PREDICTIVE ABILITY OF ORDINARY CERTIFICATE OF SECONDARY EDUCATION GRADES TO ADVANCED CERTIFICATE OF SECONDARY EDUCATION: EVIDENCE FROM TANZANIA'S NECTA RESULTS

Ulingeta O.L. Mbamba*

ABSTRACT

Normally, selection of students for higher education is based on previous performance. Whether or not previous performance is a good predictor varies for different examining bodies. This study evaluates the predictive ability of Tanzania's Ordinary Certificate of Secondary Education (OCSE) performance to Advanced Certificate of Secondary Education (ACSE) performance using linear regression and chi-square tests. By linking OCSE of 2003 and 2004 with ACSE of 2006 and 2007, respectively, research results indicate that the OCSE is a good predictor for ACSE. The predictive ability of OCSE improves as data are disaggregated based on whether or not students studied at the same school and based on subject combinations, especially for science-based subjects. This indicates that admission criteria, generally employed for ACSE, serve the academic purposes they are assumed to be serving.

INTRODUCTION

One of the areas that has received attention by many educationists is students' academic success. The main aim is to be able to predict a student's performance before he/she joins any further academic programmes. The process of predicting is a challenging one because students' performance could be based on pre-entry qualifications (which can be easily established), as well as their performance during the course and external factors (which cannot be established *a priori*). This complicates the process of selecting students for various further education programmes.

This paper explores whether or not ordinary certificate of secondary education (OCSE) results, used mainly as selection criteria for advanced certificate of secondary education (ACSE), are a good predictor of ACSE.

Background Information

In the Tanzanian education system, a pupil joins non-compulsory nursery school at three or four years of age for three to four years. Thereafter, a pupil enters standard one to seven for seven years. The pupil may then join form one for OCSE

Business Management Review 10(2) pp 61- ISSN 0856-2253

* Dr. Ulingeta O.L. Mbamba; Lecturer in the Department of General Management, Faculty of Commerce and Management, University of Dar es Salaam, P.O. Box 35046, Dar es Salaam, Tanzania; e-mail: mbamba@fcm.udsm.ac.tz

for four years, where he/she is supposed to take seven plus subjects. The student may proceed to form five and six for ACSE for two years, where he/she will take three principal subjects. However, it is possible to find a student having more than three principal subjects before joining a university or any institution of higher learning for a degree programme that may last between three to five years, depending on the degree programme selected. There are two main subsidiary subjects at ACSE, namely, basic applied mathematics (BAM) and general studies. All ACSE are expected to take general studies, while BAM is for all students who are taking science and few arts-related subjects like economics, provided that the student is not taking advanced mathematics as one of the principal subjects.

At OCSE and ACSE, students' performance is aggregately measured using divisions (performance level). There are five divisions in both cases. These are divisions one, two, three, four and failed (commonly known as division zero), in a descending chronology.

Examinations at OCSE and ACSE are administered by the National Examination Council of Tanzania (NECTA). NECTA was established by Act of Parliament in ….

Statement of the Problem

Students are selected to join ACSE based on their performance at the OCSE. The assumption is that performance at OCSE is a good predictor of ACSE. Studies indicate that there are a number of factors that influence students' performance (Hanushek,1986; Todd and Wolpin, 2003; Hough,1981; Bradley and Taylor,1998; Sawkins, 2002, Peterson and Barrett,1986; Petripin and Johnson,1991, Tiggemann and Crowley,1993; Mbamba1993). This paper explores whether or not performance at OCSE is a good predictor of ACSE.

For a country like Tanzania, only a third of all students who do OCSE are admitted to high schools for ACSE. All of these are selected, based on their OCSE results. As stated before, the assumption is that OCSE results are a good predictor of ACSE. Schools scramble to get high-performing students in order to increase their ranking in performance. The high ranking will normally attract more students in the future. With an admission ratio of approximately 3:1, one would like to know the answers to questions that arise about admission criteria for the programme. Would a high-performing student in OCSE perform better? To what extent does admission decision variables, which is OCSE, predict a student is academic performance? This is based on the observed failure rate. For example, in 2006, out of 21,001 students who sat the ACSE examinations, 768 failed. This is 3.7%. In 2007, out of 26,261 who sat the ACSE examinations, 2,216 students failed. This is nearly 10%. One may wonder whether or not the criteria used for selecting students were appropriate (the data is based on schools candidates only, it ignores private students).

The importance of accurate estimation of a student's future performance based on current performance is essential in order to provide the student with adequate assistance in the learning process. The research will provide as a good input into the advisory services for students, parents, school administrators and regulatory authorities in education. Better use of results will improve ACSE performance. The research will contribute to the body of knowledge regarding administration of examinations and selection of students for further studies.

Methodology

The research data used a sample of all students who completed OCSE in 2003 and 2004. These students thereafter completed ACSE in 2006 and 2007. These years were selected as their data were readily available (convenience sampling). OCSE results for 2003 (2004) and ACSE 2006 (2007) results were downloaded from the NECTA website. Results were on the hypertext markup language. All results, together with additional information, were exported to an electronic spreadsheet. While data were on the spreadsheet, the linking process was done. In that process, it was found that 60,222 and 63,982 students sat Tanzania's OCSE in 2003 and 2004, respectively, as school candidates. Some of these were selected to join form five and sat the ACSE. Thus, some of these students sat the ACSE in both 2006 and 2007. In 2006, there were 21,001, while in 2007 there were 26,460 ACSE students. Using spreadsheets, these students were linked together in order to know how many students were the same in both categories, that is, OCSE in 2003 and 2004 with ACSE in 2006 and 2007 respectively. The linking field was the names of students. The number of times OCSE names appeared in ACSE is shown in Table 1. From Table 1, it is evident that there are some names that appeared more than once. All names that did not appear in ACSE results, as well as names that appeared more than once, were deleted from the sample.

Table1: Counting Number of Times OCSE names appeared in ACSE

Number of times the name appears	Total number of Names in 2006	Total number of Names in 2007
0	42671	42399
1	16901	20572
2	455	676
3	105	211
4	56	61
5	13	63
6	10	0
7	0	0
8	11	0
	60,222	63982

Source: Data analysis (2007)

The same process was carried out for ACSE names numbering 21,001 and 26,460 for the two years which were linked with the remaining names in OCSE (there were 16,901 and 20,572 names, which appeared once in 2006 and 2007, respectively). The number of times the ACSE names appeared in the remaining OCSE results is shown in Table 2. All the names that did not appear in OCSE or appeared more than once were deleted. The total number of candidates that remained was 15,510 and 18,764 names in ACSE for the two years.

Table 2: Counting Number of Times ACSE names appeared in OCSE

Number of times the name appears	Total number of names in 2006	Total number of names in 2007
0	4916	6934
1	15510	18764
2	421	568
3	98	132
4	35	41
5	14	15
6	6	5
7	0	1
8	0	
9	1	
	21001	26460

Source: Data analysis (2007)

By this time, there were 16,901 (20,572 for 2004) names in OCSE while for ACSE there were 15,510 (18,764 for 2007) names. Again, by cross-linking (by names comparisons), 1,392 names were deleted from OCSE and 1 name was also deleted from the ACSE list. Therefore, in both cases, 15,509 (18,735) names remained. In order to ensure that the names really were the same, the indicated sexes in results for OCSE and ACSE were compared. For names with sexes in ACSE not in agreement with sexes in OCSE were deleted from the sample. The final sample remained with 15,487 (18,726). Therefore, the study had a sample of 34213. Furthermore, 123 students were removed as their ACSE points were 0 and their divisions were not indicated for 2007, making a usable sample of 34,090 students.

Detailed Data Analysis

This section was guided by several research questions. The main data analysis techniques were chi-squire tests and regression analysis.

Measures of Association: Chi-square and Contingency Tables

Measures of association for normal data do not depend on a particular order in which categories are listed. These are several measures of association. However, most of them depend on the Chi-square statistic. Liebetrau (1976) summarises the main common measures of association. Some of the measures are Pearson's coefficient of mean square, Pearson Contingency Coefficient and Sakoda Modification, Tschuprow's Contingency Coefficient (ibid.). Generally, the chi-square test statistic checks whether or not the two data sets are related, that is, associated. This means that after knowing one characteristic, one can know the second one. This is also known as test of independence, whereby two data sets are independent of one another.

There are two ways to check whether or not conclusions made from chi-square should be interpreted. These are based on minimum value in each cell. One argument is that all values in each cell in a contingency table should be greater than 5. The second line of thought is that all values in the contingency table should have the expected value greater than one for each cell when either of the number of rows or columns is two (Everett,1977; Slakter,1966; Lewontin and Felsenstein,1965). This study presents both of the values. If any of the values are not met, the test is discarded.

However, a major weakness of the chi-square test is its dependence on sample size. If the sample is extremely small the chance of failing to reject null hypotheses increases. On the other hand, if the sample is extremely big, the chance of always accepting competing (alternative) hypotheses increases. There are several modifications proposed in order to rectify this problem (see also Joreskog and Sörbom,1982). They include goodness of fit index (GFI), adjusted goodness of fit index (AGFI), and root mean square residual (RMS). However, even these indices are not without problems because they all stem from the chi-square statistic (see Bone et al.,1989). Some scholars present a rule of thumb in concluding whether or not there is an association or not (for example Browne and Cudeck, 1992; Hu and Bentler, 1999).

Regression Analysis

Regression analyses are sets of techniques that allow one to assess the relationship between one dependent variable and several independent variables. There are several assumptions regarding the use of regression analysis. These assumptions have to be observed during analysis. Unfortunately, some assumptions are based on the rule of thumb, especially on sample size.

The first assumption is about sample size. Green (1991) provides two rules of thumb in determining the number of variables to be used. These are $N \geq 50 + 8m$ (for medium-size effect relationship between independent and

dependent variables) and $N \geq \frac{8}{f^2} + (m-1)$ where N is the sample size, m is the number of the independent variable and f is effect size (f^2 is equal to 0.01, 0.15 and 0.35 to account for small, medium and large effects of the sample). Specifically $f^2 = \frac{R^2}{(1-R^2)}$, where R^2 is the expected squared multiple correlation. A stepwise regression requires the ratio $N \geq 40m$ (Tabachnick and Fidell, 1996). The study nearly met these requirements for all performed analyses.

The study met all the conditions for using regression as well as chi square tests. The sample size was large enough as required by regression analysis. The large sample in chi-squire normally makes it reject null hypotheses. The sample size was 34,090. Even with data sample segregation, the sample was not reduced to less than 2,000.

Research Questions

a. Is there any association between divisions obtained in OCSE and ACSE?

In order to check whether there is any association between divisions obtained in OCSE to ones obtained at ACSE, divisions obtained in OCSE and ACSE were cross-tabulated. Table 3a provides a cross-tabulation of divisions in these two levels. In order to establish whether or not there is an association, chi-squire test was conducted. At 1% significance level, there was an association between the two divisions, ACSE and OCSE (Table 3b). Given that it is known one has to sit the OCSE before ACSE, it may be an intelligent guess that OCSE results predict ACSE results. From the same tables (table 3a and 3b), the strength and direction of association is positive (likelihood ratio).

A low significance value (typically less than 0.05) for Kendall's tau indicates that there is a relationship between the two variables. This has also been observed in this research.

Table3a Cross-tabulation of Division Obtained at OCSE versus ACSE

		F6_DIV					
		I	II	III	IV	FLD	Total
F4_DIV	I	1720	2052	2095	385	76	6328
	II	798	1724	3383	1487	447	7839
	III	1883	4182	6561	3367	1338	17331
	IV	128	506	1132	575	207	2548
	FLD	4	11	21	7	1	44
	Total	4533	8475	13192	5821	2069	34090

Source: Research Data (2007)

Table3b Cross-tabulation of Division Obtained at OCSE versus ACSE Chi-Square Tests

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	2412.383	16	.000
Likelihood Ratio	2496.098	16	.000
N of Valid Cases	34090		

a 1 cells (4.0%) have expected count less than 5. The minimum expected count is 2.67.

Symmetric Measures

		Value	Asymp. Std. Error	Approx. T	Approx. Sig.
Ordinal by Ordinal	Kendall's tau-b	.102	.004	22.933	.000
	Kendall's tau-c	.088	.004	22.933	.000
	Gamma	.147	.006	22.933	.000
N of Valid Cases		34090			

a Not assuming the null hypothesis.
b Using the asymptotic standard error assuming the null hypothesis.
Source: Data analysis (2007)

One of the interesting results from the table above is that 76 students who managed to get Division one at OCSE failed at ACSE, while 43 out of 44 students who failed at OCSE did not fail at ACSE.

b. Is the strength of Prediction using Regression Analysis Strong Enough?

In order to further establish the predictive ability and strength of the prediction, a regression analysis was conducted based on points obtained in OCSE to points obtained at ACSE. Tables 4a and 4b provide results of regression analyses. The established model was not strong enough as R was .34 and adjusted r squared was .11. Despite its weak adjusted r squared, the model was significant at 1%, indicating that OCSE results could be used to predict ACSE. The model developed was

$$ACSE_{Points} = 7.515 + 0.298OCSE_{Points}$$

Both coefficients, for constant and OCSE points were significant at 1%. Given the weak adjusted r squared this model should be cautiously interpreted. Table 4 provides s detailed regression analysis.

Table 4a: Model Summary

Model	R	R Square	Adjusted R Square
1	0.337775	0.114092	0.114066

a Predictors: Constant, OCSE points
b Dependent Variable: ACSE points
Source: Research Data (2007)

Table 4b: Coefficients of Regression Analysis

Coefficients

		Unstandardized Coefficients		Standardized Coefficients	T	Sig.
Model		B	Std. Error	Beta		
1	(Constant)	7.519672	0.094763		79.35278	0
	F4_PTS	0.29798	0.004497	0.337775	66.25725	0

a Dependent Variable: F6_PTS
b Dependent Variable: ACSE points
Source: Research Data (2007)

c. Does the same school at OCSE and ACSE Matter in Predicting Performance?

Students were classified into two groups, namely, those who either studied at the same or different schools for OCSE and ACSE. Tables 5a and 5b provide the cross-tabulation of OCSE and ACSE results based on whether one studied at the same school or not. There were 30731 students who studied at different schools, while 3359 continued at the same school. Chi-squire test at 1% indicated that there is an association between ACSE and OCSE division. On critical analysis of the divisions obtained, there were no significant differences whether one changes school or not. Kendall's tau results indicated there is a relationship.

Table 5a: Cross-tabulation of OCSE versus ACSE results for students whether they studied at the same school or not

			F4_DIV					
Same School			I	II	III	IV	FLD	Total
		I	1459	726	1762	124	4	4075
		II	1755	1561	3802	463	11	7592
		III	1822	3095	5955	1035	20	11927
		IV	336	1364	3069	509	7	5285
No		FLD	66	408	1205	172	1	1852
		Total	5438	7154	15793	2303	43	30731
		I	261	72	121	4		458
		II	297	163	380	43		883
		III	273	288	606	97	1	1265
		IV	49	123	298	66		536
		FLD	10	39	133	35		217
Yes	F6_DIV	Total	890	685	1538	245	1	3359

Source: Data analysis

Table 5b: Chi square test of OCSE versus ACSE results for students whether they studied at the same school or not

Chi-Square Tests

Same School?		Value	df	Asymp. Sig. (2-sided)
No	Pearson Chi-Square	2008.380	16	.000
	Likelihood Ratio	2082.321	16	.000
	N of Valid Cases	30702		
Yes	Pearson Chi-Square	441.254	16	.000
	Likelihood Ratio	455.328	16	.000
	N of Valid Cases	3359		

a1 cells (4.0%) have expected count less than 5. The minimum expected count is 2.59.
b 5 cells (20.0%) have expected count less than 5. The minimum expected count is .06.

On checking the strength of the linear regression models developed by these two groups, the r squared for the two group: for the different schools decreased from

the combined one while for the same school it increased. The opposite trend was observed when looking at the strength of coefficients. For changing schools, the constant increased while the coefficient for OCSE points decreased. On the other hand, while for the same school the constant decreased, the coefficient for constant increased. Both coefficients were significantly different from zero, indicating that the factors were relevant.

Table 5c: Model Summary of OCSE versus ACSE results for students whether they studied at the same school or not

Same School	R	R Square	Adjusted R Square
No	0.3222	0.1038	0.1038
Yes	0.4619	0.2134	0.2132

a Predictors: Constant, f4_pts
b Dependent Variable: f6_pts

Model		Unstandardised Coefficients		Standardised Coefficients	t	Sig.
Same school		B	Std. Error	Beta		
No	(Constant)	7.699	0.102		75.255	0.000
	F4_PTS	0.289	0.005	0.322	59.672	0.000
Yes	(Constant)	6.248	0.249		25.121	0.000
	F4_PTS	0.369	0.012	0.462	30.177	0.000

a Dependent Variable: f6_pts
Source: Data analysis

d. Does predictive ability differ for different combinations?

In order to further establish the predictive ability of individual subjects, all combinations with more than700 students were included in the analysis (this is due to regression analysis constraints and assumptions). The subjects taken into consideration were Economics, Commerce and Accountancy, with basic applied mathematics (BECA), Geography, Advanced Mathematics and Economics (GAE), Geography, Chemistry and Biology with basic applied mathematics (GCBB), History Geography and English (HGE), History Geography and Kiswahili (HGK), History, Kiswahili and English (HKE), Physics, Chemistry and Advanced Mathematics (PCA), Physics, Chemistry and Biology with basic applied mathematics (PCBB) as well as History, Geography and Economics (HGBE). The number of students for each combination is given in Table6. From this table it is clear that many students at ACSE are taking arts-related subjects.

On cross-tabulating the divisions obtained at OCSE and ACSE, for each combination the chi-square tests are presented below. For all combinations, there is an association between OCSE and ACSE divisions obtained.

Table6: Chi-Square Tests (F4DIV to F6DIV)

ACSECOM		Value	Df	Asymp. Sig. (2-sided)
	Pearson Chi-Square	571.9758	12	0.0000
	Likelihood Ratio	599.5928	12	0.0000
BECA	N of Valid Cases	1821		
	Pearson Chi-Square	844.8639	16	0.0000
	Likelihood Ratio	874.3875	16	0.0000
GAE	N of Valid Cases	2523		
	Pearson Chi-Square	792.7491	16	0.0000
	Likelihood Ratio	678.0684	16	0.0000
GCBB	N of Valid Cases	2016		
	Pearson Chi-Square	575.5957	16	0.0000
	Likelihood Ratio	551.0332	16	0.0000
HGBE	N of Valid Cases	2586		
	Pearson Chi-Square	935.2626	16	0.0000
	Likelihood Ratio	819.6774	16	0.0000
HE	N of Valid Cases	5575		
	Pearson Chi-Square	351.8456	16	0.0000
	Likelihood Ratio	351.3917	16	0.0000
HGK	N of Valid Cases	2589		
	Pearson Chi-Square	400.0352	16	0.0000
	Likelihood Ratio	378.0544	16	0.0000
HKE	N of Valid Cases	3529		
	Pearson Chi-Square	1340.094	16	0.0000
	Likelihood Ratio	1465.143	16	0.0000
PCA	N of Valid Cases	3988		
	Pearson Chi-Square	1625.265	16	0.0000
	Likelihood Ratio	1764.514	16	0.0000
PCBB	N of Valid Cases	4035		

Source: Data analysis

On testing the strength of the model, regression analysis was conducted for each subject combination. The strength of the model (the models are significant for all subject combinations) is higher for science subjects, followed by business subjects and lastly arts subjects. This implies that the OCSE results are a better predictor for ACSE science subjects than for business and arts subjects. Table7a shows the models strength in terms of r, r squared and adjusted r squared. Detailed observations for coefficients of these regression analyses are presented in Table7b.

Table7a: Model Summary for regression analysis by subject combinations

Model	R	R Square	Adjusted R Square
BECA	0.576	0.331	0.331
GAE	0.586	0.344	0.344
GCBB	0.573	0.328	0.328
HGBE	0.464	0.215	0.215
HGE	0.495	0.245	0.245
HGK	0.459	0.211	0.210
HKE	0.392	0.153	0.153
PCA	0.663	0.440	0.440
PCBB	0.704	0.496	0.496

a Predictors: Constant, f4_pts
b Dependent Variable: f6_pts

Table7b:Coefficients for Linear Regression for Different Combinations

		Unstandardised Coefficients		Standardised Coefficients	t	Sig.
Model		B	Std. Error	Beta		
BECA	(Constant)	3.218	0.362		8.879	0.000
	F4_PTS	0.549	0.018	0.576	30.025	0.000
GAE	(Constant)	5.634	0.238		23.718	0.000
	F4_PTS	0.446	0.012	0.586	36.371	0.000
GCBB	(Constant)	2.965	0.417		7.109	0.000
	F4_PTS	0.607	0.019	0.573	31.368	0.000
HGBE	(Constant)	3.094	0.379		8.166	0.000
	F4_PTS	0.468	0.018	0.464	26.634	0.000
HGE	(Constant)	1.863	0.266		7.005	0.000
	F4_PTS	0.510	0.012	0.495	42.505	0.000
HGK	(Constant)	1.153	0.426		2.707	0.007
	F4_PTS	0.484	0.018	0.459	26.285	0.000
HKE	(Constant)	1.765	0.405		4.352	0.000
	F4_PTS	0.443	0.018	0.392	25.286	0.000
PCA	(Constant)	4.051	0.195		20.777	0.000
	F4_PTS	0.601	0.011	0.663	55.928	0.000
PCBB	(Constant)	4.524	0.174		25.952	0.000
	F4_PTS	0.583	0.009	0.704	63.026	0.000

a Dependent Variable: f6_pts
Source: Data analysis

Discussion and Conclusions

From the above results, this study appears to support the hypothesis that OCSE a good predictor of ACSE. However, the strength of the linear relationship between the two performances is not strong (through linear regression). It is equally insignificant whether one continues at the same school or not. However, the strength of the model increases when a student remains at the same school. Interestingly, those who performed poorly have a better chance of closing well they change school, while those who performed well have a good chance of performing better if they remain at the same school.

When the results were disaggregated by combinations, the OCSE results seem to be a good predictor for ACSE. The predictive ability of the model increased for all combinations tested, although science combinations seem to have a high linear relationship between OCSE and ACSE.

This study shows that there is a significant relationship between student OCSE grades and their achievement in ACSE.

From the above discussion, it is concluded that, to a great extent, there is a significant relationship between students' entry into OCSE and their overall academic achievement at ACSE. This indicates that the admission criteria, generally employed in the ACSE, serve the academic purposes they are assumed to be serving, besides being used for.

Since this study shows that academic achievement in the ACSE is largely dependent on entry grades, conditions favourable to proper learning should be provided for students at OCSE. Given the weak predictive ability concerning arts related subjects, while strong regarding science, this matter has to be further looked into.

Study Limitations

The examination results used by this study have been set by the same board. Given that the examination is set at the higher level, this may influence the predictive ability of their results.

Of course, there are many other influences upon individual students' school performance. Among others are the school is effectiveness like school leadership, (Sanday, 1990), the systems adopted (Gorard, Taylor, and Fitz, 2003), and school size (Pack and Peck, 2006). However, these and other factors were not easily measurable. This could also reduce the reliability of the conclusions derived but not the data themselves. Other factors which cannot be ignored are academic potential, student support systems and motivation.

REFERENCES

Bone, P.F.; Sharma, S.; and Shimp, T.A. (1989): "A Bootstrap Procedure for Evaluating Goodness of Fit Indices of Structural Equation and Confirmatory Factor Models", *Journal of Marketing Research*, February, pp.105-111.

Bradley, S.; Crouchley, R.; Millington, J. and Taylor, J. (2000): "Testing for quasi-market forces in secondary education", *Oxford Bulletin of Economics and Statistics,* Vol.62, No.3, pp.357–390.

Browne, M.W. and Cudeck (1992): "Alternative Ways of Assessing Model Fit" in K.A. Bollen and J.S. Long eds. *Testing Structural Equation Models,* Sage Publication, Newbury, Part, CA, pp132-162.

Chan, K. C., & Shum, C. (1997). "Class attendance and student performance in principles of finance". *Financial Practice and Education*,*7*(2), 58–65.

Everett, B.S. (1977): *The Analysis of Contingency Tables,* John Wiley and Sons, London.

Frawley, W.; Piatetsky-Shapiro, G. & Matheus, C. (1992). "Knowledge Discovery in Databases: An Overview." *AI Magazine*, pp. 213-228.

Gorard, S.; Taylor, C. and Fitz, J. (2003) Schools, Markets and Choice Policies, Falmer, Routledge (2003).

Green, S.B. (1991): "How Many Subjects does it take to do regression analysis", *Multivariate Behavioural Research*, Vol. 26, No.3, pp.499-510.

Hand, D.; Mannila, H. &Smyth, P. (2001). Principles of Data Mining. MIT Press, Cambridge, MA, ISBN0-262-08290-X

Hanushek, E.A. (1986). "The economics of schooling: Production and efficiency in public schools", *Journal of Economic Literature* Vol. 24 No.3, pp.1141–1177.

Hough, J. R. (1981). *A study of school costs*. Windsor, NFER-Nelson.

Hu and Bentler (1999) "Estimates and Tests in structural Equation Modelling" in R.H. Hoyle ed. *Structural Equation Modelling*, Sage Publications, Thousand Oaks, CA, pp37-55.

Joreskog, K.G. and Sörbom, D. (1982): "Recent Developments in Structural Equation Modelling", *Journal of Marketing Research,* Vol.19, No.4,

Kelly, L. (1999). "A longitudinal study measuring the ability of two South African Mathematics tests to predict mathematics performance of Grade9 high school pupils", *South African Journal of Education*,19(2),100.108.

Lewontin, R.C. and Felsenstein, J. (1965): "The Robustness of Homogeneity Tests in 2 X N Tables", *Biometrics,* Vol.21, pp.19-33.

Liebetrau, (1976): *Measures of Associations,* (Sage University Paper Series on Quantitative Applications in the Social Sciences, Series No.07-0032), Newbury Park, California: Sage.

Mbamba, U.O.L. (1993): "The Use of Expert Systems In Selecting Students into Different Specialization Options", Unpublished MBA Dissertation, University of Nairobi.

Peck, James Foreman and Peck, Lorraine Foreman (2006): "Should schools be smaller? The size-performance relationship for Welsh schools", *Economics of Education Review*, Vol. 25, No. 2 , April, pp157-171.

Peterson, C. and Barrett, L. (1987): "Explanatory style and academic performance among university freshmen", *Journal of Personality and Social Psychology,* Vol.53, pp.603–607.

Petripin, G.L. and Johnson, M.E. (1991): "Effects of gender, attributional style, and items difficulty on academic performance", *Journal of Psychology*, Vol.125, pp.45–50.

Sanday, A. (1990). *Making schools more effective*. CEDAR Paper no 2, University of Warwick.

Sawkins, J.W. (2002): "Examination Performance in Scottish secondary schools: An ordered logit approach", *Applied Economics,* Vol.34, No.16, pp.2031–2041.

Slakter, M.J. (1966): "Cognitive Validity of the Chi-square and the two modified chi-square goodness-of-fit tests for small but equal expected frequencies" *Biometrica*, Vol.53, pp.619-623.

Tabachnick, B.G and Fidell, L.S. (1996): *Using Multivariate Statistics*, Third Edition, Harper Collins College Publishers, New York.

Tiggemann, M. and Crowley, J.R. (1993): "Attributions for academic failure and subsequent performance", *Australian Journal of Psychology* Vol.45, pp.35–39.

Todd, P.E. and Wolpin, K.I. (2003) On the specification and estimation of the production function for cognitive achievement, *Economic Journal* Vol.113, pp. F3–F33.