The Impact of Learning in Mother Tongue First: Evidence from a Natural Experiment in Ethiopia

Yared Seid[a]

[a]*International Growth Center, LSE, Addis Ababa, Ethiopia; e-mail: Y.Seid@lse.ac.uk*

January 2017

Abstract

This study explores the effect of learning in mother tongue first on students' academic achievement later after they transition to English instruction. Even if Ethiopia has adopted mother-tongue instruction in primary school, its states have discretion to choose when students transition to English instruction. This results in a variation in the timing of the transition to English instruction across states in Ethiopia. Southern Nations, Nationalities, and People's (SNNP) state, for instance, has legislated for students to transition to English instruction in grade 5 whereas students in other states in Ethiopia, except those in Gambella, do so either in grade 7 or 9. Due to the ethno-linguistic diversity of SNNP state, however, when students in the state progress from grade 4 to 5, the medium of instruction changes from mother tongue and second language to English for language-majority and language-minority students, respectively. This results in a variation in the intensity of the impact of the transition to English instruction by language group. Exploiting these two plausibly exogenous sources of variations (across state and language group) and using data from Young Lives' 2012-2013 Ethiopian school survey, we provide empirical evidence on the causal effect of learning in mother tongue first on students' academic achievement later after they transition to English instruction by estimating triple-differences model. The estimates from our preferred specification suggest that learning in mother tongue first (in grades $1-4$) increases students' mathematics and literacy tests scores later after they transition to English instruction (in grade 5) by 0.269 and 0.089 standard deviations, respectively.

Keywords: Medium of instruction, Primary school education, Triple-differences model, Ethiopia

JEL: I25, O10, O12

1. Introduction

A large number of countries in the developing world have made tremendous effort to make education more and more inclusive and accessible to historically marginalized groups. Adopting mother-tongue instruction in primary school has played a role in this regard as it motivates children from language-minority groups to attend school (Cummins, 1999). Though the adoption of mother-tongue instruction in primary school has increased enrollment in primary school and performance at school (e.g., Seid, 2016), it has a potential to limit students' labor market opportunities later in life as it makes students less proficient in both national and international languages (Angrist and Lavy, 1997). As a result, many governments in developing countries design their education language policies in such a way that students transition from mother-tongue to English (or other foreign-language) instruction after completing few years of primary schooling.

Consider Ethiopia, which is also the focus of the present study, as an example. Ethiopia has adopted mother-tongue instruction in primary school following the signing of the Education and Training Policy into law in 1994 (Ministry of Education, 1994). The same policy document, on the other hand, states that mother-tongue languages should be used as media of instruction up to only a certain grade, after which students have to transition to English instruction.[1] However, we have a limited understanding of the effect of learning in mother tongue first (relative to learning in non-English second language first) on students' academic achievement later after they transition to English instruction.

Exploring the effect of learning in mother tongue first on the performance of students from different language groups (i.e., those taught in their mother tongue first versus their peers taught in their second language first) is particularly important for a multilingual country like Ethiopia where, in some states,[2] a large number of ethnic groups live in close geographic proximity, but speak different languages. This implies that it is difficult to ensure that a great majority of students learn in their mother tongue as there are practical limitations on the number of languages that can feasibly be adopted as media of instruction, particularly in a resource-constrained country like

[1]The 1994 education reform also prescribes the introduction of teaching English as a subject starting from grade 1 in all schools in Ethiopia.

[2]Ethiopia is a federal country with three levels of government: federal, state (or regional), and local. Currently, the country has nine states and two chartered cities.

Ethiopia.

By 2007, about 25 languages were adopted as media of instruction in primary schools in Ethiopia (Seidel and Moritz, 2007), which is a huge improvement, particularly considering that Amharic[3] was the only medium of instruction in primary school in 1990. Comparing the number of languages used as media of instruction with the number of languages that are being spoken in the country,[4] however, reveals that a large number of students are still learning in their second language in primary school.

The challenge in ensuring that a great majority of students learn in their mother tongue is sever in states that are ethnically more diverse. In the most ethnically diverse state of Ethiopia, Southern Nations, Nationalities, and People's (SNNP) state,[5] for instance, it is estimated that about 56 languages are being spoken while only 13 languages have been adopted as media of instruction in the first cycle of primary education[6] by 2007 (Heugh et al., 2007). Partly due to the extraordinary ethno-linguistic diversity of people in SNNP state, the state, immediately after the 1994 Ethiopian education reform, has legislated that students have to transition to English instruction in grade 5.[7] This is contrary to the fact that Addis Ababa, Afar, Amhara, Oromiya, Somali, and Tgray states (hereafter other states in Ethiopia) have legislated that students have to transition to English instruction either in grade 7 or 9.

The fact that students in SNNP state transition to English instruction earlier (relative to their peers in other states in Ethiopia) makes SNNP state unique. However, it is reasonable to assume that there is a variation in the intensity of the impact of the transition to English instruction among grade 5 students in SNNP state who come from different language groups. This is primarily because language-majority students are taught in their mother tongue first (in grades $1-4$)[8] whereas their language-

[3]Amharic has been the only official language of the federal government of Ethiopia since the Ethiopian history has been recorded.

[4]It has been estimated that more than 90 languages are being spoken in Ethiopia (Bamgbose, 1991).

[5]The 2007 Ethiopian population census shows that SNNP state is the third largest state in Ethiopia in terms of population size, with a total population of about 14 million and accounts for about 19 percent of the population in Ethiopia.

[6]Primary education in Ethiopia covers 8 years of schooling which are equally divided into two: the first (i.e., grades $1-4$) and the second (i.e., grades $5-8$) cycles of primary education.

[7]Students in Gambella state also transition to English instruction in grade 5. However, we do not focus on Gambella state in this paper since the state is not surveyed in the Young Lives' 2012-2013 Ethiopian school survey which is used as the primary source of data in this paper.

[8]Hereafter, we use the word "first" in a phrase like "students are taught in their mother tongue

minority peers are taught in their second language first. That is, when students in SNNP state progress from grade 4 to grade 5, it is a transition from mother-tongue to English instruction for language-majority students while it is a transition from second-language to English instruction for language-minority students. This provides a natural experiment to explore whether learning in mother tongue first (relative to learning in second language first) help students learn better after they transition to English-instruction classrooms.

From a policy perspective, understanding whether the advantage that mother-tongue students enjoy in gains in academic achievement while learning in their mother tongue will carry over to later years in school after they transition to English instruction is crucial. The findings from this kind of study, for instance, highlight whether the choice of medium of instruction in primary school sets students from different language groups to different trajectories in their academic achievement in later years in school and, hence, in their labor market outcomes later in life.

Though there is a growing evidence on the role mother-tongue instruction plays in performance in primary school, we have a limited understanding on whether students taught in their mother tongue first learn in English (or other foreign languages) better than their peers taught in their second language first. A priori it is not clear whether students taught in their mother-tongue first will be at disadvantage later when they transition to English-instruction classrooms.

On the one hand, students taught in their mother tongue first seem to be better off since subject contents/concepts that are first learned through mother tongue can easily be transferred to English (Brock-Utne, 2007) as long as students have reached a certain threshold in their proficiency in English (Cummins, 2000). On the contrary, students taught in their second language first seem to be better off since they have had experience in learning in a language different from their mother tongue by the time they transition to English-instruction classrooms. Thus, it might be easier for these students to quickly adapt to English-instruction classrooms relative to their peers taught in their mother tongue first.

The literature on the role of English instruction in primary school focuses on comparing the relative effectiveness of bilingual education and English-immersion programs. The vast majority of prior studies on the topic come from the US, partly

(or second language) first" to refer to grades $1-4$.

because many primary schools in the US enroll a large number of immigrant students with limited English proficiency. It is typical for US primary schools to adopt either bilingual education or English-immersion program as a solution (Slavin et al., 2011).[9]

The findings from prior studies on the relative effectiveness of the two programs, however, are mixed, ranging from documenting no significant differences in the two programs (e.g., Rossell and Baker, 1996) to the superiority of bilingual education program (e.g., Cheung and Slavin, 2012; Slavin and Cheung, 2005). Some studies, on the other hand, has indicated that English-immersion program improves educational outcomes of students with limited English proficiency (e.g., Kuziemko, 2014). Even if prior studies differ in the spectrum of their findings on the relative effectiveness of bilingual education and English-immersion programs, the majority of the studies on the topic share a common feature: they consider students with limited English proficiency as homogeneous. In the US, the primary focus is on Spanish-speaking students even if a reasonable number of language-minority students from different language groups enroll in the US primary schools every year.

On the other hand, the literature on alternative language-of-instruction regimes from developing countries, especially those from Sub-Saharan Africa, is limited. However, it is not uncommon for students in developing countries to transition to English instruction after completing few years of primary schooling. The transition to English instruction is believed to be necessary to prepare students for further education since the medium of instruction in high school and college is, for the most part, English.

The limited studies from developing countries find out that mother-tongue instruction improves performance in primary school (e.g., Hynsjö and Damon, 2016; Piper et al., 2016; Seid, 2016) and mother-tongue instruction in early grades improves English acquisition later in grades $4-6$ (e.g., Taylor and von Fintel, 2016). However, we are not aware of studies that empirically document whether students taught in their mother tongue first learn in English better (than their peers taught in their non-English second language first) after they transition to English-instruction classrooms, except anecdotal evidence that suggests concepts that are first learned in mother tongue can be transferred to English (e.g., Brock-Utne, 2007).

The present study, therefore, builds on the literature on the role of mother-

[9] In bilingual education program, students with limited English proficiency learn in their native language first whereas in English-immersion program they are expected to learn in English from the beginning.

tongue instruction and attempts to fill this gap in the literature by exploring whether learning in mother tongue first improves students' academic achievement (measured by mathematics and literacy tests scores) later after they transition to English-instruction classrooms. It is worth mentioning that our econometric analysis takes into account, in fact exploits, students' heterogeneity. As discussed earlier, this is in contrast to the vast majority of studies from developed countries, particularly those from the US, that assume language-minority students who are expected to pass through bilingual/English-immersion programs are homogeneous. This paper, thus, also highlights the possibility that a transition to English instruction can have differential impact on academic achievement of students from different language groups.

Our identification strategy in this paper relies on two sources of plausibly exogenous variations. We first exploit the fact that students in SNNP state transition to English instruction in grade 5 whereas their peers in other states in Ethiopia do so when they progress to either grade 7 or 9. Hence, we consider SNNP state as *experimental* state and the other states in Ethiopia as *non-experimental* states.

The second source of variation comes from the difference in the intensity of the impact of the transition to English instruction across grade 5 students in SNNP state who come from different language groups, where language-majority students transition from mother-tongue to English instruction and language-minority students transition from second-language to English instruction. Thus, we assign students taught in their mother tongue first into *treated* group and students taught in their second languages first into *control* group. As part of our identification strategy, we further assign grades 4 and 5 students into *pre-treatment* and *after-treatment* groups, respectively, since English is used as the medium of instruction for only grade 5 students in SNNP state.

Using data from the Young Lives' 2012-2013 Ethiopian school survey, which administers mathematics and literacy tests to grades 4 and 5 students both at the beginning and end of the 2012-13 school year[10] and the assignment of states into experimental and non-experimental along with the assignment of students into treated, control, pre-, and after-treatment groups, we provide empirical evidence on the causal effect of learning in mother tongue first on students' academic achievement later after they transition to English instruction (in grade 5) by estimating triple-differences

[10] In Ethiopia, the school year begins in September and ends in June.

model for a sample of grades 4 and 5 students in Ethiopia.[11]

Estimates from our preferred specification suggest that learning in mother tongue first increases the performance of students later after they transition to English instruction (in grade 5). This is revealed by the coefficient estimates that suggest that learning in mother tongue first increases the standardized mathematics and literacy test scores of grade 5 students by 0.269 and 0.089 standard deviations, respectively. We also find that these effects are stronger for students in rural areas relative to those in urban areas. Falsification tests, on the other hand, suggest that our results are not confounded by other factors. This finding is consistent with the argument that, compared to their peers taught in their second language first, students taught in their mother tongue first learn in English better after they transition to English-instruction classrooms.

The remainder of the paper is organized as follows. The following section provides a brief background on schooling and language in Ethiopia. Section 3 describes the data, while Section 4 discusses the estimation strategy and presents the econometric results. The final section concludes the study.

2. Schooling and Language in Ethiopia

The Ethiopian education sector has gained the attention of the government since the change in government in May 1991. Among the many changes the sector has experienced in the 1990s, the most notable changes include restructuring the education system and adopting mother-tongue instruction in primary school following the signing of the Education and Training Policy into law in 1994.

Prior to the 1994 education reform, the education system consisted of six years of primary education (i.e., grades $1-6$) and two years of junior secondary education (i.e., grades $7-8$) where students seat for national school exit exams at the end of grades 6 and 8. After the 1994 education reform, primary education covers 8 years of schooling which are equally divided into two: the first (i.e., grades $1-4$) and second (i.e., grades $5-8$) cycles of primary education. The restructuring has abolished the national school exit exam that students used to take at the end of grade 6. Students,

[11] Since some researcher have estimated test scores production function using Ordinary Least Squares (OLS) model (e.g., Hynsjö and Damon, 2016), we also estimate test scores production function for a *subsample* of grade 5 students in SNNP state using OLS model so that it becomes relatively straightforward to compare our results with those from prior studies.

however, still have to seat for a national school exit exam at the end of grade 8.

In addition to restructuring the education system, the 1994 education reform has provided opportunity to states in Ethiopia to adopt as many languages as they choose as media of instruction in primary schools located in their jurisdictions (Ministry of Education, 1994). Following this discretion, states in Ethiopia have adopted mother-tongue instruction in primary schools, resulting in an increase in the number of languages used as media of instruction from using Amharic as the only medium of instruction in 1990 to about 25 languages by 2007 (Seidel and Moritz, 2007).[12] This, of course, does not necessarily guarantee that all primary-school students learn in their mother tongue. This is particularly because of the ethno-linguistic diversity of people in Ethiopia, and there are practical limitations on the number of languages that can feasibly be adopted as media of instruction, particularly in a resource-constrained country like Ethiopia.

Due to the variation in the extent of ethno-linguistic diversity across states in Ethiopia, the challenges of ensuring that language-minority students learn in their mother tongue in primary school are sever in relatively more diverse states. Consider SNNP state, the most diverse state in Ethiopia, as an example. In SNNP state, it is estimated that about 56 languages are being spoken within its geographic boundary whereas the state has adopted only 13 languages as media of instructions in primary school by 2007 (Heugh et al., 2007). This suggests that a large number of students in primary school in SNNP state have continued to learn in languages that are different from their mother tongue even long after the 1994 education reform.

Even though the 1994 education reform has provided states the opportunity to adopt mother-tongue instruction in primary school, it is important to note that it also mandates that mother-tongue languages should be used as media of instruction up to only a certain grade, after which students have to transition to English instruction. As a result of this specific mandate and partly due to the extraordinary ethno-linguistic diversity of people in SNNP state, the state has legislated that students have to transition to English instruction in grade 5.[13] This is contrary to the fact that other states in Ethiopia have legislated that students have to transition to

[12]See Seid (2016) for further discussion on the 1994 Ethiopian education reform and its effect on educational outcomes in primary school in Ethiopia.

[13]As mentioned earlier, Gambella state also mandates students to transition to English instruction in grade 5. In this paper, however, we do not focus on Gambella state since the state is not surveyed in Young Lives' 2012-2013 Ethiopia school survey which is the primary source of data in this paper.

English instruction either in grade 7 or 9.

Differences in the timing of the transition to English instruction between students in SNNP state and other states in Ethiopia provide a natural experiment to explore the causal effect of learning in mother-tongue first on the performance of students later after they transition to English instruction.

Transition to English Instruction as Exogenous Source of Variation

As discussed earlier, students in SNNP state transition to English instruction in grade 5 whereas students in other states in Ethiopia do so either in grade 7 or 9. This means that the medium of instruction changes for students in SNNP state when they progress from grade 4 to grade 5 whereas it continues to be the same for students in other states in Ethiopia when they progress from grade 4 to grade 5. Since grade 5 students in SNNP state transition to English instruction, which is a plausible exogenous shock,[14] we consider SNNP state as *experimental* state and the other states in Ethiopia as *non-experimental* states.

Though all students in SNNP state transition to English instruction in grade 5, there is a variation in the intensity of the impact of the transition to English instruction across students from different language groups. This is primarily because language-majority students are taught in their mother tongue first while their language-minority peers are taught in their second language. More specifically, when students in SNNP state progress from grade 4 to 5, the medium of instruction changes from mother tongue to English for language-majority students and from second language to English for language-minority students.

In order to capture the potential variation in the intensity of the impact of the transition to English instruction across students from different language groups, we assign language-majority students into *treated* group and language-minority students into *control* group. Note that this assignment of students into treated and control

[14] If there is endogenous migration (say, for instance, parents who value their kids' education more move to states to ensure that their kids are taught in their mother tongue first), then the coefficient estimates of the effect of learning in mother tongue first on students' performance later after they transition to English instruction will be biased upward. However, this is not a serious concern in our case since internal migration is not a common phenomenon in Ethiopia, particularly in states in our sample. Data from the 2007 Ethiopian population census, for instance, reveal that about 90 percent of the respondents have reported that they have lived in their current village/kebele (which is equivalent to urban neighborhood) for more than 10 years. Moreover, anecdotal evidence suggests that looking for better economic opportunities (rather than looking for better primary schools) is the primary reason for inter-state migration in Ethiopia.

groups is based on the medium of instruction they were exposed to while they were in grades $1-4$. Even if only language-majority students have received the "treatment" of learning in mother tongue first, it must be noted that the medium of instruction has changed for both group of students when they progress from grade 4 to 5. This suggests that it is reasonable to assume that the intensity of the impact of the transition to English instruction is different for students in the treated and control groups.

As part of our identification strategy, we further assign grades 4 and 5 students into *pre-treatment* and *after-treatment* groups, respectively, since English is used as the medium of instruction for only grade 5 students in SNNP state.

The assignment of states into experimental and non-experimental states along with the assignment of students into treated, control, pre-, and after-treatment groups enables us to identify the causal effect of learning in mother tongue first on the performance of students later after they transition to English instruction using triple-differences approach.

It is worth mentioning that the adoption of mother-tongue instruction in primary school as well as the legislation passed by SNNP state that mandates students in SNNP state to transition to English instruction in grade 5 were implemented immediately after the 1994 education reform. Since the data used in this paper[15] were collected long after the 1994 education reform, the possibility that the temporary disruption associated with implementing the 1994 education reform may bias our estimates is not a concern here.

It is well known that one of the identifying assumptions of the triple-differences approach is the absence of differential macroeconomic trends between SNNP state (i.e., experimental state) and other states in Ethiopia (i.e., non-experimental states) during the period of analysis. If this assumption is violated, the triple-differences estimates confound the effect of learning in mother tongue first with the effect of differential macroeconomic trends on students' performance that would have been observed even in the absence of the treatment, i.e., even when language-majority students are not taught in their mother tongue first.

In the literature, this concern is referred to as common/parallel trend assumption. This, however, does not create a major concern in the present study as the

[15]The data used in this paper come from the Young Lives' 2012-2013 Ethiopian school survey which administers mathematics and literacy tests to grades 4 and 5 students. See Section 3 for further discussion on the data used in this paper.

cross-section data used in this paper were collected in the 2012-2013 school year. Specifically, information was gathered on grades 4 and 5 students in Ethiopia in the 2012-2013 school year where, for the purpose of this study, we assign grades 4 and 5 students into *pre-treatment* and *after-treatment* groups, respectively. Since pre- and after-treatment groups are constructed using a cross-section data collected in the 2012-2013 school year, the bias due to potential differences in macroeconomic trends between experimental and non-experimental states is not a serious concern in our paper.

By the same token, we do not have to worry about potential biases of our estimates from the introduction of new education policy that could potentially affect schools in experimental and non-experimental states differently during the period of analysis. This is partly because, as discussed earlier, we use a cross-section data collected in a given school year and partly due to the fact that no new education policy was introduced in the 2012-2013 school year that could potentially affect schools in experimental and non-experimental states differently.

A related concern could be the fact that we do not observe the same set of students when they are in grades 4 and 5. Even if we do not observe the same set of students when they are in grades 4 and 5, the method employed in this paper can be considered as triple-differences as long as the mean baseline response of grade 4 students in experimental and non-experimental states is the same as the mean baseline response of grade 5 students in experimental and non-experimental states. If the experimental and non-experimental states attract the "same type" of grades 4 and 5 students, then the estimates from our triple-differences model capture the causal effect of learning in mother tongue first on the performance of students later after they transition to English-instruction classrooms (see Lee and Kang, 2006, for furher discussion on using cross-section data in the difference-in-differences approach).

If, however, the treatment (i.e., the fact that language-majority students learn in mother tongue in grades $1-4$) is substantial enough to change the composition of grades 4 and 5 students in the experimental and non-experimental states, then our results from the triple-differences model will be biased. Since there is no evidence that suggests that the treatment effect has substantially altered the composition of grades 4 and 5 students in the experimental and non-experimental states in Ethiopia, it is appropriate to employ triple-differences approach here – see Section 4 for further detail on the identification strategy used in this paper.

3. Data

The data used in this paper come from Ethiopian school survey which was administered by Young Lives (YL), an international research project based in the University of Oxford. In its household surveys, YL has collected data on children from four low income countries – Ethiopia, India (in the Andhra Pradesh state), Peru, and Vietnam. During the first household survey round of data collection in 2002, two thousand one-year-old children (hereafter "younger" cohort) and one thousand eight-years-old children (hereafter "older" cohort) were surveyed in each country. In follow-up surveys conducted in 2006, 2009, and 2013 the same children were tracked and surveyed when the younger cohort children turned to five, eight, and twelve years old, and the older cohort children turned to twelve, fifteen, and nineteen years old, respectively.

In the Ethiopian part of the survey, children were randomly sampled from 20 semi-purposively selected sentinel sites in the largest five states of the country (see Wilson et al., 2006, for a discussion on the sampling design). In addition to the longitudinal household surveys, YL conducted a school survey in Ethiopia in 2012-2013 school year – which is used as a primary source of data in this paper.

In the school survey, information was gathered on YL's younger cohort children and their peers who were in grades 4 and 5 in 2012-2013 school year. The survey was conducted in two 'Waves,' at the beginning of the school year (i.e., Wave 1) and towards the end of the same school year (i.e., Wave 2). In Wave 1, YL's younger cohort children and their non-YL's peers who attend schools located in YL's sentinel sites were surveyed. In Wave 2, a follow-up survey was administered on all children who have been surveyed in Wave 1.

A total of 11,982 students (of whom 493 are YL's younger cohort children[16]) in 94 schools who were in grades 4 and 5 in 2012-2013 school year were surveyed in Wave 1. Students who were surveyed in Wave 1 but were not present in school on days where the survey fieldwork of Wave 2 was conducted was dropped from Wave 2, resulting in a total of 10,030 students surveyed in both waves.

As mentioned earlier, five states (i.e., Addis Ababa, Amhara, Oromiya, SNNP,

[16] Only a fraction of YL's younger cohort children surveyed in the longitudinal household surveys were also surveyed in the school survey. This is mainly because some of these children attend schools located outside of the YL's sentinel sites and some other YL's younger cohort children were attending neither grade 4 nor grade 5 in the 2012-2013 school year.

and Tigray states) were surveyed in YL's Ethiopian household surveys. Similarly, YL has surveyed the same five states and additional two states (i.e., Afar and Somali states) in its 2012-2013 Ethiopian school survey. In this paper, however, we drop observations from Addis Ababa from the sample of analysis. This is mainly because Addis Ababa is uniquely heterogeneous where people from almost all ethnic groups live together. Moreover, the majority of its residents speak Amharic well,[17] implying that it is not reasonable to group students in Addis Ababa by language group and exploit this grouping in the identification strategy as suggested in this paper.

Moreover, our sample of analysis is restricted to students who attend public schools. This restriction is mainly because the 1994 education reform, which has introduced mother-tongue instruction in primary schools in Ethiopia, applies only to public schools. This restriction should not cause a serious concern in our paper since data from the YL's 2012-2013 Ethiopian school survey show that a great majority (about 89 percent) of students in Ethiopia attend public schools. On the other hand, it is important to note that public schools in Ethiopia generally perform poorly (relative to private schools) and, hence, their students receive low quality education and face weak labor market prospects later in life. Therefore, focusing on public schools and exploring factors that improve the quality of education in public schools in Ethiopia is an important contribution.

Finally, we further restrict the sample of analysis to students who have attended the same school since grade 1, which constitutes about 91 percent of students in YL's data. This sample restriction is imposed because, for students who have changed schools, we do not observe the media of instruction they were exposed to in their former schools. If students who have changed schools have been exposed to different media of instruction in their former and current schools, including these students in the sample of analysis may bias our estimates. For instance, if some students have changed their schools to ensure that they are taught in their mother tongue, then this will bias our estimates upward. Since we do not observe why students in our data change their schools, it is not possible to rule out the possibility that these students self-select themselves into schools that could teach them in their mother tongue. Thus, we have dropped students who have not attended the same school

[17] This, along with the fact that Addis Ababa is the seat of the federal government, explains why Addis Ababa has adopted Amharic as its official language as well as the medium of instruction in primary schools located within its boundary.

since grade 1. However, it is fair to say that dropping these students from our sample of analysis should not cause a serious concern in our paper since our data do not show any systematic difference in the fraction of students who have changed schools in experimental and non-experimental states.[18]

These sample restrictions leave us with a final sample size of 3,197 grade 4 students and 3,057 grade 5 students in 167 classrooms across 65 schools in Ethiopia.

The outcome variables used in this paper are students' test scores on mathematics and literacy tests that were administered to grades 4 and 5 students in the second wave of the 2012-2013 school year. We standardize the test scores to Z-scores with mean zero and standard deviation of one.[19] A key feature of the design of the tests was assessments of students' competency in mathematics and reading comprehension, which is directly linked to Ethiopia's Ministry of Education's Minimum Learning Competencies (MLCs).

The mathematics and literacy tests each consists of 25 questions where the mathematics test includes items that assess number operations, measurement, and so on, to the extent that these were reflected in the Ministry of Education's MLCs. The literacy test, on the other hand, focuses on reading comprehension, with sections in matching words and pictures, sentences and pictures, asking children to fill in the blanks with the correct word, and reading comprehension passage.

Table 1 presents the descriptive statistics for a sample of students used in the econometric analysis. The table shows that there are no striking differences (along a range of student-, school-, and household-level characteristics) by language group and the state's experimental status. However, Table 1 confirms our a priori expectation that mother-tongue students (i.e., those taught in their mother tongue in grades $1-4$) perform better both in mathematics and literacy tests relative to their peers taught in their second language.[20] In the next section, we assess whether learning in mother tongue first has played any role in this difference in tests scores.

[18]Our data shows that about 7 and 10 percent of students in experimental and non-experimental states have, respectively, changed schools, but this difference is not statistically significant.

[19]For the purpose of this paper, we record a blank answer as an incorrect answer.

[20]See A.1 in Appendix A for descriptive statistics of students' tests scores by experimental state, language group, and grade.

Table 1: Summary Statistics of Variables used in the Econometric Analysis by Experimental State and Language Group

	Experimental State		Non-experimental States	
	MT‡ students	Non-MT students	MT students	Non-MT students
Math Z-score - Wave 2	0.193	-0.231	0.093	-0.276
	(0.946)	(0.976)	(0.997)	(0.928)
Math Z-score - Wave 1	0.136	-0.219	0.040	-0.254
	(0.918)	(0.996)	(1.003)	(0.986)
Literacy Z-score - Wave 2	0.312	-0.062	0.291	0.105
	(0.899)	(0.947)	(0.906)	(0.829)
Literacy Z-score - Wave 1	0.277	-0.058	0.208	-0.092
	(0.251)	(0.240)	(0.203)	(0.281)
Student's age (in years)	11.689	11.713	11.209	12.265
	(1.900)	(1.848)	(1.576)	(2.769)
Dummy for female student	0.501	0.524	0.498	0.505
	(0.500)	(0.498)	(0.500)	(0.500)
Dummy for student's preschool attendance	0.479	0.439	0.458	0.453
	(0.500)	(0.481)	(0.479)	(0.498)
Dummy for grade repetition	0.266	0.249	0.287	0.224
	(0.342)	(0.327)	(0.319)	(0.392)
Dummy for student's participation in paid work	0.222	0.193	0.179	0.240
	(0.445)	(0.351)	(0.366)	(0.439)
Dummy for female math teacher	0.341	0.359	0.386	0.371
	(0.355)	(0.400)	(0.387)	(0.315)
Dummy for female literacy teacher	0.448	0.444	0.412	0.423
	(0.309)	(0.397)	(0.370)	(0.327)
Dummy for post-secondary educ - math teacher	0.668	0.686	0.643	0.653
	(0.496)	(0.466)	(0.479)	(0.404)
Dummy for post-secondary educ - literacy teacher	0.561	0.574	0.631	0.610
	(0.500)	(0.497)	(0.483)	(0.487)
Years of experience - math teacher	12.646	12.161	11.026	11.698
	(3.893)	(3.019)	(3.935)	(3.390)

Years of experience - literacy teacher	11.774	11.255	11.877	11.080
	(3.134)	(3.432)	(3.901)	(3.140)
Dummy for post secondary educ - principal	0.310	0.328	0.292	0.291
	(0.313)	(0.359)	(0.406)	(0.343)
Years of experience - principal	3.596	2.888	3.162	3.513
	(1.886)	(1.563)	(2.140)	(2.063)
Dummy for the school has library	0.525	0.514	0.483	0.464
	(0.380)	(0.301)	(0.365)	(0.282)
Dummy for the school has access to electricity	0.518	0.514	0.488	0.474
	(0.386)	(0.401)	(0.363)	(0.369)
Dummy for the school is in urban area	0.404	0.410	0.388	0.404
	(0.357)	(0.401)	(0.400)	(0.395)
Number of siblings	4.124	4.502	4.067	4.817
	(2.125)	(2.079)	(1.981)	(2.143)
Dummy for literate mother	0.365	0.348	0.317	0.323
	(0.400)	(0.386)	(0.395)	(0.378)
Dummy for literate father	0.492	0.482	0.471	0.484
	(0.392)	(0.394)	(0.389)	(0.398)
Number of rooms in the household	2.199	2.291	2.538	2.758
	(1.265)	(1.394)	(1.283)	(1.396)
Dummy for household has access to electricity	0.541	0.534	0.469	0.457
	(0.497)	(0.477)	(0.499)	(0.497)
Number of meals a student eats per day	2.656	2.484	2.778	2.742
	(0.606)	(0.648)	(0.504)	(0.573)
Observations	1642	698	3122	792

Notes: Standard deviations are reported in parentheses.

‡ MT and non-MT denote language-majority and language-minority students, respectively, where the former are taught in their mother tongue in grades $1-4$ whereas the later are taught in their second language in grades $1-4$.

4. The Effect of Learning in Mother Tongue First

Our primary objective in this paper is estimating the causal effect of learning in mother tongue first (in grades $1-4$) on academic achievement later after the student transition to English instruction (in grade 5). The unique nature of the Ethiopian education system and data from the YL's Ethiopian school survey, which administers mathematics and literacy tests to grades 4 and 5 students in the 2012-2013 school year, help us identify the causal effect of learning in mother tongue first on students' performance later after they transition to English instruction. In this section, we discuss the empirical strategy, present the results from the econometric analysis, run falsification tests, and, finally, explore whether there is heterogeneity in treatment effect.

4.1. Empirical Strategy

As discussed earlier, we exploit two sources of plausibly exogenous variations. First, we employ the fact that students in SNNP state transition to English instruction in grade 5 while their peers in other states in our sample do not transition to English instruction until they progress to either grade 7 or 9. Second, we exploit the variation in the intensity of the impact of the transition to English instruction across grade 5 students in SNNP state who come from different language groups.

Using these two sources of variations and data from the YL's Ethiopian school survey, we estimate the following triple-differences model:

$$\begin{aligned} Score_{igs} = {} & \beta_0 + \beta_1 Exp_s + \beta_2 After_{igs} + \beta_3 Treated_{igs} \\ & + \beta_4 (Exp_s * After_{igs}) + \beta_5 (Exp_s * Treated_{igs}) + \beta_6 (After_{igs} * Treated_{igs}) \\ & + \beta_7 (Exp_s * After_{igs} * Treated_{igs}) \\ & + \beta_8 Score_{igs\_lagged} + \mathbf{X}_{igs} \boldsymbol{\gamma} + \phi + \psi + \epsilon_{igs}, \end{aligned} \tag{1}$$

where i, g, and s index individuals, grades, and states, respectively. $Score$ is a student's contemporaneous standardized test scores (in mathematics and literacy tests) that were administered at the end of the 2012-2013 school year; Exp is a binary indicator for whether a state is experimental (i.e., equals 1 for SNNP state, and 0 for other states in our sample); $After$ is a binary indicator for whether a student is in grade 5 (i.e., equals 1 for grade 5 students, and 0 for grade 4 students); $Treated$ is a binary indicator for whether a student is taught in her/his mother tongue in grades $1-4$;[21] $Score_{igs\_lagged}$ is a student's lagged (or baseline) test scores (i.e., scores from

[21] As discussed earlier, we assign students into treated and control groups based on the medium of instruction they were exposed to while they were in grades $1-4$. Even if only language-majority students have received the "treatment" of learning in mother tongue first, it must be noted that the medium of instruction changes for both group of students when they progress from grade 4

tests that were administered at the beginning of the 2012-2013 school year); $\mathbf{X}_{igs}$ is a vector of contemporaneous control variables; ϕ and ψ are class and school fixed effects, respectively; and ϵ_{igs} is idiosyncratic error term.

The other fixed effects control for the characteristics of the experimental state (β_1), changes in test scores as students progress from grade 4 to grade 5 (β_2), and the characteristics of the treated group (β_3). The second-level interactions control for changes in the experimental state as students progress from grade 4 to grade 5 (β_4), characteristics of the treated group in the experimental state (β_5), and changes as students progress from grade 4 to grade 5 for the treated group nationwide (β_6). The third-level interaction (β_7) captures all variations in test score specific to grade 5 students (relative to grade 4 students) who were taught in their mother tongue in grades $1-4$ (relative to students taught in their second language in grades $1-4$) in the experimental state (relative to the non-experimental state). This is a triple-differences estimate of the causal effect of learning in mother tongue first on academic achievement later after the student transition to English instruction (in grade 5).

It must be noted that Equation 1 estimates the education production function. Prior studies that estimate test scores employ education production function by using either contemporaneous, cumulative, or value-added specifications (see Todd and Wolpin, 2003, for a survey of the literature on estimating production functions for cognitive achievement). In this paper, however, we employ the value-added specification which is proved to be less restrictive compared to the other two specifications.

If there are any systematic differences in students' innate abilities and other unobservable characteristics by language group that may not be differenced out by estimating the triple-differences equation, the coefficient estimates of the third-level interaction (β_7) will be biased. This is not, however, a serious concern in our triple-differences model since we expect any potential systematic differences in students' abilities and other unobservable characteristics by language group to be constant across states and grade levels, which the triple-differences approach is proved to effectively control for. However, as additional layer of protection and robustness check, we control for lagged test scores ($Score_{igs\_lagged}$) in the triple-differences model to ensure that our results are not derived by any systematic potential differences in

to 5. That is, it changes from mother tongue to English for language-majority students and from second language to English for language-minority students. Thus, it is reasonable to assume that the intensity of the impact of the transition to English instruction is different for students in the treated and control groups.

students' innate abilities and other unobservable characteristics by language group. Controlling for lagged test scores help mitigate potential biases due to omitted information on historical education inputs and family- and student-level characteristics (including students' innate abilities).[22] This is a common estimation technique that assumes a lagged/baseline test score provides a sufficient statistics for all historical inputs and student's ability (Todd and Wolpin, 2003).[23]

Finally, it is worth mentioning that the lagged/baseline and contemporaneous tests were administered at the beginning and end of the same school year, respectively. Thus, it is less likely for parents and, to some extent, schools to adjust their behavior after observing the student's baseline test score as the school year was already in session by the time the baseline test scores became observable. This mitigates possible biases from potential endogeneity of lagged/baseline test scores.

4.2. Results

As discussed earlier, some researchers have estimated the production function for cognitive development using OLS model arguing that the estimates would not be biased if the lagged test scores are controlled for in the OLS regression. Thus, we first present the results from the OLS regression of the test score production function in columns $1-3$ of Table 2. Then, we present the results from the triple-differences model in columns $4-6$ of Table 2. In all columns of Table 2, the upper and lower panels present results from both the OLS and triple-differences estimations where the dependent variables are standardized mathematics test scores in the upper panel and

[22]Controlling for lagged test scores also ensures that Equations 1 and 2 are value-added specifications without setting the parameter on lagged test scores to one.

[23]Following this argument by Todd and Wolpin (2003), some researchers have estimated test scores production function using OLS model (e.g., Hynsjö and Damon, 2016). To make our results somehow comparable to results from prior studies, we also estimate test score production function for a *subsample* of grade 5 students in SNNP state using the following OLS model:

$$Score_i = \beta_0 + \beta_1 MotherTongue_{i\_base} + \beta_2 Score_{i\_lagged} + \mathbf{X}_i \boldsymbol{\gamma} + \phi + \psi + \epsilon_i, \tag{2}$$

where $Score_i$ is a student's contemporaneous standardized test scores (in mathematics and literacy tests) that were administered at the end of the 2012-2013 school year; $MotherTongue_{i\_base}$ is a binary indicator for whether a student was taught in her/his mother tongue first (i.e., equals 1 for students taught in their mother tongue in grades $1-4$, and 0 for students taught in their second language in grades $1-4$); $Score_{i\_lagged}$ is a student's lagged (or baseline) test scores (i.e., scores from mathematics and literacy tests that were administered at the beginning of the 2012-2013 school year); $\mathbf{X}_i$ is a vector of contemporaneous control variables; ϕ and ψ are class and school fixed effects, respectively; and ϵ_i is idiosyncratic error term.

literacy test scores in the lower panel.[24]

Impact on Mathematics Test Scores

In column 1 of the upper panel of Table 2, the coefficient estimate of the binary indicator for whether grade 5 students in SNNP state were taught in their mother tongue in grades $1-4$ (i.e., the coefficient estimate of "Dummy for learning in MT first" variable[25]) is positive and statistically significant at 1 percent level, suggesting differences in performance between students taught in their mother tongue first and their peers taught in their second language first. That is, students taught in their mother tongue first perform better in mathematics test which was administered after they transition to English instruction (i.e., at the end of grade 5) relative to their peers taught in their second language first.

The results from the OLS regression presented here can be interpreted as the causal effect of learning in mother tongue first on students' achievement later after they transition to English instruction under the assumption that students taught in their mother tongue first and those taught in their second language first are not inherently different. If this assumption is not satisfied, the results from the OLS regression presented here cannot be interpreted as a "true" causal effect of learning in mother tongue first. If, for instance, parents who value education more move to states that teach their kids in their mother tongue first, then the coefficient estimates of the effect of learning in mother tongue first will be biased upward.

However, this is not a serious concern in our case since internal migration is not a common phenomenon in Ethiopia, particularly in SNNP state from where the data used to estimate the OLS model is obtained. Data from the 2007 Ethiopian population census, for instance, reveal that only 7 percent of the population in SNNP state have migrated to SNNP state in the previous 10 years. Moreover, as mentioned earlier, anecdotal evidence suggests that looking for better economic opportunities (rather than looking for better primary schools) is the primary reason for inter-state migration in Ethiopia.

Similarly, if there are systematic differences in students' innate abilities and unobservable characteristics by language group, this will bias our OLS estimates.

[24]The complete regression results from models presented in Table 2 are reported in the online supplementary appendix.

[25]In our sample, about 66 percent of grade 5 students in SNNP state was taught in their mother tongue first.

Table 2: Estimates of the Effect of Learning in Mother Tongue First on Students' Performance Later after they Transition to English Instruction

	OLS			Triple Differences		
	(1)	(2)	(3)	(4)	(5)	(6)
Math Z-score Equation						
Dummy for learning in MT‡ first	0.412***	0.326***	0.310***			
	(0.106)	(0.101)	(0.093)			
$Exp * After * Treated$				0.299***	0.286***	0.269***
				(0.037)	(0.029)	(0.036)
Lagged math Z-score		0.393***	0.381***		0.021*	0.016*
		(0.019)	(0.013)		(0.011)	(0.009)
Student- & school-level controls	Yes	Yes	Yes	Yes	Yes	Yes
Household-level controls	No	No	Yes	No	No	Yes
Observations	1136	1136	1136	6254	6254	6254
R-squared	0.611	0.624	0.665	0.505	0.591	0.627
Literacy Z-score Equation						
Dummy for learning in MT‡ first	0.221***	0.164*	0.117*			
	(0.061)	(0.095)	(0.062)			
$Exp * After * Treated$				0.088***	0.082***	0.089***
				(0.017)	(0.013)	(0.015)
Lagged literacy Z-score		0.215***	0.193***		0.009**	0.011
		(0.026)	(0.033)		(0.004)	(0.017)
Student- & school-level controls	Yes	Yes	Yes	Yes	Yes	Yes
Household-level controls	No	No	Yes	No	No	Yes
Observations	1136	1136	1136	6254	6254	6254
R-squared	0.425	0.508	0.448	0.623	0.653	0.721

Notes: *p < 0.10, ** p < 0.05, *** p < 0.01.
‡ MT denotes mother tongue.
Robust standard errors are reported in parentheses. In the triple-differences regressions, we control for dummies for whether the state is experimental, the student is in grade 5, and the student is in treated group. Besides, we control for the second-level interactions of the above three dummy variables. The additional regression controls (in both OLS and triple-differences models) are student-level characteristics (i.e., student's age in years and binary indicators for whether the student is female, has attended preschool, has ever repeated grade, and has participated in paid work), school-level characteristics (i.e., years of experience of mathematics/language teacher and the principal; binary indicators for whether the mathematics/language teacher and the principal have post-secondary education and whether the mathematics/language teacher is female; binary indicators for whether the school has library, access to electricity, and is found in urban area), household-level characteristics (i.e., binary indicators for whether the student's mother and father are literate, binary indicator for whether the household has access to electricity, number of rooms in the household, number of siblings, and the number of meals the student eats in a typical day), and classroom and school fixed effects.

Since it is difficult to observe and measure students' innate abilities, we mitigate the possible systematic differences (if there are any) in students' abilities and other unobservable characteristics by language group in two ways. First, we control for lagged test scores, which, in our case, is students' test scores from the test administered at the beginning of the 2012-2013 school year. As mentioned earlier, this is a common estimation technique that assumes a lagged/baseline test score provides a sufficient statistics for all historical inputs and the student's ability (Todd and Wolpin, 2003). The results from this approach are presented in columns 2 and 3 of Table 2.

Second, we employ triple-differences approach which is proved to control for any time-invariant differences by language group (including differences in abilities) and state. The results from the triple-differences approach are presented in columns $4-6$ of Table 2.

Let us first discuss the results from the OLS regression that controls for lagged mathematics test scores (see column 2 of the upper panel of Table 2). As can be seen from the table, controlling for lagged mathematics test scores in the OLS regression does not affect the sign and significance level of the coefficient estimate of the "Dummy for learning in MT first" variable. This, however, has substantially decreased its magnitude, suggesting that naive OLS regression, as expected, may bias the effect of learning in mother tongue first upward.

One of the major challenges of estimating production functions for cognitive achievement is its heavy data requirement where, in an ideal world, requires information on both historical and contemporaneous school-, family- and student-level characteristics, including students' innate abilities. However, existing datasets are deficient in one or more of these domains since the majority of the surveys are exclusively focusing on gathering extensive information either on school or family characteristics. As in many other school surveys, the YL's Ethiopian school survey, which is the data source for the present study, primarily focuses on collecting information on school inputs. However, the YL's Ethiopian school survey collects limited information on family characteristics by directly asking students about the demographic characteristics of their family.[26]

[26]Even if YL has also administered separate longitudinal household surveys that can provide a wide range of information on household-level characteristics in Ethiopia, we decided not to use information gathered in the YL's household survey. This is because only 493 students were surveyed both in the YL's Ethiopian household and school surveys, implying the small sample size does not allow precise estimation.

Exploiting the limited household demographic information gathered in the YL's Ethiopian school survey, we present results from the OLS regression that further controls for selected household-level characteristics in column 3 of Table 2. Again, the results reported in column 3 of the upper panel of Table 2 confirms what has been documented in columns 1 and 2 of the upper panel of Table 2 since the coefficient estimate for "Dummy for learning in MT first" is still positive and statistically significant at 1 percent level. To be exact, the results reported in column 3 of the upper panel of Table 2, which controls for lagged test scores, student-, school-, and household-level characteristics, suggest that learning in mother tongue first increases mathematics test scores later after students transition to English instruction (in grade 5) by 0.310 standard deviations.

So far, we have documented that the estimates from the OLS regressions suggest that learning in mother tongue first improves students' mathematics test scores after they transition to English-instruction classrooms. Now, we explore whether the estimates from the triple-differences model – which are presented in columns $4 - 6$ of the upper panel of Table 2 – confirm this conclusion. As in the case of the OLS regressions, the triple-differences estimation presented in column 4 of Table 2 controls for student- and school-level characteristics while the estimation reported in column 5 includes lagged mathematics test scores as additional control. Similarly, the last column of Table 2 further controls for selected household-level characteristics.

In the triple-differences model, which is summarized in Equation 1, the primary (explanatory) variable of interest is the third-level interaction, i.e., "$Exp * After * Treated$," where β_7 captures the effect of learning in mother tongue first on academic achievement (measured by mathematics and literacy tests scores) later after students transition to English instruction (in grade 5).

The results reported in columns $4 - 6$ of the upper panel of Table 2 show that the coefficient estimates of the treatment effect, i.e., "$Exp * After * Treated$" variable, are uniformly positive and significant. This suggests that students taught in their mother tongue first perform better (than their peers taught in their second language first) after they transition to English-instruction classrooms.

Comparing the coefficient estimates of the treatment effect reported in columns 4 and 5 of the upper panel of Table 2 shows that controlling for lagged (mathematics) test scores does not affect the coefficient estimates of the treatment effect, except slightly decreasing its magnitude. This is contrary to what we have seen in the OLS regressions reported in columns 1 and 2 of the upper panel of Table 2, where controlling for lagged (mathematics) test scores has substantially decreased the magnitude of the coefficient estimate of the treatment effect. This implies that the estimates of the treatment effect from the triple-differences model are, unlike those from the OLS regression, less likely to be biased from omitted information on historical education inputs and family- and student-level characteristics (including students' innate abilities).

On the other hand, controlling for basic household demographic characteristics in column 6 of the upper panel of Table 2 does not change the sign and significance level of the treatment effect, suggesting that learning in mother tongue first improves the performance of students later after they transition to English-instruction classrooms. Our preferred specification, presented in column 6 of the upper panel of Table 2, shows that students taught in their mother tongue first have scored 0.269 standard deviations higher than their peers taught in their second language first in the mathematics test that was administered after students transition to English instruction (i.e., towards the end of grade 5).

Impact on Literacy Test Scores

The lower panel of Table 2 is the counterpart of the upper panel of the same table where the specifications across columns (and their justifications) are exactly the same for both panels. The only difference between the upper and lower panels of Table 2 is the dependent variable is standardized mathematics test scores in the upper panel and it is standardized literacy test scores in the lower panel.

As in the case of mathematics test scores, the lower panel of Table 2 shows that the coefficients estimates of the treatment effects (i.e., "Dummy for learning in MT first" variable in the OLS model and "$Exp * After * Treated$" variable in the triple-differences model) are positive and statistically significant. This suggests that learning in mother tongue fist increases students' literacy test scores later after they transition to English instruction (in grade 5).

It is interesting to see that the magnitude of the coefficient estimates of the treatment effects reported in the lower panel of Table 2 are uniformly smaller relative to those reported in the upper panel of the same table. In our preferred specification, i.e., the triple-differences model presented in column 6 (of the upper and lower panels) of Table 2, for instance, the treatment effect is 0.269 and 0.089 standard deviations in mathematics and literacy tests scores, respectively. This suggests that the gain in academic achievement from learning in mother tongue first seems to come from more of a larger gain in mathematics skills than the gain in literacy skills. This is consistent with the findings from prior studies that document stronger effect of learning in mother tongue on mathematics test scores relative to its effect on literacy test scores (e.g. Hynsjö and Damon, 2016).

4.3. Falsification Tests

In the main analysis presented here above, we have documented the positive effect of learning in mother tongue first on students' performance later after they transition to English-instruction classrooms. This implicitly assumes that the triple-differences estimates that are presented in this paper pick up the effect of learning in mother tongue first, not the effect of potential factors that may have *differential* effect on the performance of students from different language groups across states in Ethiopia.

To assess the validity of this claim, we conduct falsification tests by restricting our sample to students in *non-experimental* states and randomly assigning them into *placebo* experimental and non-experimental states. As mentioned earlier, the YL's Ethiopian school survey samples from 6 states (i.e., Afar, Amhara, Oromiya, SNNP, Somali, and Tigray states) and Addis Ababa, the federal capital, but we have excluded Addis Ababa from our sample of analysis. In the falsification test analysis, we further drop SNNP state (the true experimental state) and randomly assign the remaining 5 non-experimental states into placebo experimental and non-experimental states.

To ensure that the assignment of states into placebo experimental and non-experimental states is random, we exploit the administrative numbers (such as 01, 02, 03, etc) that states are assigned to by the federal government for administrative convenience. Specifically, we assign odd-numbered states into placebo experimental state and even-numbered states into non-experimental states.[27] Since such assignment of states into experimental and non-experimental is random, we expect to find no treatment effect in the falsification test analysis if the positive treatment effect documented in Table 2 is driven by the fact that language-majority students are taught in their mother tongue first, and not by other confounding factors.

The results from these falsification tests of the triple-differences estimations are

[27] This leads to assigning Amhara, Somali, and Tigray states into placebo experimental states and Afar and Oromiya states into non-experimental states.

presented in Table 3, where the upper and lower panels present results from where the dependent variables are standardized mathematics and literacy tests scores, respectively. As can be seen from both the upper and lower panels of Table 3, the coefficient estimates of the third-level interaction,"$Exp * After * Treated$," are insignificant in all specifications. This confirms that the positive treatment effect presented in the main analysis is driven by the fact that language-majority students in SNNP state are taught in their mother tongue first, and not by other confounding factors.

4.4. Heterogeneity

So far we have documented the positive effect of learning in mother tongue first on the performance of students later after they transition to English-instruction classrooms. While our research is unable to identify the exact mechanisms of the treatment effect, prior studies suggest that mother-tongue instruction improves performance at school because mother-tongue instruction facilitates adjustment between home and school (e.g., Trudell, 2005); mother-tongue instruction helps students express themselves well which, in turn, helps them develop higher level of cognitive skills relatively quickly (e.g., Sonaiya, 2002); students strongly identify themselves with teachers who come from the same language group (e.g., Klaus, 2003); and teachers who come from the same language group are more trustworthy and more subject to social control, reducing the risk that they will abuse their students sexually or otherwise (e.g., Benson, 2005).

Given the documented positive effect of learning in mother tongue on students' performance in primary school in the literature, the finding documented in the present study that suggests that learning in mother tongue *first* improves students' performance later after they transition to English-instruction classroom supports the argument that subject contents/concepts that are first learned through mother tongue can be transferred to English.

In addition to exploring the effect of learning in mother tongue first, it is interesting to further explore whether the results documented in the present study are heterogeneous across different groups. One way to do this is to look at the differences in the results for rural and urban subsamples. Since there is evidence that supports that the 1994 education reform (which has led to mother-tongue instruction for many primary school students in Ethiopia) has had stronger positive effect for kids in rural areas relative to those in urban areas (Seid, 2016), we would expect that the coefficients for the rural subsample would be larger.

Table 3: **Falsification Test:** Triple-differences Estimates of the Effect of Learning in Mother Tongue First on Students' Performance Later after they Transition to English Instruction

	(1)	(2)	(3)
Math Z-score Equation			
$Exp, placebo * After * Treated$	0.016	-0.012	-0.009
	(0.107)	(0.164)	(0.135)
Lagged math Z-score		0.004**	0.003
		(0.002)	(0.008)
Student- & school-level controls	Yes	Yes	Yes
Household-level controls	No	No	Yes
Observations	3914	3914	3914
R-squared	0.279	0.288	0.312
Literacy Z-score Equation			
$Exp, placebo * After * Treated$	0.117	0.109	0.089
	(0.139)	(0.142)	(0.106)
Lagged literacy Z-score		0.012***	0.008
		(0.003)	(0.007)
Student- & school-level controls	Yes	Yes	Yes
Household-level controls	No	No	Yes
Observations	3914	3914	3914
R-squared	0.167	0.189	0.264

Notes: *p < 0.10, ** p < 0.05, *** p < 0.01.
Robust standard errors are reported in parentheses. In the triple-differences regressions, we control for dummies for whether the state is experimental, the student is in grade 5, and the student is in treated group. Besides, we control for the second-level interactions of the above three dummy variables. The additional regression controls are student-level characteristics (i.e., student's age in years and binary indicators for whether the student is female, has attended preschool, has ever repeated grade, and has participated in paid work), school-level characteristics (i.e., years of experience of mathematics/language teacher and the principal; binary indicators for whether the mathematics/language teacher and the principal have post-secondary education and whether the mathematics/language teacher is female; binary indicators for whether the school has library, access to electricity, and is found in urban area), household-level characteristics (i.e., binary indicators for whether the student's mother and father are literate, binary indicator for whether the household has access to electricity, number of rooms in the household, number of siblings, and the number of meals the student eats in a typical day), and classroom and school fixed effects.

Table 4 presents the coefficient estimates of selected variables from the triple-differences regressions that are run separately for rural and urban subsamples. The upper and lower panels of Table 4 present results from where the dependent variables are standardized mathematics and literacy tests sores, respectively. Across all specifications in the table, the treatment effect is positive and statistically significant, but the magnitude of the coefficient estimates of the "$Exp * After * Treated''$ variable is higher for the rural subsample. This suggests that learning in mother tongue first, as expected, has stronger effect for kids in rural areas relative to those in urban areas.

We have also conducted a similar experiment to investigate whether the results are heterogeneous by gender by running the triple-differences regressions separately for the boys and girls subsamples. The coefficient estimates of the "$Exp * After * Treated''$ variable (which are not reported here for the interest of space) are not systematically different for the boys and girls subsamples, implying that there is no heterogeneity in treatment effect by gender.

5. Conclusions

Achieving universal primary education has been a priority to developing countries for a long time. As part of this objective, many developing countries have put in a lot of effort to make primary education accessible to traditionally marginalized groups. Adopting mother-tongue instruction has played its own role in improving primary school enrollment among kids from marginalized groups and performance at school. Though mother-tongue instruction has a positive effect on improving enrollment and students' performance in primary school, it is not clear whether the gain in academic achievement carries over to later years in school after students transition from mother-tongue to English-instruction classrooms.

In this paper, thus, we attempt to fill this gap in the literature by exploring whether learning in mother tongue first improves students' academic achievement (measured by mathematics and literacy tests scores) later after they transition to English-instruction classrooms. To document the causal relationship between learning in mother tongue first and academic achievement later, we exploit the unique nature of the Ethiopian primary education system and the country's ethno-linguistic diversity.

Ethiopia has adopted mother-tongue instruction in primary school following the signing of the 1994 education reform into law, but states in Ethiopia have discretion to choose when students transition to English instruction. Following this

Table 4: Triple-differences Estimates of the Effect of Learning in Mother Tongue First on Students' Performance Later after they Transition to English Instruction by Location of Residence

	Rural Subample			Urban Subsample		
	(1)	(2)	(3)	(4)	(5)	(6)
Math Z-score Equation						
$Exp * After * Treated$	0.314***	0.299***	0.281***	0.291***	0.280***	0.257***
	(0.023)	(0.016)	(0.017)	(0.025)	(0.024)	(0.022)
Lagged math Z-score		0.014**	0.019		0.026*	0.017*
		(0.006)	(0.016)		(0.015)	(0.009)
Student- & school-level controls	Yes	Yes	Yes	Yes	Yes	Yes
Household-level controls	No	No	Yes	No	No	Yes
Observations	3943	3943	3943	2311	2311	2311
R-squared	0.405	0.421	0.461	0.365	0.371	0.401
Literacy Z-score Equation						
$Exp * After * Treated$	0.107***	0.105***	0.096***	0.092***	0.086***	0.088***
	(0.018)	(0.016)	(0.016)	(0.021)	(0.023)	(0.025)
Lagged literacy Z-score		0.004**	0.012		0.010**	0.009
		(0.002)	(0.027)		(0.004)	(0.006)
Student- & school-level controls	Yes	Yes	Yes	Yes	Yes	Yes
Household-level controls	No	No	Yes	No	No	Yes
Observations	3943	3943	3943	2311	2311	2311
R-squared	0.537	0.588	0.596	0.503	0.553	0.562

Notes: *p < 0.10, ** p < 0.05, *** p < 0.01.
‡ MT denotes mother tongue.
Robust standard errors are reported in parentheses. In the triple-differences regressions, we control for dummies for whether the state is experimental, the student is in grade 5, and the student is in treated group. Besides, we control for the second-level interactions of the above three dummy variables. The additional regression controls are student-level characteristics (i.e., student's age in years and binary indicators for whether the student is female, has attended preschool, has ever repeated grade, and has participated in paid work), school-level characteristics (i.e., years of experience of mathematics/language teacher and the principal; binary indicators for whether the mathematics/language teacher and the principal have post-secondary education and whether the mathematics/language teacher is female; and binary indicators for whether the school has library and access to electricity), household-level characteristics (i.e., binary indicators for whether the student's mother and father are literate, binary indicator for whether the household has access to electricity, number of rooms in the household, number of siblings, and the number of meals the student eats in a typical day), and classroom and school fixed effects.

discretion, SNNP state has legislated for students to transition to English instruction in grade 5 whereas students in other states in Ethiopia transition to English instruction either in grade 7 or 9. It is important to note that SNNP state is the most diverse state in Ethiopia. Thus, a large number of students in SNNP state still learn in their second language in the first cycle (i.e., grades $1-4$) of primary education. As a result, grade 5 student in SNNP state are composed of two groups of students: those who were first (i.e., in grade $1-4$) taught in their mother tongue and their second language. This implies a variation in the intensity of the impact of the transition to English instruction across grade 5 students in SNNP state by their language group.

Our identification strategy in this paper relies on these two sources of plausibly exogenous variations. Specifically, we first exploit the fact that students in SNNP state transition to English instruction in grade 5 whereas their peers in other states in Ethiopia do so when they progress either to grade 7 or 9. Therefore, we consider SNNP state as *experimental* state and the other states in Ethiopia as *non-experimental* states.

The second source of variation comes from a variation in the intensity of the impact of the transition to English instruction across grade 5 students in SNNP state who come from different language groups, where language-majority students transition from mother-tongue to English instruction and language-minority students transition from second-language to English instruction. Hence, we assign students who are first taught in their mother tongue and second language into *treated* and *control* groups, respectively. As part of our identification strategy, we further assign grades 4 and 5 students into *pre-treatment* and *after-treatment* groups, respectively, since English is used as the medium of instruction for only grade 5 students in SNNP state.

Using data from YL's 2012-2013 Ethiopian school survey, which administers mathematics and literacy tests to grades 4 and 5 students both at the beginning and end of the 2012-2013 school year, and the assignment of states into experimental and non-experimental along with the assignment of students into treated, control, pre-, and after-treatment groups, we provide empirical evidence on the causal effect of learning in mother tongue first on students' academic achievement later after they transition to English instruction (in grade 5) by estimating triple-differences model for a sample of grades 4 and 5 students in Ethiopia. As alternative specification and to make comparison of our estimates with those from prior studies straightforward, we also estimate test scores production function for a *subsample* of grade 5 students in SNNP state using OLS model.

The estimates from our preferred specification suggest that learning in mother tongue first increases the performance of students later after they transition to English instruction (in grade 5) in mathematics and literacy tests by 0.269 and 0.089 standard deviations, respectively. We also find that these effects are stronger for students in rural areas relative to those in urban areas. Falsification tests, on the other hand, suggest that our results are not confounded by other factors.

The findings in our paper are consistent with the argument that, compared to their peers taught in their second language first, students taught in their mother tongue first learn in English better after they transition to English-instruction classrooms. However, these findings should be treated carefully since they only document the short-term effect of learning in mother tongue first on students' academic achievement later after they transition to English-instruction classroom. It is not clear whether the gain in academic achievement due to learning in mother tongue first fades away as students progress through grades. Research on whether gains in academic achievements in primary school due to learning in mother tongue first are translated to better labor market outcomes later in life would be a valuable contribution to the literature.

Bibliography

Angrist, Joshua D and Victor Lavy, "The Effect of a Change in Language of Instruction on the Returns to Schooling in Morocco," *Journal of Labor Economics*, 1997, pp. S48–S76.

Bamgbose, Ayo, *Language and the Nation: The Language Question in Sub-Saharan Africa*, Edinburgh University Press, 1991.

Benson, Carol, "Girls, Educational Equity and Mother Tongue-Based Teaching.," *UNESCO Bangkok*, 2005.

Brock-Utne, Birgit, "Learning Through a Familiar Language versus Learning Through a Foreign Language – A Look into Some Secondary School Classrooms in Tanzania," *International Journal of Educational Development*, 2007, *27* (5), 487–498.

Cheung, Alan CK and Robert E Slavin, "Effective Reading Programs for Spanish-dominant English Language Learners (ELLs) in the Elementary Grades: A synthesis of Research," *Review of Educational Research*, 2012, pp. 351–395.

Cummins, Jim, "Alternative Paradigms in Bilingual Education Research: Does Theory have a Place?," *Educational Researcher*, 1999, pp. 26–41.

_ , *Language, Power, and Pedagogy: Bilingual Children in the Crossfire*, Vol. 23, Multilingual Matters, 2000.

Heugh, Kathleen, Carol Benson, Berhanu Bogale, and Mekonnen Alemu, "Study on Medium of Instruction in Primary Schools in Ethiopia," *Ministry of Education, Federal Democratic Republic of Ethiopia*, 2007.

Hynsjö, Disa and Amy Damon, "Bilingual Education in Peru: Evidence on how Quechua-medium Education Affects Indigenous Children's Academic Achievement," *Economics of Education Review*, 2016.

Klaus, David, "The Use of Indigenous Languages in Early Basic Education in Papua New Guinea: A Model for Elsewhere?.," *Language and Education*, 2003, *17* (2), 105.

Kuziemko, Ilyana, "Human Capital Spillovers in Families: Do Parents Learn from or Lean on their Children?," *Journal of Labor Economics*, 2014, *32* (4), 755–786.

Lee, Myoung-jae and Changhui Kang, "Identification for Difference in Differences with Cross-section and Panel Data," *Economics letters*, 2006, *92* (2), 270–276.

Ministry of Education, "Education and Training Policy," *Ministry of Education, Federal Democratic Republic of Ethiopia*, 1994.

Piper, Benjamin, Stephanie S Zuilkowski, and Salome Ongele, "Implementing Mother Tongue Instruction in the Real World: Results from a Medium-Scale Randomized Controlled Trial in Kenya," *Comparative Education Review*, 2016, *60* (4), 776–807.

Rossell, Christine H and Keith Baker, "The Educational Effectiveness of Bilingual Education," *Research in the Teaching of English*, 1996, pp. 7–74.

Seid, Yared, "Does Learning in Mother Tongue Matter? Evidence from a Natural Experiment in Ethiopia," *Economics of Education Review*, 2016, *55*, 21 – 38.

Seidel, Katrin and Janine Moritz, "Changes in Ethiopia's Language and Education Policy–Pioneering Reforms?," in "Proceedings Paper: 16th International Conference of Ethiopian Studies, Trondheim" 2007, pp. 2–7.

Slavin, Robert E and Alan Cheung, "A Synthesis of Research on Language of Reading Instruction for English Language Learners," *Review of Educational Research*, 2005, *75* (2), 247–284.

_ , Nancy Madden, Margarita Calderón, Anne Chamberlain, and Megan Hennessy, "Reading and Language Outcomes of a Multiyear Randomized Evaluation of Transitional Bilingual Education," *Educational Evaluation and Policy Analysis*, 2011, *33* (1), 47–58.

Sonaiya, Remi, "Autonomous Language Learning in Africa: A Mismatch of Cultural Assumptions," *Language, Culture and Curriculum*, 2002, *15* (2), 106–116.

Taylor, Stephen and Marisa von Fintel, "Estimating the Impact of Language of Instruction in South African Primary Schools: A Fixed Effects Approach," *Economics of Education Review*, 2016, *50*, 75–89.

Todd, Petra E and Kenneth I Wolpin, "On the specification and estimation of the production function for cognitive achievement," *The Economic Journal*, 2003, *113* (485), F3–F33.

Trudell, Barbara, "Language Choice, Education and Community Identity," *International Journal of Educational Development*, 2005, *25* (3), 237–251.

Wilson, Ian, Sharon R. A. Huttly, and Bridget Fenn, "A Case Study of Sample Design for Longitudinal Research: Young Lives," *International Journal of Social Research Methodology*, 2006, *9* (5), 351–365.

Appendices

Appendix A. Additional Tables

Table A.1: Summary Statistics of the Outcome Variables used in the Econometric Analysis by Experimental State, Language Group, and Grade

	Experimental State				Non-experimental States			
	MT‡ Students		Non-MT Students		MT Students		Non-MT Students	
	Grade 4	Grade 5	Grade 4	Grade 5	Grade 4	Grade 5	Grade 4	Grade 5
Math Z-score - Wave 2	0.073	0.317	-0.193	-0.277	0.073	0.118	-0.311	-0.241
	(0.897)	(0.915)	(0.836)	(0.967)	(0.611)	(0.945)	(0.936)	(0.911)
Literacy Z-score - Wave 2	0.242	0.378	-0.035	-0.087	0.270	0.309	0.065	0.149
	(0.921)	(0.965)	(0.900)	(0.989)	(0.991)	(0.962)	(0.933)	(0.937)
Observations	841	801	363	335	1593	1529	400	392

Notes: Standard deviations are reported in parentheses.
‡ MT and non-MT denote language-majority and language-minority students, respectively, where the former are taught in their mother tongue in grades $1-4$ whereas the later are taught in their second language in grades $1-4$.