RISE Working Paper 21/071
May 2021

Understanding Achievement in Numeracy Among Primary School Children in Ethiopia: Evidence from RISE Ethiopia Study

Dawit T. Tiruneh, John Hoddinott, Caine Rolleston, Ricardo Sabates, and Tassew Woldehanna

Abstract

Ethiopia has succeeded in rapidly expanding access to primary education over the past two decades. However, learning outcomes remain low among primary school children and particularly among girls and children from disadvantaged backgrounds. Starting with a systematic review of quantitative studies on the determinants of learning outcomes among primary school children in Ethiopia, this study then examined key determinants of students' numeracy achievement over the 2018-19 school year. The study focused on Grade 4 children (N=3,353) who are part of an on-going longitudinal study. The two questions that guided this study are: what are the key determinants of numeracy achievement at Grade 4 in primary schools in Ethiopia, and how does our current empirical study contribute to understanding achievement differences in numeracy among primary school children in Ethiopia? We employed descriptive and inferential statistics to examine factors that determine differences in numeracy scores at the start and end of the school year, as well as determinants of numeracy scores at the end of the school year conditional on achievement at the start of the school year. We examined differences across gender, region, and rural-urban localities. We also used ordinary least squares and school 'fixed effects' approaches to estimate the key child, household and school characteristics that determine numeracy scores in Grade 4. The findings revealed that boys significantly outperformed girls in numeracy both at the start and end of the 2018/19 school year, but the progress in numeracy scores over the school year by boys was similar to that of girls. Besides, students in urban localities made a slightly higher progress in numeracy over the school year compared to their rural counterparts. Students from some regions (e.g., Oromia) demonstrated higher progress in numeracy over the school year relative to students in other regions (e.g., Addis Ababa). Key child (e.g., age, health, hours spent per day studying at home) and school- and teacher-related characteristics (e.g., provision of one textbook per subject for each student, urban-rural school location, and teachers' mathematics content knowledge) were found to be significantly associated with student progress in numeracy test scores over the school year. These findings are discussed based on the reviewed evidence from the quantitative studies in Ethiopia.

Understanding Achievement in Numeracy Among Primary School Children in Ethiopia: Evidence from RISE Ethiopia Study

Dawit T. Tiruneh
Research for Equitable Access and Learning (REAL) Centre, University of Cambridge

John Hoddinott
Cornell University

Caine Rolleston
Institute of Education, University College London

Ricardo Sabates
Research for Equitable Access and Learning (REAL) Centre, University of Cambridge

Tassew Woldehanna
Department of Economics, Addis Ababa University

Acknowledgements:

Funding for the RISE Ethiopia project comes from the RISE Programme, which is funded by the UK's Foreign, Commonwealth and Development Office, Australia's Department of Foreign Affairs and Trade, and the Bill and Melinda Gates Foundation. Research partners for the RISE Ethiopia include Addis Ababa University, the Policy Studies Institute in Addis Ababa, Cornell University, Stockholm University, University College London and the REAL Centre at the Faculty of Education, University of Cambridge.

This is one of a series of working papers from "RISE"—the large-scale education systems research programme supported by funding from the United Kingdom's Foreign, Commonwealth and Development Office (FCDO), the Australian Government's Department of Foreign Affairs and Trade (DFAT), and the Bill and Melinda Gates Foundation. The Programme is managed and implemented through a partnership between Oxford Policy Management and the Blavatnik School of Government at the University of Oxford.

Please cite this paper as:
Tiruneh, D., Hoddinott, J., Rolleston, C., Sabates, R. and Woldehanna, T. 2021. Understanding Achievement in Numeracy Among Primary School Children in Ethiopia: Evidence from RISE Ethiopia Study. RISE Working Paper Series. 21/071. https://doi.org/10.35489/BSG-RISE-WP_2021/071

Use and dissemination of this working paper is encouraged; however, reproduced copies may not be used for commercial purposes. Further usage is permitted under the terms of the Creative Commons License.

The findings, interpretations, and conclusions expressed in RISE Working Papers are entirely those of the author(s) and do not necessarily represent those of the RISE Programme, our funders, or the authors' respective organisations. Copyright for RISE Working Papers remains with the author(s).

Contents

Understanding Achievement in Numeracy Among Primary School Children in Ethiopia: Evidence from RISE Ethiopia Study

1. Introduction

The primary education system in Ethiopia has rapidly expanded from three million learners in early 1990s to over 20 million in 2018/19 (Ministry of Education, 2019). However, despite the tremendous progress in expanding access to primary education, poor education quality persists as a major challenge in Ethiopia (Le Nestour et al., 2021; Ministry of Education, 2010, 2015; World Bank, 2017). A large share of children completes their primary education lacking basic literacy and numeracy skills (e.g., NEAEA, 2016; Tesfay, 2012; USAID, 2019; Woldehanna et al., 2016).

Recognising the inadequacy of the primary education system to sufficiently equip children with the required knowledge and skills, major efforts began in 2008 to address issues of quality education through the introduction of government- and donor-supported programs. Among the government-supported programs is the Education Sector Development Programs (ESDP) IV and V (Ministry of Education, 2010, 2015), and one of the most prominent donor-supported programs focusing on quality education is the General Education Quality Improvement Program (GEQIP) (World Bank, 2008).

The GEQIP program aimed to improve the learning and teaching environments in schools by increasing the supply of qualified primary school teachers, providing continuous in-service training for teachers, and distributing textbooks, learning materials and school grants (World Bank, 2013, 2017). Despite the implementation of two rounds of the GEQIP program, GEQIP-I (2008-2012) and GEQIP-II (2012-2018), learning outcomes for many primary school children have not yet reached to the minimum expected standards set by the Ministry of Education (Ministry of Education, 2015; NEAEA, 2016; USAID, 2019). For example, children's average composite achievement in previous five successive national learning assessments (NLAs) at Grades 4 and 8 were far below the 50% expected minimum standards set by the Ministry of Education (GEQAEA, 2008b; NEAEA, 2013, 2016). More specifically, the 2015 NLA revealed that about 44% of Grade 4 students tested nationally were placed at the "Below Basic"[1] level in oral reading fluency. Moreover, the proportion of children performing at advanced levels is very low, in most cases below 10% in Grades 4 and 8 for both mathematics and reading (NEAEA, 2016). Relative to international norms set, for example, by the Trends in International Mathematics and Science Study (TIMSS), about half of the Ethiopian children at the age of 12 fail to reach the low achievement benchmark for children aged 10 years internationally (Singh, 2014). Findings from the 2015 NLA also disclosed major inequality in student achievement across gender and region of residence. In both Grades 4 and 8, female students scored significantly lower than male students. Additionally, students' scores showed disparity across regions, those in Addis Ababa being the highest achievers and those in the so-called "emerging regions" such as Afar and Somali being the lowest achievers (GEQAEA, 2008b; NEAEA, 2016; NOE, 2004a).

There is evidence from the extensive international literature that differences in student learning outcomes are generally attributed to several child, school and household characteristics (e.g., Glewwe

[1] There are four competency levels for oral reading fluency: "None" (non-reader); "Below Basic" (reading slowly with limited comprehension); "Basic" (reading with some fluency and comprehension); and "Proficient" (reading fluently with full comprehension).

et al., 2017; Hungi et al., 2017; Iyer et al., 2020). For example, more advantaged students tend to make more learning progress than the disadvantaged (Glewwe et al., 2017); some schools could make a bigger contribution than others on students' learning progress over an academic year (Rolleston et al., 2013), and that nutritional status, parental education, hours spent studying and doing homework contribute to higher learning outcomes (Iyer et al., 2020). However, there is a dearth of empirical evidence in the context of primary school education in Ethiopia on factors associated with differential learning progress among students within an academic year. Particularly, studies that jointly examined the child, household and school factors that predict primary school children's learning progress over an academic year by using nationally representative longitudinal data are limited.

This study is our initial approach to understand more deeply the trends that are emerging in the RISE Ethiopia longitudinal data. In doing so, we map our findings to the broader quantitative literature on the determinants of learning outcomes among primary school children in Ethiopia. The study is exploratory and aims to achieve two main goals. First, we aim to take stock of the current empirical evidence from quantitative studies on the determinants of learning outcomes in primary schools in Ethiopia. To do this, we undertook a systematic review of the relevant literature. Second, we aim to provide evidence on the associations between learning outcomes using numeracy test scores and key child, household and school characteristics. We then map the findings of the RISE Ethiopia research to current quantitative evidence from the systematic review to understand whether there are persistent or changing factors with respect to what is known in the empirical evidence and/or whether there are any new areas of research enquiry. The following research questions are addressed in the present study:

Regarding the systematic review of the literature, we address the following research question:

(1). What is the current quantitative evidence on the determinants of learning outcomes (and learning *progress* over time) among primary school children in Ethiopia?

Based on the RISE Ethiopia quantitative data, we address the following research questions:

(2). To what extent do Grade 4 students' progress in numeracy test scores over an academic year differ by gender, region, and rural-urban locations?

(3). To what extent are Grade 4 students' numeracy test scores over an academic year associated with key child, household, and school factors?

2. Understanding determinants of learning in primary schools in Ethiopia: Evidence from a systematic review

To situate the present study within the broader Ethiopian context, a systematic review of the relevant literature on the determinants of learning outcomes for primary school children in Ethiopia was conducted. Published official government reports available from the website of the Ministry of Education and other websites from donors were searched. Two databases (ERIC and Web of Science) were searched to look for journal articles, working papers and master's/doctoral theses. The reference sections of previous journal articles were also scanned for relevant articles. We used the following set of keywords (or possible synonyms) to search the relevant articles from the databases: determinants of learning outcomes* (achievement*, progress*, gain*), primary education, Ethiopia, quality primary education, primary school, differences in learning outcomes. We limited our search to journal articles and other official government resources published between 2000 and 2020 (May).

The search from the various sources mentioned above resulted in a total of 25 articles/working papers/theses and nine government/donor reports. The nine government/donor research reports focusing on learning outcomes of primary school children in reading and numeracy were included directly into the review. The abstracts of the 25 journal articles/working papers/theses were read to decide whether the full text of an article should be retrieved or not. To decide whether an article/working paper/theses should be included in the review, the following inclusion criteria were set: (a) the paper should focus on primary school education in Ethiopia, (b) the paper should include quantitative analysis of learning achievement/outcomes of students based on some quantitative measures of learning, (c) the paper should compare learning outcomes of students in relation to several variables that may influence learning outcomes. Overall, our systematic search of the literature resulted in the inclusion of only six journal articles/working papers (most of them are based on the rich Young Lives dataset), and nine government/donor national-level research reports[2].

We present below the main findings from the quantitative empirical literature that focuses on primary school children learning outcomes and the factors that are associated with differences in learning outcomes among primary school children. We first present our review of the findings based on the national-level reports by the National Organization for Examinations (NOE), the National Education Assessment and Examination Agency (NEAEA), and the USAID. Next, the findings from journal articles/working papers based on the Young Lives dataset and other small-scale studies are presented.

2.1. Trends in primary school children's learning outcomes in Ethiopia: Findings from the National Learning Assessments (NLAs) and Early Grade Reading Assessment (EGRA)

2.1.1. Findings from the National Learning Assessments (NLAs)

Recognising the fact that successes in expanding access to primary education were not accompanied by quality learning outcomes, the NOE under the Ministry of Education launched the first national learning

[2] Although we found several studies that focus on primary education in Ethiopia, the topics are mainly related to enrolment (e.g., Dendir, 2014), grade progression/automatic promotion (e.g., Abafita & Kim, 2015; Ahmed & Mihiretie, 2015), dropout (e.g., Woldehanna & Hagos, 2015), and gender inequalities (e.g., Rose & Al-Samarrai, 2001). Quantitative studies that deal with primary school children's learning outcomes and its determinants are relatively scant, perhaps due to the lack of large-scale learning achievement data in Ethiopia. We continue to search for this evidence using our networks and other experts in the field.

assessment (NLA) in 2000 to obtain system-level information on student learning outcomes (NOE, 2000). The NLAs continued to be conducted at four-year intervals until the 2015 academic year. The second NLA was conducted in 2004, the third in 2007, the fourth in 2011, and the fifth and the latest in 2015. The learning assessments were administered to nationally representative children who are at the end of the first cycle (Grade 4) and second cycle (Grade 8) of their primary school education. The key subjects covered were: Reading in mother-tongue, English, mathematics, and environmental science (for Grade 4), and English, mathematics, biology, chemistry, and physics (for Grade 8). Achievement tests for all the subjects were developed based on the national curriculum using the minimum learning competencies[3]. In addition to achievement tests to students, student background questionnaire, school supervisor questionnaire and questionnaire for Grades 4 and 8 subject teachers were administered to collect data on key child, school and household factors[4]. The main objectives of the NLAs were to determine the level of achievement of primary school children upon completion of the first cycle (Grade 4) and second cycle (Grade 8) of the primary education system in view of the Minimum Learning Competencies (MLCs) set by the Ministry of Education, and explore the key child, school, and household factors associated with learning outcomes.

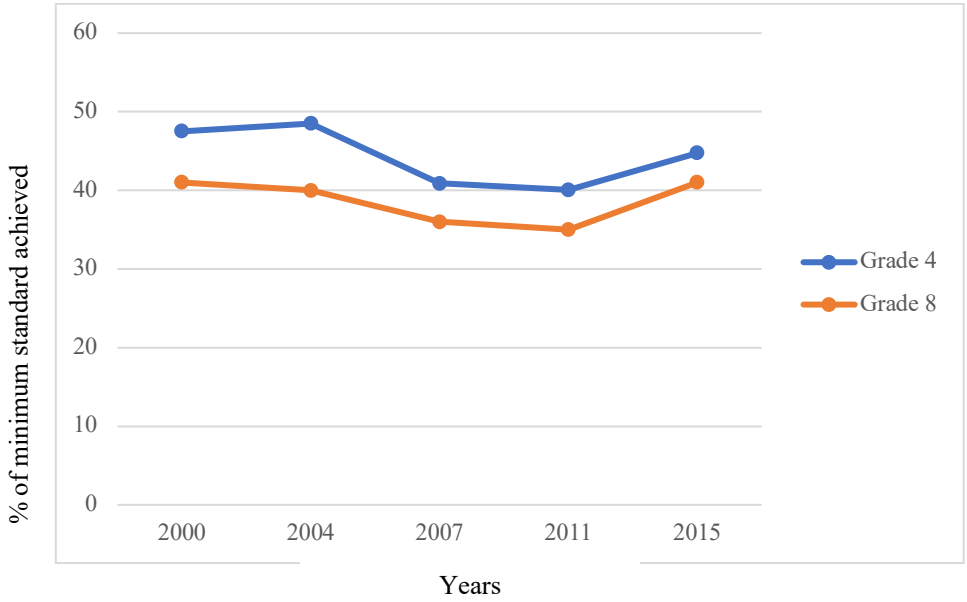
Although it is not explicitly mentioned in all the NLA reports, the decision on the selection of key child, school and household data to be collected and included in the regression models appears to be made based on evidence from the international literature on determinants of student learning outcomes. A brief section in all the NLA reports discusses what is known from the literature on which key factors are strongly associated with learning outcomes. The key background factors included in the reports largely correspond to the identified factors from the literature review section of the reports, and those background factors included are largely consistent across the various NLAs.

The major finding of the five consecutive NLAs was that the composite average scores at the national level for both Grades 4 and 8 were less than the expected 50% minimum standards set by the Ministry of Education. See Figure 1 for the trends in the average composite scores across the five NLAs. Although there is a slight improvement in 2004 on the composite average score, learning outcomes overall declined between 2000 and 2011. There is, however, a slight improvement in 2015 (but remains below the 50% minimum standard).

[3] The different rounds of NLA are not strictly comparable over time, as different sampling strategies have been employed across rounds, and achievement tests are not linked over time. However, we feel the results are comparable in a rough and ready way.

[4] It should be noted that household-related data such as parents' level of education, household wealth, and support from parents on homework were collected from the students, unlike RISE Ethiopia data where household data were collected from primary caregivers in the household.

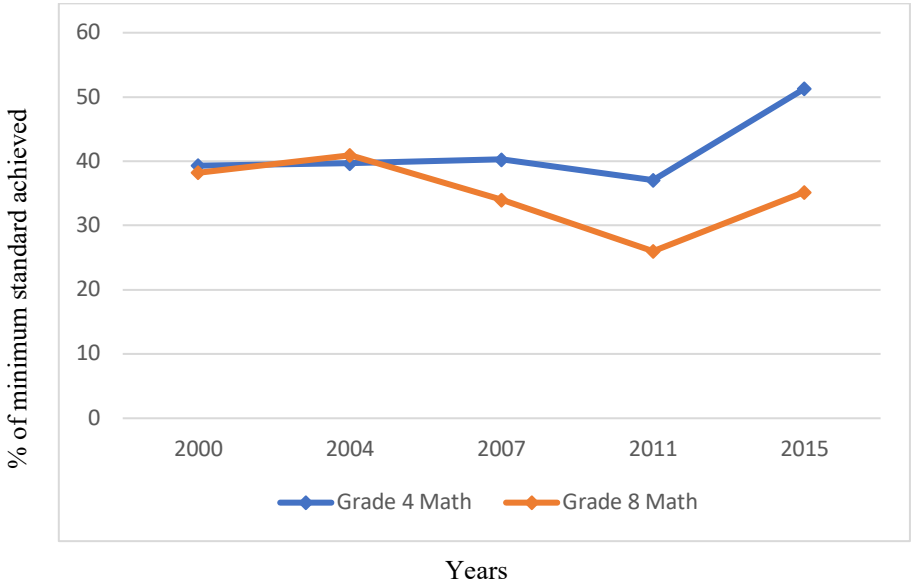
Figure 1. Trends in average composite scores across the five NLAs for Grades 4 and 8



Source: Compiled from the data available on these sources: NEAEA (2013, 2016); GEAQAEA (2008b; 2000, 2004a)
Notes: The standard deviations (SDs) of the students for composite average scores for the different years are presented below in the parentheses:
Grade 4: 2000, 2004 (15.1), 2007 (11.8), 2011 (15.0), & 2015 (15.8).
Grade 8: 2000, 2004 (11.4), 2007 (11.0), 2011 (8.6), & 2015 (13.9).

We particularly examined the trends in mathematics achievement across the five NLAs for both Grades 4 and 8 (see Figure 2). Although students in both grade levels made substantial improvement in mathematics in 2015, the average mathematics scores across the five consecutive NLAs were less than the expected 50% minimum standards (see, NEAEA, 2016).

Figure 2. Trends in mathematics mean scores across the five NLAs for Grade 4 and Grade 8



Source: Compiled from the data available on these reports: NEAEA (2013, 2016); GEQAEA (2008b; 2000, 2004a).
Notes: The SDs of the students for composite average scores for the different years are presented below in the parentheses:
Grade 4: 2000, 2004 (17.5), 2007 (17.0), 2011 (16.8), & 2015 (19.9).
Grade 8: 2000, 2004 (16.2), 2007 (14.7), 2011 (11.4), & 2015 (16.5).

Moreover, the findings from the consecutive NLAs revealed differences in students' learning outcomes across gender, region, and rural-urban locations. Boys significantly outperformed girls in composite average scores at both Grades 4 and 8 in all the consecutive NLAs. When we specifically look at mathematics scores at both grade levels, again boys significantly outperformed girls across the five NLAs[5] (GEQAEA, 2008b, 2008a; NEAEA, 2013, 2016; NOE, 2000, 2004a, 2004b). Across regions, both Grades 4 and 8 students in the major urban centres, such as Addis Ababa and Dire Dawa were mostly the highest achievers, whereas those in emerging regions such as Gambella and Benishangul-Gumuz regions were mostly the lowest achievers (e.g., NEAEA, 2016).

Learning outcome differences across urban-rural localities are mixed across grade levels and the year of NLA administration. For example, the 2004 Grade 4 NLA findings disclosed that students in urban locations significantly outperformed their rural counterparts both in composite average and mathematics scores. However, findings from the same year (2004) but from Grade 8 NLA disclosed that students in rural locations significantly outperformed their urban counterparts both in composite average and mathematics scores (NOE, 2004a, 2004b). Similarly, the 2007 NLA indicated that both Grades 4 and 8 students in rural locations significantly outperformed their urban counterparts in composite average scores. For mathematics scores, again Grade 8 students in rural locations significantly outperformed their urban counterparts, but the difference was not statistically significant for Grade 4 students. The last two NLAs in 2011 and 2015, however, consistently showed that the trend has changed in favour of students in urban locations. Both Grades 4 and 8 students in urban locations significantly outperformed their rural counterparts in composite average scores both at the 2011 and 2015 NLAs. This significant difference also holds true for mathematics scores in the 2015 NLA (NEAEA, 2013, 2016).

Looking at the achievement differences across regions, composite average scores across the five NLAs mostly indicated that students in urban-populated regions including Addis Ababa and Harari, and the so-called established regions including Tigray, Amhara, Oromia and SNNP regions scored significantly higher than those students in the so-called emerging regions including Afar, Gambella and Be-Gu (see, NEAEA, 2013, 2016).

Various child, household and school characteristics that may explain differences in learning outcomes between children were explored in the NLA studies. Table 1 presents the child, household and school characteristics that were reported to be significantly associated with either Grade 4 or Grade 8 students' learning outcomes in the 2015 NLA study.

[5] The only exception is in the 2008 NLA where the mathematics mean score difference between boys and girls in Grade 4 was not statistically significant (although boys scored higher than girls).

Table 1. Findings from the fifth NLA on child, household and school factors* that were reported as significantly associated with student learning outcomes

Child characteristics	*Household characteristics*	*School characteristics*	*Teacher characteristics*
Age	Parental education	Class size	Gender
Gender	Support for child in homework	Distance from home to school	Age
Student absence	Household wealth	Lack of school resources/facilities	Frequency of contact with child parents
Access to additional reading materials at home	Access to electricity at home	Textbook availability	Teachers' qualification
Living with mother and father	Similarity between language spoken at home and language of instruction		Teaching experience
Frequency of meal per day	Urban-rural locality		Teacher absence
			Teachers' teaching assignment per week
			Supervision by principals
			Teachers' in-service training
			Content knowledge
			Frequency of supervision by principal

Source: Compiled from the findings reported on NEAEA (2016).
Note: *The factors explored throughout the consecutive NLAs are similar and we assume that giving an outline of the child, school and household factors examined in the 2015 NLA is sufficient for the purpose of this study.

2.1.2. Findings from the Early Grade Reading Assessment (EGRA)

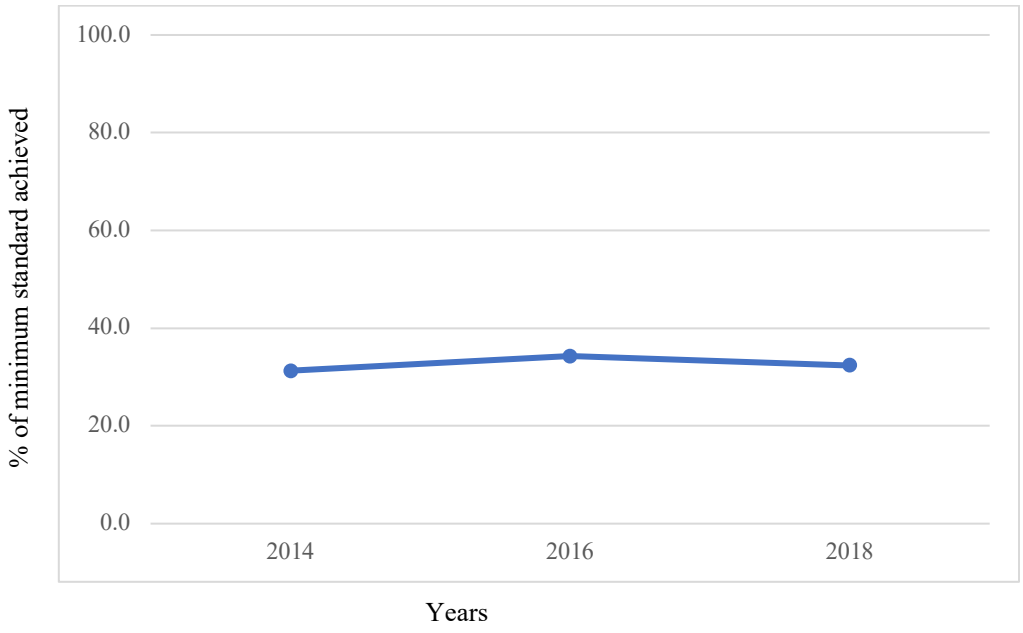
EGRA has been administered in Ethiopia since 2010 by the USAID in collaboration with the Ministry of Education targeting Grades 2 and 3 in eight regions and six local languages. The assessment focuses on reading and language skills identified as being critical for students to become fluent readers. The EGRA instrument consists of a variety of subtasks designed to assess foundational reading skills crucial to becoming a fluent reader: phonological awareness, decoding, reading fluency, reading comprehension, and listening comprehension (USAID, 2010, 2019). The EGRA assessment also includes student, teacher, and principal background questionnaires to explore associations between reading outcomes and various contextual factors. Several student, school, teacher and household factors that were assumed to have some associations with reading performance were included in the EGRA analysis, but it is not clear from the reports how and why those factors were selected.

The 2010 EGRA results disclosed that nearly 34% of students in Grade 2 were unable to read a single word of a 'grade-appropriate' story; 48% of students were unable to answer a single comprehension question, and only 5% of students were able to reach 60 words per minute in reading fluency (USAID, 2010). The 2018 EGRA results showed that the overall percentage of Ethiopian students in Grades 2 and 3 combined, who demonstrated functional reading literacy is 32.4%. The overall percentage-correct for reading comprehension was only 20%. Comparison of EGRA performance across three consecutive EGRA administrations indicated only little improvement in overall reading performance. The percentage of students with Functional Reading Fluency is shown in Figure 3. In terms of gender

differences, boys performed significantly higher than girls in all the sub-tasks for the six local languages except Amharic. Girls scored higher than boys in Amharic, but with a negligible size of difference in scores (USAID, 2019).

The EGRA 2018 study also identified several child, school, teacher, and household factors associated with students' reading performance. Brining the mother-tongue textbook to class every day, reading books in languages other than the mother tongue, borrowing supplementary reading materials, and students' use of the school library were positively associated with students' oral reading fluency. Availability of mother-tongue textbooks for children and guide books for mother-tongue teachers, the number of mother tongue teachers at the school, school library, presence of a reading corner at the school, and the availability of supplementary reading materials were among the school-related factors reported as having a positive association with an increase in Oral Reading Fluency scores. Teacher-related factors including gender, frequency and duration of training received, years of experience, qualification, proper use of student textbook and teacher guide every time they teach, and discussion with parents regarding students' reading progress were reported to have significant positive associations with students' reading performance. Household-related factors[6] including having books at home, having literate family members, getting support from a family member to read, and having enough time for children to practice reading at home were reported as showing significant and positive associations with children's reading performance (USAID, 2019).

Figure 3. Trends in Functional Reading Fluency across three EGRA tests in Ethiopia



Source: USAID (2019).

[6] It should be again noted that household data were collected from the students

2.1.3. Findings from studies based on the Young Lives Ethiopia dataset[7]

Using the longitudinal data from the Young Lives Ethiopia study, Tesfay (2012) examined the effects of ethnicity on primary school children's math and literacy learning outcomes, and the extent to which child, home, and school characteristics predict learning outcomes. Tesfay explicitly mentioned in her report that the key child, home and school variables included in the model were selected based on a review of the relevant literature and availability of data in the Young Lives dataset. The study revealed that learning outcomes are generally low for all the study participants across the ethnic groups, and particularly for certain ethnic minority students including Hadiya and Sidama. Students in Addis Ababa scored significantly higher than the other ethnic groups. On top of ethnicity, Tesfay (2012) identified gender, years of pre-school education, and school start age among the child characteristics that significantly predict both mathematics and literacy scores at the age of 15. On the effect of preschool education, Woldehanna and Hagos (2015), also using the Young Lives Ethiopia dataset, noted that preschool attendance improves the cognitive performance of children at age 8.

Tesfay (2012) further revealed that the availability of services at the household level including access to electricity and water were associated with higher mathematics scores than household ownership of consumer durables such as radio, television and mobile phone. For school characteristics, learning in schools located in urban areas predicted higher math and literacy scores than learning in schools located in rural areas. Principal's qualification significantly predicted literacy scores but not math scores. Other household (e.g., mother's education, number of older siblings) and school (e.g., half-day/full-day school, availability of library, availability of toilets, play areas) factors did not significantly predict either math or literacy score for children aged 15 (Tesfay, 2012).

Another study based on the Young Lives dataset is by Eigbiremolen (2017), which examined the determinants of learning among primary school children in Ethiopia using Round 2 (Older Cohort, aged 12) and Round 3 (Younger Cohort, aged 8) Young Lives datasets. Analysis of the learning outcomes based on the Peabody Picture Vocabulary Test (PPVT) and mathematics tests for the older cohort revealed several findings on the associations between learning outcomes and various child and household characteristics. Concerning child characteristics, Eigbiremolen concluded that the PPVT and mathematics learning outcomes of female children were significantly lower than their male counterparts. Besides, child's age was found to have a significant positive effect on PPVT scores, but not on mathematics. Time spent by a child in a typical school day, however, significantly increased both PPVT and mathematics scores (Eigbiremolen, 2017). In terms of household characteristics, household income was found to have a significant positive effect on both mathematics and PPVT scores; father's education was reported to have a positive effect on PPVT scores, but no effect was found for mathematics score; mother's education was found to have no effects on neither a child's PPVT nor mathematics scores. Eigbiremolen concluded that there is weak evidence that parental education improves children's learning achievements based on the data from the Older Cohort aged 12. Large household size increased PPVT scores, but this effect was not observed for mathematics (Eigbiremolen, 2017).

Again, based on the Young Lives dataset, Woldehanna and colleagues (2016) examined the disparity in mathematics and reading scores between two cohorts of children of the same age (12 years old), but seven years apart (Older Cohort vs. Younger Cohort). They found a substantial and statistically significant difference in mathematics and reading scores between the two cohorts showing a decline

[7] It was our impression that several studies that examine the relationships between learning progresses over time and key child/household/school characteristics were conducted based on the rich Young Lives Ethiopia longitudinal dataset, but we could not find as many quantitative studies as we had expected. We will continue to review the literature on the topic and please contact us if you have published or read any study relevant to the topic based on the Young Lives Ethiopia dataset.

from the older to the younger cohort. Woldehanna and colleagues further explained that lower score for the Younger Cohort was associated with factors including rural-urban residence of the children, primary caregivers' level of education, household wealth and type of school children attended (private vs. government-owned schools).

A recently published study by Iyer, Rolleston, Rose and Woldehanna (2020), based on the Young Lives Ethiopia 2016/17 School Survey, examined learning outcomes of "First Generation Learners" (FGLs - defined as neither parent of a learner has never been to school) who have entered the school system in recent years compared to their peers who have at least one parent with some education. More specifically, Iyer and colleagues examined, among other issues, the effects of FGL status and additional key child background characteristics including gender, age, household wealth, number of meals per day, distance to school, and family size on Grades 6 and 7 students' end-of-year mathematics learning outcome and learning progress in mathematics over the 2016/17 academic year. Some of the main findings of the study are that (i) FGLs scored significantly lower than their peers in their end-of-year mathematics scores after controlling for the afore-mentioned key child background characteristics, and (ii) FGLs make less learning progress in mathematics than their peers within a school year when accounting for prior achievement. Other interesting findings of this study are that background characteristics including gender, age, household wealth, number of meals per day, and family size all significantly affect learning progress over a school year when accounting for prior achievement.

James (2018), based on the 2012/13 Young Lives School Survey, examined the relationship between language of instruction in primary schools and learning outcomes in Ethiopia. A couple of aims of the study were to explore (i) whether there is an advantage to being a mother tongue learner in terms of learning outcomes, and (ii) whether there are inequalities in learning progress between students learning in different languages of instruction. The findings broadly indicated that there is an advantage to being a mother tongue learner in Amharic language classes, and that between-language of instruction inequality of learning outcomes are prevalent, with students learning in many of the 'newly introduced' languages of instruction making less progress in mathematics than their counterparts in Amharic language of instruction classrooms.

2.1.4. Findings from other small-scale studies

Asfaw (2015) examined numeracy skills of Grades 1 and 2 children in Tigray Regional State adopting the Early Grade Mathematics Assessment approach with a focus on counting and number sense. Data were collected from 834 children, 55 teachers and 21 headteachers from 21 randomly selected primary schools and seven districts. Asfaw found that the numeracy skills of Grade 1 and Grade 2 children as measured by the Early Grade Mathematics Competency was 39%, far less than the 50% benchmark set by the Ministry of Education. Besides, Asfaw indicated that children in urban schools significantly outperformed their rural counterparts, and boys scored significantly higher than girls. Among the child and school characteristics, age, gender, availability of functional library and student-textbook ratio predicted higher numeracy learning outcomes for both grades.

Raju and Asfaw (2009) looked at the predictive nature of several variables such as test anxiety, socio-economic status, study habits and parental involvement on academic achievement. A sample of 497 Grade 6 students was randomly selected from seven schools within a suburb of Addis Ababa. A stepwise multiple regression analysis indicated study habits and parental involvement in the child's education as significant and positive contributors to academic achievement. Although test anxiety correlated with academic achievement, it was found to be a non-predictor of achievement in the presence of the other variables.

2.2. Summary of the systematic review

The systematic review overall revealed that numeracy and literacy learning outcomes of primary school children in Ethiopia have been consistently below the minimum competency standards set by the Ministry of Education. It is particularly noted that the targets set by the Ministry of Education in ESDP IV (Ministry of Education, 2010) on increasing the shares of students scoring 50% and above in all subjects by 2015 were not met at all. In terms of variation in learning outcomes across gender, regions, and location, the reviewed studies appeared to consistently indicate that (i) boys significantly outperform girls both in average scores for all subjects and particularly in mathematics and oral reading fluency, (ii) students from urban centres including Addis Ababa, Dire Dawa and Harari significantly outperformed those students in the emerging regions of Be-Gu and Gambella, and (iii) the gap in learning outcomes between students in rural and urban locations is increasing recently in favour of those located in urban areas.

Overall, in terms of child-related characteristics, indicators including gender (male), age (older), number of meals per day, long years of preschool education, and a large amount of time spent at school found to be consistently associated with higher learning outcomes. Child background characteristics including health status, mother-tongue instruction, and height were consistently found to be not significantly associated with learning outcomes. Household characteristics including child support in homework/study at home, access to supplementary reading materials at home, urban location/residence, and living with both biological mother and father were consistently found to be significantly associated with higher learning outcomes. The association between learning outcomes and some household characteristics including parental education, household wealth, and number of older siblings appear to be inconsistent. Besides, when two learning outcomes are measured at the same time (for example, PPVT and mathematics), father's education was found to be significantly associated with PPVT scores but not with mathematics, and that having large household size was found to be significantly associated with higher PPVT scores, but not with mathematics. Most of the factors related to the school (e.g., access to electricity, access to tap water, principal's qualification, availability of student textbooks and teacher guides) and the teacher (e.g., qualification, experience, frequency of in-service training, receiving frequent supervision from the principal, content knowledge, and frequent contact with parents on child's learning) found to be significantly associated with student learning outcomes. The only inconsistency we found on school characteristics was by Tesfay (2012) in which being a half or full-day school, availability of a library, and toilets were reported to be not significantly associated with higher learning outcomes.

Drawing on the findings of the systematic review of the literature on the factors that were found to consistently predict learning outcomes, our knowledge of the context in Ethiopia and data availability in the RISE Ethiopia dataset, we selected key child, household and school characteristics to examine associations with learning outcomes (and learning progress over time). Table 2 presents the key child, household and school characteristics that are included to be explored in the present exploratory study. We used the RISE Ethiopia data to map those factors which were examined in previous studies to understand how they relate to the RISE Ethiopia data.

It should be noted that most of the studies we included in the systematic review and that examined associations between learning outcomes and child/school/household factors followed a cross-sectional approach. There is a learning level measured at the end of a certain grade level, which is considered as the outcome measure of interest, and several child, school, and household factors were included as explanatory variables. The findings normally represent the relative importance of each key variable in explaining the variance in the outcome measure, other variables remaining constant. The results can be interpreted as the factors that explain variation in achievement at a particular grade level. If we look at

gender, for example, most of the studies we reviewed above indicated that boys significantly outperformed girls in mathematics achievement measured at a certain point in time. A major limitation of this type of analysis is that we get very little information on whether this significant gender difference in achievement relates to issues that have happened at school or home. It is possible that this advantage for higher achievement for boys began at home and continued in the school, or it is possible that boys started school lagging behind girls, and that attending school in an academic year has enabled them to catch up and later outperform girls. A major limitation of existing empirical evidence in Ethiopia is therefore related to the dearth of studies that attempt to separate the role of school factors from children backgrounds (see, James, 2018).

The present study employed the RISE longitudinal data on Grade 4 students' mathematics scores collected at the start and end of the 2018/19 academic year. We present empirical estimates for child, household and school factors that determine achievement in Grade 4 mathematics at the start and then at the end of the academic year. In addition, and unlike most existing quantitative studies reviewed here, we also focus on the determinants of Grade 4 mathematics achievement at the end of the school year conditional on the achievement of children at the start of the academic year. It is important to highlight that we are focusing on estimates on the level of numeracy attainment in Grade 4 as this is the estimation method used by most of the existing empirical studies reviewed above. Hence, we are able to make more direct comparisons of the results obtained with the RISE data with respect to other studies, particularly those at national level. A complementary study, which we are currently undertaking, is to estimate the change in numeracy scores between baseline and endline as an outcome variable. This is what is known in the literature as "value added" models.

Table 2. Selected key child, household and school characteristics for the present study

Child characteristics	**Household characteristics**	**School characteristics**
Sex*	Wealth/income*	Principal qualification
Age*	Parental education	Principal's experience
Preschool education	Child support in homework/study	Time spent supervising teachers' lessons
Number of meals per day*	Urban-rural location	Student: Textbook ratio
Health status**	Family size**	Availability of a resource centre in the school
Hours spent a school per day		Availability of library
Hours spent studying per day		Teachers' experience
Hours spent on domestic tasks		Teacher qualification**
Speaks at least one language in addition to mother tongue**		Frequency of in-service training
		Math content knowledge**

Notes: *These variables were found to significantly predict learning *progress* although only based on one study. **These variables were included in just one previous study, and we have included those variables in our study irrespective of the findings because we assumed that sufficient studies have not yet been conducted to draw firm conclusions on those particular variables.

3. Method

3.1. The RISE Ethiopia Sample

RISE Ethiopia adopts a longitudinal design to understand the impact of GEQIP-E on equitable access to quality primary education for all children. The target population of RISE Ethiopia included primary school children, their parents/primary caregivers, school principals, and mathematics and literacy/reading teachers. Participants are from seven regions: Addis Ababa, Amhara, Benishangul Gumuz (Be-Gu), Oromia, Southern Nations, Nationalities and People's (SNNP), Somali and Tigray.

The longitudinal design of RISE Ethiopia involves collecting data from two cohorts of children (Cohort A and Cohort B) at four points in time: at the start and the end of the 2018/19 academic year (Round 1); and at the start and the end of the 2021/22 academic year (Round 2). At Round 1, Cohort A were at Grade 1 and Cohort B at Grade 4. At Round 2, Cohort A will be Grade 4 and Cohort B Grade 7[8]. Round 1 data collection was already completed in the 2018/19 academic year. This study uses part of the Round 1 RISE Ethiopia data. The sample selection followed the RISE Ethiopia identification strategy as described in Hoddinott, Iyer, Sabates, & Woldehanna (2019).

Data presented in this study include data sources from Grade 4 children in 166 schools, their primary caregivers, school principals, and Grade 4 mathematics teachers. Student data were collected at the start and end of the 2018/19 academic year from the same Grade 4 children. Table 3 presents an overview of the sample across the regions. The 166 schools selected capture the regional diversity within the country (see, Hoddinott et al., 2019 for RISE Ethiopia sampling strategy). Across the regions, 4137 Grade 4 students participated at the Round 1 baseline test (R1-B), and 3536 Grade 4 students participated at the Round 1 endline test (R1-E). Data from primary caregivers, school principals and mathematics teachers from the targeted 166 schools were collected at R1-B only.

Table 3. Distribution of the sample Grade 4 students by region, gender and baseline-endline survey

Region	R1-B			R1-E		
	F	M	Total	F	M	Total
Addis Ababa	253	277	530	241	257	498
Amhara	304	323	627	266	290	556
Be-Gu	224	244	468	172	199	371
Oromia	517	565	1082	404	444	848
SNNP	286	282	568	226	208	434
Somali	162	178	340	167	168	336
Tigray	260	262	522	250	225	475
TOTAL	**2006**	**2131**	**4137**	**1639**	**1714**	**3536**

Source: RISE Ethiopia Quantitative Study.

Table 4 presents Grade 4 sample distribution across the seven regions for female and male students who completed both the R1-B and R1-E tests. The overall sample reflects roughly the same number of female and male students, with females making nearly 49% and males 51%. Disaggregated by region, the

[8] See Hoddinott et al. (2019), one of the published RISE Ethiopia working papers, for a detailed explanation on RISE Ethiopia sampling and identification strategy.

pattern is largely reflected in Addis Ababa, Amhara, Be-Gu, Oromia, and Somali, where males are slightly higher than females by nearly 3 to 7%. In SNNP and Tigray, females are slightly higher than males by nearly 4%.

Table 4. Sample distribution across gender and region

Region	Female		Male		Total
	N	%	N	%	N
Addis Ababa	222	47.8	242	52.2	464
Amhara	249	48.3	267	51.7	516
Be-Gu	172	46.4	199	53.6	371
Oromia	404	47.6	444	52.4	848
SNNP	226	52.1	208	47.9	434
Somali	136	48.7	143	51.3	279
Tigray	230	52.2	211	47.8	441
TOTAL	**1639**	**48.9**	**1714**	**51.1**	**3353**

Source: RISE Ethiopia Quantitative Study.

3.2. Instruments

Numeracy test data administered to Grade 4 students at the start and end of the 2018/19 academic year was used to measure students' learning progress over the academic year[9]. The test items were adapted from the Grade 4 mathematics test items administered at the Young Lives Ethiopia School Survey in 2012-13 academic year. The RISE Ethiopia team revised and piloted the Young Lives items in February 2018 following guidance from test developers at the Ministry of Education and the National Educational Assessment and Examinations Agency (NEAEA) in Ethiopia. The final Grade 4 test included 25 numeracy items both at R1-B and R1-E.

The test included 15 common (anchor) items both at R1-B and R1-E. The 15 anchor items from the R1-B were replicated at R1-E because the items were found to be functioning well at the baseline, and were expected to enable the linking of test scores between R1-B and R1-E. The remaining 10 items at R1-E were new, included to measure those competencies the children would be expected to have developed over the academic year, as a result of learning Grade 4 mathematics. The replicated items at R1-E were placed at the same location as in the R1-B to easily compare how the items functioned both in the baseline and endline. See Appendix 1 for the list of R1-B and R1-E Grade 4 numeracy items.

A questionnaire was administered to school principals from the targeted 166 schools and data such as principal's qualification, experience, frequency of supervision of teachers' lessons, students' access to educational resources, boys-girls' enrolment and learning, availability of girls' clubs, availability and on-time delivery of school grants, and enrolment and learning of children with disabilities were collected. A household questionnaire was administered to the targeted students' primary caregivers, and data such as students' age, access to pre-school education, health, nutrition, daily food diversity, functional difficulty, household wealth, etc. were collected.

[9] Literacy tests were also administered to all students. Given that local language was used to capture literacy, there is less reliability on the literacy tool to capture changes across languages than with the numeracy tool. For this reason, we focus on this paper on numeracy.

A teacher content knowledge test was also administered to Grade 4 mathematics teachers at R1-E. One of the assumptions in the Ethiopian primary education system is that poor teacher quality is a major contributor to low student learning outcomes (Ministry of Education, 2015). We therefore examined the relationship between mathematics teachers' content knowledge and students' numeracy learning outcomes. A total of 20 items, adapted from the Young Lives Ethiopia 2012-13 School Survey, were administered.

3.3. Transforming raw scores from Grade 4 math test into interval scales: Item-Response Theory Modeling

A two parameter-logistic model (2PL) was fitted on the Grade 4 numeracy item responses to determine the ability of the test items to measure the latent trait and individual item functioning. The 2PL IRT model allowed us to transform the raw scores on both the R1-B and R1-E test items into a common scale for each sample, taking into account our respondents' ability, the item difficulty and item discrimination parameters of each item. The IRT generated scores reflect the number of items answered correctly by a student including the difficulty levels of the items answered correctly. Unlike percentage correct scores, scaled IRT scores allow a direct comparable estimate of the underlying skill (latent trait), and more precise and unbiased estimates of individual or group baseline-endline student achievement differences. The estimates of latent traits are extracted in the form of variables centred on zero reported in logits. To aid interpretation, and to ensure that all scores are compared in standard deviation terms, we transformed the estimates into a different interval scale, while retaining the same properties. The latent trait estimates were transformed into a scale with an average of 500 and a standard deviation of 100 for the R1-B data. Data from R1-E tests were transformed onto the R1-B scale, and thus the R1-E scores can be readily interpreted by comparison with R1-B scores in terms of learning gain[10].

3.4. Analytical strategies

We employ the following empirical strategies to answer our research questions. Grade 4 students numeracy learning levels at baseline and endline (as well as the gain) are presented for the whole sample, as well as by gender, region, and rural-urban localities. An independent sample t-test was conducted to compare the students' numeracy scores separately at baseline and endline across gender, region, and urban-rural localities. We also conducted a paired t-test to determine whether, on average, there was a change in numeracy scores from the start to the end of the school year for the whole sample and across gender, region, and urban-rural localities. In addition, we divide our sample into quartiles based on baseline mathematics scores and explore the extent to which students remain within the same quartile by endline. To do so, we adopt the technique of transition matrices by Feinstein (2003), and later employed by Carter, Rose, Sabates and Akyeampong (2020). With the view that GEQIP-E reform pays an extra emphasis on equity, we examine extreme movement patterns of low-performing children at the baseline separately for male and female students.

To examine the associations between learning levels (at baseline, endline and endline conditional on baseline attainment) and key child, household and school characteristics, we employ ordinary least squares and school 'fixed effects' approaches. Because the RISE Ethiopia 2018/19 dataset includes both baseline and endline numeracy test scores, this gave us the advantage to examine endline learning conditional on baseline learning. This model is equivalent to a value-added model, where learning gains between endline and baseline becomes the main explanatory variable, under the assumption that baseline test scores predicts endline scores with a parameter of 1. In other words, a one-point increase

[10] Please see the Data Analysis Plan, one of the deliverables submitted to the RISE programme in March 2020, for details on why and how we employed IRT modelling to measure students' learning progress over a school year.

in baseline score predicts a one-point increase in endline scores. This assumption may be too strong, and therefore we opted to estimate the parameter of baseline test in the model. For this reason, our models estimate numeracy test scores in endline conditional on test scores at baseline and we focus on the role of child, household and school characteristics in predicting conditional numeracy scores at endline. Using these models, we improve the accuracy of the estimates because the inclusion of the baseline score controls for students' prior achievement in numeracy. The interpretation of the remaining student and household background characteristics would, therefore, be how well students did on the test at the end of Grade 4, conditional on how well students did at the start of Grade 4. It should be noted that the RISE Ethiopia sample was from 166 schools across seven regions in Ethiopia. Similar to previous studies (e.g., James, 2018; Rolleston et al., 2013), it is our goal to control for the potentially powerful effects of different intakes of students and school resources/facilities in different schools. We employ therefore a school 'fixed effects' approach so that results compare students with their peers within each school.

To examine the extent to which the identified key school factors determine students' numeracy tests at the start and at the end of the academic year and tests at the end of academic year conditional on tests at the start of the academic year, we run a separate regression model using ordinary least squares approach. In Model 1, we run regression taking the R1-E test score as an outcome variable and key school characteristics as predictors (Column 1). In Model 2, we run regression taking the R1-E test score as our outcome variable and the R1-B test score and the identified key school characteristics as predictors (Column 2). These models do not use school fixed effects as our aim is to understand the school level factors which are related to numeracy test scores during one academic year.

4. Findings

The first section presents some descriptive analysis on numeracy test scores at the start and at the end of the academic year, as well as gains made during one academic year. We present these for the overall sample, as well as by gender, regions and urban-rural localities. The second section presents findings using multivariate analyses and school fixed effects to examine associations between numeracy test scores during the academic year and key child and household characteristics. To examine the key school characteristics associated with test scores we employ ordinary least squares as our analytical tool.

4.1. Numeracy: learning levels and gains across gender, regions and urban-rural localities

Table 5 presents the Grade 4 numeracy mean scores and their standard deviations for the full sample at the start of the academic year (R1-B), at the end of the academic year (R1-E) and the gains over this period. Across the sample, the R1-E mean score (536 points) is significantly higher than the R1-B mean score (500 points), $t(6704) = 13.935$, $p < 0.001$. The learning gain over the academic year is 36 points, or 0.36 standard deviation. Relative to the baseline score, the students made substantial progress over the academic year. See Figure 4 for the distribution of the scores both at R1-B and R1-E.

Table 5. *Means and standard deviations of Grade 4 numeracy scores at R1-B and R1-E, by gender, region and locality*

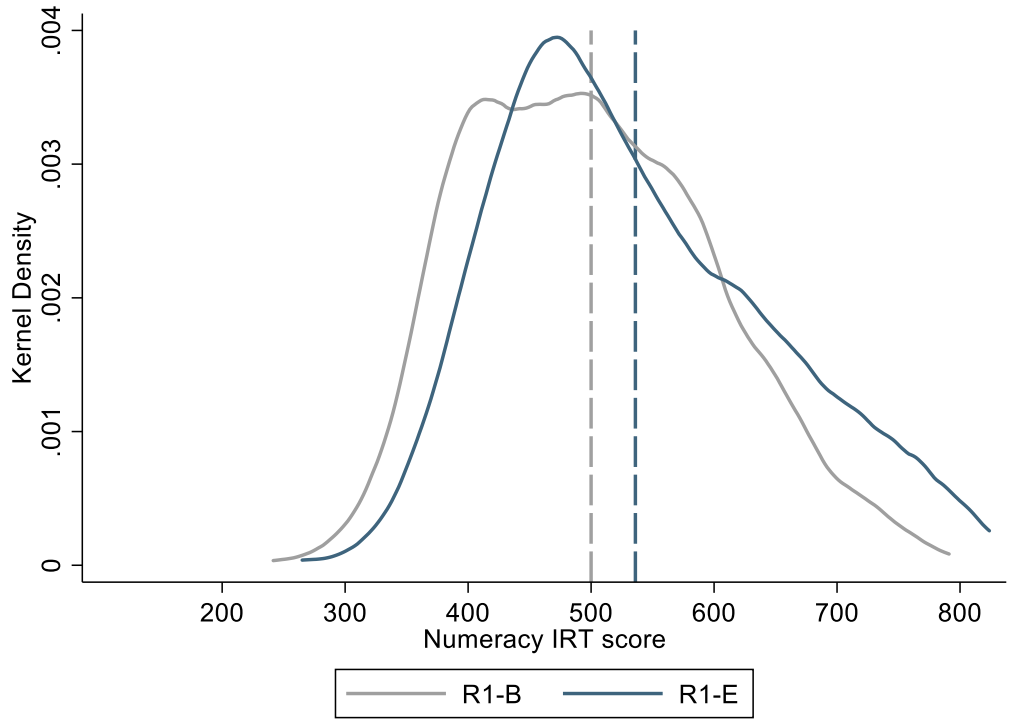
		Students present at both R1-B and R1-E						
		N	**R1-B**		**R1-E**		**Gain**	t statistics
			Mean	**SD**	**Mean**	**SD**		
Total		3353	500	100	536	111	36***	13.93
Gender	Male	1714	508	101	544	114	36***	9.76
	Female	1639	492	98	528	107	36***	10.02
	Difference		16***		16***			
Region	Addis Ababa	464	582	85	620	105	38***	6.04
	Amhara	516	516	100	558	114	42***	6.33
	Be-Gu	371	456	94	489	95	33***	4.78
	Oromia	848	481	99	533	108	52***	10.22
	SNNP	434	475	85	498	90	23***	3.91
	Somali	279	473	73	453	61	-20***	3.53
	Tigray	441	510	95	557	106	47***	7.02
Locality	Urban	1146	544	97	583	112	39***	8.99
	Rural	2207	477	94	511	103	34***	11.58
	Difference		67***		72***			

Source: RISE Ethiopia Quantitative Study. *Notes*: t-test of the learning gain is significant at $^{***}p<0.001$; $^{**}p<0.05$ $^{*}p<0.1$; Standard errors were clustered at the school level.

Males significantly outperformed females both at R1-B and R1-E, with 16 points difference in each testing phase. However, the learning gain over the school year is significant for both female and male students across the sample, and the same 36 points value for each (see Figure 5). From a very high starting point, males gained 36 points, and females, from a relatively lower starting point, gained the same 36 points over the academic year. This suggests that the gender achievement gap, across the

sample, did not reduce over the school year and it is possible that boys continue to outperform girls, on average, in the numeracy test in Grade 4.[11]

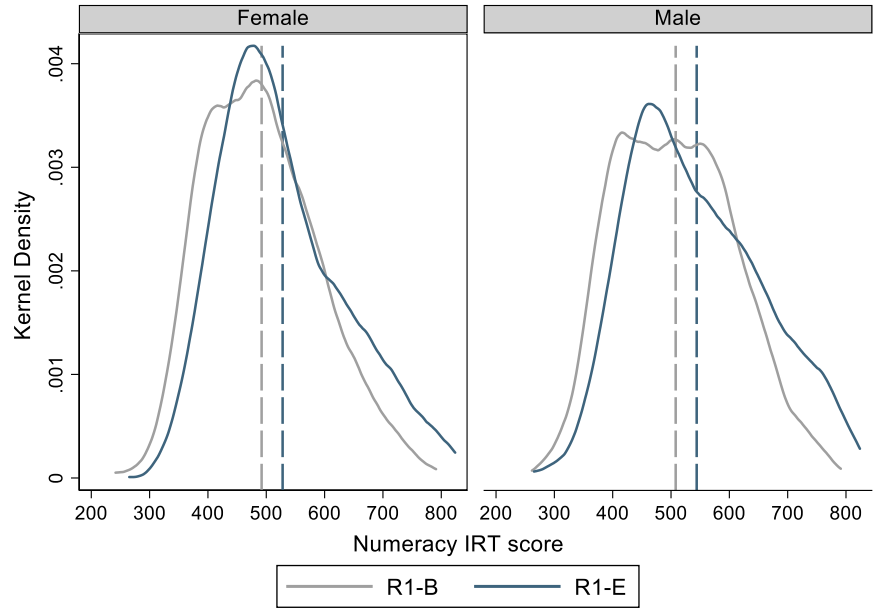
Figure 4. Distribution of numeracy scores across the Grade 4 sample at the R1-B and R1-E tests



Source: RISE Ethiopia Quantitative Study

[11] Although not explored in this working paper, there are regional differences in learning gains. Except for Addis Ababa, where females gained 10 points higher than males over the school year, a slightly higher learning gain is made by males than females in all other regions.

Figure 5. Distribution of numeracy scores across gender at the R1-B and R1-E tests

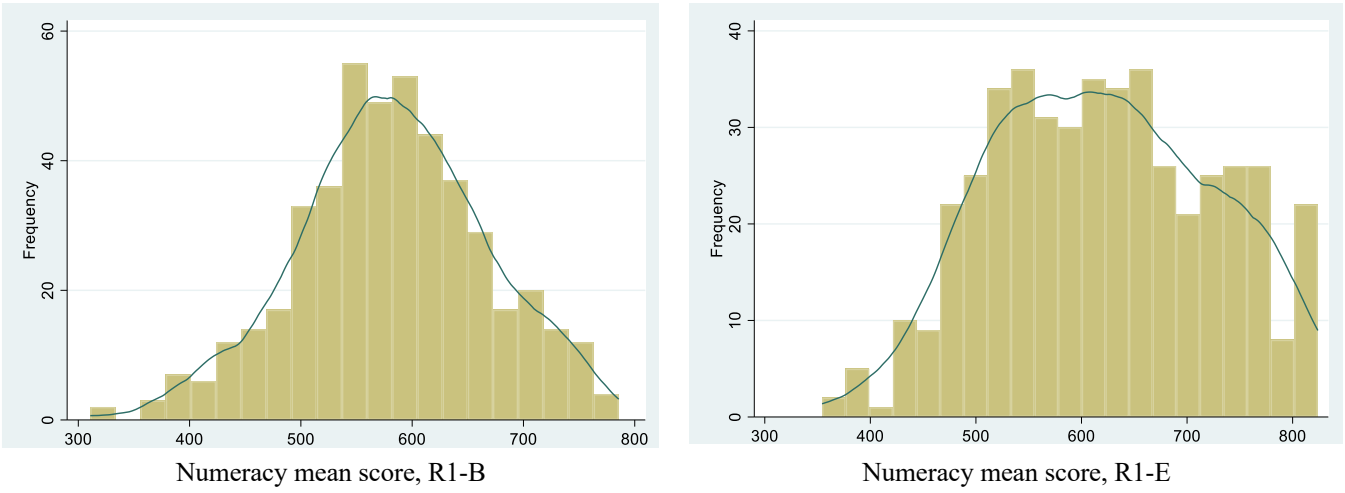


Source: RISE Ethiopia Quantitative Study

Overall, the learning gains vary across regions. Students in Addis Ababa were the top achievers both at R1-B (582 points) and R1-E (620 points). However, the highest learning gain was observed in Oromia (52 points), followed by Tigray (47 points) and Amhara (42 points) regions. The learning gain in Addis Ababa was the third lowest (38 points), possibly because they had an already very high starting point and the test design imposed a ceiling to demonstrate their potential improvement. For example, at R1-B, 33% of the students in Addis Ababa were placed at the top 25 percentile score, whereas only 19% of the students in Oromia region were at the top 25 percentile. See Figure 6 for the distribution of numeracy scores for students in Addis Ababa at R1-B and R1-E.

Students in Somali region were the second-lowest achievers at R1-B (477 points), but they became the lowest achievers at R1-E (453 points), with a decrease in 20 points. It should be noted, however, that students in Be-Gu, one of the emerging regions, gained higher than SNNP, one of the established regions, over the academic year.

Figure 6. Distribution of Grade 4 numeracy scores for students in Addis Ababa at the R1-B and R1-E tests



Source: RISE Ethiopia Quantitative Study.

Students in urban schools significantly outperformed their rural counterparts both at R1-B and R1-E. The difference in numeracy scores between students in rural and urban localities at R1-B was 67 points, and this had slightly increased to 72 points at R1-E. An important point to note is that the average numeracy score of students living in rural localities at R1-E (511 points) was lower than that of children in urban areas at R1-B (544 points). In other words, after one year in school, children in rural areas could not reach the level that children from urban areas had at the start of Grade 4.

The progress in numeracy scores over the academic year for students both in rural and urban localities was also statistically significant. Students in urban schools, however, made a slightly higher learning gain (39 points) compared to their rural counterparts (34 points) (see Figure 7). It should be noted again that the slightly higher learning gain for urban students was from an already higher starting point.

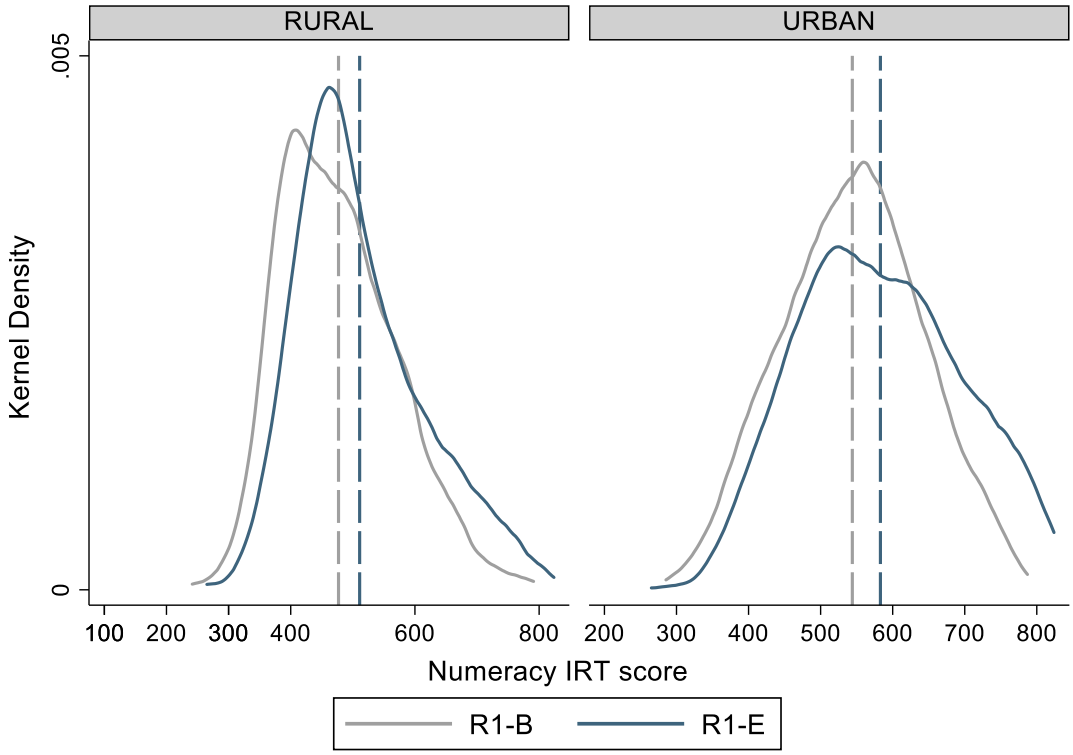
Through the application of transition matrices, we also explored the relationship between baseline and endline numeracy scores of initial low- and high-performing students. The goal is to understand the extent to which lowest performing students at baseline differ in their ability to progress from their relatively weak initial positions. As indicated in Table 6, out of the 25% of students scoring the lowest at the baseline (Bottom Quartile), 53% were still in the bottom quartile at the end of the academic year. Only a few of the lowest performing (nearly 3%) entered the top quartile at the end of the academic year. Similarly, among the top scoring group, 66% of them remained in the top quartile at the end of the academic year. Overall, the findings indicate that the lowest-performing students at the baseline largely remain in their lowest position at the endline without considerable movement within the distribution over the academic year. Although a slightly higher number of lowest-performing boys shifted to the top quartile at the end of the academic year, the movement appears to be largely stable for both boys and girls.

Table 6. Quartile transition matrices for lowest, medium and top achievers in numeracy at baseline and endline

Quartile at Baseline	**Quartile at Endline**			
	Bottom Quartile	Second Quartile	Third Quartile	Top Quartile
Bottom Quartile	53.1	33.3	11	2.7
Second Quartile	31.9	36.5	25.9	5.6
Third Quartile	12.4	22.8	39.4	25.4
Top Quartile	2.6	7.4	23.8	66.3
Total	100	100	100	100

Source: RISE Ethiopia Quantitative Study.

Figure 7. Distribution of Grade 4 numeracy scores across rural and urban localities at the R1-B and R1-E tests



Source: RISE Ethiopia Quantitative Study.

4.2. Regression Analysis

We explore in this section the extent to which key child, household and school factors determine students' numeracy scores, particularly numeracy achievement at the end of the academic year conditional on the achievement at the start of the academic year. As described in Section 3.5., we examine these issues using multivariate statistical models with the inclusion of school fixed effects. For models looking into school factors, we only employed multivariate analyses (excluding school fixed effects).

4.2.1. Key child characteristics

Regression results are presented in Table 7 and are interpreted as conditional associations on differences in child characteristics (the explanatory variables) on numeracy test scores (outcome variable) when holding all other variables constant. We first estimate key child characteristics that determine baseline numeracy test score (R1-B) (Column 1). We then estimate the same model for endline numeracy test scores (R1-E) (Column 2). The final model measures the key child characteristics associated with numeracy achievement at the end of the school year conditional on achievement at the start of school year. This final model is equivalent to value added and could be interpreted as progress over the academic year (results presented in Column 3).

As shown in Table 7, males scored significantly higher than females on both R1-B and R1-E numeracy tests (columns 1 and 2). When measuring achievement at the end of the academic year conditional on

achievement at the start of the academic year (for simplicity, progress in achievement)[12], we also find that males scored significantly higher than females, yet only significant at 10%. Taken together, these results indicate that boys do outperform girls in Grade 4 numeracy tests, both at the start and at the end of the school year. There is a slight indication also that boys made more progress, however this result is only significant at 10% level.

Being older, spending many hours per day studying at home, and speaking at least one language in addition to mother tongue are significantly associated with higher numeracy test scores both at baseline and endline. In addition, age, health and hours spent per day studying at home are significantly associated with progress in numeracy test scores over the school year. Another child variable associated with numeracy test scores at baseline shown in Table 7 is food consumption, but this is only significant at 10% level.

Nearly 40% of the Grade 4 sample reported to have attended preschool education. However, attending a pre-school education (whether it is 1-year or 3-years long duration) does not have any significant association with numeracy scores on both R1-B and R1-E tests. Attending preschool education was not significantly associated with progress in numeracy scores over the academic year neither. Hours spent per day at school and hours spent per day on domestic tasks were also not significantly associated with both baseline and endline numeracy test scores.

As shown in Table 7, the R-squared values for individual numeracy R1-B (3.3%) and R1-E (3.1%) models are low, which imply that key child characteristics explain relatively little of the variation in test scores within classes. However, the R-squared value increases to nearly 41% for the model where numeracy achievement at the end of the academic year is conditional on achievement at the start of the academic year. This indicates a high predictive power of test scores at the start of the academic year, with the estimated parameter indicating that a one standard deviation increase in baseline scores predicts 0.69 standard deviation in numeracy tests at the end of the academic year. As previously mentioned, we do not have a 1-to-1 correspondence between test scores at the start and end of the academic year, as assumption required for some value-added models.

4.2.2. Key household characteristics

Table 8 presents the association between numeracy test scores at both baseline and endline and progress in numeracy scores over the academic year with key household characteristics, conditional on child factors. Following from the previous section, we first estimate separately the model for baseline and endline numeracy achievement conditional on child and household factors (columns 1 and 2, respectively). We then estimate the progress in numeracy scores over the course of an academic year conditional on achievement at the start of the school year and key child and household factors (column 3).

[12] For simplicity, we will refer to the models that measure achievement at the end of the academic year conditional on achievement at the start of the academic year as "progress in numeracy test scores". We are careful here not to confuse the reader with estimations of value-added models as our dependent variable is not the change in test scores over one academic year.

Table 7. Key child characteristics and learning progress over an academic year

	(Model 1) R1-B Math	(Model 2) R1-E Math	(Model 3) R1-E Math|R1-B Math
Prior learning (R1-B Math)			0.687***
			(0.022)
Gender: Male	16.53***	15.98***	4.630*
	(3.695)	(3.974)	(2.863)
Age	4.713***	5.554***	2.318**
	(0.919)	(1.071)	(0.823)
Pre-school education (Base group: No preschool)			
1year	0.165	0.270	0.156
	(4.475)	(4.218)	(3.701)
2 years	-5.839	-6.219	-2.210
	(5.632)	(5.920)	(4.889)
3 years	-15.42*	-8.128	2.459
	(6.996)	(7.860)	(6.342)
Food Consumption Score (Base group: Below acceptable)			
Acceptable	0.212*	0.201	0.056
	(0.129)	(0.143)	(0.110)
Health (Base group: Poor/Average)			
Healthy	4.647	12.65*	9.459**
	(6.939)	(7.706)	(4.770)
Hours spent in school (per day)	1.352	2.697	1.768
	(1.935)	(2.045)	(1.706)
Hours spent studying at home (per day)	7.059***	7.765***	2.918**
	(1.942)	(2.020)	(1.466)
Hours spent on domestic tasks (per day)	0.388	-0.294	-0.561
	(1.560)	(1.627)	(1.172)
Speaks at least one language in addition to mother tongue (Base group: No)			
Yes	14.63**	12.79**	2.744
	(5.315)	(5.505)	(4.339)
Constant	403.6***	414.8***	137.7***
	(18.765)	(22.10)	(21.019)
Observations	3300	3300	3300
R^2	0.033	0.031	0.408

Source: RISE Ethiopia Quantitative Study.

Notes: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.001$; Robust standard errors in parentheses; The standard errors were clustered at the school level; A school 'fixed effects' approach was employed so that results compare students with their peers within each school. Parameters in Model 1 and Model 2 are comparable as they are measuring stocks; Parameters in Model 3 measure flows. To compare the magnitude of the parameters in Models 1 and 2 with Model 3, parameters in the latter must be transformed by *parameter/(1-prior learning parameter).*

Table 8. Regressions for key child and household characteristics and learning outcomes

	(Model 1) R1-B Math	(Model 2) R1-E Math	(Model 3) R1-E Math|R1-B Math
Prior learning (R1-B Math)			0.689***
			(0.021)
Gender: Male	17.31***	16.73***	4.804*
	(3.695)	(4.087)	(3.005)
Age	5.170***	6.077***	2.514**
	(0.950)	(1.129)	(0.870)
Pre-school education (Base group: No preschool)			
1year	-0.165	-0.464	-0.350
	(4.670)	(4.518)	(3.841)
2 years	-8.119	-7.786	-1.987
	(5.955)	(6.133)	(5.125)
3 years	-19.39**	-11.92	1.549
	(7.463)	(8.523)	(6.715)
Food Consumption Score (Base group: Below acceptable)			
Acceptable	0.101	0.0478	-0.0220
	(0.134)	(0.151)	(0.114)
Health (Base group: Poor/Average)			
Healthy	6.129	15.80**	11.58**
	(7.256)	(8.017)	(4.895)
Hours spent at school by a child per day	1.409	2.243	1.272
	(2.063)	(2.300)	(1.879)
Hours spent studying at home per day	6.662***	7.812***	3.221**
	(1.891)	(2.057)	(1.492)
Hours spent on domestic tasks per day	0.708	-0.661	-1.149
	(1.537)	(1.652)	(1.216)
Speaks at least one language in addition to mother tongue (No)			
Yes	13.84**	10.13*	0.595
	(5.387)	(5.924)	(4.749)
Household characteristics			
Primary caregiver's education (can't read at all)			
Able to read part or full sentence	2.011	2.712	1.326
	(3.812)	(4.127)	(3.306)
Wealth Index	36.49**	34.34	9.190
	(16.60)	(18.39)	(16.10)
Household help with child's homework (never or hardly ever)			
A few times a month	11.40**	6.563	-1.297
	(3.809)	(4.406)	(3.537)
Household's schooling aspiration for the child (Secondary/Post-secondary)			
Undergrad degree	2.341	-0.533	-2.146
	(4.585)	(5.308)	(3.806)
Postgrad degree	7.561	11.62*	6.409
	(6.088)	(7.264)	(5.440)
Household satisfaction with quality education at school (Dissatisfied)			
Satisfied	2.034	-6.189	-7.591
	(5.074)	(5.822)	(4.875)
Constant	380.3***	407.2***	145.1***
	(20.03)	(23.53)	(21.26)
Observations	3134	3134	3134
R^2	0.041	0.040	0.418

Source: RISE Ethiopia Quantitative Study.

Notes: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.001$; Robust standard errors in parentheses; The standard errors were clustered at the school level; A school 'fixed effects' approach was employed so that results compare students with their peers within each school; Parameters in Model 1 and Model 2 are comparable as they are measuring stocks. Parameters in Model 3 measure flows. To compare the magnitude of the parameters in Models 1 and 2 with Model 3, parameters in the latter must be transformed by *parameter/(1-prior learning parameter).*

The regression coefficients indicated that being a child from a relatively wealthy household is significantly associated with higher numeracy scores only at the start of the school year. We found no significant difference in numeracy learning progress over the school year between children from wealthier and poorer households even though household wealth is significantly associated with past numeracy scores (as evidenced by the strong correlation between household wealth and numeracy test scores at baseline). Moreover, household support with child's homework is significantly associated with higher numeracy test scores at the start of the school year, and having a parent with a higher schooling aspiration (postgraduate degree) is significantly associated with higher numeracy test scores at the end of the school year. However, none of the household indicators were associated with progress in numeracy scores over the 2018-19 academic year.

In these models, involving both the child and household factors, the R-squared values for each numeracy R1-B (4.1%) and R1-E (4.0%) scores are low, which again implies that a combination of key child and household characteristics explain relatively little of the variation in the baseline and endline numeracy test scores between students across schools. As before, the inclusion of past test scores as predictors of value-added increases the explanatory power of the model (41.8%) and the estimated parameter of 0.69 remains unchanged.

4.2.3. Key school and teacher characteristics

We run a separate regression model using ordinary least square approach to examine the extent to which key school factors determine students' end of year numeracy test scores and progress in numeracy scores over the course of the school year[13]. As shown in Table 9, we first estimate the model taking the endline numeracy achievement scores as an outcome variable conditional on key school, child and household factors (Model 1). We then estimate the progress in numeracy achievement over the course of the school year taking the end of school year numeracy test scores as an outcome variable, conditional on the start of year numeracy test scores and the same school, child and household factors as Model 1. These models do not use school fixed effects as our aim is to understand the school level factors which are related to numeracy test scores during one academic year. The school- and teacher-level factors included in our model are those that are related to the GEQIP-E reform in Ethiopia.

As shown in Table 9, principal's characteristics including years of experience as a school principal and frequency of supervision of teachers' lessons are significantly associated with higher levels of student numeracy scores at the end of the school year (Table 9, column 1). Teacher-related characteristics including higher frequency of participation in in-service teacher training and higher mathematics content knowledge are significantly associated with higher levels of student numeracy test scores at the end of the school year. School resources/facilities including the provision of one textbook per subject for each student and having a functional library are all significantly associated with higher student numeracy test scores at the end of the school year. In relation to the urban-rural location of the school, learning in a school located in urban areas is significantly associated with higher numeracy test scores at the end of the school year.

With respect to progress in numeracy scores over the course of the school year, we found that schools' provision of one textbook for one child, being located in urban areas, and higher teacher's mathematics content knowledge were all significantly associated with progress in numeracy scores over the course

[13] We decided to run the regression separately from the previous child and household factor regression models because we did not employ school 'fixed effects' approach in this model.

of the academic year (Table 9, Column 2). Having a functional resource centre in the school and a higher frequency of teachers' participation in in-service training are also associated with progress in numeracy test scores over the school year, but only significant at 10% level.

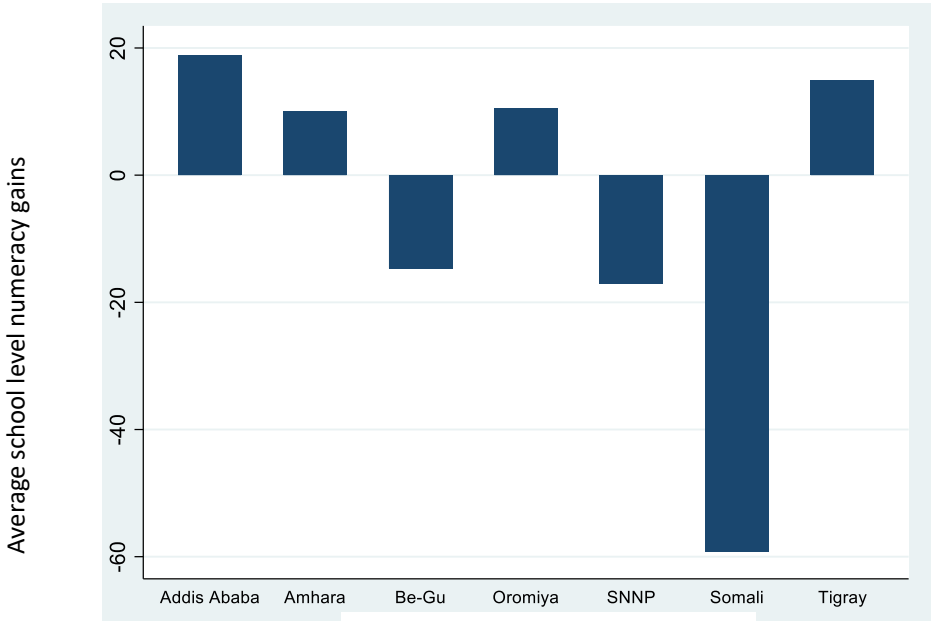
Table 9. Regressions for key school/teacher characteristics and learning outcomes

	(1) Math R1-E	(2) Math R1-E|R1-B
Baseline Math Score (R1-B)		0.723***
		(0.0152)
Principal's qualification (Diploma or below)		
University teaching degree	1.885	-1.095
	(4.289)	(3.244)
Experience as a principal	1.155**	0.446
	(0.372)	(0.282)
Time spent by principal supervising teachers' lessons (less frequently)		
Frequently	12.59**	3.777
	(4.422)	(3.347)
Student:Textbook ratio (1:2 or higher)		
1:1	40.91***	9.575**
	(5.349)	(4.128)
Resource Centre in the school (Yes)	7.768	7.541*
	(3.979)	(3.009)
School location (Rural)		
Urban	42.62***	16.69***
	(5.855)	(4.461)
Library (Yes)	14.05**	4.004
	(4.872)	(3.691)
Teachers' participation in in-service training (base group: never or almost never)		
A few times a year	4.284	3.819
	(5.078)	(3.840)
Once or more per month	15.16**	9.758*
	(5.577)	(4.219)
Teacher's mathematics content knowledge measured at endline	0.600***	0.443***
	(0.117)	(0.088)
Child and household factors	YES	YES
Constant	280.0***	69.70***
	(19.53)	(15.03)
Observations	3058	3058
R^2	0.191	0.537

Source: RISE Ethiopia Quantitative Study.
Notes: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.001$; Robust standard errors are in parentheses; The Child-level factors are gender, age, hours spent at school by a child per day, hours spent studying by a child per day, hours spent on domestic tasks by index child per day; The household-factors are primary caregiver's literacy, wealth index, and household help with child's homework; Parameters in Model 1 measure stocks. Parameters in Model 2 measure flows. To compare the magnitude of the parameters in Models 1 with Model 2, parameters in the latter must be transformed by *parameter/(1-prior learning parameter).*

To examine more deeply the extent to which school factors are associated with progress in numeracy scores, we also computed learning gains (as we did in Section 4.1). First, it is essential to note that Grade 4 students across regions and localities started (based on their R1-B scores) at different levels of mastery of the numeracy curriculum. It may, therefore, be more difficult to make larger progress in numeracy at the end of the school year for those students who already started at a high level of numeracy knowledge at the start of the academic year. We addressed this issue by predicting students' expected scores at the end of the academic year based on their baseline scores. A new score was then generated by subtracting expected scores from the actual R1-E numeracy scores. Figure 8 shows the differences between actual and expected scores on the R1-E test accounting for R1-B by region. The graph provides an estimate of the gains in numeracy over the course of the academic year across regions: the average over-performance or under-performance of students within a school relative to a school consisting of average students. Schools in Addis Ababa, as expected, made the highest numeracy progress followed by schools in Tigray, Oromia, and Amhara regions. Schools in Somali, Be-Gu and SNNP regions appeared to contribute very little to students' numeracy progress over the school year. Except for the SNNP region, the finding related to Be-Gu and Somali regions are not surprising as the two regions are homes to schools that are among the most historically disadvantaged. The present GEQIP-E reform, with its focus on equity is expected to improve the resources for the schools in the most disadvantaged regions. The finding in relation to SNNP region is unexpected because it is one of the so-called established regions in Ethiopia together with Amhara, Tigray and Oromia regions, which are assumed to be relatively better-resourced.

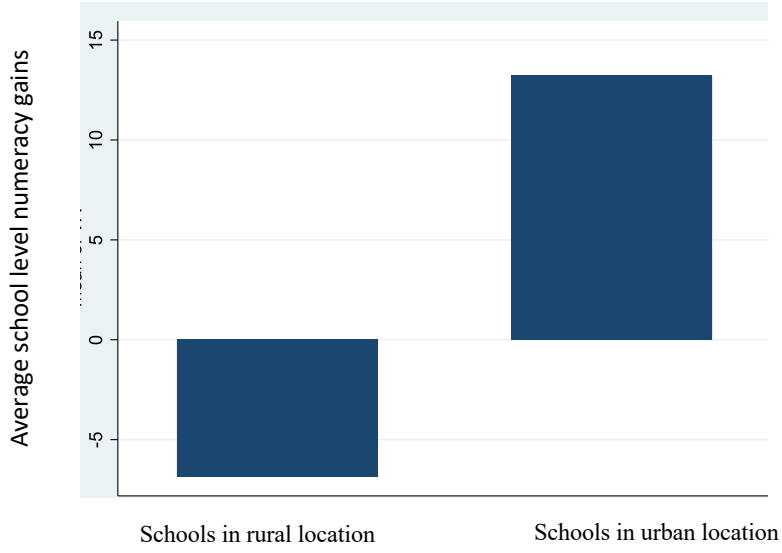
Figure 8. Student differences between actual and expected scores on the R1-E test based on R1-B, by region



Source: RISE Ethiopia Quantitative Study.
; Notes: Total number of schools per region: Addis Ababa=20, Amhara=25, Be-Gu=19 , Oromiya =41, SNNP =22, Somali=19, and Tigray=20.

Figure 9 presents an estimate of the numeracy score gains in schools over the course of the school year by urban-rural localities. Again, as expected, schools in urban locations contributed substantially higher learning gains than the schools in rural locations. This finding may suggest the prevalence of inequalities in primary school quality in Ethiopia between rural and urban locations.

Figure 9. Student differences between actual and expected scores on the R1-E test based on R1-B, by urban-rural location



Source: RISE Ethiopia Quantitative Study.
Notes: Total number of schools in rural locations=111 (67%), total number of schools in urban locations = 55 (33%).

5. Discussion and Conclusion

There are two major analyses conducted in the present study. One is exploring the associations between students' learning levels measured at the start and end of the academic year and key child, household and school factors associated with these learning levels. Most of the studies we found in the systematic review focused on this type of analysis. In other words, most evidence on the determinants of learning outcomes in Ethiopian primary schools from quantitative studies focuses on the average differences in achievement. Findings from our analysis are largely consistent with the overall quantitative evidence in Ethiopia (see Table 10). Child-related characteristics including gender (male), age, hours spent studying at home per a typical day, and speaking at least one more language in addition to mother tongue all were found to be significantly associated with average differences in numeracy achievement in grade 4 (both at the start and end of the academic year). The major mismatch is on preschool education attendance and numeracy scores. Previous studies consistently showed that attending preschool education is significantly associated with higher test scores in later classes. We did not find this to be the case for the RISE Ethiopia data using numeracy achievement for Grade 4 students in the 2018-19 academic year.

Household-related characteristics including household wealth and household help with child's homework/study were found to be significantly associated with baseline numeracy scores. The findings related to household wealth and child support in homework are consistent with the findings from our systematic review. However, unlike many of the previous studies, parental education was not found to

be significantly associated with numeracy scores in the present study. It should be noted that the empirical evidence in Ethiopia on the association between parental education and learning outcomes is inconclusive (e.g., Tesfay, 2012).

School- and teacher-related characteristics including principal's experience, time spent by the principal supervising teachers' lessons, student/textbook ratio, availability of library in schools, teachers' participation in in-service training, and teacher content knowledge were found to be significantly associated with numeracy test scores at the end of the academic year. Our finding related to principal's qualification did not match with the existing literature.

In addition, our study design and analytical approach enabled us to explore also progress in numeracy scores over a school year. Hence, we are able not just to estimate average differences in achievement but whether progress in numeracy scores in a school year is different for children living in different households or attending certain schools. As mentioned before, the quantitative evidence in the primary education system in Ethiopia on the extent to which key child, school, and household factors determine learning progress over time conditional on past test scores is scant, and we are unable to compare our findings with the literature (except for studies using the Young Lives school survey data in 2012-13 academic year). However, it is necessary to note that some of the child and school factors that we identified in the review to be significantly associated with learning levels were found to significantly determine progress in numeracy scores over an academic year in the present study. As shown in Table 10, child characteristics including age, health status, hours spent studying at home; school characteristics including urban location of a school, provision of one textbook per subject for each student, and teachers' mathematics content knowledge were found to be significantly associated with progress in numeracy scores over an academic year. None of the key household characteristics included in the present study was found to be significantly associated with progress in numeracy scores over an academic year.

Overall, a crucial advantage of the present study is that the inclusion of baseline test scores in the numeracy test scores progress estimation over an academic year has deepened our understanding of the determinants of learning gain in primary schools in Ethiopia. Average differences are important, but also whether progress is made in learning outcomes. It is important to highlight that we have not exhaustively included all the potential variables that may influence progress in numeracy scores into our models. We were driven by the systematic review, and thus more analysis is needed to engage more deeply with the determinants of learning to further explore the potential influence of GEQIP reform on learning outcomes of primary school children in Ethiopia.

Table 10. Summary of the key child, household and school factors found to be significantly associated with numeracy scores either at baseline or endline, and progress in numeracy scores over a school year

Child characteristics			**Household characteristics**			**School characteristics**		
Variables reported in the literature as positively & significantly associated with learning levels	*Does the finding from this study match concerning:*		*Variables reported in the literature as positively & significantly associated with learning levels*	*Does the finding from this study match concerning:*		*Variables reported in the literature as positively & significantly associated with learning levels*	*Does the finding from this study match concerning:*	
	Numeracy scores either at baseline or endline?	*Progress in numeracy scores over a school year?*		*Numeracy scores either at baseline or endline?*	*Progress in numeracy scores over a school year?*		*Numeracy scores either at endline?*	*Progress in numeracy scores over a school year?*
Sex	Yes	No	Wealth/Income	Yes	No	Principal qualification	No	No
Age	Yes	Yes	Parental education	No	No	Principal's experience	Yes	No
Preschool education	No	No	Child support in homework/ study	Yes	No	Time spent supervising teachers' lessons	Yes	No
Number of meals per day	No	No	Parent's schooling aspiration for the child	Yes	No	Student/Textbook ratio	Yes	Yes
Health status	No	Yes	Parent's satisfaction with quality education at school	No	No	Availability of a resource centre in the school	No	No
Hours spent at school per day	No	No				Availability of library	Yes	No
Hours spent studying per day	Yes	Yes				Urban location of a school	Yes	Yes
Hours spent on domestic tasks	No	No				Frequency of in-service training	Yes	No
Speaks at least one language in addition to mother tongue	Yes	No				Mathematics content knowledge	Yes	Yes

Source: RISE Ethiopia Quantitative Study

References

Abafita, J., & Kim, C. (2015). Determinants of children's schooling: The case of Tigray Region, Ethiopia. *Educational Research and Reviews*, *10*(8), 1130–1146. https://doi.org/10.5897/ERR2015.2095

Ahmed, A. Y., & Mihiretie, D. M. (2015). Primary school teachers and parents' views on automatic promotion practices and its implications for education quality. *International Journal of Educational Development*, *43*, 90–99. https://doi.org/https://doi.org/10.1016/j.ijedudev.2015.05.003

Asfaw, A. (2015). *Quality of primary education in Ethiopia: The case of early grade mathematics competency in Tigrai* [Doctoral Thesis. Addis Ababa University]. https://doi.org/10.1590/s1809-98232013000400007

Carter, E., Rose, P., Sabates, R., & Akyeampong, K. (2020). Trapped in low performance? Tracking the learning trajectory of disadvantaged girls and boys in the Complementary Basic Education programme in Ghana. *International Journal of Educational Research*, *100*(February). https://doi.org/10.1016/j.ijer.2020.101541

Dendir, S. (2014). Children's cognitive ability, schooling and work: Evidence from Ethiopia. *International Journal of Educational Development*, *38*, 22–36. https://doi.org/10.1016/j.ijedudev.2014.06.007

Eigbiremolen, G. O. (2017). Determinants of Learning among Primary School Children in Ethiopia: Analysis of Round 2 and 3 of Young Lives Data. *African Development Review*, *29*(2), 237–248. https://doi.org/10.1111/1467-8268.12253

Feinstein, L. (2003). Inequality in the early cognitive development of British children in the 1970 cohort. *Economica*, *70*(277), 73–97. https://doi.org/10.1111/1468-0335.t01-1-00272

GEQAEA. (2008a). *Ethiopian Third National Learning Assessment of Grade Eight Students. Addis Ababa. General Education Quality Assusrance & Examinations Agency*.

GEQAEA. (2008b). *Ethiopian Third National Learning Assessment of Grade Four Students. Addis Ababa. General Education Quality Assusrance & Examinations Agency.*

Glewwe, P., Krutikova, S., & Rolleston, C. (2017). Do schools reinforce or reduce learning gaps between advantaged and disadvantaged students? Evidence from Vietnam and Peru. *Economic Development and Cultural Change*, *65*(4), 699–739. https://doi.org/10.1086/691993

Hoddinott, J., Iyer, P., Sabates, R., & Woldehanna, T. (2019). *Evaluating Large-Scale Education Reforms in Ethiopia. RISE Working Paper* (19/034).

Hungi, N., Ngware, M., Mahuro, G., & Muhia, N. (2017). Learning barriers among Grade 6 pupils attending rural schools in Uganda: implications to policy and practice. *Educational Research for Policy and Practice*, *16*(2), 129–155. https://doi.org/10.1007/s10671-016-9199-2

Iyer, P., Rolleston, C., Rose, P., & Woldehanna, T. (2020). A rising tide of access : what consequences for equitable learning in Ethiopia ? *Oxford Review of Education*. https://doi.org/10.1080/03054985.2020.1741343

James, Z. (2018). *Language and Learning in Ethiopia*. Doctoral Thesis. University College Loondon.

Le Nestour, A., Moscoviz, L., & Sandefur, J. (2021). The long-term decline of school quality in the developing World. *Article Submitted to a Journal*, 1–60. https://doi.org/10.1007/978-94-017-8905-9_300161

Ministry of Education. (2010). *Education Sector Development Programme IV (ESDP IV): 2010/11 - 2014/15. Program Action Plan*.

Ministry of Education. (2015). *Education Sector Development Programme V (ESDP V): 2015/16 - 2019/20. Program Action Plan. Addis Ababa. Federal Ministry of Education.*

Ministry of Education. (2019). *Education Statistics Annual Abstract 2011 E.C. (2018/19). Addis Ababa. Federal Ministry of Education.*

NEAEA. (2013). *Ethiopian Fourth National Learning Assessment of Grades 4 and 8 Pupils. Addis Ababa. National Educational Assessment & Examinations Agency.*

NEAEA. (2016). *Ethiopian Fifth National Learning Assessment of Grades 4 and 8 Pupils. Addis Ababa. National Educational Assessment & Examinations Agency.*

NOE. (2000). *Ethiopian National Baseline Assessment on Grade Four Pupils' Achievement: Technical report. Addis Ababa. National Organization for Examinations.*

NOE. (2004a). *Ethiopian Second National Learning Assessment of Grade 4 Students. Addis Ababa. National Organization for Examinations.*
NOE. (2004b). *Ethiopian Second National Learning Assessment of Grade 8 Students. Addis Ababa. National Organization for Examinations.*
Raju, P. M., & Asfaw, A. (2009). Recalled test anxiety in relation to achievement, in the context of general academic self-concept, study habits, parental involvement and socio-economic status among Grade 6 Ethiopian students. *Education 3-13*, *37*(3), 269–285. https://doi.org/10.1080/03004270902734085
Rolleston, C., James, Z., Pasquier-Doumer, L., & Tam, T. N. T. M. (2013). *Making Progress. Report of the Young Lives School Survey in Vietnam. Working Paper 100*. Oxford: Young Lives.
Rose, P., & Al-Samarrai, S. (2001). Household Constraints on Schooling by Gender : Empirical Evidence from Ethiopia Published by : The University of Chicago Press on behalf of the Comparative and International Education Society Stable URL : http://www.jstor.org/stable/10.1086/447644 Your us. *Comparative Education Review*, *45*(1), 36–63.
Singh, A. (2014). *Emergence and Evolution of Learning Gaps across Countries. Working Paper 124*. Oxford: Young Lives.
Tesfay, N. (2012). *Horizontal inequalities in children's educational outcomes in Ethiopia* [Master's Thesis. Wolfson College]. https://doi.org/10.1016/j.ijedudev.2014.08.009
USAID. (2010). *Ethiopia Early Grade Reading Assessment. Data Analytic Report Language and Early Learning. Addis Ababa. USAID*.
USAID. (2019). *Early Grade Reading Assessment (EGRA) 2018 Endline Report. Addis Ababa. USAID*.
Woldehanna, T., Araya, M., & Gebremedhin, A. (2016). Assessing Children's Learning Outcomes: A Comparison of two cohorts from Young Lives Ethiopia. *The Ethiopian Journal of Education*, *XXXVI*(1), 149–187.
Woldehanna, T., & Hagos, A. (2015). Economic shocks and children's dropout from primary school: Implications for education policy in Ethiopia. *Africa Education Review*, *12*(1), 28–47. https://doi.org/10.1080/18146627.2015.1036548
World Bank. (2008). *General Education Quality Improvement Project (GEQIP-I). Project Appriasal Document. Report No.: 45140-ET*.
World Bank. (2013). *General Education Quality Improvement Project II (GEQIP-II). Project Appriasal Document*.
World Bank. (2017). *General Education Quality Improvement Program for Equity (GEQIP-E): Project Appraisal Document: Report No: 121294-ET*. The World Bank.