schools	0.176	0.176	0.176	-0.0255	-0.220
	(0.196)	(0.196)	(0.196)	(0.017)	(0.160)
Distance to school (kms)	-0.0007**	0.00003	0.00003	-0.0234	-0.222
	(0.0003)	(0.0001)	(0.0001)	(0.024)	(0.237)
Neighbourhood characteristics: Serviced by municipality	0.0218	0.0218	0.0218	0.0183***	0.172***
	(0.048)	(0.048)	(0.048)	(0.006)	(0.058)
Community economic status	-0.109*	-0.109*	-0.109*	-0.0125**	-0.122**
	(0.056)	(0.056)	(0.056)	(0.005)	(0.050)
Population density and housing quality	0.121**	0.121**	0.121**	0.0128**	0.125**
	(0.056)	(0.056)	(0.056)	(0.005)	(0.049)
Constant	-0.215	-1.467***	-2.734***	0.915***	9.224***
	(0.447)	(0.453)	(0.475)	(0.040)	(0.397)
F-statistic				10.11	45.90
Sample size	**1352**	**1352**	**1352**	**1352**	3023

Note: *(1)* Standard errors in parenthesis. *(2)* The * indicates significance difference at different levels: * p<0.1, ** p<0.05, and *** p<0.01. *(3)* Pupil-teacher ratio is as provided by the Department of Basic Education in the NIDS administration data. *(4)* Expenditure per is learner derived from matching respondent's school quintile variable in the NIDS administration data to the amount of expenditure allocated by the government to each student using school quintile every year. *(5)* For all the neighbourhood characteristics the variables are calculated by *(a)* calculating the average value in each cluster and *(b)* assigning each individual a value corresponding to his/her cluster. *(6)* The sample is of African respondents aged 15-30 years. *(7)* Both expenditure per pupil and household income are in real Rands, deflated by CPI with 2008 as base year. *(8)* See Appendix for definitions of all variables used. *(9)* Post-stratification weights are used.

5 Summary and conclusion

The objective of this study has been to investigate the determinants of educational attainment for Africans in South Africa with a specific interest on school inputs -pupil-teacher ratio, and expenditure per pupil. We considered both actual education attained and relative educational attainment index. We used the National Income Dynamics Survey Wave 1, a household survey that has been matched with national school data to identify the impacts of school quality (pupil-teacher ratio and expenditure per pupil) and socioeconomic characteristics such as parental education, home language, and location on education attainment. We also consider neighborhood characteristics such as access to piped water, availability of electricity and street lighting, availability landline telephone, collection of refuse by municipality, availability of toilet facility, average household size and house size, and community unemployment rate We focus on Africans aged 15 to 30 years who are not household heads or in a position of having dependents.

We find that school inputs and some of socio-economic factors considered are important determinant of educational attainment. Our findings suggest that a low pupil-teacher ratio increases the likelihood of attaining some secondary education and similarly a larger expenditure per pupil increases of attaining some secondary education. The pupil teacher ratio has different effects: at different values, that is, it is nonlinear, and on attainment of different education levels. We find a smaller pupil-teacher ratio has higher effects at lower education levels. The ratio is therefore more important in influencing the likelihood of educational attainment at lower level of education than at higher levels. We also find a larger expenditure per pupil increases the likelihood of attaining a higher education. our results indicate that smaller allocations of expenditure per pupil have larger effects on the likelihood of attainment and these effects vary by education level attained. The effects of larger allocations mainly increase the likelihood attaining a primary education and higher education levels but, has decreasing effects on the likelihood of attaining some secondary education and matriculation. We infer that the higher effects on primary education attainment are an indication that the increased expenditure per pupil is having significant effects in improving matriculation. We, however, caution that the fact that education is compulsory till end of primary school could be an important factor in increased attainment at this level thereby is likely to be confounding our results.

Appendix 1

Table 8: Definition of variables

Variable	Definition
Education level attained	The highest education level a respondent has attained. The categories are defined as: primary/general (grade 1-9), some secondary (grade 10 or/and 11), matriculation (grade 12), and higher (grade 9 and further education and training certificate, or and grade 12 plus certificate or diploma or degree).
Pupil-teacher ratio	Pupil-teacher ratio in each school as per the Department of Basic Education records matched to NIDS school administration data
Expenditure per learner	Expenditure per is learner derived from matching respondent's school quintile variable in the NIDS administration data to the amount of expenditure allocated by the government using school quintile every years in Rands. We use CPI with 2008 as the base year to calculate the real values.
Age	Respondents age in years as calculated using interview date less their date of birth.
Gender	Respondents gender, female=1 and male=0
Number of siblings	Respondents number of sibling
Respondents/Mothers/fathers education	No schooling=0, general education(grade 1-9)=1, and at least some secondary(grade 10-12, further education and training certificate, and bachelor's degree and other qualifications)=2
Household monthly income	Total household monthly income per capita in Rands. We use CPI with 2008 as the base year to calculate the real values.
Geographic area	Respondents geographical home area; Traditional/rural area=1, , Urban=2, and farm=3
Ex-department classification	Independent homelands=1, Self-governing territory=2, Department of education=3, House of Assembly, representative, delegates=4, and New schools=5
Neighbourhood characteristics	The neighbourhood characteristics the variables are calculated by (a) calculating the average value in each cluster and (b) assigning each individual a value corresponding to his/her cluster. They are: • average rooms per person per PSU, • average number of household members per PSU % of respondents with electricity per PSU, • respondents who share toilet with other households per PSU=1 • respondents who live in a neighbourhood refuse collected by the municipality per PSU=1 • respondents with electricity connected per PSU=1 • Respondents with street lights in their neighbourhood per PSU=1 • people employed per PSU=1 • respondent with pipe water per PSU=1

References

Bhorat, H. and Oosthuizen, M. (2008). Determinants of grade 12 pass rates in the post-apartheid south african schooling system, *Journal of African Economies* **18**(4): 634–666.

Bhorat, H., Van der Westhuizen, C. and Jacobs, T. (2009). *Income and non-income inequality in post-apartheid South Africa: what are the drivers and possible policy interventions?*, Development Policy Research Unit, University of Cape Town.

Boes, S. and Winkelmann, R. (2006). Ordered response models, *Modern Econometric Analysis*, Springer, pp. 167–181.

Case, A. and Deaton, A. (1998). School quality and educational outcomes in south africa, *Technical report*.

Case, A. and Deaton, A. (1999). School inputs and educational outcomes in south africa, *The Quarterly Journal of Economics* **114**(3): 1047–1084.

Department of Education, . (2003). A review of financing, resourcing and costs of education in public schools, *Technical report*, Department of Education, pretoria, South Africa.

Glewwe, P. W., Hanushek, E. A., Humpage, S. D. and Ravina, R. (2011). School resources and educational outcomes in developing countries: A review of the literature from 1990 to 2010, *Technical report*, National Bureau of Economic Research.

Glick, P. and Sahn, D. E. (2000). Schooling of girls and boys in a west african country: the effects of parental education, income, and household structure, *Economics of education review* **19**(1): 63–87.

Gradin, C. (2014). Race and income distribution: Evidence from the usa, brazil and south africa, *Review of Development Economics* **18**(1): 73–92.

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

Hall, K. and Giese, S. (2008). Addressing quality through school fees and school funding, *South African child gauge* **2009**: 35–40.

Hanushek, E. (2013). Economic growth in developing countries: The role of human capital, *Economics of Education Review* .

Hanushek, E. A. (1995). Interpreting recent research on schooling in developing countries, *The world bank research observer* **10**(2): 227–246.

Johnson, D. W. (1981). Student-student interaction: The neglected variable in education, *Educational researcher* pp. 5–10.

Leibbrandt, M., Woolard, I. and de Villiers, L. (2009). methodology: report on nids wave 1, *technical paper no. 1*, South African Labour and Development Research Unit.

Mani, S., Hoddinott, J. and Strauss, J. (2009). Determinants of schooling outcomes: Empirical evidence from rural ethiopia, *Fordham University, Department of Economics Discussion Paper* (2009-03).

Republic of South Africa, . (1998b). National norm and standards for school funding, south african school act, 1996, *Technical report*, Department of Education.

Selod, H. and Zenou, Y. (2003). Private versus public schools in post-apartheid south african cities: Theory and policy implications, *Journal of Development Economics* **71**(2): 351–394.

Tansel, A. (1997). Schooling attainment, parental education, and gender in cote d'ivoire and ghana, *Economic Development and Cultural Change* **45**(4): 825–856.

Taylor, N., Fleisch, B. and Shindler, J. (2008). Changes in education since 1994, *Higher Education* **32**: 3–5.

United Nation, . (2010). Keeping the promise: united to achieve the millennium development goals, *technical manual*, United Nation.

United Nation, . (2013). The millennium development goals report 2013, *Technical report*, United Nation.

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

Van der Berg, S. and Moses, E. (2011). Inertia in south african education quality: evidence from sacmeq ii and iii, *Policy brief*, department of economics, University of Stellenbosch.

Williams, R. (2006). Generalized ordered logit/partial proportional odds models for ordinal dependent variables, *Stata Journal* **6**(1): 58–82.

Wilson, K. (2001). The determinants of educational attainment: Modeling and estimating the human capital model and education production functions, *Southern Economic Journal* pp. 518–551.

Wittenberg, M. (2002). Job search in south africa: a nonparametric analysis, *South African Journal of Economics* **70**(8): 1163–1196.

Yamauchi, F. (2011). School quality, clustering and government subsidy in post-apartheid south africa, *Economics of Education Review* **30**(1): 146–156.