
Modeling Primary School Absenteeism and Academic Performance in Ethiopia: A Multivariate and Count Regression Models Approaches

Edossa Merga Terefe¹, Zeytu Gashaw Asfaw²

¹Department of Statistics, Wollega University, Nekemte, Ethiopia

²School of Mathematical and Statistical Sciences, Hawassa University, Hawassa, Ethiopia

Email address:

edossamerga@gmail.com (E. M. Terefe)

To cite this article:

Edossa Merga Terefe, Zeytu Gashaw Asfaw. Modeling Primary School Absenteeism and Academic Performance in Ethiopia: A Multivariate and Count Regression Models Approaches. *American Journal of Theoretical and Applied Statistics*. Vol. 7, No. 1, 2018, pp. 8-20.

doi: 10.11648/j.ajtas.20180701.12

Received: November 21, 2017; **Accepted:** December 1, 2017; **Published:** January 5, 2018

Abstract: School absenteeism and low academic performance at primary schools remain a big issue for developing countries like Ethiopia. Thus, this study aims to determine predicting factors influencing academic performances and school absenteeism jointly at primary schools in Ethiopia. A cross-sectional data were obtained from the Young Lives project from wave 1 (the starting month of academic year) and wave 2 (the last month of academic year). Multivariate regression model was used to investigate the predictors on the linear combination of academic performances and count regression model was also used to investigate the predictor of school absenteeism. In fact, both Poisson and Negative Binomial Regression Models were considered but the latter better fit to the data that have been used for this study than the former model. The result at national level showed mean of school absenteeism and academic performance at wave 1, respectively are 6 days and 67.64 scores and the average performance at wave 2 is also 61.44 score. It has been found out that the number of meals, number of siblings, mother's literacy, survivor-ship of mother, type of school siblings attending, pre-school attendance, time to get to school, grade repeating, school drop outing, extra class attendances and the availability of helping person with school works at home have a combined effect on the school absenteeism and academic performance. Thus, potential stakeholders should pay attention for the aforementioned factors so as to reduce school absenteeism and then maximize academic performance.

Keywords: School Absenteeism, Academic Performance, MANOVA, Canonical Correlation, Count Regression, DAG

1. Introduction

Education has long been recognized as a fundamental human right and an important building block in the development of children and broader society [1]. The primary level of education tends to play crucial role in subsequent levels of education and develop the critical skills and abilities needed to become productive and responsible adults [2], [3]. It is recognized by education scholars that effective participation and engagement in primary education is an important precursor to educational success [4]. Despite this, the level of participation and engagement even in compulsory education processes can generally be viewed on a spectrum. This has been a feature of the educational literature across a range of countries over many decades, and

underscores the fact that there is typically a segment of the population who has persistent difficulties in this area [5].

The cost of student absenteeism in terms of investment in their education is still made by governments but the benefits are not received by the student or community and become plagues both urban and rural elementary schools of low developing countries such as Ethiopia. The issue of chronic absenteeism due to different reasons such as economic reasons of their parent [6] is prevalent in many urban and a rural school usually begins in the main subjects such as mathematics and English. Once a student begins to fall behind in one of these subjects, he or she tends to gradually tends to temporary dropout of the school and the student started absenteeism in all of classes, but the net result of frequent school absenteeism is permanently dropping out of

the school. Persistent non-attendance may require additional resources to re-engage students in schooling and to help them catch up on missed learning involve learning and personal student support staff and interagency collaborations which incur additional cost to the country. This poor school attendance represents a loss of educational opportunity and also a cost to the community.

The impact of chronic absenteeism on academic achievements, however, has been indicated by different studies. For instance, [7] showed in his studies that a positive relationship of student attendance and academic achievements at primary level of education. Other scholars have also found a strong correlation between school attendance and academic achievement of student. Similarly, [8] found out that class attendance is the main predictor of academic achievements than any other variables at primary level of education.

In Ethiopia, the lowest class repetition rate (6.4%) was observed in grade 6 while the highest class repetition rate (9.0%) was registered in grade 1 mainly due to school absenteeism according to ministry of education report. The repetition rate for grades 5 and 8 were the next highest in

2011/12 academic year. The highest (18.6%) nationally registered dropout rate was in the year 2008/09 while the lowest (13.1%) was observed in the year 2009/10 [9]. The figures reported as lowest repetition or dropout rate are serious problem challenging educational outcomes particularly skilled man power production and wasting resources for any country that seeking to have efficient educational system and improved utilized limited resources hence the rates are commonly used to measure the efficiency of the educational in producing graduates of a particular education cycle or level. Thus, this study model and deeply investigate the school absenteeism and academic achievement of primary school students in Ethiopia.

2. Method and Data Description

2.1. Response and Explanatory Variables

School absenteeism and Academic performances are the two response variables in this study paper, while their expected determinants were listed in Table 1 along with their coding.

Table 1. List of explanatory variables and Coding.

Explanatory variable	Label	Value
Boygirl	Gender	boy =1, girl=2
Age	Age at the time of survey	1=8 to 10, 2=11 to 13, 3=14 and above
Nummeal	Number of meals a student normally eat per a day	1=1, 2=2, 3=3 or more
Oldrsblng	Number of older brothers and sisters	0=0, 1=1 to 3, 2=4 or more
Yngrsblng	Number of younger brothers and sisters	0=0, 1=1 to 3, 2=4 or more
Mthralve	Survivor-ship of mother	0=No, 1=Yes
Fthralve	Survivor-ship of father	0=No, 1=Yes
Rdwrnthr	Mother reads/writes	0=No, 1=Yes
Rdwrftthr	Father reads/writes	0=No, 1=Yes
Rdwrbrth	Sibling(s) read/write	0=No, 1=Yes
Describe	School type siblings attending	1=attending government school, 2=attending private school, 3=not attending any school
Attpresc	Pre-school attendance	0=No, 1=Yes
Timeschol	Time taken to get to school (in minutes)	0=0 to 15, 1=16 to 30, 2=31 and more
Reptgrde	Ever repeated grade	0=No, 1=Yes
Dropschl	Ever dropped out school	0=No, 1=Yes
Hrsfrpay	Time spend working for pay on a usual school day (hours)	0=none, 1=<1, 2=1 to 2, 3=2 to 3, 4=3 to 4, 5=>4
Extrmath	Attend extra classes of numeracy	0=No, 1=Yes
Extrlng	Attend extra class of literacy	0=No, 1=Yes
School03	Teachers care about students	1=strongly agree, 2=agree, 3=disagree, 4 =strongly disagree
School06	Teachers are open and honest	1=strongly agree, 2=agree, 3=disagree, 4 =strongly disagree
Scwrhome	have people to help with school work at home	0=No, 1=Yes
Timehmk	Time spend on homework outside school per a day (hours)	0=none, 1=1, 2=1 to 2, 3=more than 2

2.2. Source of Data and Young Lives Project

This study was based on the secondary data that comes from Young Lives Project (the school survey data that was introduced in Ethiopia in 2010). The Young Lives survey is an international innovative long-term project investigating the changing nature of childhood poverty in four developing countries (Ethiopia, India (in the state of Andhra Pradesh), Peru and Vietnam) over 15 years. Young Lives is core-funded from 2001 to 2017 by UK aid from the Department for International Development (DFID), and co-funded by the Netherlands Ministry of Foreign Affairs from 2010 to 2014.

Sub-studies are funded by the Bernard van Leer Foundation and the Oak Foundation.

The total number of students covered in this study was 13,724 at wave 1 (September, the start of Ethiopian academic year), and 11,985 at wave 2 (May, the end of Ethiopian academic year) in all sentimental sites of seven Ethiopian regional states including Addis Ababa city administration. The descriptive analysis result indicated that 12.67% of wave 1 students have dropped out the school in the study year. Among those students who have been covered at wave 1, only 11,778 (85.82%) and 11,790 (85.91%) of them have taken the literacy test and Numeracy test respectively.

Among those students who have been covered at wave 1, only 11,778 (85.82%) and 11,790 (85.91%) of them have taken the literacy test and Numeracy test respectively. The difference in number of students who have been covered and who have taken the tests is equal to the number of students who have not taken in accordance with the tests. For example, (13,724 - 11,778 = 1946) is a total number of students who have not taken the wave 1 literacy test.

Similarly, among those students who have been surveyed at wave 2, only 10,072 of them have taken the literacy test while only 10,079 of wave 2 covered students have taken the Numeracy test. In the comparison of the percentage of students who have taken the tests at both waves, the percentage of students who have taken the tests at wave 2 is higher than that of wave 1 students among the registered students at the school. There were about 50.4% female and 49.6% male respondents, and equal percentages of students for both rural and urban sites. 75.5% of the total participants are in the age group 8-12 whereas the remaining 21.7% and 2.8% are respectively, in the age groups 13-15 years and 16-20 years. The results of descriptive analysis revealed that the average age of the participants was 12 years. More than 24% of students have ever repeated the grade since they started the schooling. Among those students who have ever repeated the grade, only 39.7% of them repeated the grade because of the dropping out the school. Obviously, the remaining percentage repeated because of failure in examinations.

2.3. The Statistical Models

Different statistical methods were used to achieve the goals of this study includes multivariate multiple linear regression model, multivariate analysis of variance (MANVA), Poisson and negative binomial regression model and binary logistic regression and simple linear regression model as given below:

2.3.1. Multivariate Methods

1. Multivariate Analysis of Variance (MANOVA): The hypothesis which concerns a comparison of vectors of group means and the goal is to maximally discriminate between two or more distinct groups on a linear combination of quantitative variables.
2. Multivariate Multiple Linear Regression models: Regression analysis for predicting the values of one more response (dependent) variables from a collection of predictor (independent) variable values.
3. Canonical Correlation Analysis: An analysis that

focuses on the correlation between a linear combination of the variables in one set and linear combination of the variables in another set.

2.3.2. Count Regression Models

Poisson regression model: Assumes that the response variable has a Poisson distribution and the logarithm of its expected value can be modeled by a linear combination of unknown parameters.

Negative Binomial Regression model: regression model which is a direct extension to the Poisson model that allows for over-dispersion.

2.3.3. Directed Acyclic Graphs (DAG)/Causal Diagrams

Graphical models for causal relationship that can serve a role complementary to conventional models.

3. Results and Discussions

3.1. Bivariate Descriptive Analysis

The result showed that the average academic performance of pupils was 67.64 at wave 1 and 61.44 at wave 2, where as the average school absenteeism in one year was 6 days. The average academic performances of students differ by if the student attended pre-school or not. Accordingly, students who attended pre-school education achieved an average academic performance of 67.4 while who did not attend achieved 59.88 at wave 1 but there is no large difference at wave 2. The average academic performance differs from region to region in both waves. For example, the highest average academic performance was observed in Tigray with the mean of 71.20 followed by Oromiya region with the mean of 70.24 whereas the lowest performance was observed in Amhara region with the mean of 58.28 followed by Somali region with the mean of 60.28. At wave 2, there was a bit little change in average academic performance with the highest was observed in Addis Ababa city with the mean of 74.44 and the lowest in Affar region with the mean of 60 but the performance in the rest of five regions is nearly the same.

Both the average academic performance of students and school absenteeism do differ according to the number of meals they normally consumed per a day. Accordingly, the average academic performance at wave 1 of students who normally ate only once per aday was 53.80 while the students who normally eat three or more meals per a day was 64.88 which was consistent result reported by [10] given that similar data sets as indicated in Table 2.

Table 2. Bivariate Analysis of academic performance and school absenteeism (Ethiopia, 2013-2014).

Variables	Academic Performance						School Absenteeism		
	Wave 1			Wave 2			Mean	N	SD
	Mean	N	SD	Mean	N	SD			
Gender									
Female	63.80	5941	16.33	61.57	5243	13.46	5.48	5467	1.22
Male	63.55	5851	16.27	61.16	5056	13.47	6.09	5270	2.04
Age									
8 – 10	61.03	3617	16.43	61.24	3260	13.49	5.42	3383	1.45
11 – 13	65.15	6575	15.92	61.57	5687	13.45	5.67	6022	1.56
≥14	63.50	3532	11.54	61.33	3038	13.25	8.08	1875	2.25

Variables	Academic Performance						School Absenteeism		
	Wave 1			Wave 2			Mean	N	SD
	Mean	N	SD	Mean	N	SD			
Number of meals/day									
1	53.95	519	15.48	60.34	472	13.83	6.51	476	2.49
2	60.28	1752	16.51	61.52	1637	13.27	6.98	1568	2.62
≥3	64.88	9500	16.02	61.41	8177	13.48	5.52	8674	1.37
Siblings school type									
Government	62.61	8096	16.11	61.17	7439	13.33	6.17	7353	1.91
Private	68.00	2400	15.71	61.96	1698	13.77	4.22	2213	0.35
Not attending any	61.78	1227	17.19	61.76	1103	13.85	6.42	1106	1.66
Older siblings									
0	65.65	1947	16.50	61.81	1689	13.69	5.24	1802	1.98
1-3	63.57	6604	16.35	61.32	5816	13.34	5.74	6014	1.43
≥4	62.97	5173	12.80	61.41	4480	13.40	6.84	3464	2.96
Younger siblings									
0	66.85	2456	16.08	61.40	2177	13.53	4.83	2294	0.84
1-3	63.41	8022	16.23	61.39	7132	13.42	5.94	7318	1.73
≥4	61.78	3246	10.63	61.53	2676	13.29	5.38	2676	0.32
Mother alive									
No	63.67	771	16.612	61.79	669	13.08	5.82	682	1.50
Yes	63.66	11003	16.28	61.34	9615	13.50	5.78	10042	1.65
Father alive									
No	63.35	1999	16.30	62.09	1819	13.56	5.88	1802	1.63
Yes	63.78	9710	16.29	61.19	8414	13.44	5.78	8858	1.67
Mother reads/writes									
No	62.19	6168	16.27	60.89	5514	13.41	6.72	5532	2.33
Yes	65.31	5631	16.17	61.93	4790	13.51	4.78	5211	2.70
Father reads/writes									
No	62.87	4862	16.30	61.07	4315	13.42	6.59	4338	2.40
Yes	64.24	6937	16.27	61.59	5989	13.49	5.23	6405	2.03
Brother or Sister can reads/writes									
No	61.74	3843	16.25	61.25	3333	13.50	6.36	3462	2.24
Yes	64.62	7956	16.24	61.43	6971	13.45	5.51	7281	1.33
Attended pre-school									
No	59.86	5693	16.50	59.86	5131	16.50	6.89	5119	2.48
Yes	67.38	6033	15.19	67.38	5109	15.19	4.75	5560	0.64
Time to get to school in minutes									
0 to 15	63.18	6528	16.40	63.18	5632	16.40	5.50	5974	1.48
16 to 30	64.78	3530	16.30	64.78	3105	16.30	5.65	3213	1.34
≥31	63.36	3666	11.12	63.36	3248	11.12	7.97	2093	2.38
Have ever repeated grade									
No	65.03	8907	16.14	65.03	7667	16.14	5.39	8142	1.21
Yes	59.49	2858	16.02	59.49	2606	16.02	7.03	2572	2.78
Have ever dropped out school									
No	64.42	9626	16.01	64.42	8327	16.01	5.47	8864	1.34
Yes	60.45	2077	17.17	60.45	1906	17.17	7.38	1787	1.92
Time spend on homework outside of school (hours)/day									
Not at all	66.54	652	14.37	66.54	546	14.37	6.67	588	2.01
1	63.81	5597	16.25	63.81	4964	16.25	5.82	5143	1.81
1 – 2	62.22	3264	16.83	62.22	2866	16.83	5.78	2945	1.57
≥2	64.70	2264	15.89	64.70	1911	15.89	5.42	2046	0.86
Teachers are open and honest									
strongly agree	62.92	8024	16.10	62.92	6998	16.10	5.71	7313	1.58
Agree	65.10	2210	16.61	65.10	1943	16.61	5.77	2016	1.74
Disagree	65.55	917	16.77	65.55	811	16.77	6.17	836	1.65
Strongly disagree	65.47	614	16.13	65.47	521	16.13	6.34	545	1.15
Teachers care about students									
strongly agree	62.82	8034	16.13	62.82	7020	16.13	5.60	7305	1.21
Agree	65.44	2417	16.50	65.44	2146	16.50	6.18	2209	2.39
Disagree	65.84	812	16.43	65.84	699	16.43	6.01	736	2.28
Strongly disagree	65.63	516	16.49	65.63	423	16.49	6.48	473	2.28
Attend numeracy extra classes									
No	64.63	6905	16.83	64.63	2146	16.83	6.18	2209	1.83
Yes	62.34	4894	15.415	62.34	699	15.415	6.01	736	1.41
Attend literacy extra classes									
No	64.21	8239	16.31	64.21	7235	16.31	6.00	7448	1.72
Yes	62.46	3560	16.19	62.46	3069	16.19	5.28	3295	1.452

Variables	Academic Performance						School Absenteeism		
	Wave 1			Wave 2			Mean	N	SD
	Mean	N	SD	Mean	N	SD			
Have help with school work at home									
Never	62.35	2167	17.38	62.35	1919	17.38	7.25	1936	2.66
Always	64.55	5372	15.63	64.55	4633	15.63	5.10	4920	1.10
Sometimes	63.29	4222	16.48	63.29	3726	16.48	5.92	3853	1.68
Hours spend working for pay on usual school day									
None	63.74	9236	16.42	63.74	8111	16.42	5.63	8459	1.45
Less than 1	62.98	841	16.26	62.98	707	16.26	6.27	766	2.07
1-2	63.91	402	15.71	63.91	350	15.71	5.54	362	1.58
2-3	61.02	417	15.72	61.02	368	15.72	6.97	371	2.69
3-4	61.78	346	16.52	61.78	314	16.52	6.99	299	2.10
More than4	67.24	538	14.02	67.24	444	14.02	6.22	467	2.15

With regard to school absenteeism and eating habit, the more number of meals they normally eat per a day the less number of days a student absent from the school may imply that there are some students who absent from the school for searching for the meal. Similarly, the school absenteeism and the average academic performance of students differ by the number of periods of teaching their class receives per a day. The highest average academic performance (77.00) was observed at wave 1 for students whose school attend four periods of teaching in a day as opposed to the lowest average academic performance (64.80) was observed for students whose school attended eight periods of teaching per a day may implying that lengthy period becoming boring and the period of classes per day something in between 4 and 8 at primary level. And, the lowest average school absenteeism (2 days) was observed in school where the periods of teaching are eight per a day as opposed to the highest average school absenteeism (4.05days) was observed in a school where the periods of teaching are twice per a day may indicating that some modification made the number periods per day.

The result in Table 2 showed that both the average academic performance of students and their average school

absenteeism also vary by the type of school attending characteristics by siblings. Accordingly, the highest average performance (68.00) at wave 1 was observed for the students whose sister(s) or brother(s) is/are attending the private school and the lowest (61.76) was for the students whose sister(s) or brother(s) is/are not attending any school. With regards to school absenteeism, the lowest average school absenteeism (4.22) was observed for the students whose sister(s) or brother(s) is/are attending a private school while the highest average school absenteeism (6.42) was for the students whose sister(s) or brother(s) is/are not attending any school may indicate siblings attending characteristics influence the school attending and achievement.

3.2. Result of Multivariate Multiple Linear Regression Analysis

In this section, the overall determinants of academic performance were assessed simultaneously to identify the basic determinant factors which were important for identifying factors of primary school pupils' academic achievement.

Table 3. Estimate of Multivariate Multiple linear Regression Model.

Dependent Variable	Covariates	Estimate	Se (Estimate)	Hoteling's t-test	Significance	Partial Eta Square	Non-Parameter	Obs. Pow.
Average performance at wave 1	Intercept	18.17	0.748	24.285	0.00	0.061	24.285	1
	[boygirl=1]	0.04	0.078	0.518	0.605	0	0.518	0.081
	[Age=1]	-1.62	0.139	11.675	0.00	0.015	11.675	1
	[Age=2]	-0.384	0.127	-3.023	0.003	0.001	3.023	0.856
	[nummeal=1]	-2.235	0.188	11.883	0.00	0.015	11.883	1
	[nummeal=2]	-0.738	0.108	-6.85	0	0.005	6.85	1
	[Oldrsblng=0]	0.577	0.128	4.494	0	0.002	4.494	0.994
	[Oldrsblng=1]	0.189	0.091	2.071	0.038	0	2.071	0.544
	[Yngrsblng=0]	1.201	0.157	7.658	0	0.006	7.658	1
	[Yngrsblng=1]	0.693	0.138	5.033	0	0.003	5.033	0.999
	[mthralve=0]	0.189	0.167	1.13	0.259	0.00	1.13	0.204
	[fthralve=0]	-0.076	0.109	-0.699	0.484	0	0.699	0.108
	[rdwrnthr=0]	-0.202	0.085	-2.389	0.017	0.001	2.389	0.666
	[rdwrftthr=0]	0.046	0.088	0.523	0.601	0	0.523	0.082
	[rdwrbrth=0]	-0.667	0.088	-7.578	0	0.006	7.578	1
	[describe=1]	-2.097	0.639	-3.282	0.001	0.001	3.282	0.907
	[describe=2]	-1.496	0.643	-2.325	0.02	0.001	2.325	0.642
	[describe=3]	-2.252	0.643	-3.501	0	0.001	3.501	0.938
	[describe=77]	2.435	3.755	0.648	0.517	0	0.648	0.099
	[attpresc=0]	-1.548	0.081	-18.99	0	0.038	18.998	1
[timeschol=0]	-0.353	0.114	-3.089	0.002	0.001	3.089	0.871	
[timeschol=1]	0.017	0.123	0.136	0.892	0	0.136	0.052	

Dependent Variable	Covariates	Estimate	Se (Estimate)	Hoteling's t-test	Significance	Partial Eta Square	Non-Parameter	Obs. Pow.
Average performance at wave 2	[reptgrde=0]	1.532	0.092	16.594	0	0.029	16.594	1
	[dropschl=0]	0.708	0.106	6.686	0	0.005	6.686	1
	[hrsfrpay=0]	-0.709	0.198	-3.586	0	0.001	3.586	0.948
	[hrsfrpay=1]	-0.994	0.244	-4.074	0	0.002	4.074	0.983
	[hrsfrpay=2]	-0.71	0.288	-2.465	0.014	0.001	2.465	0.693
	[hrsfrpay=3]	-1.106	0.283	-3.909	0	0.002	3.909	0.974
	[hrsfrpay=4]	-1.074	0.299	-3.594	0	0.001	3.594	0.949
	[extrmath=0]	0.195	0.081	2.397	0.017	0.001	2.397	0.669
	[extrlang=0]	0.261	0.087	2.986	0.003	0.001	2.986	0.847
	[school03=1]	0.009	0.202	0.046	0.963	0	0.046	0.05
	[school03=2]	0.419	0.213	1.969	0.049	0	1.969	0.504
	[school03=3]	0.448	0.244	1.837	0.066	0	1.837	0.451
	[school06=1]	-0.288	0.188	-1.528	0.126	0	1.528	0.333
	[school06=2]	-0.122	0.201	-0.604	0.546	0	0.604	0.093
	[school06=3]	-0.14	0.226	-0.619	0.536	0	0.619	0.095
	[scwrhome=0]	0.048	0.113	0.425	0.671	0	0.425	0.071
	[scwrhome=1]	0.099	0.088	1.13	0.259	0	1.13	0.204
	[timehmwk=0]	0.715	0.191	3.734	0	0.002	3.734	0.962
	[timehmwk=1]	-0.199	0.108	-1.844	0.065	0	1.844	0.454
	[timehmwk=2]	-0.562	0.117	-4.819	0	0.003	4.819	0.998
	Intercept	15.45	0.677	22.808	0	0.054	22.808	1
	[boygirl=1]	-0.075	0.071	-1.061	0.289	0	1.061	0.186
	[Age=1]	-0.02	0.126	-0.157	0.875	0	0.157	0.053
	[Age=2]	0.089	0.115	0.773	0.44	0	0.773	0.121
	[nummeal=1]	-0.374	0.17	-2.2	0.028	0.001	2.2	0.595
	[nummeal=2]	0.093	0.097	0.956	0.339	0	0.956	0.16
	[Oldersblng=0]	0.045	0.116	0.391	0.696	0	0.391	0.068
	[Oldersblng=1]	-0.008	0.083	-0.101	0.919	0	0.101	0.051
	[Yngrsblng=0]	-0.036	0.142	-0.253	0.8	0	0.253	0.057
	[Yngrsblng=1]	0.022	0.125	0.179	0.858	0	0.179	0.054
	[mthralve=0]	0.22	0.151	1.46	0.144	0	1.46	0.309
	[fthralve=0]	0.319	0.098	3.247	0.001	0.001	3.247	0.901
	[rdwrnthr=0]	-0.228	0.077	-2.978	0.003	0.001	2.978	0.846
	[rdwrfthr=0]	-0.15	0.08	-1.889	0.059	0	1.889	0.472
	[rdwrbrth=0]	0.01	0.08	0.12	0.904	0	0.12	0.052
	[describe=1]	0.012	0.578	0.02	0.984	0	0.02	0.05
	[describe=2]	0.233	0.582	0.4	0.689	0	0.4	0.069
	[describe=3]	0.186	0.582	0.319	0.75	0	0.319	0.062
	[describe=77]	-3.467	3.398	-1.02	0.308	0	1.02	0.175
	[attpresc=0]	0.178	0.074	2.41	0.016	0.001	2.41	0.674
	[timeschol=0]	0.081	0.104	0.778	0.437	0	0.778	0.122
	[timeschol=1]	0.026	0.112	0.232	0.816	0	0.232	0.056
	[reptgrde=0]	0.085	0.084	1.012	0.312	0	1.012	0.173
	[dropschl=0]	0.152	0.096	1.589	0.112	0	1.589	0.355
	[hrsfrpay=0]	0.217	0.179	1.214	0.225	0	1.214	0.229
	[hrsfrpay=1]	-0.022	0.221	-0.101	0.919	0	0.101	0.051
	[hrsfrpay=2]	-0.079	0.261	-0.303	0.762	0	0.303	0.061
	[hrsfrpay=3]	-0.042	0.256	-0.164	0.87	0	0.164	0.053
	[hrsfrpay=4]	-0.588	0.27	-2.177	0.03	0.001	2.177	0.586
	[extrmath=0]	-0.068	0.073	-0.928	0.354	0	0.928	0.153
[extrlang=0]	-0.176	0.079	-2.223	0.026	0.001	2.223	0.604	
[school03=1]	-0.256	0.183	-1.396	0.163	0	1.396	0.287	
[school03=2]	-0.081	0.193	-0.422	0.673	0	0.422	0.071	
[school03=3]	-0.006	0.22	-0.027	0.978	0	0.027	0.05	
[school06=1]	-0.307	0.171	-1.801	0.072	0	1.801	0.437	
[school06=2]	-0.432	0.182	-2.37	0.018	0.001	2.37	0.659	
[school06=3]	-0.272	0.205	-1.33	0.184	0	1.33	0.265	
[scwrhome=0]	-0.266	0.102	-2.599	0.009	0.001	2.599	0.739	
[scwrhome=1]	-0.016	0.08	-0.2	0.842	0	0.2	0.055	
[timehmwk=0]	0.297	0.173	1.715	0.086	0	1.715	0.403	
[timehmwk=1]	0.239	0.097	2.45	0.014	0.001	2.45	0.688	
[timehmwk=2]	0.177	0.106	1.678	0.093	0	1.678	0.389	

The results of multivariate regression model in Table 3 indicated that the combined factors that found to have a significant effect on the linear combination of average academic performance at wave 1 and at wave 2 were: Age,

number of meals a student normally eat per aday, number of younger siblings, number of older siblings, ability of mother to read and/or write, ability of sister(s) and/or brother(s) to read and /or write, type of school siblings attending, pre-school attendance, time spent to travel to school, grade repeating, school drop outing, time spent for paid work per day, extra class attendance of numeracy, extra class attendance of literacy and availability of helping person with school works at home.

The result of β coefficients in Table 3 cannot be assembled into a regression equation directly, as in OLS regression, due to differences in computation [11]. It can be interpreted, for instance, the average performance was a continuous variable, so in the output illustrated in Table 3, the negative β coefficient for number of meals a student normally eat per a

day (nummeal) indicates that respondent’s number of meals (nummeal) is significantly associated with scoring average performance at wave 1 less. Similarly, the positive β coefficient for school dropout (dropschl) indicates that school dropout is significantly associated with scoring average performance at wave 1 more. And, the rest of beta coefficients can be discussed in similar ways.

3.3. Result of MANOVA

A MANOVA was used to see the main effects of categorical variables on multiple dependent continuous variables. It is a way to test the hypothesis that one or more independentvariables, or factors, have an effect on a set of two or more dependent variables.

Table 4. Estimate of MANOVA.

Effect		Value	F	Hyp. df	Error df	Sig.	Partial Eta Sq.	Noncent Paramet.	Obs. power
Intercept	Pillai's Trace	0.094	315.33	3	9106	0	0.094	945.994	1
	Wilks' Lambda	0.906	315.33	3	9106	0	0.094	945.994	1
	Hotelling's Trace	0.104	315.33	3	9106	0	0.094	945.994	1
	Roy's Largest Root	0.104	315.33	3	9106	0	0.094	945.994	1
Boygirl	Pillai's Trace	0.002	4.886	3	9106	0.002	0.002	14.659	0.91
	Wilks' Lambda	0.998	4.886	3	9106	0.002	0.002	14.659	0.91
	Hotelling's Trace	0.002	4.886	3	9106	0.002	0.002	14.659	0.91
	Roy's Largest Root	0.002	4.886	3	9106	0.002	0.002	14.659	0.91
Age	Pillai's Trace	0.027	41.32	6	18214	0	0.013	247.92	1
	Wilks' Lambda	0.973	41.571	6	18212	0	0.014	249.423	1
	Hotelling's Trace	0.028	41.821	6	18210	0	0.014	250.926	1
	Roy's Largest Root	0.027	81.457	3	9107	0	0.026	244.371	1
Nummeal	Pillai's Trace	0.02	30.912	6	18214	0	0.01	185.469	1
	Wilks' Lambda	0.98	31.032	6	18212	0	0.01	186.193	1
	Hotelling's Trace	0.021	31.153	6	18210	0	0.01	186.916	1
	Roy's Largest Root	0.019	58.737	3	9107	0	0.019	176.211	1
Older_sablings	Pillai's Trace	0.003	4.899	6	18214	0	0.002	29.396	0.99
	Wilks' Lambda	0.997	4.902	6	18212	0	0.002	29.413	0.99
	Hotelling's Trace	0.003	4.905	6	18210	0	0.002	29.431	0.99
	Roy's Largest Root	0.003	9.461	3	9107	0	0.003	28.383	0.99
Younger_siblings	Pillai's Trace	0.009	13.427	6	18214	0	0.004	80.563	1
	Wilks' Lambda	0.991	13.455	6	18212	0	0.004	80.729	1
	Hotelling's Trace	0.009	13.483	6	18210	0	0.004	80.895	1
	Roy's Largest Root	0.009	26.830	3	9107	0	0.009	80.491	1
Mthralve	Pillai's Trace	0.001	2.278	3	9106	0.077	0.001	6.834	0.57
	Wilks' Lambda	0.999	2.278	3	9106	0.077	0.001	6.834	0.57
	Hotelling's Trace	0.001	2.278	3	9106	0.077	0.001	6.834	0.57
	Roy's Largest Root	0.001	2.278	3	9106	0.077	0.001	6.834	0.57
Fthralve	Pillai's Trace	0.002	5.162	3	9106	0.001	0.002	15.485	0.92
	Wilks' Lambda	0.998	5.162	3	9106	0.001	0.002	15.485	0.92
	Hotelling's Trace	0.002	5.162	3	9106	0.001	0.002	15.485	0.92
	Roy's Largest Root	0.002	5.162	3	9106	0.001	0.002	15.485	0.92
Rdwmthr	Pillai's Trace	0.005	14.06	3	9106	0	0.005	42.191	1
	Wilks' Lambda	0.995	14.06	3	9106	0	0.005	42.191	1
	Hotelling's Trace	0.005	14.06	3	9106	0	0.005	42.191	1
	Roy's Largest Root	0.005	14.06	3	9106	0	0.005	42.191	1
Rdwrfthr	Pillai's Trace	0.002	5.318	3	9106	0.001	0.002	15.955	0.93
	Wilks' Lambda	0.998	5.318	3	9106	0.001	0.002	15.955	0.93
	Hotelling's Trace	0.002	5.318	3	9106	0.001	0.002	15.955	0.93
	Roy's Largest Root	0.002	5.318	3	9106	0.001	0.002	15.955	0.93
Rdwrbrth	Pillai's Trace	0.007	22.78	3	9106	0	0.007	68.34	1
	Wilks' Lambda	0.993	22.78	3	9106	0	0.007	68.34	1
	Hotelling's Trace	0.008	22.78	3	9106	0	0.007	68.34	1
	Roy's Largest Root	0.008	22.78	3	9106	0	0.007	68.34	1
Describe	Pillai's Trace	0.008	6.055	12	27324	0	0.003	72.662	1
	Wilks' Lambda	0.992	6.063	12	24092	0	0.003	64.15	1
	Hotelling's Trace	0.008	6.068	12	27314	0	0.003	72.82	1

Effect		Value	F	Hyp. df	Error df	Sig.	Partial Eta Sq.	Noncent Paramet.	Obs. power
Attpresc	Roy's Largest Root	0.006	14.36	4	9108	0	0.006	57.46	1
	Pillai's Trace	0.043	135.16	3	9106	0	0.043	405.504	1
	Wilks' Lambda	0.957	135.16	3	9106	0	0.043	405.504	1
	Hotelling's Trace	0.045	135.16	3	9106	0	0.043	405.504	1
	Roy's Largest Root	0.045	135.16	3	9106	0	0.043	405.504	1
Time_to_scho ol	Pillai's Trace	0.005	6.962	6	18214	0	0.002	41.773	1
	Wilks' Lambda	0.995	6.964	6	18212	0	0.002	41.783	1
	Hotelling's Trace	0.005	6.965	6	18210	0	0.002	41.792	1
Reptgrde	Roy's Largest Root	0.004	10.71	3	9107	0	0.004	32.129	0.99
	Pillai's Trace	0.032	98.82	3	9106	0	0.032	296.459	1
	Wilks' Lambda	0.968	98.82	3	9106	0	0.032	296.459	1
	Hotelling's Trace	0.033	98.82	3	9106	0	0.032	296.459	1
	Roy's Largest Root	0.033	98.82	3	9106	0	0.032	296.459	1
Dropschl	Pillai's Trace	0.006	18.89	3	9106	0	0.006	56.695	1
	Wilks' Lambda	0.994	18.89	3	9106	0	0.006	56.695	1
	Hotelling's Trace	0.006	18.89	3	9106	0	0.006	56.695	1
	Roy's Largest Root	0.006	18.89	3	9106	0	0.006	56.695	1
Hrsfrpay	Pillai's Trace	0.006	3.799	15	27324	0	0.002	56.983	1
	Wilks' Lambda	0.994	3.8	15	25138	0	0.002	52.446	1
	Hotelling's Trace	0.006	3.801	15	27314	0	0.002	57.009	1
	Roy's Largest Root	0.004	6.779	5	9108	0	0.004	33.893	0.99
Extrmath	Pillai's Trace	0.002	4.689	3	9106	0.003	0.002	14.066	0.89
	Wilks' Lambda	0.998	4.689	3	9106	0.003	0.002	14.066	0.89
	Hotelling's Trace	0.002	4.689	3	9106	0.003	0.002	14.066	0.89
	Roy's Largest Root	0.002	4.689	3	9106	0.003	0.002	14.066	0.89
	Pillai's Trace	0.004	12.64	3	9106	0	0.004	37.919	1
Extrlng	Wilks' Lambda	0.996	12.64	3	9106	0	0.004	37.919	1
	Hotelling's Trace	0.004	12.64	3	9106	0	0.004	37.919	1
	Roy's Largest Root	0.004	12.64	3	9106	0	0.004	37.919	1
	Pillai's Trace	0.005	5.022	9	27324	0	0.002	45.202	1
School03	Wilks' Lambda	0.995	5.028	9	22161	0	0.002	36.704	0.99
	Hotelling's Trace	0.005	5.032	9	27314	0	0.002	45.288	1
	Roy's Largest Root	0.004	13.33	3	9108	0	0.004	40.005	1
	Pillai's Trace	0.002	1.651	9	27324	0.095	0.001	14.863	0.77
School06	Wilks' Lambda	0.998	1.651	9	22161	0.095	0.001	12.057	0.66
	Hotelling's Trace	0.002	1.651	9	27314	0.095	0.001	14.863	0.77
	Roy's Largest Root	0.001	3.379	3	9108	0.017	0.001	10.137	0.76
Sewhme	Pillai's Trace	0.005	7.493	6	18214	0	0.002	44.961	1
	Wilks' Lambda	0.995	7.50	6	18212	0	0.002	45.002	1
	Hotelling's Trace	0.005	7.507	6	18210	0	0.002	45.044	1
	Roy's Largest Root	0.005	14.38	3	9107	0	0.005	43.163	1
	Pillai's Trace	0.009	9.514	9	27324	0	0.003	85.626	1
Tmhmwk	Wilks' Lambda	0.991	9.53	9	22161	0	0.003	69.555	1
	Hotelling's Trace	0.009	9.54	9	27314	0	0.003	85.856	1
	Roy's Largest Root	0.007	22.37	3	9108	0	0.007	67.135	1

The result in Table 4 revealed that the overall F test (over all three dependent variables) and contains the set of multivariate tests of significance that indicate whether there are statistically significant differences among the groups on a linear combination of the academic achievement at wave 1 and wave 2 based on Wilks' lambda (λ) test statistic, the most commonly used to carry out MANOVA [12] measure percent of variance in the dependent variables that is explained by differences in the level of the independent variable whose value varies between one and zero. In this study, all the four different multivariate tests gave the same result of significance of each explanatory variable. Accordingly, all the four test statistics explored that survivor-ship of mother, and openness and honesty of teachers are the only two explanatory variables that have no significant impact on academic performance at wave 1 and wave 2 among all the considered explanatory variables while other covariates were

significant effect on student performance.

Therefore, it has been found that gender, age, number of meals student eat, number of older siblings, number of younger siblings, survivor-ship of father, ability of mother to read/write, ability of father to read/write, ability of siblings to read/write, school type siblings attending, pre-school attendance, time taken to get to school, grade repeating, dropping out school, time spend working for pay on a usual school day, attending extra classes of numeracy, attending extra classes of literacy, teachers caring about students, existence of helping person with school work at home and time spend on homework outside school have a significant effect on the academic performances in both waves. Continuing to examine the results further, it has been found that the partial eta squared associated with the main effect of grade repeating is 0.032 (which is maximum) and the power to detect the main effect is 1. And also, the Wilks' Lambda

associated with grade repeating is 0.968 indicating that 96.8% of variance in the set of dependent variables is explained by differences in the level of the independent variable (grade repeating). These are very good results!

As a hypothetical idea, the parameters of grade repeating up would be written as the following way: "A one-way MANOVA revealed a significant multivariate main effect for grade repeating, Wilks' $\lambda = 0.968$, $F(3, 9106) = 98.82$, $p < 0.05$, partial eta squared = 0.032 and Power to detect the effect is 1. Thus, the alternative hypothesis is confirmed." And, similar discussion would hold true for the rest of significant explanatory variables. Of those twenty two investigated predictors, using partial eta squared as a criterion, grade repeating is the most important variable (though it might have been a weak predictor of the set of academic performances).

3.4. Result of Canonical Correlation Analysis

Table 5. Statistical significance tests for the full CCA model.

Test Name	Value	Approx. F	Hypoth. DF	Error DF	Sig.
Pillai's	0.18628	31.90735	57	27471	0.00
Hotelling's	0.20972	33.67851	57	27461	0.00
Wilks'	0.82081	32.79014	57	27298	0.00
Roy's	0.13601				

EFFECT - WITHIN CELLS Regression: Multivariate Tests of Significance ($S = 3$, $M = 7\ 1/2$, $N = 4576\ 1/2$).

Table 7. Canonical Correlations for Each Function Separately.

Eigen Values and Canonical Correlations					
Root No.	Eigen value	Percent	Cum. Percent	Can. Correlation	Sq. Corr
1	.15742	75.06309	75.06309	.36879	.13601
2	.04547	21.68025	96.74335	.20854	.04349
	.00683	3.25665	100.000	.08236	.00678

Table 8. Standardized Weights for All Functions for the Criterion Variable Set.

Standardized Canonical Coefficients for Dependent Variables			
Variable	Function No.		
	1	2	3
Average_	-.92746	-.43412	-.07129
Averag_1	.01250	.19616	-.98124
dysabsrt	.22070	-.97898	-.21536

Table 9. Structure Coefficients for All Functions for the Criterion Variable Set.

Correlations Between Dependent and Canonical Variables			
Variable	Function No.		
	1	2	3
Average_	-.97658	-.20826	-.05408
Averag_1	-.02308	.20982	-.97747
dysabsrt	.42843	-.88708	-.17188

Table 10. Standardized Weights for All Functions for the Predictor Variable Set.

Standardized Canonical Coefficients for COVARIATES			
COVARIATE	Canonical Variable		
	1	2	3
Boygirl	-.01337	.17765	-.05401
Age	-.34999	-.41253	-.25012
Nummeal	-.36384	-.05455	-.13884

Table 6. Statistical significance tests where only the last canonical function is tested separately.

Dimension Reduction Analysis					
Roots	Wilks L	F	Hypoth. DF	Error DF	Sig.
1 to 3	.82081	32.79014	57.00	27298.10	.00
2 to 3	.95002	13.20831	36.00	18312.00	.00
3 to 3	.99322	3.67882	17.00	9157.00	.00

As in case of multivariate multiple linear regression analysis and multivariate analysis of variance, the result of canonical correlation analysis that can be observed from Table 5 indicates that all the four multivariate test statistics are significant. The result revealed that all the coefficients are significant ($p = 0.00$) as observed under 'Dimension Reduction Analysis' part of the Table 6. And, the hypothesis that pairs of canonical variates tested simultaneously CV_x is independent of CV_y could be rejected. Next tests all remaining pairs (as a set) with the first pair removed, then all remaining pairs with the first and second pairs removed and so on. In this study, all the canonical correlations were significant ($p = 0.00$) with the first removed. However, it is possible this test would be significant even though a test for the correlation itself would not be. As it is simply a correlation coefficient, one should be more interested in the size of the effect, more than whether it is different from zero.

Standardized Canonical Coefficients for COVARIATES			
COVARIATE	Canonical Variable		
	1	2	3
Older_sa	.11689	-.10196	.00313
Mthralve	.04957	-.09449	.12391
Fthralve	.00142	-.10048	.31831
Rdwrmtmr	-.12605	.38849	-.29493
Describe	-.10150	-.10113	-.04364
Rdwrbrth	-.21997	.09997	.10069
Attprese	-.59775	.08484	.41826
Time_to_	-.08642	-.25806	.00006
Reptgrde	.45490	-.03812	.05629
Dropschl	.21945	-.11141	.14956
Hrsfrpay	-.01475	-.20801	.35569
Extrmath	.05373	.17245	-.04910
Extrlang	.06545	.30033	-.14445
Timehmwk	.02836	.18503	.42139
Scwrhome	-.00327	.21592	-.18481
school03	-.06677	-.23832	-.48975

Table 11. Structure Coefficients for All Functions for the Predictor Variable Set.

Correlations Between Covariates and Canonical Variables			
COVARIATE	Canonical Variable		
	1	2	3
Boygirl	-.01795	.25017	-.05099
Age	-.21676	-.55020	-.20651
Nummeal	-.44525	-.04166	-.11937

Correlations Between Covariates and Canonical Variables			
COVARIATE	Canonical Variable		
	1	2	3
Older_sa	.12539	-.10656	.02203
Mthralve	.00462	.00597	.12838
Fthralve	-.00503	-.04358	.28503
Rdwrnthr	-.23650	.47056	-.20562
Describe	-.10179	-.08220	-.05786
Rdwrbrth	-.21470	.15359	.02292
Attpresc	-.60550	.15661	.38700
Time_to_	-.04656	-.34117	-.03612
Reptgrde	.45763	-.10998	.07535
Dropschl	.27679	-.27780	.09695
Hrsfrpay	.00343	-.28986	.35654
Extrmath	.14433	.28703	-.09256
Extrlang	.16164	.38004	-.16904
Timehmwk	-.01193	.11639	.41744
Scwrhome	-.00864	.23011	-.16910
School03	-.10170	-.30858	-.43553

Next, looking at the number of pairs of canonical variates formed under 'Eigen values and Canonical Correlations' part indicated in Table 7 and the canonical correlations among the three pairs (the maximum) of canonical variates constructed. The first is always of most interest and here probably the only one in this study. The first has a canonical correlation of 0.37, the second 0.21. The first canonical correlation have a value at least as large as the largest correlation coefficient between one variable and the opposite set of variables [13], but that canonical correlation can be much larger than that largest correlation coefficient. Standardized canonical coefficients for dependent variables and Standardized canonical coefficients for covariates parts of the Table 7, Table 8, Table 9, Table 10 and Table 11 gave us the standardized coefficients ($a_1, a_2, \dots, b_1, b_2, \dots$) for each pair of canonical variates. These coefficients have the same

3.6. Result of Directed Acyclic Graphs (DAG): Causal Diagrams

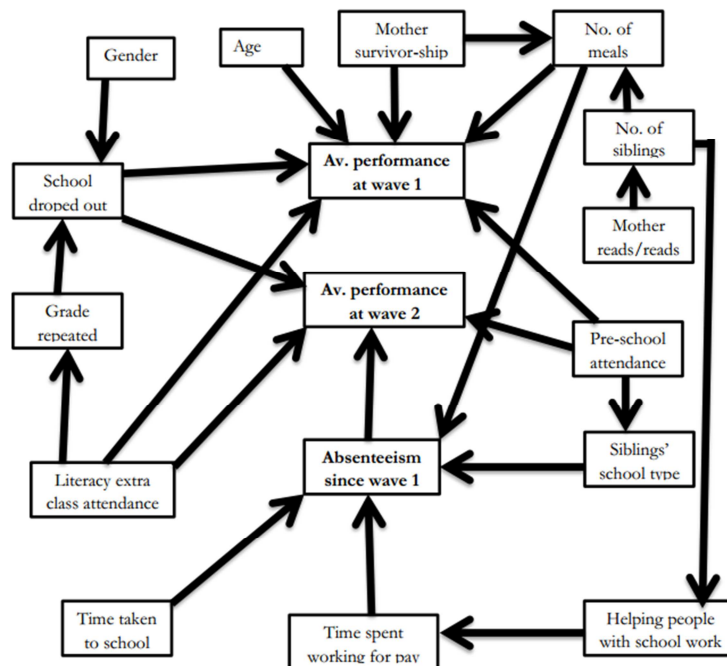


Figure 1. Direct Acyclic Graph showing the assumptions of causality in the model.

interpretation as regression coefficients, and are provided for each pair of variates created, regardless of the correlation's size or statistical significance.

3.5. Result of Simple Linear Regression Model

Simple linear regression analysis was used to investigate the effect school absenteeism on later academic performance of the pupils.

Table 12. Estimate of Regression Model.

Covariate	$\hat{\beta}$	Se ($\hat{\beta}$)	t	Sign.
Intercept	62.285	0.043	361.03	0.00
Days of school absent since wave 1	-0.074	0.40	-2.13	0.033

The result linear regression analysis indicated in the above Table 12 revealed that the school absenteeism has a significant effect on the later academic performance (Average academic performance at wave 2) of a pupil. The significance test from the ANOVA table also indicates the same result. Thus, fitting the simple linear regression of academic performance on school absenteeism as:

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X = 62.285 - 0.074X,$$

where Y and X are the average academic performance at wave 2 and the number of days absent from school since wave 1 respectively. The regression coefficient can be interpreted as "for each day a student absent from the school, his/her academic performance would be reduced by 0.074 scores. The higher number of days a student absent from the school, the lower test scores a student would achieve.

Under this section, the modeling was based on the DAG Figure 1. Consider the proposed study of the relation of mother’s survivor-ship and pre-school attendance to the average academic performance of a student at wave 1, and intending to study how these two variables (exposures) affect the average academic performance of a student at wave 1. Suppose the narrative asserts that mother’s survivor-ship and pre-school attendance are independent among the primary school children. From the DAG, it is observed that considering "mother’s survivor-ship" as the exposure and "average academic performance at wave 1" as the outcome, it is not needed to adjust for any other covariates. The same is true if “pre-school attendance" is considered as the exposure. The statistical model considered under this section is the generalized linear models (binary logistic regression with the assumption of 'no interaction effect'). The explanatory variables are categorical coded as Mother survivor-ship (0 = mother not alive, 1 = mother alive) and Pre-school attendance (0 = not attended pre-school, 1 =attended pre-school)

whereas the outcome variable, average academic performance at wave1, is also coded as binary (0 = below the mean, 1 = above the mean), where the mean score is 67.64.

The result in the diagram indicated that “holding pre-school attendance constant, the odds of student whose mother alive scored an academic performance above the mean was 1.048 times (4.8% higher than) the odds for a student whose mother not alive”. Similarly, keeping the survivor-ship of mother constant, the odds of a student who attended a pre-school education score an average academic performance above the mean was0.432 times (56.8% higher than) the odds for a student who did not attend the pre-school education.

3.7. Results of Count Regression Models

The negative binomial and the Poisson regression analysis were used to examine the effect of covariates on school absenteeism, considering absenteeism as count variable which count the number of school absents.

Table 13. Estimation of Poisson and Negative Binomial regression models.

Parameter	Poisson			Negative Binomial		
	Estimate	Std. Error	p-value	Estimate	Std. Error	P-value
(Intercept)	2.268	0.0778	0.00	2.219	0.2129	0
[boygirl=1]	0.074	0.0082	0.00	0.064	0.0218	0.003
[Age=1]	-0.042	0.0136	0.002	-0.068	0.0374	0.07
[Age=2]	-0.074	0.0121	0.00	-0.097	0.0341	0.005
[nummeal=1]	0.13	0.0189	0.00	0.145	0.0528	0.006
[nummeal=2]	0.141	0.0108	0.00	0.131	0.0305	0
[Older_siblings=0]	-0.156	0.0136	0	-0.176	0.0354	0
[Older_siblings=1]	-0.084	0.0094	0	-0.095	0.0253	0
[Younger_siblings=0]	-0.227	0.0156	0	-0.228	0.0418	0
[Younger_siblings=1]	-0.074	0.0129	0	-0.063	0.036	0.019
[mthralve=0]	-0.097	0.0173	0	-0.066	0.046	0.008
[fthralve=0]	-0.046	0.0115	0	-0.014	0.0304	0.651
[rdwrnthr=0]	0.177	0.009	0	0.173	0.0235	0
[rdwrnthr=0]	0.098	0.009	0	0.111	0.0242	0
[rdwrbrth=0]	0.125	0.009	0	0.142	0.024	0.584
[describe=1]	-0.196	0.0676	0.004	-0.054	0.1833	0.769
[describe=2]	-0.501	0.0682	0	-0.38	0.1844	0.069
[describe=3]	-0.221	0.0681	0.001	-0.101	0.1844	0.583
[attpresc=0]	0.271	0.0087	0	0.271	0.0225	0.689
[time_to_school=0]	-0.14	0.0114	0	-0.152	0.0318	0
[time_to_school=1]	-0.128	0.0124	0	-0.132	0.0344	0
[reptgrde=0]	-0.236	0.0093	0	-0.244	0.0256	0
[dropschl=0]	-0.15	0.0103	0	-0.167	0.0292	0
[hrsfrpay=0]	-0.046	0.0195	0.018	-0.078	0.0529	0.142
[hrsfrpay=1]	0.053	0.0239	0.027	0.035	0.0655	0.591
[hrsfrpay=2]	-0.057	0.0296	0.052	-0.068	0.0786	0.389
[hrsfrpay=3]	0.123	0.0275	0	0.103	0.0769	0.18
[extrmath=0]	0.05	0.0085	0	0.043	0.0222	0.035
[extrlang=0]	0.128	0.0093	0	0.133	0.0238	0
[school03=1]	-0.142	0.0198	0	-0.112	0.0538	0.038
[school03=2]	-0.055	0.0209	0.008	-0.04	0.0571	0.482
[school03=3]	-0.104	0.0241	0	-0.102	0.0656	0.118
[school06=1]	-0.074	0.0188	0	-0.092	0.0507	0.069
[school06=2]	-0.094	0.0203	0	-0.111	0.0547	0.042
[school06=3]	-0.035	0.0227	0.126	-0.05	0.0616	0.414
[scwrhome=0]	0.138	0.0111	0	0.137	0.0313	0
[scwrhome=1]	-0.055	0.0094	0	-0.053	0.0245	0.031
[timehmwk=0]	0.177	0.0189	0	0.154	0.0521	0.003
[timehmwk=1]	0.108	0.0115	0	0.074	0.0296	0.012
[timehmwk=2]	0.087	0.0124	0	0.053	0.0323	0.02

Parameter	Poisson			Negative Binomial		
	Estimate	Std. Error	p-value	Estimate	Std. Error	P-value
Df	10, 299			10, 299		
Pearson χ^2	92316.80			14204.82		
Deviance	10083.04			10046.81		
Log L	-49608.22			-28821.92		

The result in Table 13 showed that the comparison between the Poisson and the negative binomial regression models to investigate the effect of covariates contributing to school absenteeism. The regression parameters for both models give similar estimates. The coefficients in this Table 13 were nearly the same to each other in both models, but the standard errors from the Negative Binomial regression model were all approximately (for instance, 1.5times) larger than the standard errors in the Poisson model, which follows from the variance of the Negative Binomial distribution and significant dispersion.

To make the inference about the regression parameters, the Negative Binomial model give standard errors that are relatively larger than the Poisson regression model and hence, resulted in an insignificant regression parameters for some of covariates in the Poisson regression parameters. Because the only difference between the Poisson and the Negative Binomial lies in their variances; regression coefficients tend to be similar across the two models, but standard errors can be very different. When the outcome variable is over-dispersed relative to the Poisson distribution, standard errors from the Negative Binomial model was larger but more appropriate. Thus, p values in Poisson regression are artificially low and confidence intervals too narrow in the presence of over-dispersion. The deviance for the Poisson regression model in school absenteeism was relatively lower than the degree of freedom. For instance, the deviance was 0.97 times (in Table 13) smaller than the degree of freedom, and thus, indicating the absence of over dispersion. The comparison of the two models was also implemented based on AIC and BIC. Thus, based on the AIC and BIC criteria, the Negative Binomial is better fitting the data than the Poisson regression model due to smaller values.

Thus, taking into account the better fitting of the data based on AIC and BIC criteria, the inference was made on the result of Negative Binomial regression model. Accordingly, the result in Table 5 indicated that the covariates were found to be significant effect on school absenteeism were: Gender, Age, number of meals a student normally eat per day, number of older siblings, number of younger siblings, survivor-ship of mother, ability of mother to read and/or write, ability of father to read and/or write, ability of sister(s)and/or brother(s) to read and /or write, time to get to school, grade repeating, school drop outing, extra classes attendance, availability of helping person with school works at home and time spent on homework outside school.

4. Conclusions

This study aim to investigate the covariates contributing for school absenteeism and associated with average academic

performance based on young lives data using multivariate methods and count regression models. The result of this study found evidence that most of the considered variables have significant effect on school absenteeism and academic performance. The results showed that the nature of variation of school absenteeism and academic performance was different by the number of meal sa student normally eat per a day, type of school siblings attending, pre-school attendance, grade repeating and region.

The multivariate methods applied in this study showed that the covariates that affect the linear combination of academic performances at wave 1 and 2 in Ethiopia include: gender, number of meals a student normally eat per aday, number of competing siblings, ability of mother to read/write, type of school siblings attending, pre-school attendance, time to get to school, grade repeating, school drop outing, time spent working for pay on usual school day and availability of helping person with the school works at home. The Negative Binomial regression model revealed that nearly all that associated with academic performances were also contributing to the school absenteeism too, which was very interesting in this study.

To investigate the effect of school absenteeism on the average academic performance at wave 2 further, the simple linear regression model was fitted. The result indicated that school absenteeism has a significant impact on the academic performance of pupils at primary level i.e. negatively associated with student's academic performances. Moreover, this study found out that pre-school attendance is one of the significant predicting covariate of academic performances that adversely affecting. The number of meal sa student normally eats per a day is also an significantly associated covariate that predict both school absenteeism and academic performances.

It is the hope that the results obtained from this study were of great concern to the family of children and policy makers because of children are generations who are expected to change the country if and only if they are well educated. In order to formulate policies to control the problems of school absenteeism and low academic performance in Ethiopia, it is important to understand the effect of reforms on the primary school and intervention is necessary targeting the significant predictors.

References

- [1] Human Rights Education in Primary and Secondary School Systems: A Self-assessment Guide for Governments, (United Nations, 2012).
- [2] Early child development; the base for learning society, (Keating and Hertzman, 1999).

- [3] Our Children, Our Future – Achieving Improved Primary and Secondary Education Outcomes for Indigenous Students, (Zubrick et al., 2006). Characteristics; Review of Educational Research Vol. 80, No. 2 (June 2010), pp. 272-295, (Crede, 2010).
- [4] Breaking down barriers to learning: primary school-based counseling and support; believe in children, (Murray et al., 2012). [9] Ethiopian Ministry of Education Annual report; Education for All 2015 National Review, (EFMoE, 2015).
- [5] Preventing conduct problems and improving school readiness: evaluation of the Incredible Years Teacher and Child Training Programs in high-risk schools, (Reid, 2008). [10] Impact of Different Types of Schooling on Achievement in the School System: Evidence from Ethiopia, ESP Working Paper Series, University of Sussex, (B. Azubuike, 2015).
- [6] The Contributing Factors to Student Absenteeism/ Truancy and the Effectiveness of Social Services and Interventions, (Teasley, 2004). [11] GLM Multivariate, MANOVA and Canonical Correlation; Statistical Associates blue book series, (G. D. Garson, 2015 p.18).
- [7] The Effect of Attendance on Academic Performance; Mediterranean Journal of Social Sciences MCSER Publishing, Rome-Italy; Vol 5 No 23, (Colby, 2014). [12] Using Multivariate Statistics; 5th edition, (Tabachnick and Fidell, 2006).
- [8] Class Attendance in College: A Meta-Analytic Review of the Relationship of Class Attendance with Grades and Student [13] Canonical Correlation-copy write 2013, (Karl L. Wuensch, 2013).