Assessing Children's Learning Outcomes: A Comparison of two cohorts from *Young Lives Ethiopia*

Tassew Woldehanna[1], Mesele W. Araya[2] and Aregawi Gebremedhin[3]

Received: 09 October 2016; Accepted: 09 February 2017

Abstract: This paper examines the disparity in learning outcomes between two cohorts of children of the same age (12 years old) but seven years apart using data from *Young Lives Ethiopia*. Learning outcomes are measured by the differences in mathematics, common reading and receptive vocabulary tests. By applying simple statistical tests, we particularly looked at whether the level of learning outcomes declined or improved over the seven years period of time. We found a substantial and statistically significant difference in mathematics and reading scores between the two cohorts showing a decline from the older to the younger cohort. In the case of mathematics test, for example, the percentage of correct scores declined from 54.42% in 2006 to 37.17% in 2013. The 17.25 percentage point difference is significant at $P<0.01$. This decline was mainly driven by those children who were living in rural areas; whose caregivers had received little or no education; whose households' wealth tercile was at the bottom; and/or who were in government primary schools. Though the percentage of receptive vocabulary score showed a slight increase from the older to the younger cohort, a rising inequality was observed at intra-cohort level. These findings point out that there is a need to halt the deterioration of learning outcomes and rising inequalities early in life as these may result in different form of inequalities if left unchecked. The paper also proposes further research to identify the in-school and out-of-school factors that might contribute to the declining levels of learning outcomes and widening learning inequalities.

[1] Professor of Economics, Department of Economics at Addis Ababa University and Principal Investigator of Young Lives Ethiopia, based at the Ethiopian Development Research Institute (EDRI).
[2]Assistant Professor, Department of Economics, Addis Ababa University.
[3] Assistant Researcher, Young Lives, Ethiopia.

Introduction

Education plays a key role in the development process of a country. It helps break the intergenerational chains of poverty as it is intrinsically linked to all development goals, including empowering women, improving child and maternal health, reducing hunger, fighting the spread of HIV and diseases of poverty, spurring economic growth and building peace (UNESCO, 2015; Glewwe & Kremer, 2005). Evidence, however, points out that for education to enhance economic growth and reduce poverty children should be both in school and learning (Hawkes & Ugur, 2012; Psacharopoulos & Patrinos, 2004).

It is to be recalled that aiming to meet targets such as the Millennium Development Goal (MDGs) of universal primary education, the major focus of the global education community in recent years has been on getting children into school (Pritchett, Banerji & Kenny, 2013; Van Fleet, Watkins & Greubel, 2011). That effort has been a success, whereas the result the average years of schooling in developing countries increased from 2.0 years in 1950 to 7.2 years in 2010 (Barro & Lee, 2010; Pritchett, 2013). Yet, progress in learning outcomes in many of the developing countries, especially in the areas of mathematics, literacy and science was stagnant or even declining (Beatty & Pritchett, 2012; Kremer et al., 2013; Perlman-Robinson, 2011). As the promise of schooling has failed to translate into learning, measuring the educational status of children through school enrolment and completion rates per se may lead to unintended policy outcomes (Glewwe, 2013; Pritchett, 2013; Petrosino et al., 2012). Because a country could achieve a primary school enrolment rate of 100% but its children might not have acquired the necessary literacy and numeracy skills after completing their primary schooling (Pritchett, Banerji & Kenny, 2013).

There are a number of empirical works that support this argument. A recent study from UNESCO (2013) indicated that about 250 million primary school-age children in low and middle-income countries, of whom 130 million spent at least four years in school, were unable to meet the minimum learning standards in reading and mathematics. This figure could have been much worse had it been measured more widely, but the problems in many developing countries do not measure basic reading or arithmetic in primary grades (Aga Khan Foundation, 2010).

For countries with available data, the figures are however staggering. For example, a study by Pratham and Card (2011) indicated that 31% of children in third grade could not recognize simple words in India. Another study from India also showed that the proportion of children in Grade 8 who could do basic division in mathematics fell from 70 to 57 percent between 2006 and 2011 (Pritchett, Banerji & Kenny, 2013). Similarly, Pandey et al. (2006) studied learning levels and gaps in Pakistan with a focus on students' achievement as measured through test scores at the end of the third grade. They found that learning levels were low, where few students had mastered mathematics to the appropriate level and only about 30% of the students were able to supply the missing word in a sentence.

Gove & Cvelich (2010) who used cross-country data of Early Grade Reading Assessments also revealed that the percentage of students who could not read a single world of a one paragraph story after two years in school was about 18% in Senegal, 28.5% in Guyana, 34% in Liberia, 53% in Gambia; 56% in Mozambique and more than 70% in Nepal. Evidence coming from East Africa countries (Kenya, Tanzania and Uganda) by Uwezo (2012) also showed the similar phenomena, where despite important unprecedented gains in access to primary schooling in the region, large numbers of children are simply not learning. In these three countries, more than two third of children in grade 3 do not have the basic literacy and numeracy skills set at the

grade 2 level. Furthermore, utilizing data from the International Educational Assessment's Trends in International Mathematics and Science Study (TIMSS) at Grade 4 and 8, Mullis et al. (2004) indicated that the performance of children from many of the low-income nations was by far below the mean for countries belonging to the Organization for Economic Cooperation and Development (OECD). Using the same data source, Pritchett, Banerji & Kenny (2013) particularly found that the average eighth grader in Ghana has a test score that would place in the bottom 0.2 percent of US students. This implies that in addition to within the same country over time, learning outcomes vary a great deal between different countries. In relation to this, Beatty & Pritchett (2012) concluded that at their current rates countries like Colombia and Turkey would respectively take 30 and 194 years to reach the average level of learning for the OECD countries while countries such as Indonesia, Iran, Jordan, Malaysia, Thailand and Tunisia are less likely catch up, as learning levels had actually declined from one testing period to the next. Singh (2014) also studied the emergence and evolution of learning gaps across four countries using the unique child-level data obtained from Ethiopia, India, Peru, and Vietnam and found that substantial learning gaps exist across these countries, with Ethiopia at the lower end and Vietnam at the upper end. The magnitudes of the learning gap across these countries also widens as the children grow up, still with children from Ethiopia at the lower end.

The Government of Ethiopia has a vision to transform the country into a low middle-income economy by the year 2025, which demands transformation of the economy through the application of science and technology (Growth and Transformation Plan- GTP II). This calls for continued expansion of and equitable access to high-quality general education with promising foundations in science and mathematics. To this end, several Education Sector Development Programmes (ESDP I to V) have been undertaken over recent years by the Government. Under these programmes Ethiopia has made significant progress in school enrollment. For example, the

primary school net enrolment rate almost doubled in the first decade of the century, rising from just 48% in 2000/1 to more than 90% in 2013/4 (MoE, 2013/14).

Despite the remarkable progress made towards the Millennium Development Goals and the Education for All goals (narrowing down the gender gap in enrollment), declining in learning outcomes as well as high grade repetition and drop-out rates have remained obstacles to the development of the Ethiopian education sector (Woldehanna & Araya, 2016). In an effort to mitigate these problems, the Government has, in addition to the ESDPs, been introducing educational reforms such as launching the General Education Quality Improvement Package (GEQIP) in 2008.

Nonetheless, regardless of the concerted effort of the Government, studies coming from the National Education Assessment and Examinations Agency (NEAEA) do not show promising learning outcomes as expected. Instead students' achievement declined significantly over recent years, where the composite score for learning achievement in Grade 4 fell from 48% in 2000 to 41% in 2007 and has remained relatively flat until 2012. The same is true for grade 8 students, where the composite score declined from 43% to 40% between 2000 and 2007, well below the required levels, and continued to decline since 2007 and reached 35.3% in 2012 (NEAEA 2008a, 2008b, 2003a, 2003b). Another below-par outcome according to the 2010 Ethiopia Early Grade Reading Assessment (EGRA) (RTI International, 2010) was in the area of reading. About 80% of students were not able to read with fluency and disparities were also observed by gender and rural/urban location. In connection to this, Smith, Stone & Comings (2012) wrote that children in Ethiopia are not learning to read well before completing primary school, partly because of lack of clear, attainable goals for reading and English acquisition; lack of textbooks for all children; teachers are not taught the basic components of reading and how to teach them. This is

also supported by another study, where it was found that about 35% children who started grade 3 could not read a single world of a one paragraph story after two years in school (Gove & Cvelich, 2010).

What can be inferred from the above data is that there are declining levels of achievement and deterioration in the levels of learning with a winding inequality among children of different socio-economic backgrounds in Ethiopia. But to ascertain these, there is still a need to explore the learning outcomes of primary school-age children using rich longitudinal datasets that enable us to compare and contrast learning outcomes both at intra and cross-cohorts and to see whether they are declining and for whom.

Accordingly, the objective of this paper is to examine the levels of learning and disparity among children aged 12 years but seven years apart by

- evaluating and comparing mathematics, common reading and receptive vocabulary scores of two cohorts of children of the same age in 2006 and 2013; and
- identifying the main drivers of learning disparities among the children at intra and cross-cohort levels.

The significances of the study are twofold. First, the result provides a picture of the state of children's learning in Ethiopia from a unique longitudinal dataset. That is, having evidence based responses for the above objectives will be an important input for educational quality improvements that are critical for the post-2015 development agenda as many of the 2030 global sustainable development goals will be highly dependent on the quality of education of each country. Second, as far as it is known there are no other studies in Ethiopia that assessed learning outcomes of children at intra and inter-cohort level at a time. So, this study is the first of its kind to look into children's learning outcomes of the same age but with different

cohorts. The outcome of this paper will then serve as a unique empirical evidence for Ethiopia to maximize its transformational potential in education.

Method

Survey design

Young Lives survey is designed as a longitudinal study that is following the lives of 12,000 children in four low and middle-income countries - Ethiopia, India, Peru and Vietnam - over 15 years. Its aim is to improve understanding of the drivers and impacts of child poverty and development over long-term. In each country, as illustrated in Figure 1, the survey tracks 3,000 children in two age cohorts (a Younger Cohort of 2000 children who were about 1 year old during the baseline survey in 2002 and an Older Cohort of 1000 children then aged about 8 years). That is, the two cohorts are seven years apart by birth and it is this type of design that allows the researchers to compare learning outcomes of two cohorts of the same age at any point.

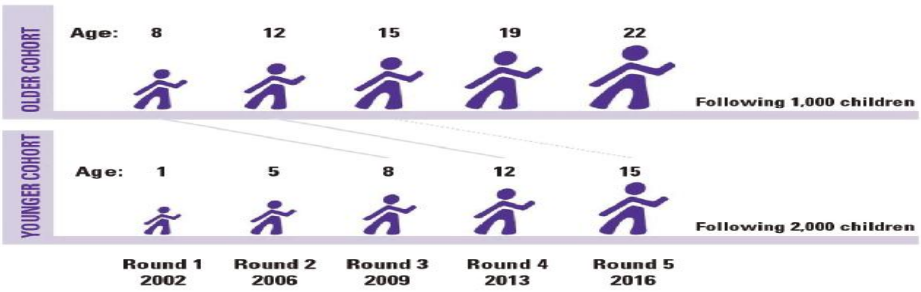

Figure 1: Young Lives design

Source: Young Lives, 2016

The Young Lives survey in Ethiopia is overseen by the Ethiopian Development Research Institute (EDRI) and University of Oxford, funded by UK Aid from the Department for International Development (DFID). The validity and reliability of the instruments used in each survey were approved by the Ethics Committee of the University of Oxford.

Sampling Strategy

Both the Younger and Older cohorts were selected from 20 sentinel sites[4] of the country's five major regions (Tigray, Amhara, Oromia, SNNP and Addis Ababa), which all together represent for more than 90% of the Ethiopian population. The methodology in the first stage was purposive because the sentinel sites in each region were chosen such that: (i) the profile of the selected sites captured Ethiopia's diversity across regions and ethnicities, in both urban and rural areas, and (ii) the cost of tracking children in the future was manageable, to reduce the probability of attrition in remote pastoralist areas (see Young Lives, 2008).

Based on these criteria, three to five sites were selected in each region, with a balanced representation of poor and less-poor households, urban and rural areas. Once the sites were identified, 100 children who were born between April 2001 and June 2002 (the Younger Cohort) and 50 children born between April 1994 and June 1995 (the Older Cohort) were selected in each sentinel site using simple random sampling.

[4] *The study sites from each region are listed as follows: from Tigray (Edaga Hamus, Tsenkanit, Aragure, and May Mekden); from Amhara (Shimshaha, Yimraha, Angote and Mekubia); from Oromia (Dodicha, Bele, Debretsige and Denkaka); from SNNP (Adedamot, Mickaelo, Hayik Dar, Dongara and Tuka); from Addis Ababa city administration(Lideta, Kaliti and Gulelie)*

At the beginning, there was no refusal from the caregivers of the children to be part of the long-term panel survey, but once the survey was started some children were untraceable and some others refused to give responses or moved to overseas and were unreachable. Yet, compared to other longitudinal surveys the attrition rate is low and even slightly lower than in the other three Young Lives study countries (Woldehanna and Araya, 2016). Disaggregated by socio-economic groups, the sample children used in this analysis for the two cohorts are provided in Table 1.

Table 1: Sample children by socio-economic groups and cohorts

	R2- OC (12 years old in 2006)		R4- YC (12 years old in 2013)	
	N	**%**	**N**	**%**
Gender				
Female	479	48.88	884	47.22
Male	501	51.12	988	52.78
Caregiver's education				
None	470	47.96	860	45.87
Adult literacy or religious	153	15.61	164	8.75
Lower primary (1–4)	154	15.71	312	16.64
Upper primary (5–8)	128	13.06	317	16.91
Grade 8 and above	75	7.65	222	11.84
Urban or rural site				
Rural	584	59.59	1,102	58.77
Urban	396	40.41	773	41.23
Region				
Tigray	201	20.51	382	20.42
Amhara	192	19.59	367	19.62
Oromia	200	20.41	392	20.95
SNNPR	244	24.90	463	24.75
Addis Ababa	143	14.59	267	14.27
Type of school				
Private	38	4.09	113	6.39
Community	68	7.33	36	2.04
Government [a]	822	88.58	1,618	91.46
Wealth tercile [b]	328	33.50	630	33.82
Bottom tercile	325	33.20	623	33.44
Middle tercile	326	33.30	610	32.74
Top tercile	328	33.50	630	33.82
Total [C]	**980**	**100.00**	**1875**	**100.00**

Source: Authors' computation

Note: *a. Government school refers to (a school partly funded by government and partly by student fees); and community school refers (NGO/charity/religious)*

b. The wealth index is constructed from housing quality, consumer durables and access to services, including electricity, clean water, and sanitation and cooking fuel.

c. As *result of attrition and deaths, total number of the sample declined from 1000 to 980 for the older cohort and from 2000 to 1875 for the Younger cohort*

There is one caveat on the representativeness of the sample, however. As the Young Lives survey's aim has been mainly to follow the lives of 3000 children and to track the impact of poverty on their overall development over a long period of time, there was some selection bias towards poor households from the beginning in 2002. If this assumption of oversampling of the poor holds, the data does not need to be representative and our interpretation of the results is under this caution. In spite of this selection bias, the sample is an appropriate and valuable instrument in analyzing causal relations and modeling child welfare in the country, where it is hard to find any kind of longitudinal dataset (Sanchez & Outes-Leon, 2008).

Procedure of data collection

The Young Lives survey was conducted for four rounds, in the years 2002, 2006, 2009 and 2013[5]. The data is of a high quality with a wide range of information gathered through child, household and community questionnaires in each round survey. Several new research topics were added to the questionnaires of successive survey to accommodate the fact that as the children grow up, new aspects have become important and relevant, mainly in livelihood, social capital and network, subjective well-being, literacy, numeracy, cognitive skills, education of household members and anthropometry.

During the baseline survey, many of the field supervisors had at least 5 years' experience in field work supervision and all were educated to a university level. Also, prior data collection of each survey a two weeks training led by a research team from the Ethiopian Development Research Institute (EDRI) and University of Oxford was given for all field supervisors in Addis Ababa. The training mainly focuses on the child, household and community questionnaires, tracking schedule,

[5] A fifth round of the survey was also administered between October and December 2016, but the data has not been disclosed for public yet (as of January 2017). Young Lives longitudinal data are available for public use by registering for a password with the UK Data Service and signing a confidentiality agreement before accessing the data.

preliminary interview, consent form, household roster and contact details. The child and household questionnaires were also translated into local languages using an adaptive translation approach, and back-translation to check for accuracy. The field supervisors then used the translated questionnaires to train their respective field enumerators in five groups for a period of one month in each region. Immediately following the trainings, pilot surveys were conducted so that necessary amendments could be made to the questionnaires and organization of the field work. The field enumerators were considered well trained if no error of administrating and scoring was reported by their immediate field supervisors during the piloting. Importantly, working with the same field supervisors and enumerators since the baseline survey in 2002 has enabled us to build stable relations with the families of the children and to minimize errors of administering the interviews and educational tests (Woldehanna, 2016).

Measures

Learning outcomes of the 12 years old children of the two cohorts were measured by the disparities in mathematics test, reading test and receptive vocabulary scores. As the children were in the range of primary school age by the time of the two surveys, all tests were conducted in their mother tongue. This is also in line with the 1994 Ethiopian Education and Training Policy (ETP) that states children of school age need to learn in mother tongue. To be much clearer with the nature of the learning outcomes of the children, it is imperative to provide operational definition for each of the three tests[6].

- *Mathematics test:* An achievement booklet was used to administer the Mathematics test. The children were left to work by themselves in the booklet without any interference unless they ask a question

[6] The validity and reliability of those learning measures in particular and the Young Lives protocol in general were approved by the Ethics Committee of Oxford University.

on how to answer. The number of questions was the same in terms of content in both rounds of the surveys.

- *Reading test***:** this was done by showing the children literacy cards and asked to read letters, words and full sentences on them. The two tests were also the same in terms of content and number of questions over the two rounds survey. The results of the reading test fell into four categories: cannot read anything, can read letters, can read words, and can read sentences.
- *Receptive vocabulary test***:** it measures the overall educational achievement of children. The test requires respondents to select the pictures that best represent the meaning of a series of stimulus words read out by the examiner. It consists of 17 sets of 12 words each and the raw score test results can take possible values from 0 to 204.

Data analysis

Aiming to be much clearer with our data interpretation and to meet the interest of wider readers, we have limited our means of data analysis to common statistical tests such as t-test, one way ANOVA and chi-squares. But this does not mean that our analysis is shallow and lacks depth. We rather presented our data quite in a very clear way by looking at various cross-cutting issues such as gender, urban versus rural and poor versus less-poor. We have also disaggregated the results by parental education, school type and regional location to have an insight on the main drivers of learning disparities among the children. The analysis was done using STATA version of 13.

Results

In this section learning outcomes of the sample children are presented. All the results are disaggregated by different socio-economic characteristics to look at the main drivers of disparity in learning outcomes of the children. We begin our analysis with the mathematics outcomes.

Mathematics test outcomes

Inter-cohort comparison of mathematics test scores

Table 2 reports the percentage of correct answers in mathematics tests for both the Older and Younger Cohorts. It appears that the percentage of correct answers declined from 54.4% in 2006 to 37.2% in 2013, resulting in about 17.3 percentage point differences (vertical educational inequality) over the seven years. This difference is also statistically significant at P<0.01.

The declines in learning outcomes are not, however, equally distributed among children of different socio-economic groups except by gender. If we see the result by geographic location, children in rural areas have higher rate of decline than those in urban areas, with a decline of 20 percentage points for rural and 15 percentage points for urban children. By school type, children who attended private school did see a lower rate of decline than those who were in community and government-funded schools. While those who were in private schools saw their test scores fall by about 11.95 percentage points, for those who were in government-funded and community schools (supported by charity or religious institutions) were about 18 and 27 percentage points, respectively. This means that those who attended community schools experienced the highest fall in their mathematics scores. The same is true for wealth tercile of the family where children from the bottom tercile showed the steepest rate of decline (about 19 percentage points) and statistically significant at P<0.01, while those in the top wealth tercile have seen their scores decline the least. In terms of caregiver's education level, children from all groups have seen noticeable decline, with statistically significant at P<0.01. Regional wise, children from all regions have seen a declining trend of scores over the seven years, but with different magnitude. The highest rate of decline is observed in Tigray (28.3 percentage points), followed by SNNPR (17.6), while the lowest is observed in Oromia (11.8), followed by Addis Ababa (13.4).

In general, one can see from the results that the percentage of correct answers in mathematics test scores declined for all groups, with the highest for children from Tigray (a decline of 28.32 percentage points) followed by children from community schools (26.86) and children whose caregivers had been educated to lower primary level (23.36). On the contrary, the lowest percentage point differences were observed for children who attended private school (11.95) and for children located in Addis Ababa and Oromia regional state.

Table 2: Percentage of correct scores in mathematics test in 2006 and 2013 (%)

	R2- OC (12 years old in 2006)	R4- YC (12 years old in 2013)	Difference (R4-R2)
Gender			
Female	53.05	37.32	-15.74***
Male	55.76	37.09	-18.67***
Caregiver's education			
None	49.69	31.98	-17.71***
Adult literacy or religious	52.18	34.74	-17.44***
Lower primary (1–4)	57.91	34.55	-23.36***
Upper primary (5–8)	62.12	43.26	-18.85***
Grade 8 and above	67.11	51.39	-15.72***
Urban or rural site			
Rural	48.04	28.05	-19.99***
Urban	63.42	48.58	-14.85***
Region			
Tigray	61.47	33.15	-28.32***
Amhara	49.72	34.73	-14.99***
Oromia	45.80	33.97	-11.83***
SNNPR	49.74	32.17	-17.57***
Addis Ababa	69.93	56.57	-13.36***
Type of school			
Private	73.68	61.74	-11.95***
Community	73.20	46.34	-26.86***
Government	53.57	35.96	-17.61***
Wealth tercile			
Bottom	46.24	26.94	-19.30***
Middle	51.40	34.04	-17.36***
Top	65.03	48.55	-16.48***
Total	54.42	37.17	-17.25***
Sample size	950	1638	

*Note: *** p<0.01, ** p<0.05, * p<0.1; R2-Round 2; R4-Round 4; community school refers to (NGO/Charity/Religious).*

Intra-cohort comparisons of mathematics test scores

In addition to the vertical educational inequality of the children as seen in Table 2, it is important to have a look at the horizontal educational inequality of each cohort. To do this, Table 3 presents the percentage point difference of correct answers in mathematics test for each of the Older and Younger cohort children. It appears that the difference in percentage points was not found to be statistically significant for boys and girls in both cohorts, though the difference was positive in 2006 and negative in 2013, which might entail that boys outperformed girls in 2006 but girls outperformed boys in 2013.

There is, however, a huge and statistically significant disparity between children of rural and urban sites. Among the Older cohort the difference was about 15.38 percentage points but this rose to 20.52 percentage points for the Younger cohort. One can also see that the differences in the caregiver's education are correlated with significant differences scores, but the horizontal educational inequality was much wider between children of educated and non-educated primary caregivers (17.47 percentage points for the older cohort versus 19.41 percentage points for the Younger cohort). Also, all regions experienced considerable horizontal educational inequality scores, with SNNPR faring the worst compared to Addis Ababa in 2013. Again, the type of school the children attended was found to be related with significant horizontal differences in their scores. For example, the horizontal difference between children attending government-funded and private schools was about 20.11 percentage points for the Older Cohort and 25.78 for the Younger cohort (statistically significant at $P<0.01$). This means that horizontal educational inequality got wider for the Younger cohort than in the older cohort as the percentage point difference is slightly higher for the former group of children. There is also a statistically significant difference between those in the top and bottom wealth terciles in both cohorts, with children at the top performing better than those in the bottom and middle wealth terciles. Overall, comparing the two cohorts, the horizontal inequality is much higher for the Younger Cohort than for the Older cohort and all horizontal disparities are statistically significant at $P<0.01$.

Table 3: Intra-cohort difference of mathematics scores (Percentage points)

	R2- OC (12 years old in 2006)	R4- YC (12 years old in 2013)
Gender		
(Male) – (Female)	2.71	-0.23
Urban or rural site		
(Urban) – (Rural)	15.38***	20.52***
Caregiver's education		
(Adult literacy or religious) - (None)	2.49	2.75
(Lower primary (1–4) – (None)	8.22***	2.57
(Upper primary (5–8)) – (None)	12.43***	11.28***
(Grade 8 and above) – (None)	17.42***	19.41***
(Lower primary (1–4)) – (Adult literacy or religious)	5.73	-0.18
(Upper primary (5–8)) – (Adult literacy or religious)	9.94**	8.53***
(Grade 8 and above) – (Adult literacy or religious)	14.93***	16.66***
(Upper primary (5–8)) – (Lower primary (1–4))	4.21	8.71***
(Grade 8 and above) – (Lower primary (1–4))	9.20	16.84***
(Grade 8 and above) – (Upper primary (5–8))	4.99	8.13***
Region		
(Amhara) – (Tigray)	-11.75***	1.58
(Oromiya) – (Tigray)	-15.67***	0.82
(SNNPR) – (Tigray)	-11.74***	-0.99
(Addis Ababa) – (Tigray)	8.46**	23.42***
(Oromiya) – (Amhara)	-3.92	-0.76
(SNNPR) – (Amhara)	0.02	-2.57
(Addis Ababa) – (Amhara)	20.21***	21.84***
(SNNPR) – (Oromiya)	3.93	-1.81
(Addis Ababa) – (Oromiya)	24.13***	22.60***
(Addis Ababa) – (SNNPR)	20.19***	24.41***
Type of school		
(Community) – (Private)	-0.48	-15.39***
(Government) – (Private)	-20.11***	-25.78***
(Government) – (Community (NGO/charity/religious))	-19.63**	-10.38**
Wealth tercile		
(Middle) – (Bottom)	5.17**	7.10***
(Top) – (Bottom)	18.79***	21.61***
(Top) – (Middle)	13.63***	14.51***

*Note: *** p<0.01, ** p<0.05, * p<0.1; R2-Round 2 ; R4-Round 4 ; community school refers to (NGO/Charity/Religious);*

Reading outcomes

Similar to the mathematics scores, we also compared the reading outcomes of the two cohorts and explored whether the differences by family wealth index, school type, residence area, gender and region are correlated with significant horizontal and vertical differences in the reading outcomes. We first look at the inter-cohort comparison (vertical educational inequality).

Inter-cohort comparison of reading outcomes

Table 4 presents the percentage of 12-year-olds who could not read anything for both cohorts. It appears that the percentage of children who could not read anything increased from 10% in 2006 to 14% in 2013. Also, even though the percentage of children who could read sentences was higher for the Younger Cohort than for the Older cohort, the total level of reading (including letters and words) had declined from 90% in 2006 to 86% in 2013 (Table 5). This shows that the level of reading of the sampled 12-year-olds showed a declining trend of reading abilities from 2006 to 2013, with a percentage point difference of 3.87 and that is statistically significant at $P<0.01$. By gender, the percentage of girls who could not read anything increased from 9.39% to 14.14% over the seven years. In fact, the proportion of boys who could not read anything also showed an increasing trend, but at a lower rate of increase than that of girls. The results are also statistically significant at $P<0.01$ for girls and at $P < 0.05$ for boys. The reading scores are also disaggregated by caregiver's education level, but the difference in the percentage of children who could not read anything between the two cohorts of children was not found to be statistically significant for all levels of the caregiver's education except for those who are with no schooling at all. In terms of residence area, there are differences in the percentages of reading outcomes between the two cohorts in both urban and rural areas, but the disparity is statistically significant only for the rural children. Regional location was also found

to be with variations, where the levels of reading have increased for some of the regions (Tigray, Amhara and Addis Ababa), but decreased for others (Oromia and SNNPR). The steepest decline in this regard was particularly observed in SNNPR, where the percentage of children who could not read anything almost doubled for the Younger cohort (statistically significant at $P<0.01$). An increase in the proportion of children who could not read anything is also observed in Oromia but not statistically significant at all. By school type, the percentage of children who could not read anything decreased among children attending private schools, while the opposite was the case for those who attended government and community-funded schools. The scores were also disaggregated by wealth terciles and it was found that the percentage of children who could not read anything had increased for all wealth terciles but this increase was only statistically significant for the bottom tercile, and the percentage difference was also the highest for this low-income group (7.44 percentage point difference).

Table 4: Percentage of 12-year-olds who could not read anything in 2006 and 2013

	R2- OC (12 years old in 2006)	R4 - YC (12 years old in 2013)	Difference (R4–R2)
Gender			
Female	9.39	14.14	4.82**
Male	10.57	13.66	3.09*
Caregiver's education			
None	11.70	18.84	7.14***
Adult literacy or religious	7.84	7.32	-0.53
Lower primary (1–4)	9.74	12.50	2.76
Upper primary (5–8)	10.94	11.67	0.73
Grade 8 and above	2.67	4.50	1.84
Urban or rural site			
Rural	12.67	18.06	5.39***
Urban	6.06	7.89	1.83
Region			
Tigray	9.45	6.54	-2.91
Amhara	8.85	8.45	-0.41
Oromia	10.00	13.52	3.52
SNNPR	15.98	31.53	15.55***
Addis Ababa	2.10	1.87	-0.23
Type of school			
Private	5.26	3.54	-1.72
Community	5.88	14.29	8.40
Government	8.39	11.93	3.53***
Wealth tercile			
Bottom	14.94	22.38	7.44***
Middle	9.85	12.68	2.83
Top	5.21	6.07	0.85
Total	10.00	13.87	3.87***
Sample size	980	1875	

*Note: *** p<0.01, ** p<0.05, * p<0.1; R2=Round 2 ; R4=Round 4 ; public school refers a school partly funded by government and partly by student fees)*

Intra-cohort comparison of reading outcomes

In addition to the inter-cohort disparity (vertical inequality in reading outcomes as reported in Table 4 above), we also analyzed the reading outcomes at intra-cohort (horizontal inequality in reading) for both the Younger and Older cohort in relation their socio-economic factors. Since the reading outcomes are categorical we used a chi-square test to assess if there are associations between the reading outcomes and the different socio-economic factors.

Table 5 presents the results of the chi-square tests for both intra-cohorts. While the results under the 2006 represent for the older cohort, the ones under 2013 stand for the younger cohort. It seems that gender does not have a differential effect in each of the cohorts as the chi-square tests (Chi2(3) =4.24; P-value = 0.236 for the Older cohort and chi2(3) =1.001; P-value=0.801 for Younger cohort) are not statistically significant to be different between boys and girls.

There is, however, a clear horizontal inequality in reading outcomes for both cohorts when we look at the data by other socio-economic and geographic variables. In terms of primary caregiver's education, the chi-square test ($Chi^2(3)=27.97$ and P=0.006) shows that there is a statistically significant relationship between reading outcomes and caregiver's education level, where the proportion of children who could not read anything was highest for children whose primary caregivers had no education at all, followed by those whose caregivers were educated to lower primary level and those whose caregivers had received adult literacy or religious education. Similarly, a statistically significant relationship was found between the reading outcomes of children and their residence in urban or rural sites, where children who lived in rural sites performed worse than children in urban areas, and the difference was found to be highly statistically significant ($Chi^2(3)=40.88$ and P=0.01). There was also a statistically significant difference between reading outcomes and the regions in which children lived. The impact of school type

was, however, only marginally significant, while a considerable differential effect was seen by wealth tercile. In conclusion, it seems that the horizontal disparities in reading outcomes of the children are similar by all socio-economic categories for both cohorts, where noticeable differences are observed by all categories except gender.

Table 5: Percentage of children in terms of reading ability in 2006 and 2013 (%)

		2006				2013			
		Can't tread anythi	Reads letters	Reads words	Reads sentences	Can't tread anything	Reads letters	Reads words	Reads sentences
Gender	Male	10.8	15.6	16.0	57.6	13.8	10.5	10.8	65.0
	Female	9.5	13.5	13.1	64.0	14.3	9.1	11.1	65.5
	Difference	1.3	2.2	3.0	-6.4	-0.5	1.4	-0.3	-0.5
		Pearson chi2(3)=4.24; P-value = 0.236				Pearson chi2(3) =1.001; P-value = 0.801			
Caregiver's education	None	11.9	17.1	15.6	55.3	19.2	11.1	11.0	58.8
	Adult literacy or religious	8.0	11.9	14.6	65.6	7.4	6.8	9.3	76.5
	Lower primary (1–4)	10.3	13.2	14.6	61.9	12.1	10.9	12.0	65.1
	Upper primary (5–8)	2.7	9.5	6.8	81.1	6.0	4.8	9.6	79.6
	Grade 8 and above	0.0	0.0	100.0	0.0	0.0	3.6	7.3	89.1
		Pearson hi2(12)=27.97; P-value =0.006				Pearson chi2(12) =68.74;P-value = 0.000			
Urban or rural site	Urban	6.1	9.6	11.7	72.6	7.9	4.9	6.0	81.1
	Rural	12.9	17.9	16.6	52.6	18.3	13.3	14.4	54.0
	Difference	-6.8	-8.3	-4.9	20.0	-10.4	-8.4	-8.5	27.2
		Pearson chi2(3) =40.88;P-value = 0.000				Pearson chi2(3) = 147.17;P-value = 0.000			
Region	Tigray	9.5	19.4	13.9	57.2	6.6	7.9	4.2	81.3
	Amhara	9.1	8.1	11.8	71.0	8.5	4.1	8.0	79.3
	Oromia	10.2	13.7	20.3	55.8	13.6	16.9	20.2	49.4
	SNNPR	16.2	19.1	11.2	53.5	32.3	14.2	15.3	38.3
	Addis Ababa	2.1	9.8	16.8	71.3	1.9	3.0	3.4	91.8
		Pearson chi2(12)=49.36; P-value = 0.00				Pearson chi2(18) = 408.52;P-value = 0.00			
Type of school	Private	5.3	2.6	10.5	81.6	3.5	2.7	5.3	88.5
	Community	5.9	5.9	5.9	82.4	14.3	14.3	11.4	60.0
	Government	8.4	15.3	15.2	61.1	12.0	10.0	11.4	66.6
		Pearson chi2(9) =15.08;P-value = 0.089				Pearson chi2(12) =26.35;P-value = 0.010			
Wealth tercile	Bottom tercile	15.2	18.6	16.2	50.0	22.8	14.7	13.9	48.5
	Middle tercile	10.0	15.1	13.5	61.4	12.8	9.4	10.4	67.5
	Top tercile	5.2	10.1	14.1	70.6	6.1	5.4	8.4	80.1
		Pearson chi2(6) =36.28;P-value = 0.000				Pearson chi2(6) = 146.14;P-value = 0.000			
	Total	**10.1**	**14.6**	**14.6**	**60.7**	**14.0**	**9.9**	**10.9**	**65.2**

*Note: *** p<0.01, ** p<0.05, * p<0.1; R2-Round 2 ; R4-Round 4 ; community school refers to (NGO/charity/religious (Source: own computation)*

Receptive Vocabulary Scores

Inter-cohort comparisons

It is also important to have a further look at the children's learning outcomes beyond mathematics and reading abilities such as on language capabilities. To do this, we have chosen to make analysis on the receptive vocabulary scores of the children, which is a proxy for their language capabilities. Table 6 presents the average percentage of children with correct receptive vocabulary scores in 2006 and 2013. We can see that, overall, the percentage of corrected receptive vocabulary score increased by 2.41 percentage points and was statistically significant ($P<0.01$), showing that unlike in the mathematics and reading tests stated above, the performance of the 12-year-olds in terms of receptive vocabulary score improved over the seven years of time.

However, there are still percentage differences by the various socio-economic factors, where some of the increases in receptive vocabulary scores are not statistically significant at all. For example, when the results are disaggregated by gender we can see that the scores of both boys and girls increased, but the increase is not statistically significant for girls. Likewise, while the increase in score between urban children of the two cohorts is statistically significant ($P<0.01$), the difference for the rural children of the two cohorts is not statistically significant to be different.

Disaggregated by caregivers' education levels, we can also see that the scores increased with the caregiver's level of education for the older cohort but not for the Younger Cohort. However, comparing the scores of the two cohorts, the increasing rate is the highest and statistically significant for children whose caregivers had no education. By region, the receptive vocabulary scores of children living in all regions except SNNPR increased and that the increases

were statistically significant. The scores of Younger cohort from the SNNPR have deteriorated, being about 22 percentage points lower than those of the Older cohort children at the same age, and the difference between the cohorts was found to be statistically significant ($P<0.01$). In terms of school ownership, the average PPVT score increased slightly over the seven years and the difference was statistically significant for both private and government-funded school attendees. The results for the wealth terciles also show that the percentage receptive vocabulary scores decreased by about 7.6% for children in the bottom wealth tercile, which was statistically significant (at $P<0.01$). The result however increased for children in both the middle and top wealth terciles, which may entail that a growing inequality of learning outcomes among children of the poor and rich families.

Table 6: Percentage of Receptive Vocabulary Scores in 2006 and 2013 (vertical comparison)

	R2- OC (12 years old in 2006)	R4- YC (12 years old in 2013)	Difference (R4-R2)
Gender			
Female	61.31	63.49	2.17
Male	62.14	64.87	2.72*
Caregiver's education			
None	56.81	61.66	4.85***
Adult literacy or religious	63.21	65.70	2.49
Lower primary (1-4)	64.43	60.85	-3.58
Upper primary (5-8)	67.37	63.11	-4.25
Grade 8 and above	74.32	78.70	4.38
Urban or rural site			
Rural	54.49	53.74	-0.75
Urban	72.41	78.98	6.57***
Regions			
Tigray	62.28	69.62	7.34***
Amhara	54.64	64.5	9.86***
Oromia	59.80	75.43	15.63***
SNNPR	57.90	36.04	-21.86***
Addis Ababa	79.76	87.82	8.05***
Type of school			
Private	80.37	87.47	7.10***
Community	75.03	69.19	-5.84
Government	60.53	63.43	2.90***
Wealth tercile			
Bottom	53.63	49.54	-4.09**
Middle	58.84	62.50	3.66**
Top	72.62	81.05	8.43***
Total	**61.74**	**64.15**	**2.41****
Sample size	**964**	**1,872**	

*Note: *** p<0.01, ** p<0.05, * p<0.1; R2-Round 2 ; R4-Round 4 ; and community school refers to (NGO/Charity/Religious*

Source: own computation

Intra-cohort comparisons of receptive vocabulary scores

Table 7 presents the relationships between the receptive vocabulary scores and different characteristics of the children for both cohorts, just similar to the horizontal inequality we have discussed in the mathematics and common reading tests. For both the Older and Younger cohorts, gender does not show any statistical significant variation in terms of receptive vocabulary scores.

There are, however, noticeable differences by the other household characteristics and geographic location. For example, the percentage of receptive vocabulary scores of children living in urban sites and those in rural ones for both cohorts of children are quite different. While the horizontal inequality was about 17 percentage points for the older, this was about 25 percentage points for the Younger cohort, suggesting that not only have children from rural areas seen their scores decline but also the gap between their scores and that of urban children has become much wider over time. Except in a few of the categories, caregiver's education level is found to be correlated with considerable differences in percentage point of receptive vocabulary scores as well. All regions also saw statistically significant horizontal disparities for both cohorts, where the highest disparity (25.12%) for the older cohort was seen between children of Amhara regional state and Addis Ababa, in favor of the latter. The is also true for the Younger, where the highest disparity (51.78%) was observed between SNNPR and Addis Ababa, with still in favor of children from Addis Ababa. The same is also true with the type of school attended by the children, where a statistically significant relationship with the percentage point of receptive vocabulary scores of children and the school type was found in both 2006 and 2013. The gaps for the Younger Cohort was particularly higher among the children in government-funded schools and those who went to private schools. Family wealth was also found with significant horizontal inequality of receptive vocabulary for both cohorts. Especially, the disparity between children of the bottom and top family wealth terciles are very high, with 19.99 and 31.51 percentage points of receptive vocabulary scores for the older and the younger cohort children, respectively.

Table 7: Percentage point of difference in Receptive Vocabulary Scores in 2006 and 2013 (horizontal comparison)

	R2- OC (12 years old-2006)	R4- YC (12 years old-2013)
Gender		
(Male) – (Female)	0.83	1.38
Urban or rural site		
(Urban) – (Rural)	17.92***	25.24***
Caregiver's education		
(Adult literacy or religious) – (None)	6.40***	4.04
(Lower primary (1–4) – (None)	7.62***	-0.81
(Upper primary (5–8)) – (None)	10.56***	1.45
(Grade 8 and above) – (None)	17.51***	17.04***
(Lower primary (1–4)) – (Adult literacy or religious)	1.22	-4.85
(Upper primary (5–8)) – (Adult literacy or religious)	4.15	-2.58
(Grade 8 and above) – (Adult literacy or religious)	11.11***	13.00***
(Upper primary (5–8)) – (Lower primary (1–4))	2.93	2.26
(Grade 8 and above) – (Lower primary (1–4))	9.89***	17.85***
(Grade 8 and above) – (Upper primary (5–8))	6.95*	15.58***
Region		
(Amhara-) – (Tigray)	-7.64***	-5.12**
(Oromiya-) – (Tigray)	-2.48	5.81***
(SNNP) – (Tigray)	-4.38*	-33.58***
(Addis Ababa) - (Tigray)	17.48***	18.20***
(Oromiya) – (Amhara)	5.16**	10.93***
(SNNP) – (Amhara)	3.26	-28.46***
(Addis Ababa) – (Amhara)	25.12***	23.32***
(SNNP) – (Oromiya-)	-1.90	-39.39***
(Addis Ababa) – (Oromiya)	19.97***	12.39***
(Addis Ababa) – (SNNP)	21.86***	51.78***
Type of school		
(Community) – (Private)	-5.34	-18.27***
(Government) – (Private)	-19.84***	-24.04***
(Government) – (Community)	-14.50***	-5.77
Wealth tercile		
(Middle) – (Bottom)	5.21***	12.95***
(Top) – (Bottom)	18.99***	31.51***
(Top) – (Middle)	13.78***	18.55***

*Note: *** p<0.01, ** p<0.05, * p<0.1; R2-Round 2 ; R4-Round 4 ; R2- OC =(12 Years Old-2006); R4- YC =(12 years old-2013); community school refers to (NGO/charity/religious (Source: own computation)*

Discussion

Although school enrollment has seen unprecedented improvement in Ethiopia over the last decade or so, children's achievement in school so far is not encouraging as it was supposed to be (Gove & Cvelich, 2010; Smith, Stone & Comings, 2012; Woldehanna and Araya, 2016). To verify this argument with evidence and in an effort to improve learning outcomes of children we feel that undertaking an empirical analysis using suitable longitudinal dataset could be an important policy input. This study is conducted to assess the trend of learning outcomes among children aged 12 years, but seven years apart, making use of longitudinal data from Young Lives Project in Ethiopia.

The overall results indicate that children's learning outcomes, in the areas of mathematics competency and common reading scores deteriorated over the seven years (2006 - 2013). The decline in mathematics achievement was particularly noticeable, where the percentage of correct scores significantly declined from 54.42% in 2006 for the Older Cohort to 37.17% in 2013 for the Younger cohort. The same is true with the common reading scores, in which the percentage of children who could not read anything increased from 10% in 2006 to 14% in 2013. In fact, it is not only the percentage of children who could not read anything that has seen an increase over the seven years, but also the total level of reading (including letters and words) declined from 90% in 2006 to 86% in 2013.

After looking at the trend of learning outcomes, we also tried to identify what drives the declines in score both at intra-and inter-cohort levels focusing on socio-economic factors and geographic locations of the Young Lives sites. It seems that learning level of children was found to be more affected by institutional factors such as type of school attended by the children and their geographic location and their household's characteristics such as family wealth and parental education than by child characteristics such as gender. This may signify that although there could be different institutional and household

factors associated with the declining scores, these phenomena may be explained by the increased enrolment and intake rates in the country over recent years. Between 2006 and 2013, the apparent intake[7] rate increased from 124% to 144% and the net intake rate increased from 63% to 96%. Similarly, the gross enrolment rate increased from 92% to 95% and the net enrolment rate increased from 79% to 86% (MoE, 2006/07; 2012/3). Therefore, the deterioration might partially be explained by the high enrolment and intake rates without the necessary equipment and qualified personnel available at primary level, which might be reflected by an increasing grade repetition and drop-out rates, and ultimately a decline in learning outcomes.

Such argument can be manifested by the fact that the type of school that the children attended is highly correlated with the levels of learning both in 2006 and 2013, where those who attended private schools performed better than those who were in government and community schools, may be because private schools are equipped with better service provision that enable them to deliver better-quality education (Seboka, 2003). This argument is also in line with other evidence that private schools in Ethiopia emerge primarily in response to the high demand for high-quality education by parents, and hence parents preferred private schools to government schools because of their quality and the low level of bureaucracy involved (Hoot, Szente and Mebratu, 2004; Azubuike, 2015; Woldehanna and Araya, 2016). Azubuike (2015) particularly argued that while individual and household characteristics the children enroll into school with are the key drivers of learning outcomes, school factors explain a large share of the educational achievement gaps of primary school pupils in Ethiopia.

[7] *Apparent intake rate refers to the percentage of new entrants (irrespective of age) in Grade 1 out of the total number of children of the official primary admission age (age 7) in a given year*.

Similar to the school type, location of residence (rural or urban; regional site versus Addis Ababa sites) also appeared to make a difference to the learning outcomes of the children. Those who were located in urban sites in general and in Addis Ababa sites in particular scored better learning outcomes than children who were in the rural sites. The reason as to why children from urban sites scored better educational achievements might not be different from the reasons mentioned in relation to the partial impact of school type as it still could be due to the relatively better social infrastructure and facilities in urban sites than in rural sites of each region.

Among the household characteristics, the education level of the caregiver and household's wealth level are found to be correlated with the learning outcomes of the children, while child gender is not statistically significant to show differential learning scores among the primary age children. The possible reason why boys and girls score similar results might be due to the balance in the enrolment and participation of girls in education over recent years. A recent study from Ministry of Education (2012/13) indicates that the primary-level Gender Parity Index, defined as the ratio of female to male gross enrolment is close to 1, which might imply that boys and girls have an equal chance of accessing school lately. This result of gender differential effect is however in contrast to a result indicated by Seleshi Zeleke (2001) who found that a gender difference in favor of boys among sixth graders.

Concluding Remarks

The study used child data from Young Lives project in Ethiopia and employed simple descriptive statistics to compare the levels of learning of two cohorts of children at the same age (12 years old), but seven years apart. Children's learning outcomes are measured by the differences in mathematics, common reading and receptive vocabulary tests. The results show that there is a noticeable decline of children's

learning outcome both in terms of mathematics and common reading scores from the older cohort in 2006 to the Younger cohort in 2013.

The decline in learning scores are, however, quite different for children of different socio-economic backgrounds. Mathematics competency and reading scores have largely fallen to rural children, children of the poorest, children attending community schools, children with uneducated primary caregivers and children from SNNPR as compared to those from other regions in the sample. However, comparing to children in the two rounds overall, the receptive vocabulary score has increased, although rural children and children from SNNPR have still shown a declining trend. It also seems that inequality of learning outcome between children of different socioeconomic groups and regions, already apparent in 2006 increased markedly by 2013. In addition to this vertical inequality of learning outcomes, inequality among the same cohort of children is noticeable in favor of urban children, children of the less poor, children of more educated families, children who attended private schools and those who are from Addis Ababa sites.

In conclusion, there is a need to halt the deterioration of learning outcomes and increasing learning inequality among children of the same age but different socio-economic characteristics. As the different socio-economic groups may have differential access to in-school and out-of-school resources that are inputs for education, it seems that there is a need for a public education programme that can reduce the imbalances among children in accessing the inputs required. Those measures may include (1) education grants for children from disadvantaged groups, and (2) equipping government schools and rural schools with the necessary infrastructure and staff to maintain the level of service quality required. A more comprehensive study however is crucial in order to identify the in-school and out-of-school factors that affect the quality of learning in mathematics and reading, in order to inform the development of research-based intervention programmes.

References

Azubuike, B. (2015). *Impact of Different Types of Schooling on Achievement in the School System: Evidence from Ethiopia, ESP Working Paper Series No. 67.* **Privatisation in Education Research Initiative (PERI)**.

Aga Khan Foundation. (2010). *Improving Learning Achievement in Early Primary in Low- Income Countries: A Review of the Research*, Geneva

Barro, R.J. and Lee, J.W. (2010). A New Data Set of Educational Attainment in the World, 1950–2010, NBER Working Paper 15902, Cambridge, MA: National Bureau of Economic Research.

Beatty, A. and Pritchett, L. (2010). From Schooling Goals to Learning Goals: How Fast can Student Learning Improve? CDG Policy Paper 012, Washington, DC: Center for Global Development.

DFID. (2013). *Education Position Paper. Improving Learning, Expanding Opportunities*. London: Department for International Development.

Federal Democratic Republic of Ethiopia (2015). *The Second Growth and Transformation Plan (GTP II) (2015/16-2019/20), September 2015,* Addis Ababa: National Planning Commission

Glewwe, P. (2013).*What Education Policies Work Best to Increase Student Learning? Lessons from Three Recent Reviews of the Evidence.* **Health and Education Advice and Resource Team (HEART)***.*

Gove, A. and Cvelich, P. (2010). **Early Reading: Igniting Education for All. A Report by the Early Grade Learning Community of Practice** (Research Triangle Park NC: Research Triangle Institute, 2010).

Hanushek, E., and Woessmann, L. (2008). *The Role of Cognitive Skills in Economic Development.* Journal of Economic Literature 46 (3), 607–68.

Hoot, J., J. Szente, and Mebratu, B. (2004). *Early Education in Ethiopia: Progress and Prospects.* **Early Childhood Education Journal** 32 (I), 3-8.

Kremer, M. et al. (2013). *The Challenge of Education and Learning in the Developing World*. ***Science*** 340, 297. DOI: 10.1126/science.1235350.

MoE. (2013/14). *Education Statistics Annual Abstract:* Ministry of Education, Addis Ababa, Ethiopia.

MoE. (2012). *General Education Quality Improvement Project Phase 2: Project Document*, Addis Ababa: Ministry of Education.

MoE. (2010). *Education Sector Development Program IV (ESDP IV)*, Addis Ababa: Ministry of Education.

MoE. (2006/07). *Education Statistics Annual Abstract: 1999 EC /2006-07*, Addis Ababa: Ministry of Education.

MoE. (2001). *Education Statistics Annual Abstract: 1993 EC /2000-01/*, Addis Ababa: Ministry of Education.

Mullis, I.V. *et al.* (2004). *Findings from IEA's Trends in International Mathematics and Science Study at the Fourth and Eighth.* **TIMSS and PIRLS International Study Center**, Boston College.

NEAEA (2008a). *Ethiopian Second Learning Assessment for Grade 4* (2007), Addis Ababa: National Education Assessment and Examinations Agency

NEAEA (2008b) *Ethiopian Second Learning Assessment for Grade 8* (2007), Addis Ababa: National Education Assessment and Examinations Agency.

NEAEA (2003a) *Ethiopian Second Learning Assessment for Grade 4* (2002), Addis Ababa: National Education Assessment and Examinations Agency.

NEAEA (2003b). *Ethiopian Second Learning Assessment for Grade 8* (2002), Addis Ababa: National Education Assessment and Examinations Agency.

Outes-Leon, I. & Sanchez, A. (2008). An Assessment of the Young Lives Sampling Approach in Ethiopia, Technical Note 1, Oxford: Young Lives.

Perlman-Robinson, J. (2011). **A Global Compact on Learning: Taking Action on Education in Developing Countries.** Washington, DC: Brookings Institution.

Petrosino, A. *et al* (2012). Interventions in developing nations for improving primary and secondary school enrolment of children: A systematic review. *Campbell Systematic Reviews,* 2012:19

Pritchett, L. (2013). **The rebirth of education: Schooling ain't learning.** Washington, D.C.: Center for Global Development.

Pritchett, L., Banerji, R., & Kenny, C. (2013). *Schooling is not education! Using assessment to change the politics of non-learning.* Retrieved from Center for Global Development website: http://www.cgdev.org/sites/default/files/schooling-is-not-learning-WEB.pdf

Pratham, E. &. Card, L. R. (2011) Annual Status of Education Report (2011), http://images2.asercentre.org/homepage/ Conference_Pack/enrollment_and_learning.pdf.

Psacharopoulos, G. & Patrinos, H. (2004). *Returns to Investment in Education: A Further Update*, **Education Economics** 12 (2), 111–34.

RTI International. (2010). **Ethiopia Early Grade Reading Assessment. Data Analysis Report: Language and Early Learning.** Addis Ababa: USAID, Ethiopia.

Seaman, M., Levin, J., & Serlin, R. (1991). *New developments in pair wise multiple comparisons: Some powerful and practicable procedures.* **Psychological Bulletin***, 110*, 577-586.

Seleshi Zeleke. (2001). *Gender differences in mathematics performance in the elementary grades: implications for women's participation in scientific and technical occupations.* **EASSRR**, vol. XVII, no. 2.

Smith, C., Stone, R., & Comings, J. (2012). **Literacy policy and practice in Ethiopia: Building on the TELL program and EGRA results.** Washington, DC: American Institutes for Research

UNESCO. (2015). **Education for All 2000–2015: Achievements and Challenges.** Paris: UNESCO.

UNESCO. (2013). *Teaching and Learning for Development.* **Education for All Global Monitoring Report**. UNESCO, Paris.

UNICEF. (2012). **Study on Situation of Out-of-school Children in Ethiopia.** Addis Ababa: ATEM Consultancy Service

Uwezo. (2012). *Are our children learning? Literacy and numeracy across East Africa*. Retrieved from http://twaweza.org/uploads/files/Uwezo%20EA%20Report%20FINAL.pdf

Van Fleet, J., Watkins, K., & Greubel, L . (2011). Africa learning barometer. Center for Universal Education at Brookings Institute. Retrieved from Brookings website: http://www.brookings.edu/research/interactives/africa-learning-barometer#

Woldehanna, T. and Araya, M. (2017). *Early Childhood Investment and Completion of Secondary Education in Ethiopia: Lessons from Young Lives Longitudinal Study*. Working Paper 168, Oxford: The Young Lives

Woldehanna, T and Araya, M. (2016). *Educational Inequalities among Children and Young People in Ethiopia.* Country Report, Oxford: The Young Lives.

Woldehanna T. (2016). *Inequality, preschool education and cognitive development in Ethiopia: Implication for public investment in pre-primary education.* **International Journal of Behavioral Development***,* 40(6),509-16.

Young Lives. (2014). *Young Lives Survey Design and Sampling in Ethiopia.* Preliminary Findings from the 2013 Young Lives Survey (Round 4) Oxford: Young Lives.