Working Paper 177

Young Lives

Inequality in Attainment From Early Childhood to Adolescence

Longitudinal Evidence From Ethiopia

Milo Vandemoortele

www.younglives.org.uk

MARCH 2018

Inequality in Attainment From Early Childhood to Adolescence: Longitudinal Evidence From Ethiopia

Milo Vandemoortele

© Young Lives 2018
ISBN 978-1-912485-02-4

A catalogue record for this publication is available from the British Library.
All rights reserved. Reproduction, copy, transmission, or translation of any part of this publication may be made only under the following conditions:

- with the prior permission of the publisher; or
- with a licence from the Copyright Licensing Agency Ltd., 90 Tottenham Court Road, London W1P 9HE, UK, or from another national licensing agency; or
- under the terms set out below.

This publication is copyright, but may be reproduced by any method without fee for teaching or non-profit purposes, but not for resale. Formal permission is required for all such uses, but normally will be granted immediately. For copying in any other circumstances, or for re-use in other publications, or for translation or adaptation, prior written permission must be obtained from the publisher and a fee may be payable.

Core-funded by

UKaid
from the British people

Young Lives, Oxford Department of International Development (ODID), University of Oxford, Queen Elizabeth House, 3 Mansfield Road, Oxford OX1 3TB, UK

Tel: +44 (0)1865 281751 • Email: younglives@younglives.org.uk

Contents

The authors

Milo Vandemoortele is a PhD candidate at the London School of Economics (LSE). Her research examines the links between parental resources and attainment over time, and preschool education and attainment over time. In her cross-country comparative study she uses Young Lives data from Ethiopia, India, Peru and Vietnam.

Acknowledgements

I would like to thank the Young Lives team for their support in this work; Paul Dornan, Zoe James, and Alula Pankhurst for their thoughtful comments, and Roo Griffiths and Adam Houlbrook for their copy-editing work. All remaining errors are my own.

Summary

As the Ethiopian government moves to increase enrolment in preschool education from 34 per cent in 2013/14 to 80 per cent by 2020, this working paper provides some country-specific evidence on the predictors of preschool attendance and its association to attainment in Ethiopia. Using four rounds of Young Lives data, it examines: who typically attends preschool education; whether children who attend preschool education have higher attainment than children who do not; and how the association between preschool education and attainment evolves over time. Findings indicate that the key predictors of preschool education are being a first-born child, speaking Amharic as a first language, having a more educated caregiver, belonging to a household with higher wealth, living in an urban area, and having a mother who is not employed. Although being male and better nourished are positively associated with preschool attendance, they are only modestly so.

The paper also examines the association between preschool education and children's attainment, while controlling for a rich array of child-, household- and primary school-level characteristics and variations between communities. Analysis finds that preschool education has a direct and substantial association with attainment, and points to an indirect one through on-time enrolment, grade progression and retention rates (captured in highest grade completed). It also shows that preschool education has a positive and statistically significant role at 5 and 8 years old, where preschool attendees score about 5 percentage points higher on maths tests, compared to non-attendees. This association, however, fades by 12 years old. This is perhaps because exposure to schooling becomes more important. At 8 years old, exposure to schooling emerges as an intermediary outcome associated with attainment, and supersedes that of preschool education by age 12. Further analysis shows that preschool education has a direct and substantial association with school progression by age 12. At 12 years old, children who attended preschool complete 0.6 more years of education than their non-attending counterparts. These findings indicate that investments in preschool education may doubly pay off – both directly towards attainment and indirectly through exposure to schooling.

About Young Lives

Young Lives is an international study of childhood poverty, following the lives of 12,000 children in four countries (Ethiopia, India, Peru and Vietnam) over 15 years. **www.younglives.org.uk**

The views expressed are those of the authors. They are not necessarily those of, or endorsed by, the University of Oxford, Young Lives, DFID or other funders.

1. Motivation

Early childhood education for children aged 4 to 6 years old has recently come to the fore as a policy priority in Ethiopia. In its 2015 Education Sector Development Program (ESDP V), the Ethiopian government prioritised rolling out at least one year of pre-primary education to all children by 2020 (MoE 2015). Why focus on early childhood education? Five reasons stand out.

First, exposure to good-quality preschool education can improve school readiness, reduce repetition rates and avoid dropout, thus contributing to a more effective and efficient school system. This is the main reasoning behind the Ethiopian government's actions (MoE 2015). Second, there is evidence that preschool education enhances children's cognitive development, producing more productive adults. Third, the evidence also shows that it is disadvantaged children in particular who stand to benefit the most from preschool education. Therefore, governments interested in addressing inequalities, such as the Ethiopian government (MoE 2015), require appropriate interventions at early ages when disparities in cognition emerge. Fourth, cost–benefit analyses of educational interventions across childhood make a strong case for early childhood interventions. Fifth, preschool education programmes enable higher female workforce engagement and children may indirectly benefit from increased household income, although evidence on this is mixed.

When looking for evidence to inform the policy debate in Ethiopia, the main concern is that most of the research that examines the role of preschool education in children's cognitive development stems from high-income country contexts. Evidence on the role of preschool education on children's cognitive development in low-income countries, particularly in Africa, is limited. There are important differences in this role between high- and low-income settings. First, malnutrition is a significant hindrance to children's development in low-income countries. Second, early childhood education infrastructure and quality of preschool education in high-income settings are, on average, substantially better than those in low-income settings (see Woodhead et al. (2009) for examples in low-income countries). Some children live in remote areas, with no road access, so accessing the nearest preschool is impossible. Additionally, the training of teachers, curriculum development and quality of materials in preschools are usually quite poor (Woodhead et al. 2009).

Third, the conditions in which children grow up are for the most part vastly different in low-income countries to those in high-income countries. As Hoot et al. (2004) discuss, young boys in rural Ethiopia are usually expected to support their parents herding animals, even before primary school age. Meanwhile, girls have more homemaking and caring responsibilities. In addition to increased child work responsibilities, opportunities for child development within communities differ: in low-income settings, there are far fewer libraries, children's centres, educational playgrounds and activities.

As the Ethiopian government works to expand access, it is important to examine the role of preschool education in cognitive development in the country context. Using longitudinal data from the Young Lives study on approximately 2,000 Ethiopian children who were of preschool age between 2005 and 2007, this working paper examines the association of preschool education on children's outcomes, as measured by maths test scores. Young Lives collects data on preschool attendance from age 3. Cognitive development is measured in many ways; in this paper I use maths test scores to assess cognitive development, and therefore use the term *attainment*, rather than cognitive development, when discussing results.

This paper examines: who typically attends preschool; whether preschool attendance is associated with short and medium term advantages in attainment; and how the relationship between preschool education and attainment evolves over time.

2. The role of preschool education on children's cognitive development

While many factors contribute to children's cognitive development, including nutrition, this section focuses on the role of preschool education on children's cognitive development.

2.1. Preschool education: essential to children's development

Ruhm and Waldfogel (2012) review evidence of universal preschool education programmes across both high- and middle-income countries. They find that preschool education has short, medium and long-term benefits on children's cognitive development. A systematic review by Engle et al. (2007) focuses on low- and middle-income countries and finds that of 14 studies assessing the role of preschool education on children's development, all but one find a positive role played by preschool education. These findings are supported by a meta-analysis by Nores and Barnett (2010) that reviews non-US interventions, and includes studies from Europe, Asia, and Central and South America. However, evidence, particularly peer-reviewed evidence, from low-income countries is very limited.

The quality of preschool education plays an important role in the extent to which preschool attendance is beneficial.[1] A systematic review by Engle et al. (2011) focused on, but not constrained by, low- and middle-income evidence, confirms that higher quality preschool education programmes are associated with better learning outcomes.[2] In the anomalous case of Quebec (Canada), negative effects observed in vocabulary test scores after the introduction of a universal child care subsidy are explained by low-quality informal care (Ruhm and Waldfogel 2012).

2.2. The importance of early childhood education in reducing inequalities

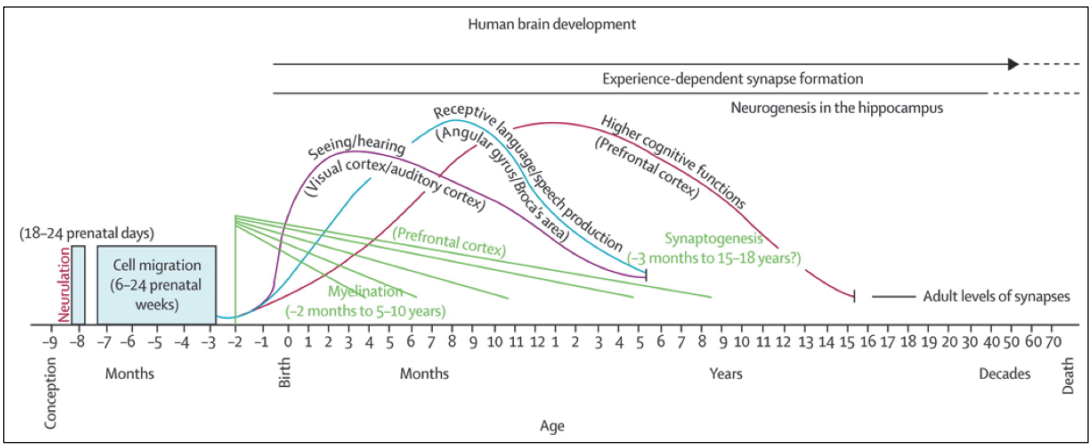
There is a growing consensus in the literature on high-income countries that disadvantaged children are not prepared for university because they were not prepared for primary school (Heckman and Raut 2013: 367). Figure 1 depicts major brain development processes, all of which build on each other. By the age of 6, the pace of visual, auditory and language development has reached its peak. Given the sensitive nature of the first few years of brain development and the cumulative nature of learning, the reasons for disparities in attainment

1 Quality is defined in many different ways. Some hard measures include pupil to teacher ratios, teacher qualifications, and variety and appropriateness of material and activities provided to the children. There are also soft aspects such as how responsive and engaged teachers are, and whether they actually use the materials and activities available to them.

2 Higher quality as defined by the articles in the review.

in adolescence and adulthood may lie in early childhood. As disadvantaged children are more likely to start school 'behind' their more advantaged peers, preschool education should help to narrow gaps in attainment before children start formal schooling. Evidence from both high- and middle-income countries is encouraging, as it shows that the benefits of preschool education are relatively greater for children from more disadvantaged households, than for those from more advantaged households (Engle et al. 2011; Ruhm and Waldfogel 2012).

Figure 1. *Human brain development*



Source: Thompson and Nelson (2001) cited by Grantham-McGregor et al. (2007).

2.3. Cost-benefit analysis of education interventions for children from 3 years old to primary school age

Building on this, Heckman and others argue that returns to investments made in early childhood are much larger than those made in later life (Heckman 2008; Heckman and Masterov 2007). They contend that early investments promote greater efficiency in learning and increase the productivity of future investments. This synergetic cycle is often referred to as '*dynamic complementarity*'. Preschool education programmes are an efficient way to help protect against costly negative outcomes in later life (e.g. illiteracy or innumeracy), and enhance the effectiveness of investments in primary and secondary school.

2.4. The indirect role of preschool education: maternal employment

Enabling maternal employment is an important consideration in many countries' preschool education programmes, particularly as they aim to enhance early childhood development indirectly through increased income (Nores and Barnett 2010; Ruhm and Waldfogel 2012). However, the relationship between maternal employment and child development is complex, and there is no clear consensus on its long-term benefits. Evidence from the United States shows that benefits of maternal employment are not equally shared: children of the most educated women are gaining with regards to their parent's time and money, while those of the least educated mothers are losing these resources (McLanahan 2004). In low- and middle-income countries this relationship is also complicated, and evidence is limited. On the one hand, children may benefit from maternal employment in terms of increased household resources, as maternal income is more likely than that of the father to be invested in children's well-being. On the other hand, the mother's work may also involve children in their work, or result in children taking on domestic chores and caring responsibilities.

2.5. The importance of a high-quality environment to children's development

Preschool education only partly contributes to children's cognitive development. For example, parents who provide their children with preschool education are also more likely to engage in other behaviours that will contribute to their children's development (Havnes 2012). These may include reading more frequently to the child, or being better informed on the types of games and attention a child needs for cognitive development. Beyond the home environment are a whole host of influences on the child, including: socio-economic status, family health, the quality of their dwelling, access to services (e.g. water, sanitation, health and education), and gender roles in society (Irwin et al. 2007).

2.6. Mechanisms through which preschool education contributes to cognitive development

There are several mechanisms through which preschool education can contribute to cognitive development. Two prominent causal mechanisms emerge from the review by Ruhm and Waldfogel (2012). The first causal pathway through which preschool education can affect child development relates to direct investment in human capital. As emphasised by Heckman and others, returns to investment in early years may be higher than those made in later years. Preschool education is therefore important because of the direct and early investment in human capital it provides. Its effect may differ across groups depending on the nature of the investment and the characteristics of the child. For example, poor-quality preschools may have a neutral or detrimental effect compared with home-based care. Alternatively, the language component of preschool education may be particularly beneficial for children whose mother tongue is distinct from the language of primary school tuition.

The second causal pathway relates to indirect investments in human capital. One of the most important documented such investments in high-income countries is parental employment, particularly of the mother. A positive result of parental employment is its potential to raise family income; a negative consequence may be a less responsive and stimulating care environment away from home.

These causal pathways are relevant to low-income settings. However, they may differ in several aspects. Human capital investment for school readiness is likely to be even more important in low-income settings, where the caregiver's literacy and numeracy skills are on average lower. The transition to primary school will be even more challenging if the child has had little or no exposure to reading and writing numbers. This is an important consideration in Ethiopia, as one of the highest dropout rates is in the first year of primary school – at 21 per cent, with 14 per cent in subsequent grades (MoE 2015: 9).[3] Good-quality preschool education in Ethiopia is likely to offer a substantial contribution to human capital if it prepares children for primary school education, reducing dropout rates.

The quality of preschools in low-income settings, however, is not 'high'. While there are high-quality preschools in the country, catering to children from privileged households, most

3 This may be reflective of reality or a statistical artefact, as in Ethiopia dropout rates are calculated by subtracting the number of children enrolled in Grade 2 from the number enrolled in Grade 1 the previous year, divided by the latter. This calculation does not take into consideration the fact that children may have been retained in Grade 1. If children are not prepared for school, high repetition in the first year is likely. Thus, a larger number of children enrolled in Grade 1 will inflate dropout rates. Further, enrolment does not reflect attendance. If a child is enrolled but does not attend, dropout rates will not include her.

preschools in Ethiopia are of poor quality (Orkin et al. 2012). Therefore, the effect on child outcomes may not be as substantial.

The potential indirect benefits of attending preschool education are multiple. A mother or an older female child may be relieved of caring responsibilities while the child attends preschool, and therefore be able to attend school or work herself. This may have a positive or negative knock-on effect on the child in preschool education, depending on whether and how this affects the mother's time at home, as well as its influence on other factors, such as household resources. Another important indirect channel is physiological (Walker et al. 2011): where free meals are provided at the preschool (as is the case in some other middle-income countries but not Ethiopia), this enhances children's cognitive development through the reduction in nutritional deficit. Preschool education may also represent a 'gateway' into better subsequent schooling (Dornan and Woodhead 2015). For example, a child living in an urban area attending a good-quality preschool education will in turn improve her chances of entering a good primary school and adjust well, and so on into secondary school. A child from a minority group living in a slum who does not access preschool education has missed this gateway, and will be less likely to then enter a good primary school or adjust well to any primary school, and so on.

2.7. Gaps in the literature

There is evidence that good-quality preschool education programmes improve children's outcomes in the short, medium and long term; that investments in early years can reduce later life inequalities, and are more productive than investments in later years; and that although this may enable maternal employment it does not necessarily benefit children's development. The preponderance of evidence is on high-income countries, and evidence from middle-income countries is growing. Evidence from low-income countries, particularly in Africa, is scant. This is perhaps a result of a lack of preschool education programmes, or lack of quality data to examine the role of preschool education on cognitive development. Regarding Ethiopia in particular, only a handful of papers by the same author address the role of preschool education from a quantitative approach, and these come from previous rounds of Young Lives data (i.e. Woldehanna 2011; Woldehanna 2016; Woldehanna and Gebremedhin 2012) or from the Older Cohort of children (i.e. Woldehanna and Araya 2017) rather than the Younger Cohort, which I examine here.[4]

4 Woldehanna et al. (2008) and Woldehanna (2011) examine similar questions using previous rounds of Young Lives data when the children were 5 years old, and Woldehanna and Gebremedhin (2012) and Woldehanna (2016) when they were 5 and 8 years old. This paper compares findings reported in the latter, as it addresses the same questions as previous papers but for 5 and 8 year olds, and is the most up-to-date analysis. Woldehanna and Araya (2017) examine long-term associations of early childhood education on successful completion of secondary education and the likelihood of transitioning to higher education, using data from the Older Cohort.

3. The case of Ethiopia

3.1. Policy relevance

Ethiopia's policy regarding preschool education has progressed quickly. In 2002, the Ethiopian Ministry of Education explicitly stated that it did not have the resources to focus on preschool education, preferring rather to consolidate the primary school system:

'[…] from the perspective of Ethiopia's economic capacity, the opening of kindergartens involving massive expenditure cannot be a top priority, as regular universal primary education has not yet been achieved. Thus, the opening of kindergartens is an area that has been left for private investors and religious organisations, and for parents who can afford to pay the fees' (MoE 2002).

Thirteen years later, in 2015, the Ethiopian government set a target to achieve an 80 per cent gross enrolment ratio (GER) in preschool education by 2020 (a substantial increase from 34 per cent in 2013/14) (MoE 2015). This target is meant to contribute to a more efficient and effective school system, reduce dropout and repetition rates, and address inequalities in attainment (MoE 2015).

3.2. Preschool education in Ethiopia in 2007

Formal compulsory primary education in Ethiopia starts at the age of 7. Therefore, preschool education targets children aged 4 to 6 years old. Most Young Lives children were of preschool age between 2005 and 2007.

Qualitative field research by Orkin et al. (2012) identifies four main types of preschools in Ethiopia. Private preschools require a fee and are mostly located in urban areas. 'Public' pre-schools are fee-paying but also funded by the government, and are also mostly located in urban areas. The quality and service provision in these schools resemble those of private preschools. Orkin et al. (2012: 25) therefore group public with private preschools. The third type consists of government preschools, established and run by the *kebele* (neighbourhood administration), but with teachers' salaries usually paid for by parents or community contributions. These schools are located mostly in urban areas, although there are a few in rural areas. Some of these government preschools also have O-classes.[5] The fourth type is mostly not-for-profit schools run by non-governmental or religious organisations. NGO schools tend to be located in urban areas, to be of relatively high quality, to target poor communities and therefore to require low or no fees. Examples of NGO preschools are those run by Save the Children, Christian Children's Fund and the SOS Children's village, and also local NGOs.

In 2007, the preschool education GER in Ethiopia was 3.1 per cent (MoE 2008). Table 1 provides information on preschool attendance and infrastructure across the five regions included in Young Lives. The highest GER was in the capital city (46.7 per cent); the

5 O-classes prepare children aged 5 to 6 who do not otherwise have access to preschool for entry into primary school (Grade 1). These classes are facilitated either by teachers selected from the primary school the child is likely to attend (MoE 2012) (during the teacher's free periods) or by other community members contracted by the school or the community (MoE 2015). The quality of instruction is therefore prone to large variations. No agreed O-class curriculum or learning materials were available, as ESDP V aims to establish these (MoE 2015). In 2012, the GER for O-classes was 13.7 per cent (MoE 2012).

remaining regions had a GER below 3 per cent. Because 83 per cent of Ethiopians live in rural areas (UN 2013), the national average remains low, despite the substantially larger GER for the capital. Teaching infrastructure varied substantially between Addis Ababa and other regions, with 4.2 preschools for every 1,000 children in Addis Ababa and only 0.17 in Amhara. A substantial portion of preschool teachers were not trained – 70 per cent in Amhara region and 23 per cent in Tigray region.

Table 1. *National statistics on preschool education in the five regions covered by Young Lives, 2006/07*

	GER* (%)	Untrained teachers (%)	No. of preschools per 1,000 children*	Pupil–teacher ratio
Tigray	2.3	23	0.34	26
Amhara	1.5	70	0.17	43
Oromia	2.3	45	0.27	42
SNNP	2.6	33	0.26	32
Addis Ababa	46.7	26	4.20	19

Notes: * Based on the school age population (4 to 6 years old). SNNP = Southern Nations, Nationalities and Peoples region.
Source: MoE (2008: 20).

In short, access to preschool education in 2007 was mostly limited to advantaged urban children. Few children attended preschool in rural areas because there were few preschools; parents could not afford the fees; the benefits of preschool education were not known to them; and/or children worked.[6] In what is a predominantly agricultural economy that uses traditional farming practices and suffers frequent droughts, children's support in herding animals and cultivation is often essential (Hoot et al. 2004). It is not uncommon for rural preschool-aged children to be working on the family smallholding.[7]

The Ministry of Education conducted a situation analysis in 2007, cited in a 2010 government report. This identified some key weaknesses in the preschool system:

'[…] high fees, lack of teacher training, lack of a standard curriculum, lack of culturally relevant storybooks, low teacher salaries and thus high turnover, the use of English as a medium of instruction, and, most importantly, a lack of awareness of the importance of ECCE [(early childhood care and education)]' (MoE (2010: 13) cited by Orkin et al. 2012).

At the time, the government's main role in the sector was limited to curriculum development, training teachers, and providing supervisory support to existing preschools (MoE 2008). Even this limited remit was severely constrained by a very low budget allocation to preschool education (MoE 2008).

6 Based on Young Lives data, financial limitations were not the foremost barrier to preschool access in Ethiopia reported by children's caregivers; rather, it appears to be infrastructure and information. When caregivers were asked why they do not send their child to preschool, the top three reasons were lack of preschools in the area (48 per cent); lack of knowledge about preschool (42 per cent); and that it was not considered necessary (17 per cent). Only 15 per cent cited lack of resources to pay for preschool among their top three reasons.

7 The opportunity cost of sending children to preschool, particularly when knowledge of its benefits is limited, contributes to lowering the likelihood of preschool enrolment. Not to mention the opportunity cost of education in general when schooling and the opportunities it is meant to create do not match up to local labour market realities.

3.3. Ambitions in the provision of preschool education

The preschool education scene in Ethiopia is expanding rapidly, but it is still predominantly an urban phenomenon that caters to more privileged children. Attendance in preschool education for children aged 4 to 6 years old reached 34 per cent in 2013/14 (MoE 2015), but the seemingly stellar growth from 3.1 per cent in 2007 is to a large extent a statistical artefact: it owes mostly to changes in the definition of preschool education to include O-classes and 'child-to-child' programmes.[8]

There has been a clear policy shift in Ethiopia to prioritise preschool education, as is evident in the ESDP V (MoE 2015), where the government outlined clear targets for achieving 80 per cent enrolment by 2020. Among these are:

- Increase the number of trained preschool teachers with a one-year certificate or a three-year 'early learning' diploma.
- Put in place a preschool teacher training curriculum, with the first cohort due to graduate in 2016.
- Produce standards for learning material and a specific curriculum for O-classes.
- Prepare a system to monitor the quality of preschools and hold them to account.
- Expand preschool provision, but also share the burden, as the government plans to provide half of all kindergartens by 2020.

Preschool education provision in 2005-07 (when the Young Lives children were preschool-aged) was and continues to be concentrated in urban areas and to cater to more privileged children. There was no curriculum for preschool education, preschool teacher training requirement or quality control system in place. In 2015, the Ethiopian government committed to increasing enrolment to 80 per cent – an ambitious goal.

4. Data

This analysis uses four rounds of Ethiopian Young Lives data, for the Younger Cohort only, for which information on attendance in preschool education was collected when the children were preschool age.[9]

Given the challenges of conducting child-focused longitudinal research in low-income countries, Young Lives collected data from 20 sites from a range of regions, urban and rural areas, and various language groups, with relatively good road access (to keep sampling costs manageable). Within these sites 100 children were randomly selected.[10] Young Lives is

8 Child-to-child programmes appear simply to be a formalisation of children playing. They entail older siblings from Grades 5 or 6 playing with younger siblings and teaching them numeracy, the alphabet and colours in order to prepare them for primary school (MoE 2012). They cater to children aged 4 to 5 years old. The GER for child-to-child programmes in 2012 was 2.6 per cent (MoE 2012). The government report does not address whether the older siblings facilitate this play at the cost of their own school attendance. This definition was revised in 2012, explaining the sudden increase in enrolment from 5.2 to 21.6 per cent between 2011 and 2012 (MoE 2012: 21).

9 Young Lives started collecting data for the Older Cohort after preschool age (when they were 7.5 to 8.5 years old), and there may be some recall problems.

10 See Appendix 1 for a detailed discussion of the Young Lives sampling strategy.

therefore not a nationally representative sample, and generalisation to the national Ethiopian population is not appropriate. Rather, Young Lives is a longitudinal study of child poverty that provides a unique and rich dataset on a population of interest – Ethiopian children born at the beginning of the millennium from the twenty sites described above.

5. Modelling framework

5.1. Determinants of attendance in preschool education

As a first step, this paper outlines how attendance in preschool education differs across various socio-economic and demographic groups. Then, to determine how much each factor predicts attendance in preschool education independently from other variables, a logistic regression model is fitted for child i in household j:[11]

$$P_{ij} = a_0 + \lambda \boldsymbol{X}_{ij} + e_{ij} \quad (1)$$

where P_{ij} represents attendance in preschool education, $\boldsymbol{X}_{ij}$ is a vector of observed child and household-level characteristics and e_{ij} is the error term. Attendance in preschool education is measured at 5 years old. Determinants of preschool education included in the model are height-for-age, gender, whether the child is a first-born child, age, caregiver's education, whether a child-carer (grandparent or older sibling) was present in the household, first language of the child, household size, parental resources (wealth index), maternal employment, and type of site (urban or rural) (see Appendix 1). All these are measured at age 5, except for parental resources which are measured in the previous round, at 1 year old.

The model does not include region of residence, given its high correlation with child's language group (0.78) which removes the statistical power of language group in the analysis.[12] The model includes one interaction: gender and first born, as the decision to provide a first born with preschool education may depend on the child's gender.

An assumption of the logistic and OLS model is that each observation is independent of all other observations in the data. However, there is likely to be correlation between children and households within communities and this would affect the standard errors. For example, households within communities may have similar socio-economic status and the children may attend the same preschool. Standard errors in Models (1), (2) and (3) therefore allow for intra-community correlation, relaxing the assumption that the observations are independent (i.e. robust standard errors).

Throughout this paper, only coefficients that are statistically significant at or below the 10 per cent level are interpreted.

11 A logistic model, as the outcome variable is binary: preschool attendance.

12 Correlation is 0.75 between the Amharic dummy and region.

5.2. The role of preschool education on schooling and maths scores

Many factors shape children's cognitive development. In order to isolate, as much as possible, the association between preschool education and cognitive development, several child- and household-level factors need to be controlled for, as well as community-level fixed effects to address between-community variations. This paper uses maths test scores to assess cognitive development, and henceforth the term *attainment*, rather than cognitive development, is used.

To examine whether attending preschool education does offer an advantage in terms of attainment, an OLS model is fitted with community-level fixed effects, to control for various child- and household-level factors associated with cognitive development. The model for child i in household j at age a is as follows:

$$Y_{ij,a} = b_0 + \beta_1 P_{ij,5} + \beta_2 \boldsymbol{X}_{ij,a} + \beta_2 \boldsymbol{Z}_{ij,a-1} + \alpha_{j,a} + \varepsilon_{ij,a} \quad (2)$$

where:

$Y_{ij,a}$ is the maths test score (per cent correct) measured at ages 5, 8, and 12 years;

$P_{ij,5}$ is the binary preschool education variable measured at age 5;

$\boldsymbol{X}_{ij,a}$ is a vector of child- and household-level characteristics;

$\boldsymbol{Z}_{ij,a}$ is a vector of household resources lagged one time period;

$\alpha_{j,5}$ represents the community-level fixed effects at age 5;

$\varepsilon_{ij,a}$ is the error term that captures the effect of unobserved factors, measurement error and random error.

P is measured in the same manner as in Model (1). Child-level characteristics include height-for-age, gender, age (in months), test taken in mother tongue, being a first-born child, the child's mother tongue, and, when the children are primary school aged, amount of time spent studying.

Household-level characteristics include caregiver's level of education, household size, parental resources (wealth quintile groups) and maternal employment. Schooling-level characteristics, which occur after preschool education, include primary school type and exposure to schooling.

Model (3) examines the role of preschool education on highest grade completed at 12 years old. The model includes the same variables and specification as Model (2), only the outcome variable changes to exposure to schooling (i.e. highest grade completed) and is no longer an explanatory variable. Missing dummies are included in these models to ensure the same sample size across regressions within rounds.

To eliminate differences between communities, community-level fixed effects are used. This ensures comparisons are being made within communities, not between them. Therefore, this model accounts for between-community differences.[13]

The error term comprises measurement unobserved variable bias and random error. Of concern is how much these unobserved variables are correlated with the control and

13 As community-level fixed effects are used, variations outside the community level, such as urban–rural and regional differences, are controlled for.

outcome variables. If they are highly correlated with controls, then the controls can act as a proxy and absorb the unobserved bias. If explanatory and unobserved variables are not correlated, the concern for the unobserved variable bias remains. If the unobserved variables are highly correlated with variables of interest, then the model will be inadequate to explain attainment at the hands of preschool education, as the association measured will be a proxy for something unobservable. In other words, it may not be preschool attendance, rather an unobservable or unmeasured benefit, which is correlated with attainment. Based on previous research in this area, albeit not in Ethiopia, it can be accepted that no single other factor will explain all the benefits of preschool education on attainment. However, some unobserved factors that will shape attainment, such as parental motivation to promote the child, may be correlated with preschool education. These cannot be measured. Coefficients should therefore be interpreted as indicating associations.

To examine how and to what extent the advantage evolves over time, we fit Model (2) at each age the children were tested: Round 2 (2006/07), Round 3 (2009/10) and Round 4 (2013/14), when the children were 5, 8 and 12 years old, respectively. Model (3) is fit only at 12 years old.

6. Results

6.1. Descriptives

The Young Lives average is higher than that of the national statistics, at 25 per cent (n=476) in 2007 versus a GER of 3.1 in 2008. This discrepancy likely owes to most of the Young Lives sample having better access to services than nationally representative samples. It is only in Addis Ababa that Young Lives households have lower access to services than the average capital city resident (see Appendix 1, and Outes-Leon and Sanchez 2008).

That preschool education in Ethiopia is predominantly an urban and private phenomenon is reflected in the Young Lives data. Table 2 shows that 23 per cent of Young Lives children attended preschool education in urban areas and 2 per cent in rural areas, and 78 per cent of those who attended preschool education attended a private school. Attendance is predominantly in Addis Ababa, representing 14 per cent of those children who attended preschool. There is also a clear socio-economic gradient, as of the 25 per cent of the children who attended preschool, the majority (13 per cent) are from the wealthiest quintile group and the next three quintiles represent the other 12 per cent. Almost none are from the poorest quintile group. Most of the children who attend preschool education speak Amharic as their first language (21 per cent), with those speaking Oromiffa and Tigray representing the other 2 and 1 per cent, respectively. Small cell sizes are a limitation in subsequent analysis.

Those who attend preschool education tend to live in smaller households than those who do not (5.5 and 6.2 people, respectively). They are also less undernourished and have caregivers with substantially higher levels of education (6.0 and 1.4 years of schooling, respectively). By age 12, they have also completed more years of education and score substantially better than their non-preschool-attending peers in all three maths tests. Although tests are not comparable across rounds, at ages 5, 8 and 12, those who attended preschool education consistently score higher than their peers who did not attend preschool education.

Figure 2 represents the kernel density estimated of students' maths scores by preschool education attendance in each round. The curve of those who did not attend preschool education is consistently to the left of those who did attend preschool education.

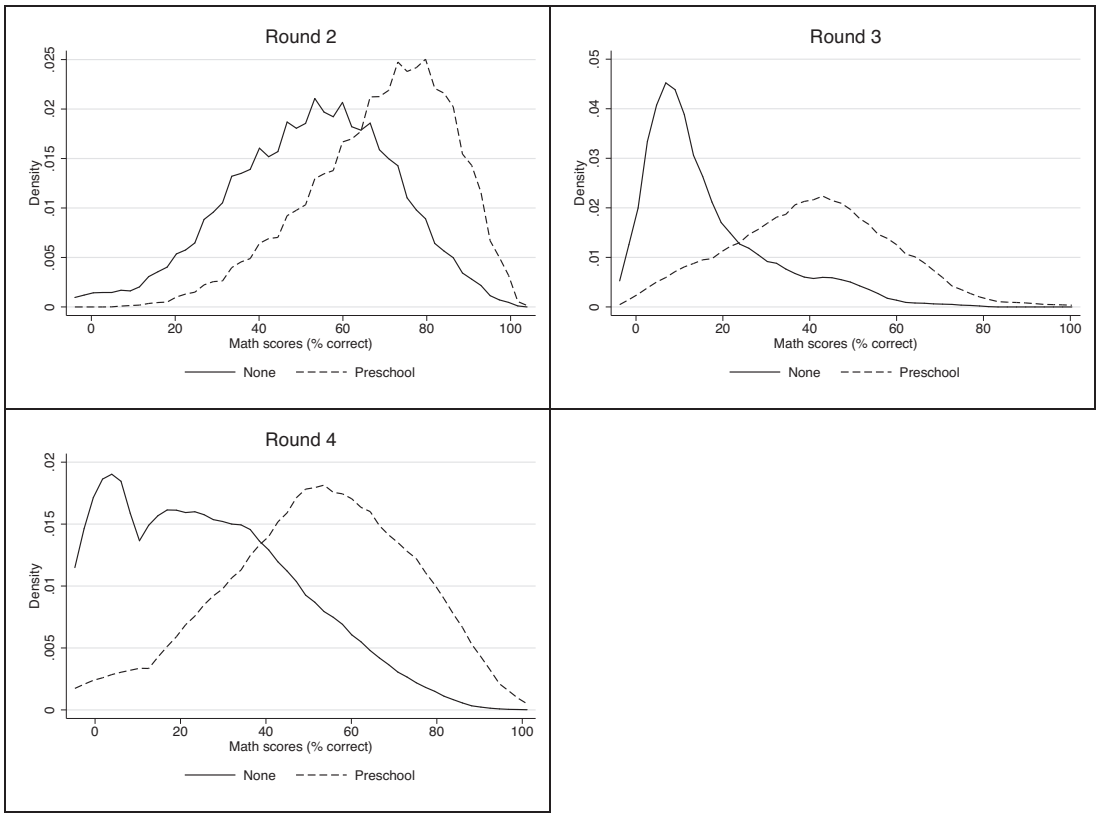
Table 2. *Descriptive statistics on preschool education attendance in 2006/07^*

	None	Preschool		None	Preschool
n	1436	476			
Gender (%)			**Preschool type (5 years, %)**		
Male	39	14	Private		78
Female	36	11	Public (Government)		5
Urban/rural (5 years, %)			Other		17
Urban	17	23	**Preschool duration (5 years, %)***		
Rural	58	2	< 6 months		15
Region (%)			6-12 months		43
Addis Ababa	1	14	> 12 months		37
Amhara	18	2	Still attending		5
Oromia	17	3	**Primary school type (12 years, %)**		
SNNP	19	6	Not enrolled	72	0
Tigray	20	0	Private	0	17
Child's first language (5 years, %)			Public (Government)	3	1
Afar	0	0	Other	0	6
Amhara	22	21	**Household size (age 5)**		
Gurage	5	0		6.2	5.5
Hadiya	5	0		(1.99)	(2.24)
Oromiffa	14	2	**Height for age (z-score)**		
Sidama	5	0	Age 1	-1.7	-1.1
Silte	0	0		(2.00)	(1.74)
Tigray	20	1	Age 5	-1.6	-1.0
Wolayta	4	0		(1.12)	(1.02)
First born (%)			Age 8	-1.3	-0.7
No	61	13		(1.18)	(1.13)
Yes	14	12	Age 12	-1.6	-1.0
Mother works (5 years, %)				(0.97)	(0.98)
Not employed	7	4	**Caregiver's education (years)**		
Agriculture	45	2	Age 5	1.4	6.0
Non-agriculture	24	18		(2.57)	(4.27)
Wealth quintile groups (1 year, %)			**Highest grade completed by child**		
Poorest	20	0	Age 12	3.1	4.6
Poorer	19	1		(2.13)	(1.28)
Middle	17	3	**Maths scores (% correct)**		
Wealthy	12	8	Age 5	52.4	68.9
Wealthiest	7	13		(18.99)	(16.49)
Expenditure quintile groups (5 years, %)			Age 8	16.6	40.5
Poorest	19	1		(14.66)	(18.07)
Poorer	18	2	Age 12	26.4	51.4
Middle	16	4		(20.95)	(21.48)
Wealthy	13	7			
Wealthiest	9	11			

Notes: * School history data used. ^ Mean reported unless otherwise indicated. s.d. in parentheses.

On average, those who did not attend preschool scored worse on the exams (see Table 2). Most worrying is the curve in Round 3, when the children were 8 years old. The peak on the left side of the maths score distribution indicates that most of those who did not attend preschool education answered more than 80 per cent of the questions incorrectly. Eight years old, in Ethiopia, is when children should be in their first years of primary school. Those who attended preschool education scored consistently better.[14] This may be partly explained by the design of the test. In Round 3 tests were the same across Young Lives countries, and were not tailored to the Ethiopian context, therefore rural children scored poorly. Round 4 tests were tailored to the Ethiopian context, perhaps explaining the differences in distributions.

Figure 2. *Kernel density estimates of students' maths scores by preschool education attendance at 5, 8 and 12 years old*



Note: The Round 2 test had 15 items, which explains the lumpy nature of the curve.

6.2. Who typically attends preschool education?

Table 3 reports the results from Model (1). It reports the log odds, which do not provide an intuitive interpretation, but do provide an indication of the direction of the relationship and its statistical significance. Given log odds unintuitive nature, here I use marginal probabilities to facilitate an interpretation of results.

14 These curves do not control for selection bias into preschool. This is further examined in Section 6.2.

Marginal probabilities reflect the probability of the outcome happening. They are similar to predicted values in OLS models, and here are calculated when all other predictors are at their mean values.[15] We can interpret the marginal probability associated with the caregiver's education, which is 0.01, as follows: for every additional year of education for the caregiver of a child with average characteristics, the probability of attending preschool increases by 1 percentage point. As the statistical significance of the marginal probabilities varies according to the value at which the other variables are set, this paper relies on the statistical significance of the log odds results, and use marginal probabilities to *illustrate* these results in an intuitive manner. When interpreting results, it is important to remember that these are marginal probabilities of children with average characteristics.

A difference of one standard deviation in height-for-age, while controlling for other factors, is associated with an increase in the child's probability of attending preschool education by 2 percentage points. While the size of this marginal probability is quite small, an increase of 1 standard deviation in height-for-age is substantial. Caregiver's decision to enrol a child into preschool education does not appear to be substantially associated with children's height-for-age at 5 years old, controlling for other factors.

Being a girl with average characteristics is associated with a 3 percentage point lower probability of attending preschool than being a boy. Being a first-born child with average characteristics is associated with, on average, a 9 percentage point higher probability of attending preschool education than a non-first-born. Although Seid and Gurmu (2015) use Young Lives data and find no evidence that indicates birth order affects the probability of primary school attendance, motivations for sending children to an optional and mostly privately provisioned preschool are likely to differ from those involved when sending a child to a mandatory and mostly publicly provided primary school.

Through an interaction factor, it is possible to examine whether first-born boys are more likely to attend preschool than first-born girls. However, the interaction is not statistically significant, indicating that in this sample it does not matter.

As is expected, a child living in a rural area is 13 percentage points less likely to attend preschool education than a child living in an urban area. Most provision of preschool education is concentrated in urban areas.

One's first language can be an important predictor of advantage in Ethiopia. However, the number of non-Amharic speaking children who attend preschool is too small to produce meaningful results (see Table 2), so I compare children who do speak with those who do not speak Amharic. Speaking Amharic as a first language, while controlling for other factors, is

15 Predicted probabilities, or marginal effects, can be calculated as 'average marginal effects' or as 'marginal probabilities at the average'. They can also be calculated as fitted probabilities, setting each category to a specific value. Here, theory does not support the selection of specific fixed values, thus average predicted probabilities are used. Average marginal effects, or average predicted probabilities, are the expected probability of a person with average characteristics. In other words, the average predicted probability of attending preschool for females is calculated for each child using that child's values of all the other explanatory variables, when sex is set equal to female. Marginal probabilities at the average are calculated holding all predictors fixed at their means. For gender, for example, rather than treating it as (1, 0) or (0, 1) depending on the observation, categorical variables are treated as (0.45, 0.55), which are the average values of females and males; in this case they represent sample proportions. In both cases, the predicted probability is compared with an abstract person. Just as 'the average person' is an abstraction of reality, it is not assured that a real person will share the mean value of explanatory variables. Here, 'marginal probabilities at the average' are reported, for the intuitive reason that, if 45 per cent of the population is female, this population ratio should be used to calculate the marginal probabilities.

associated with an increase in a child's probability of attending preschool education by 14 percentage points, compared to those who do not speak Amharic as a first language.

Being a child (with average characteristics) from the least poor expenditure quintile group (at 1 year old) is associated with a 15 percentage point higher probability of attending preschool education than one from the poorest quintile group (while holding all other variables constant, including expenditure). However, being from the other three quintile groups does not offer an advantage. This perhaps indicates that most preschool education in Ethiopia is costly and only the urban elite can access and afford it. As the wealth measure includes a component on access to services, one would expect wealth to be positively associated with preschool education attendance. However, other aspects of wealth (such as housing quality, which is not very liquid) may outweigh the access to services component in poorer quintile groups.

Controlling for other variables, a child with average characteristics with a mother who is not employed is more likely to receive preschool education, compared to a child with a mother who is employed (both in agriculture and in non-agricultural activities). As preschool education in Ethiopia is largely fee-based, the inverse results were expected, where children from employed mothers would be more likely to attend preschool. This merits further investigation.

None of age, the presence of a child-carer at home (grandparent or older sibling with caring responsibility) or household size appear to be a predictor of attendance in preschool education, when controlling for all other variables. While more people at home may limit the resources available to invest in a child, it may also mean the child is relieved of caring, domestic chores, farming or herding tasks.[16]

Factors that predict attendance in preschool education over and above other factors include being a first-born child, having a more educated caregiver, living in an urban area, speaking Amharic as a first language, belonging to the wealthiest households and having a mother who is not employed. Although being better nourished, and male are positively associated with preschool attendance, they are only modestly so. In the Young Lives Ethiopia sample, age, household size, and having a caregiver present do not appear as strong predictors of attendance in preschool education.

16 Woldehanna and Gebremedhin (2012) and Woldehanna (2016) examine determinants of preschool in urban areas and use a different set of determinants. They limit their sample to urban residents. Distinct from results from Model (1) in this paper, the authors include urban site dummies (i.e. site level fixed effects) father's education, whether the child has long-term health problems, or whether the household suffered from shocks. Unlike Woldehanna and Gebremedhin (2012) and Woldehanna (2016), Model (1) includes height-for-age, and urban-rural dummy, maternal employment, first language of the child, and a dummy for first borns. Unlike Woldehanna (2016), this model also includes maternal education. Common results are that wealth is an important predictor of preschool attendance, but Woldehanna (2016) find that parental education (measured by father's highest grade completed) has a small association (i.e. 2.3-5 per cent).

Table 3. *Determinants of attendance in preschool education between 3 and 5 years old*

Preschool attendance	(1) Log odds		(2) Marginal probability
Height-for-age (5 years, z-score)	0.30***	(0.09)	0.02
Female	-0.44*	(0.25)	-0.03
First born	0.99***	(0.30)	0.09
Age (months)	0.06	(0.04)	0.00
Caregiver's education (years)	0.17***	(0.04)	0.01
Grandparent or older sibling as carer	0.18	(0.29)	0.01
Rural (vs. urban)	-1.52**	(0.77)	-0.13
Child's first language (vs. other)			
Amharic	1.74***	(0.67)	0.14
Household size (vs. small (0-4))			
Medium (5-6)	0.06	(0.23)	0.00
Large (7+)	-0.37	(0.28)	-0.02
Wealth quintile groups – 1 year old (vs. poorest)			
Poorer	-0.19	(0.68)	-0.01
Middle	-0.41	(0.82)	-0.02
Less poor	0.83	(0.85)	0.06
Least poor	1.46*	(0.80)	0.15
Maternal employment (vs. not employed)			
Agriculture	-0.80*	(0.48)	-0.06
Non-agriculture	-0.38*	(0.20)	-0.03
Interaction (female and first born)	0.07	(0.47)	
Constant	-5.91*	(3.12)	
Observations	1,533		

Notes: Robust standard errors in parentheses. ^ All predictors at their mean value. *** p<0.01, ** p<0.05, * p<0.10.

6.3. Is preschool attendance associated with higher attainment?

The first three rows of results in Table 3 report the association between preschool education and maths scores from Model (2) at ages 5, 8 and 12. Preschool attendance is measured as a binary variable, so the coefficient reported represents the gap in maths scores between preschool attendees and non-attendees.[17] Column (1) reports essentially the correlation between preschool education and maths scores, without controlling for any other factors. In subsequent columns, it reports this association, controlling for child- and household-level controls and primary schooling-level controls (Columns (2) and (3), respectively). As the primary school enrolment age in Ethiopia is 7 years old, primary schooling-level controls are not included in the model that examines 5 year olds.[18] All results are reported with community-level fixed effects.

As expected, the size of the association decreases with additional controls but remains significant at 5 and 8 years old. At age 5, children who attended preschool education score 6 percentage points higher in their maths tests than those who do not, controlling for child- and

17 Full regression results are available in Appendix 2.

18 Including children attending primary school at age 5 does not substantially change the results (see Appendix 3).

household-level variables as well as between-community variations. This is a sizable difference, as a 5 percentage point increase is associated with having a caregiver who has ten additional years of education, controlling for other variables. Therefore, preschool appears to have an important association with 5-year-olds' maths test results.

At 8 years old, without any controls except for community-level fixed effects, preschool education offers a 12 percentage point advantage over those who do not attend preschool. After controlling for child-, household- and primary schooling-level characteristics, preschool education offers a 5 percentage point advantage. It continues to have a positive and statistically significant association with maths scores.[19]

At 12 years old any advantage disappears (i.e. it loses its statistical significance) in the full model. This is in line with what Engle et al. (2011) find in their systematic review, where the benefits of preschool education often decrease during primary school.

Specifically, the advantage of preschool education disappears when exposure to schooling is included in the model, as statistical significance appears to shift from preschool education to school exposure. This indicates that preschool education may have a direct association with maths scores, but also an indirect one through school exposure. School exposure, measured by highest grade completed, captures whether the child enrolled late in primary school, repeated grades or dropped out.

Model (3) examines the association between preschool education and exposure to schooling. The fourth row in Table 4 reports these results. Children who attended preschool education completed, on average, 0.6 more grades than those who did not attend preschool education, controlling for various child- and household-level characteristics, primary school type and between-community variation. This is in line with Woldehanna and Araya (2017), who examine the association between preschool education and secondary school completion among urban children. They find that preschool attendance is associated with a 25.7 per cent higher likelihood of completing secondary education, when compared to those who do not attend preschool.

Full regression results in Appendix 2 show that exposure to schooling is also positive and has a statistically significant association with maths scores at 8 years old, pointing to a possible shift in the importance of preschool education to exposure to schooling as children move into adolescence. Preschool education appears to be important for attainment at age 5 and 8, but the importance diminishes over time, while other factors such as exposure to schooling become more important.

19 Woldehanna and Gebremedhin (2012) and Woldehanna (2016) examine the role of preschool education on attainment at ages 5 and 8 years old (but not at 12 years old) among urban children. As an outcome variable they log maths and PPVT scores (see Appendix A1.3 for why this paper's analysis does not). They fit an OLS regression model, use propensity score matching and instrumental variable methods and use similar controls. Woldehanna (2016) also conducts mediation analysis. Unlike Model (2), they do not use community level fixed effects. Results from Woldehanna and Gebremedhin (2012) and Woldehanna (2016) are in line with those from Model (2).

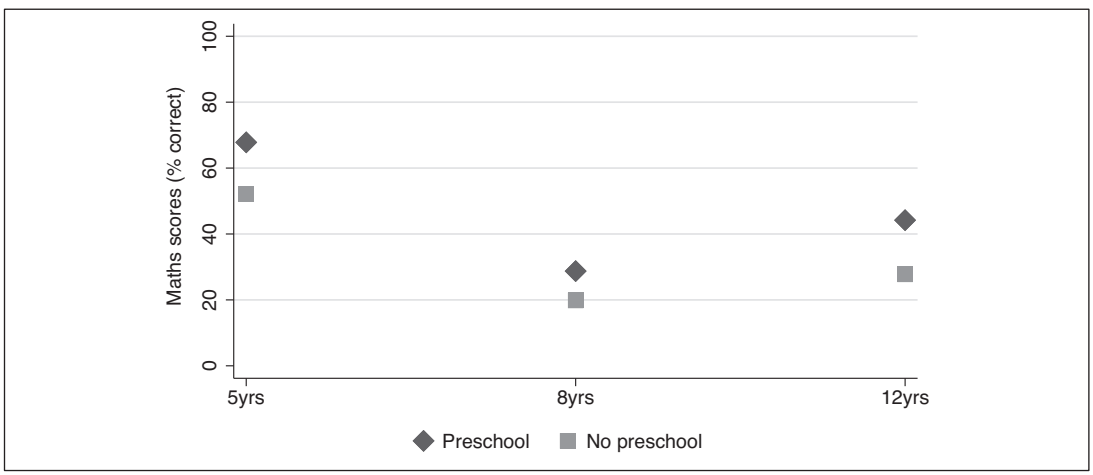
Table 4. *The role of preschool education in maths scores at 5, 8 and 12 years old*

	(1) Preschool	(2) + child and household	(3) +schooling
Gap in maths scores between preschool attendees and non-attendees (% correct)			
5 years old	8.58***	5.89***	
	(2.29)	(1.93)	
8 years old	11.87***	7.88***	5.28***
	(1.69)	(1.64)	(1.60)
12 years old	10.34***	5.19**	3.15
	(2.75)	(2.29)	(2.19)
Gap in school exposure (highest grade completed)			
12 years old	0.83***	0.55***	0.56***
	(0.16)	(0.13)	(0.13)
Observations			
Model (2) at 5 years old	1,829	1,829	
Model (2) at 8 years old	1,763	1,763	1,763
Model (2) at 12 years old	1,810	1,810	1,810
Model (3) at 12 years old	1,814	1,814	1,814
Number of communities	24	24	24

Notes: Controlling for child-, household- and primary schooling-level variables and between-community variation. Fewer children took the test at 8 years old, hence the smaller sample size. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.10.

Figure 3 reports the results in Table 3 graphically. Predicted values are the expected maths scores of a child, given her characteristics. These predicted values are calculated using the parameters estimated from the full regression model. Note that as different tests were administered to children at different ages, and the distribution of the test scores vary across rounds, results across rounds are not comparable (see Appendix 1 for a more complete discussion). Nonetheless, superiority is maintained across the age groups when controlling for child-, household- and primary schooling-level characteristics. The gaps are statistically significant only at 5 and 8 years old.

Figure 3. *Predicted maths scores by attendance in preschool education at 5, 8 and 12 years old*



Notes: These are based on predicted values and control for child-, household- and primary school-level characteristics and between-community variation. Tests across rounds are not comparable.

An interaction factor would reveal whether children from the poorest households benefited more than those from wealthier households.[20] However, as most children who attend preschool education are from wealthier backgrounds, there are not sufficient data available to address this question.

6.4. Limitations

Research from high-income countries shows that high-quality preschool education offers children benefits in terms of attainment. In the Young Lives data no objective measure of the quality of preschool education was available.[21] Also, those who attend preschool education in the Young Lives sample are distinct from the broader Ethiopian population. On average, they are more advantaged, with most living in urban areas. This raises issues of external validity – we cannot assume that if preschool education were rolled out to all children the manifest associations would be replicated. Indeed, it is hard to predict how children in rural areas or from less privileged backgrounds may benefit from preschool. Furthermore, the quality of the provision may differ from that of the private preschools. The quality of instruction in private urban preschools is likely to be higher than that in public preschools or in the O-classes rolled out in rural areas in 2015, and may therefore have a weaker association with attainment. Alternatively, if, as in high-income countries, good-quality preschools benefit the least advantaged children the most, reaching more marginalised groups may strengthen the association between preschool education and cognitive development. An opportunity exists to collect information on children attending these O-classes now, so as to be able to examine their role in children's attainment in the future.

20 The interaction factor is between preschool attendance and wealth quintile groups (lagged by one time period).

21 Carers were asked 'In your opinion, how good is the quality of the care and teaching at this preschool?' Self-reporting on the quality of a service one is paying for and sends one's child to, is likely to be biased upwards. Over 95 per cent found the school to be of excellent, good, or reasonably okay quality.

7. Conclusions

The Ethiopian government has set out to increase enrolment in preschool education from 34 per cent in 2013/14 to 80 per cent by 2020. In light of this ambitious endeavour, this paper aimed to provide some country-specific evidence on the predictors of preschool education and its role in attainment in Ethiopia. It examined who typically attends preschool education; whether children who attend preschool education have higher attainment than children who do not; and how the association between preschool education and attainment evolves over time.

The paper finds that the key predictors of preschool education are being a first-born child, speaking Amharic as a first language, having a more educated caregiver, belonging to a household with higher wealth, living in an urban area, and having a mother who is not employed. Although being male and better nourished are positively associated with preschool attendance, they are only modestly so. To policymakers intent on expanding access to preschool education, these findings provide insight into the barriers to accessing preschool education faced by families. Most point to structural inequalities: for example, wealth inequality, inequality between rural and urban areas, structural advantages of speaking Amharic as a first language, gender discrimination, and preference towards first-borns. But the sheer lack of access to preschool education outside of urban areas also points to an imbalance in provision.

The paper also examined the association between preschool education and children's attainment, while controlling for a rich array of child-, household- and primary school-level characteristics and variations between communities. The analysis finds that preschool education has a direct and substantial association with attainment, as well as potentially an indirect one through early enrolment, grade progression and dropout rates (captured in highest grade completed).

Preschool education has a positive and significant role at 5 and 8 years old, but this association fades away by 12 years old. At 8 years old, exposure to schooling emerges as an intermediary outcome associated with attainment, and supersedes that of preschool education by age 12. Further analysis shows that preschool education has a direct association with school progression by age 12. Official documents discussing Ethiopia's preschool education policy identify improvements in school exposure as a reason for investing in preschool education. Policymakers can therefore be reassured that these results indicate that preschool education would indeed support this. These findings indicate that investments in preschool education may pay off doubly – both directly towards attainment and indirectly through exposure to schooling.

References

Alemu, T., G. Asgedom, J. Liebenberg, A. Mekonnen, J. Seager, B. Tefera, and T. Woldehanna (2003) 'Young Lives Preliminary Country Report: Ethiopia', Oxford: Young Lives.

Cueto, S., and J. León (2012) *Psychometric Characteristics of Cognitive Development and Achievement Instruments in Round 3 of Young Lives*, Technical Note 25. Oxford: Young Lives.

Dornan, P., and M. Woodhead (2015) *How Inequalities Develop through Childhood: Life Course Evidence from the Young Lives Cohort Study,* Innocenti Discussion Paper 2015-01. Florence: UNICEF Office of Research.

Engle, P.L., M.M. Black, J.R. Behrman, M. Cabral de Mello, P.J. Gertler, L. Kapiriri, R. Martorell, and M.E. Young (2007) 'Strategies to Avoid the Loss of Developmental Potential in More than 200 Million Children in the Developing World', *Lancet* 369.9557: 229-242.

Engle, P.L., L.C.H. Fernald, H. Alderman, J. Behrman, C. O'Gara, A. Yousafzai, M.C. de Mello, M. Hidrobo, N. Ulkuer, I. Ertem, and S. Iltus (2011) 'Strategies for Reducing Inequalities and Improving Developmental Outcomes for Young Children in Low-Income and Middle-Income Countries', *The Lancet* 378.9799: 1339-1353.

Espinoza, P. (2014) 'Young Lives Rounds 1 to 3 Constructed Files', Oxford: Young Lives.

Grantham-McGregor, S., Y.B. Cheung, S. Cueto, P. Glewwe, L. Richter, and B. Strupp (2007) 'Developmental Potential in the First 5 Years for Children in Developing Countries', *The Lancet* 369.9555: 60-70.

Havnes, T. (2012) 'Comment on Ruhm and Waldfogel: Long-term Effects of Early Childhood Care and Education', *Nordic Economic Policy Review* 1.2012: 53-60.

Heckman, J.J. (2008) 'Schools, Skills and Synapses', *Economic Inquiry* 46.3: 289-324.

Heckman, J.J., and D.V. Masterov (2007) 'The Productivity Argument for Investing in Young Children', *Applied Economic Perspectives and Policy* 29.3: 446-493.

Heckman, J.J., and L.K. Raut (2013) *Intergenerational Long Term Effects of Preschool - Structural Estimates from a Discrete Dynamic Programming Model*, NBER Working Paper No. 19077, Cambridge: NBER.

Hoot, J.L., J. Szente, and B. Mebratu (2004) 'Early Education in Ethiopia: Progress and Prospects', *Early Childhood Education Journal* 32.1: 3-8.

Howe, L.D., J.R. Hargreaves, S. Gabrysch, and S.R.A. Huttly (2009) 'Essay: Is the Wealth Index a Proxy for Consumption Expenditure? A systematic review', *Journal of Epidemiology and Community Health* 63.11: 871-877.

Irwin, L.G., A. Siddiqi, and C. Hertzman (2007) 'Early Child Development: A Powerful Equalizer', Vancouver: Human Early Learning Partnership (HELP).

McLanahan, S. (2004) 'Diverging Destinies: How Children are Faring Under the Second Demographic Transition', *Demography* 41.4: 607-627.

MoE (Ministry of Education) (2002) 'The Education and Training Policy and its Implementation', Addis Ababa: Ministry of Education.

MoE (Ministry of Education) (2008) 'Education statistics Annual Abstract 2000 E.C. (2006/07)', Addis Ababa: Ministry of Education.

MoE (Ministry of Education) (2012) 'Education Statistics Annual Abstract 2005 E.C (2011/12)', Addis Ababa: Ministry of Education.

MoE (Ministry of Education) (2015) 'Education Sector Development Program (ESDP) V', Addis Ababa: Ministry of Education.

NICHD (National Institute of Child Health Human Development) Early Child Care Research Network, and G.J. Duncan (2003) 'Modeling the Impacts of Child Care Quality on Children's Preschool Cognitive Development', *Child Development* 74.5: 1454-1475.

Nores, M., and W.S. Barnett (2010) 'Benefits of Early Childhood Interventions Across the World: (Under) Investing in the Very Young', *Economics of Education Review* 29.2: 271-282.

Orkin, K., W.A. Yadete, and M. Woodhead (2012) *Delivering Quality Early Learning in Low-Resource Settings: Progress and Challenges in Ethiopia*, Working Papers in Early Childhood Development 59, The Hague: Bernard van Leer Foundation.

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

Ruhm, C., and J. Waldfogel (2012) 'Long-Term Effects of Early Childhood Care and Education', *Nordic Economic Policy Review* 1: 23-51.

Seid, Y., and S. Gurmu (2015) 'The Role of Birth Order in Child Labour and Schooling', *Applied Economics* 47.49: 5262-5281.

Thompson, R.A., and C.A. Nelson (2001) 'Developmental Science and the Media: Early Brain Development', *American Psychologist* 56.1: 5-15.

UN (United Nations) (2013) 'Ethiopia Country Profile (Social Indicators: Urban Population %)', https://data.un.org/CountryProfile.aspx?crName=Ethiopia (accessed 1 June 2015).

Walker, S.P., T.D. Wachs, S. Grantham-McGregor, M.M. Black, C.A. Nelson, S.L. Huffman, H. Baker-Henningham, S.M. Chang, J.D. Hamadani, B. Lozoff, J.M. Gardner, C.A. Powell, A. Rahman, and L. Richter (2011) 'Inequality in Early Childhood: Risk and Protective Factors for Early Child Development', *The Lancet* 378.9799: 1325-1338.

Wilson, I., S.R.A. Huttly, and B. Fenn (2006) 'A Case Study of Sample Design for Longitudinal Research: Young Lives', *International Journal of Social Research Methodology* 9.5: 351-365.

Woldehanna, T. (2011) 'The Effects of Early Childhood Education Attendance on Cognitive Development: Evidence from Urban Ethiopia', paper presented at the CSAE Conference on Economic Development in Africa, Oxford, 20-22 March.

Woldehanna, T. (2016) 'Inequality, Preschool Education and Cognitive Development in Ethiopia: Implications for Public Investment in Pre-primary Education', *International Journal of Behavioral Development* 40.6: 509-516.

Woldehanna, T., and M.W. Araya (2017) *Early Investment in Preschool and Completion of Secondary Education in Ethiopia: Lessons From Young Lives*, Working Paper 168*,* Oxford: Young Lives.

Woldehanna, T., and L. Gebremedhin (2012) *The Effects of Pre-school Attendance on the Cognitive Development of Urban Children aged 5 and 8 Years: Evidence from Ethiopia*, Working Paper 89, Oxford: Young Lives.

Woldehanna, T., A. Mekonnen, and T. Alemu (2008) 'Young Lives: Ethiopia Round 2 Survey Report', Oxford: Young Lives.

Woodhead, M., P. Ames, U. Vennam, W. Abebe, and N. Streuli (2009) *Equity and Quality? Challenges for Early Childhood and Primary Education in Ethiopia, India and Peru*, Studies in Early Childhood Transitions, The Hague: Bernard van Leer Foundation.

Young Lives (2006) 'The Young Lives Study Context Instrument (Questionnaire – Round 2)', Oxford: Young Lives.

Young Lives (2009) 'Procedures and Instructions for the Administration of Cognitive Development and Achievement Instruments - Round 3 - Fieldworker's Manual (Peru)', Oxford: Young Lives.

Appendices

Appendix 1. Data and measures

A1.1. Data

Wilson et al. (2006) provide a detailed justification of the Young Lives sampling approach. The approach entailed a multi-stage sampling strategy of purposive then random sampling. First, 20 sites were purposefully selected to over-sample areas with food deficiency; capture Ethiopia's diversity across regions and ethnicities, in urban and rural areas; and keep sampling costs manageable, which entailed selecting accessible, rather than remote, sites (Outes-Leon and Sanchez 2008). Figure A4.1 shows the locations of these sites.

Second, within the sites eligible households were randomly selected. Eligibility criteria were having a child aged 6-17 months old who was a singleton.[22] The Young Lives household questionnaire collects basic information on all household members, as well as information on the household (expenditure, housing quality, etc.); a child questionnaire collects information on child-related topics of interest.[23]

Cognitive development is affected by several factors, including child- and household-level variables and the broader social, cultural and political environment. Some of this can be measured, some cannot. The Young Lives dataset provides a wealth of information on children's trajectories across various domains (including cognitive, socio-economic, demographic and access to services domains) and at the child, household and community level. This richness in the data and their longitudinal nature allows us to control for many factors and to address some kinds of omitted variable bias.

Comparisons between the 2002 Young Lives data and the 2000 Demographic and Health Survey (DHS) show Young Lives households in all rural and urban areas, barring Addis Ababa, were wealthier and had better access to services. In Addis Ababa, Young Lives households were poorer and had less access to services. The DHS wealth measure is based on an index that includes access to services. Comparisons with the 2000 Welfare Monitoring Survey (WMS) show Young Lives households were poorer based on several assets that did not necessarily include access to services: land, home and livestock ownership. In terms of access to services, comparisons with the 2000 WMS were similar to those with the 2000 DHS.[24] These comparisons are in line with the sampling design, which prefers poorer areas with better road access.

Young Lives is not, nor was it envisioned as, a randomised control trial. Attendance in preschool education was not randomly allocated, nor was there an exogenous preschool-associated policy that would enable causal estimates to be drawn. Results discussed in this paper are therefore associations between preschool education and cognitive development.

22 Twins and triplets were excluded.

23 Young Lives also uses a community questionnaire that collects community-level information. We do not use this in the data analysis.

24 Young Lives households in Addis Ababa had less access to services than the average resident, but rural households had better access than the average rural resident, according to the WMS.

A1.2. Measures

Child level

Preschool education: Caregivers were asked whether the child had attended preschool education since the age of 3.[25] They were also asked what type of preschool education the child attended and for how long. A few children repeatedly attended two preschools, and data on both preschools were collected. This paper uses data on the preschool most recently attended.

The types of preschool education used include private, public (i.e. government) and other. As Orkin et al. (2012) suggest, public fee-paying schools and private schools are grouped together under the heading 'private'. Government preschools are renamed 'public preschools'. The 'other' category includes NGO, charity, religious, community and other types of preschools. In the Young Lives sample, 78 per cent of children attended private preschools, with only 5 per cent in public preschools and the remaining 17 per cent in other types of preschools (see Table 2).

The binary variable, whether the child has attended preschool education or not, is the most useful for analysis. The private–public–other categorisation seems less useful, given the small attendance at public preschool. Information in the data is not sufficient to distinguish between the role of public preschool education and that of private preschool education. Additionally, while it does provide enough information to distinguish between the role of private preschool education and that of 'other' types of preschool education, the policy relevance of this distinction is limited, given that 'other' is such a heterogeneous group.

Quality: Preschool education quality would be an important measure to consider in cognitive development, yet it is not easy to gauge. When parents are asked about the quality of the preschool education the child receives, in high-income countries, the evidence shows parents almost always report their child as receiving high-quality care (NICHD and Duncan 2003). However, most parents are not qualified to judge the difference between high- and low-quality care. Further, if parents are paying for preschool education, it is in their interests to believe the quality of education is high. Given these problems with this measure of preschool education quality, this paper does not use data on quality.

Maths scores: These measure the percentage of correct answers on three different maths tests given at ages 5, 8 and 12 years old.

Maths test: Round 2 – Cognitive Development Assessment-Quantity (CDA-Q)

Children are asked to choose an image from a selection of three or four that best represents a concept communicated by the examiner (e.g. few, most, nothing, etc.). This component has 15 items administered to the child. CDA-Q was not continued into Round 3 as it would have been too easy for most children, who would have been in primary school at the time of testing (Cueto and León 2012).

25 The specific question is, 'Since the age of 36 months, has the child regularly attended a formal or informal preschool, i.e. for a whole morning, afternoon, evening or night almost every week?'

Maths test: Rounds 3 and 4

The maths test applied to the Younger Cohort in Rounds 3 and 4 had two aims: to measure basic quantitative and number notions and to measure ability to perform basic mathematics operations. The first section included items on counting, number knowledge, number discrimination and using basic operations (e.g. 2 x 4).[26] The second section included items using numbers for addition, subtraction, multiplication and division.[27] There is some overlap of questions between the Round 3 and the Round 4 tests. Although each child took the test at her own pace, the test was discontinued at a predetermined time.

Interpreting results

The percentage of correct answers is the most useful way to measure cognitive development based on the maths test (see Section A1.3). In a regression with this measure as the outcome variable, the regression coefficient can be interpreted as changes in percentage points. Between-group comparisons within rounds are valid. However, given the tests in all rounds are distinct, comparisons across rounds are not.[28]

Height-for-age (z-score): Under-nutrition is reflected in a child's height-for-age. Each child's height is measured then standardised against a World Health Organization (WHO) defined reference population.[29] A child is classified as stunted when she lies below minus 2 standard deviations from median height-for-age of the reference population. Height-for-age is a reflection of physical development, and is linked inextricably to cognitive development. Grantham-McGregor et al. (2007) use a physical development indicator as a proxy for cognitive development in their analysis. A measure of height-for-age is therefore included in the regression model, as measured within the round.

Age: As data collection occurred over a period of several months, the child's age may explain variation in maths scores, as an older child has had more time to learn maths skills. Thus, a variable representing the child's age within each round is included in the model. The child's age is calculated by subtracting the date of the interview from the child's reported date of birth and rounding it to the nearest months.

Gender: Gender is an important aspect in children's development across the globe, and Ethiopia is no exception.

First born: There are several reasons why parents may invest more in the human capital of their first-born child. Two key reasons are having more disposable income, and cultural reasons. Using data for Round 1, we mark whether the child is a first-born child or not.

First language: Base on Round 2 data, defined by the child's first tongue. The cell sizes are very small for children attending preschool at age 5, who do not speak Amharic as a first

26 In Round 3, non-numerical questions were read to the child, to ensure maths skills were being tested rather than reading skills. These constituted nine out of the 29 questions (Young Lives 2009). In Round 4 children who could not read were not administered the maths test (personal correspondence).

27 The Rounds 3 and 4 tests had 29 and 28 questions, respectively.

28 Including the previous period's test score, as this would essentially control for factors contributing to cognitive development up until that point, and therein in preschool. It is therefore inappropriate to include the previous period's test scores in this analysis.

29 Although underweight is sometimes used to measure under-nutrition, stunting (height-for-age) is a more stable indicator, as it varies less across time. A child's weight-for-age may vary across seasons: she may gain weight after the harvest season but go hungry right before the harvest.

tongue (see Table 2). For Model (2), I therefore construct a binary variable that identifies those who speak Amharic as a first language. This is only needed for Model (2), as the outcome variable is preschool attendance, and therefore cell sizes are a concern. For Model (3) the outcome variables are maths test scores and completed schooling, where cell sizes are less of a concern.

Language of maths test: The Young Lives sample includes children speaking nine different languages as their mother tongue, and the test was administered in seven different languages.[30] Children were allowed to choose the language of the test. To control for differences across test, results are compared within language groups. For example, in translation, one question may be easier in one language than in the other.[31] There may also be cultural differences between language groups that make one questions easier for one group than another. No maths tests were available in Afar, Gurage, Silte and Wolayta.

Took maths test in mother tongue: If a child is not comfortable with the test language, this will be a disadvantage. To address this, a dummy variable is included to capture when the test language corresponded with the mother tongue of the child, for each test.

Time spent studying: Information is not available on how supportive the child's home environment is to learning and schooling, or how interested the child is in learning. While time spent studying will not fully capture how supportive a child's home environment is or their interest in learning, it is likely to be correlated to these, hence its inclusion in the model. This variable is used in the regressions when the child is primary-school age (8 and 12 years old).

Household level

Caregiver's level of education: Caregiver's level of education is likely to influence the child's cognitive development in several direct ways – such as the quantity and quality of stimulation the child receives at home and the level of vocabulary the child is exposed to – as well as in indirect ways, such as contributing to the child's physical development by being better informed on services available and having the wherewithal to seek them out. The Young Lives data report on the level of education (in years) of the child's primary caregiver. Data from Round 2 (2006/07) are used, at the time the children were of preschool age. While this does not reveal the quality of education received, whether the quality of primary education in one part of the country is equivalent to that in another part, or whether caregivers actually attended for the full academic year, it is the best measure available in the Young Lives dataset.

Household size: Household size may be negatively or positively associated with cognitive development. In high-income countries, larger households are generally associated with lower cognitive development. Household size is measured within each round, by simply summing up the number of people reported to be residing in the household.

Urban–rural and region: Type of site and the region the child lived in as reported in Round 2.

Household resources: Attending preschool is likely to be affected by parental resources, particularly in Ethiopia, given the preponderance of fee-paying preschool education. Young

30 Seven test languages are listed, with an additional 'other', which represents 5, 0.05, and 0 per cent of the Rounds 2, 3 and 4 tests, respectively.

31 This is more important for reading and vocabulary tests.

Lives collects data on household expenditure and a set of assets belonging to the household. The asset data are used to construct a wealth index. While a wealth index may represent a household's long-run material well-being, and a consumption index a shorter-run measure of material well-being,[32] in this sample they also appear to capture similar aspects of material well-being – with a very high correlation, greater than 0.90.

I therefore use the wealth index in this analysis as I contend it is a more stable and robust measure of material well-being in Ethiopia. While aspects of the wealth index are directly observable, expenditure relies on the respondent to remember everything consumed in the last two weeks. Additionally expenditure among poor and subsistence-farming communities can be very seasonal, so high expenditure at the time of the survey may not reflect the household's reality throughout the year.[33]

Wealth quintile groups

The Young Lives wealth index is essentially a sum-score, made up of three equally weighted elements: housing quality, consumer durables and service quality. Each element is made up of a subset of assets and scaled to have a range of 0-1. The wealth index is an average of the scores across these three elements, each representing one-third of the final index.

Maternal employment: One reason for sending children to preschool education may be maternal employment. Maternal employment may also benefit the cognitive development of the child through indirect channels, for example increased income and/or increased decision-making power of the mother. To separate the role of maternal employment in cognitive development from that of preschool education, we include maternal employment in the regression model. Maternal employment is measured in Round 2, and grouped into not employed, agriculture and non-agriculture. Agricultural employment includes the following categories: self-employed (food crops), self-employed (non-food agriculture), self-employed (aquaculture), self-employed (livestock), wage employment (agriculture), annual farm servant and other agriculture. Non-agricultural employment includes: self-employed (manufacturing), self-employed (services), self-employed (business), self-employed (business), wage employment (non-agriculture), regular salaried employment and house maid. No employment includes: unemployed, household chores, other unpaid activity, household dependent, begging and other unpaid activity.

Child carer in household (grandparent or sibling): If there is a carer at home, other than the mother and father, then it may be the case that the parents choose not to send the child to preschool given there is already childcare at home. Young Lives collects information on the three main people who cared for the child before she was 3 years old. I am interested in the main alternative to the mother and father as the carer in the household, and therefore only code those children where a grandparent or an older sibling took care of the child in the

32 Long-term improvements in consumption levels are likely to result in improved material circumstances of the household, and would thus be captured in the asset index. A consumption measure is more likely to be immediately affected by shocks and adverse events.

33 To measure expenditure the Young Lives surveys ask for information on food and non-food household expenditure. Datasets include total real consumption per capita at 2006 prices. Assessing a household's consumption is quite time-consuming; the Living Standards Measurement Survey (LSMS) consumption module usually takes over an hour to complete (Howe et al. 2009: 876) and is often omitted in low-income country education and health questionnaires. Having both household expenditure and a good developmental outcome variable is unique in the Young Lives databases. We use logged total per capita expenditure in real terms (2006) to group children into expenditure quintile groups. See Espinoza (2014) for a detailed description of how expenditure data were calculated.

first or second instance. In the first instance, it is mostly the mother who cared for the child (89.38 per cent), followed by grandmother (3.71 per cent) and father (3.14 per cent). In the second instance, it was mostly the father (43.57 per cent) followed by an older sister (21.6 per cent). This is a better measure than whether there is an older sibling or grandparent in the household, because we are not sure whether this older sibling or grandparent cares for the child.

Primary schooling level

Primary school type: Attainment is likely to be shaped by whether a child attends primary school, and, if she does, the quality of the schooling. In Ethiopia, public sector primary schools dominate, with a growing private sector. Type of primary school is: none, private, public or other. This variable therefore covers dropout rates, as it includes the 'none' category. In Round 4, at 12 years old, 87 per cent of children attended public primary schools, 6 per cent private schools, and 2 per cent other schools (including NGO, community and religious schools), while 5 per cent were not enrolled.

Exposure to schooling: Cognitive development will also be influenced by exposure to schooling: whether the child enrolled late to primary school, repeated grades or dropped out. This is reflected in the highest grade completed by the child. Although the variation of highest grade completed is at the child level, this analysis includes highest grade completed under primary schooling-level factors (along with primary school type) as these are intermediary factors that occur after attendance in preschool education and are related to primary schooling.

Community level

In Ethiopia, communities were defined as 'peasant association' in rural areas and *kebele* in urban areas (Young Lives 2006). As differences between communities may account for differences in cognitive development, community-level fixed effects are used to account for between-community variations. For example, whether one community has a health centre and the other does not may benefit the cognitive development of children in the community with the centre, and community-level fixed effects accounts for this.

Community variables from Round 2 are used, the age at which the child attended preschool education. In the Young Lives Ethiopia data, there are 24 communities (see Appendix 4 for a brief description of communities). Those who migrated between Round 1 and Round 2 are excluded from the analysis.

A1.3. Why use percentage correct, and not z-scores or logged scores?

Z-scores are calculated by subtracting the mean from the percentage correct then dividing by the standard deviation. When using a z-score, regression results should therefore be interpreted in changes of a standard deviation. Several authors use z-scores to compare across countries and time (Engle et al. 2011). Doing so assumes the denominators (i.e. standard deviation) in each round are the same. However, examining the distribution of test scores by round, it is apparent that these are distinct (see Figure A1.1 and Table A1.1). The standard deviation of the 12-year-old's test is larger than that of the 8-year-old's test (23.7 and 18.7, respectively). Therefore, the same absolute difference at 5 years old will look less effective at 12 years old. For example, a 10 percentage point difference between two respondents is equivalent to a change in the z-score of 0.53 and 0.42 for Round 3 and

Round 4, respectively. Results are therefore only comparable within rounds. The rationale for standardising the outcome variable is therefore obsolete. Hence the use of percentage correct as the outcome variable of the analysis here.

Based on the right-hand skew of the Round 3 and Round 4 maths score distributions in Figure A1.1, one may be inclined to log the maths results, as done by Woldehanna and Gebremedhin (2012). However, splitting the distribution between attendance and non-attendance in preschool education shows it is only those who did not attend preschool who scored poorly on the maths test; those who did attend have a relatively normal distribution. If one were to log the maths scores, this transformation would result in notable left-hand skewness (see Figure A1.2). Therefore, the percentage correct is used as the outcome variable, as logging the maths scores does not add value to the analysis. Additionally, using the percentage of correct answers rather than z-scores or logged scores offers an intuitive interpretation to results.

Figure A1.1. *Distribution of maths scores (% correct) compared with the normal distribution at 5, 8 and 12 years old*

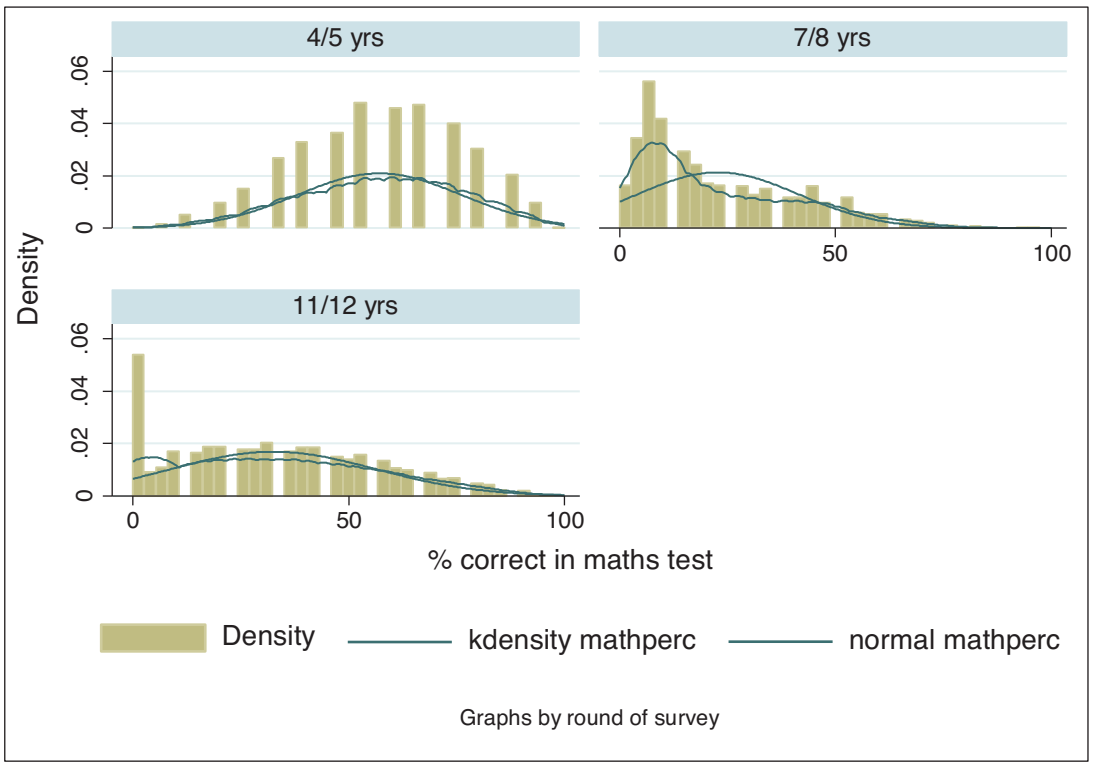
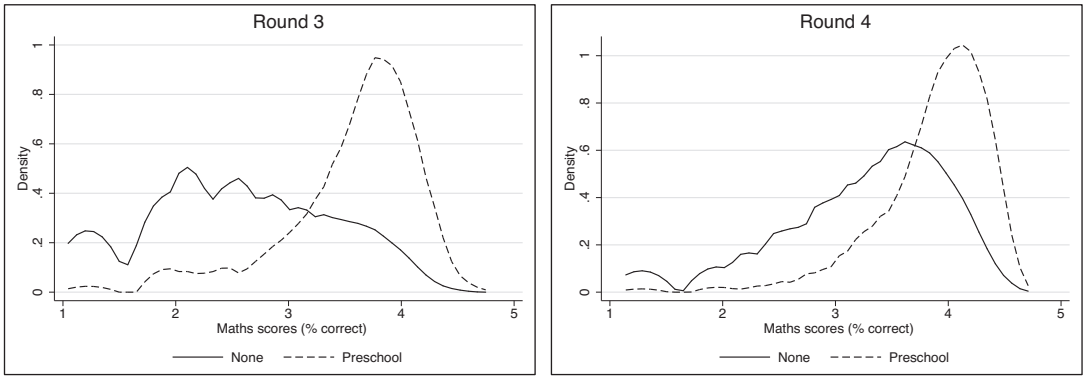


Table A1.1. *Descriptive statistics of maths scores, by round at 5, 8 and 12 years old*

	Mean	SD
5 years old	57.1	19.03
8 years old	22.8	18.69
12 years old	32.6	23.67

Figure A1.2. *Kernel density estimates of logged student's maths scores, by preschool education attendance at 8 and 12 years old*



Appendix 2. Full regression results

Table A2.1. *Full regression results – does attending preschool education offer an advantage in attainment?*

	Maths scores (% correct)								Highest grade completed		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	Pre-school	5 years old + child and hhld	Pre-school	8 years old + child and hhld	+sch	Pre-school	12 years old + child and hhld	+sch	Pre-school	12 years old + child and hhld	+ sch type
Preschool attendance	8.58***	5.89***	11.87***	7.88***	5.28***	10.34***	5.19**	3.15	0.83***	0.55***	0.56***
	(2.29)	(1.93)	(1.69)	(1.64)	(1.60)	(2.75)	(2.29)	(2.19)	(0.16)	(0.13)	(0.13)
Height-for-age		1.44**		1.26***	0.62***		1.76***	0.93**		0.25***	0.25***
		(0.53)		(0.31)	(0.17)		(0.32)	(0.42)		(0.05)	(0.05)
Female		0.03		-0.21	-1.00		-0.25	-1.00		0.12	0.12
		(0.78)		(0.68)	(0.58)		(1.08)	(1.00)		(0.10)	(0.09)
First born		-1.13		-1.74***	-1.67***		-0.44	-0.56		-0.03	-0.04
		(0.94)		(0.59)	(0.57)		(1.08)	(1.01)		(0.09)	(0.09)
Age (months)		0.71***		0.52***	0.17*		0.39***	0.20		0.06***	0.06***
		(0.10)		(0.11)	(0.09)		(0.10)	(0.13)		(0.01)	(0.01)
Ethnicity (vs. Amhara)											
Afar		2.46		2.04	6.17		2.77	3.27		-0.64*	-0.61*
		(3.28)		(4.41)	(5.07)		(10.25)	(8.97)		(0.32)	(0.32)
Gurage		-3.89***		0.29	0.83		-2.74***	-0.88		-0.31***	-0.27***
		(0.51)		(0.60)	(0.50)		(0.51)	(0.58)		(0.07)	(0.05)
Hadiya		6.24		-4.50*	-1.89		10.06***	7.09**		0.63***	0.61***
		(3.84)		(2.57)	(1.98)		(3.17)	(3.13)		(0.19)	(0.19)
Oromiffa		-2.77**		1.86	0.75		1.66	1.41		0.00	-0.00
		(1.15)		(2.66)	(2.17)		(3.96)	(3.62)		(0.21)	(0.22)
Sidama		-7.76*		-10.68***	-6.99***		-3.96	-3.82		-0.20	-0.25
		(3.83)		(1.74)	(1.36)		(2.65)	(2.56)		(0.22)	(0.21)
Tigray		-22.13***		28.55***	31.27***		2.49	5.22**		0.03	-0.03
		(4.88)		(2.18)	(2.26)		(2.62)	(2.21)		(0.22)	(0.23)
Wolayta		1.17		-10.15	-8.14**		8.71	12.03		-0.70***	-0.71***
		(5.07)		(6.42)	(3.38)		(8.62)	(8.22)		(0.16)	(0.17)
Test in mother tongue		-2.04		0.53	1.23		-3.22	-3.04		-0.10	-0.08
		(3.49)		(2.12)	(1.49)		(2.49)	(2.15)		(0.13)	(0.14)
Missing		-43.75***		-0.72	-0.40		-22.90***	-9.08**		-2.01**	-1.70
		(10.74)		(0.56)	(0.41)		(3.63)	(3.53)		(0.84)	(1.05)
Hours/day spent studying outside school				3.31***	1.73***		4.12***	2.24***		0.29***	0.26***
				(0.59)	(0.56)		(0.57)	(0.47)		(0.05)	(0.05)
Caregiver's education (years)		0.46**		0.44***	0.38***		0.52***	0.33**		0.02**	0.03**
		(0.17)		(0.11)	(0.08)		(0.17)	(0.14)		(0.01)	(0.01)
Household size		-0.20		0.01	0.10		-0.04	-0.08		0.01	0.01
		(0.20)		(0.16)	(0.16)		(0.23)	(0.17)		(0.02)	(0.02)
Wealth quintile groups^ (vs. poorest)											
Poorer		-0.53		0.00	-0.44		0.69	0.22		0.09	0.09
		(1.06)		(0.92)	(0.74)		(1.22)	(1.13)		(0.15)	(0.14)
Middle		-2.47		2.01**	1.49**		4.91***	3.90**		0.25	0.24
		(1.53)		(0.81)	(0.62)		(1.67)	(1.82)		(0.15)	(0.14)
Less poor		-0.37		2.90**	1.72		4.99**	3.65*		0.27**	0.26**
		(1.96)		(1.21)	(1.19)		(1.93)	(2.07)		(0.12)	(0.11)
Least poor		0.77		5.65***	2.55*		12.08***	8.38***		0.48**	0.52***
		(2.17)		(1.43)	(1.44)		(1.82)	(2.14)		(0.18)	(0.18)
Missing		-1.21		-2.21	-0.76		14.73	11.37*		-0.18	-0.00
		(2.84)		(1.70)	(1.95)		(9.41)	(6.51)		(0.34)	(0.37)

	Maths scores (% correct)								Highest grade completed		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	Pre-school	5 years old + child and hhld	Pre-school	8 years old + child and hhld	+sch	Pre-school	12 years old + child and hhld	+sch	Pre-school	12 years old + child and hhld	+ sch type
Maternal employment (vs. not employed)											
Agriculture		3.09		1.34	1.51*		1.68	0.55		0.24*	0.22
		(2.27)		(0.87)	(0.86)		(1.11)	(1.06)		(0.13)	(0.14)
Non-agriculture		1.01		3.93***	3.61***		3.17**	2.45**		0.21**	0.18*
		(1.60)		(0.93)	(0.95)		(1.14)	(1.12)		(0.09)	(0.09)
Missing		0.25		2.49*	2.42*		2.31	2.11		0.02	0.02
		(1.81)		(1.30)	(1.27)		(1.61)	(1.52)		(0.11)	(0.10)
Primary school type (vs. not enrolled)											
Private					11.91***			24.44***			0.00
					(1.74)			(3.03)			(0.34)
Public					4.00***			12.28***			0.37
					(0.95)			(2.87)			(0.31)
Other					6.54***			18.25***			0.07
					(2.11)			(3.35)			(0.35)
Missing					0.40			10.21**			-0.81***
					(0.85)			(4.51)			(0.18)
Grade completed					7.30***			3.20**			
					(0.97)			(1.25)			
Constant	54.27***	19.18**	19.36***	-42.81***	-15.29*	29.77***	-35.25**	-26.44	3.29***	-6.40***	-6.79***
	(0.57)	(8.18)	(0.42)	(12.13)	(8.60)	(0.67)	(16.90)	(19.33)	(0.04)	(1.18)	(1.04)
Observations	1,829	1,829	1,763	1,763	1,763	1,810	1,810	1,810	1,814	1,814	1,814
R-squared	0.01	0.08	0.05	0.17	0.30	0.02	0.13	0.23	0.02	0.10	0.11
Number of communities	24	24	24	24	24	24	24	24	24	24	24

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.10. ^ Lagged one time period, expenditure data was not collected at 1 year old.

Appendix 3. Sensitivity analysis results

In Ethiopia, early enrolment in primary school is not uncommon. Sixty-two children did not attend preschool between 3 and 6 years old but did attend primary school (primary school in Ethiopia begins at 7 years old). Of these, two-thirds have caregivers with no education, there are more than twice as many children from the poorest wealth quintile than the wealthiest, and all of their mothers work. In short, they are poorer, from poorly educated households, and have working mothers.

Hence, the sample is expanded to include both children who attended preschool education and primary school at preschool age, and a variable is included in the model to identify the difference in maths tests between early primary school attendance and preschool education attendance. Table A3.1 presents the results. The coefficients on preschool education stay essentially the same. Interestingly, however, children who enrol early in primary school score 9 percentage points lower in the maths test at 8 years old than those who do not, controlling for other variables. Irrespective of whether they attend preschool education, early enrolment in primary school does not appear to offer an advantage, when controlling for wealth, caregiver's education, maternal employment in agricultural or non-agricultural, and more.[34] Conversely, it is associated with substantially worse maths scores at 8 years old. These results suggest that it is important that the rollout of preschool education in Ethiopia be implemented with care, so that preschool education is not seen as equivalent to early primary school enrolment.

In Round 4, a full school history was collected for each child, listing by year, from birth until 12 years old: their age in years; whether they attended school; and if they did attend school, the grade and what type of school they attended. School history data were cross-checked with Model (2) results on cognitive development. The results (Table A3.1) point to a substantial difference. This may be for two reasons. First, the child may have received preschool education at age 6, after the Round 2 survey was conducted. Based on the school history data, 111 children who had not attended preschool education in Round 2 were reported to have started preschool education at or after the age of 6. The second reason is an increase in the measurement error. Caregivers were required to think back nine years ago, to when the child was 3 years old, and therefore may have had problems correctly recalling whether their child had attended preschool or not.[35]

Two additional sensitivity analyses are conducted, examining only the urban population as well as including expenditure (lagged by one time period) in the model. In both instances the results remained substantively unchanged.

34 The early enrolment variable is not significant at age 5 or 12.

35 Comparing preschool attendance data collected in 2007 with that collected in 2014, there is a discrepancy of 248 children. Some children who were reported attending preschool in 2007 reported not attending in 2014, and others who reported not attending preschool in 2007, reported attendance in 2014.

Table A3.1. *Sensitivity analysis results on maths scores (% correct)*

Math scores	(1) Preschool	(2) 5 years old + child and hhld	(3) Preschool	(4) 8 years old + child and hhld	(5) +sch	(6) Preschool	(7) 12 years old + child and hhld	(8) +sch
Preschool (main results)	8.58***	5.89***	11.87***	7.88***	5.28***	10.34***	5.19**	3.15
	(2.29)	(1.93)	(1.69)	(1.64)	(1.60)	(2.75)	(2.29)	(2.19)
Including expenditure (lagged)	8.58***	5.89***	11.87***	7.71***	5.22***	10.34***	5.03**	3.13
	(2.29)	(1.93)	(1.69)	(1.70)	(1.63)	(2.75)	(2.27)	(2.17)
Preschool attendance from school history data	2.82**	1.52	4.74***	2.48*	2.09*	7.28***	3.60	3.19
	(1.34)	(1.29)	(1.51)	(1.31)	(1.15)	(2.57)	(2.28)	(1.98)
Urban population	8.50**	5.95**	13.28***	8.94***	5.56**	12.40***	6.47	1.66
	(3.31)	(2.62)	(2.04)	(2.04)	(2.26)	(3.95)	(3.74)	(2.73)
School attendance (3-5 years)	8.71***	5.74***	10.62***	7.69***	5.26***	9.30***	5.07**	3.35
	(1.75)	(1.72)	(1.49)	(1.35)	(1.50)	(2.32)	(2.19)	(2.04)
Preschool (main results) Observations	1,829	1,829	1,763	1,763	1,763	1,810	1,810	1,810
Preschool attendance Observations	1,829	1,829	1,763	1,763	1,763	1,810	1,810	1,810
Preschool attendance from school history data Observations	1,829	1,829	1,763	1,763	1,763	1,810	1,810	1,810
Urban population Observations	717	717	694	694	694	702	702	702
School attendance (3-5years) Observations	1,829	1,829	1,763	1,763	1,763	1,810	1,810	1,810
No. of communities	24	24	24	24	24	24	24	24

Notes: Outcome variable: maths score (% correct). Number of observations and communities are the same for all analyses. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

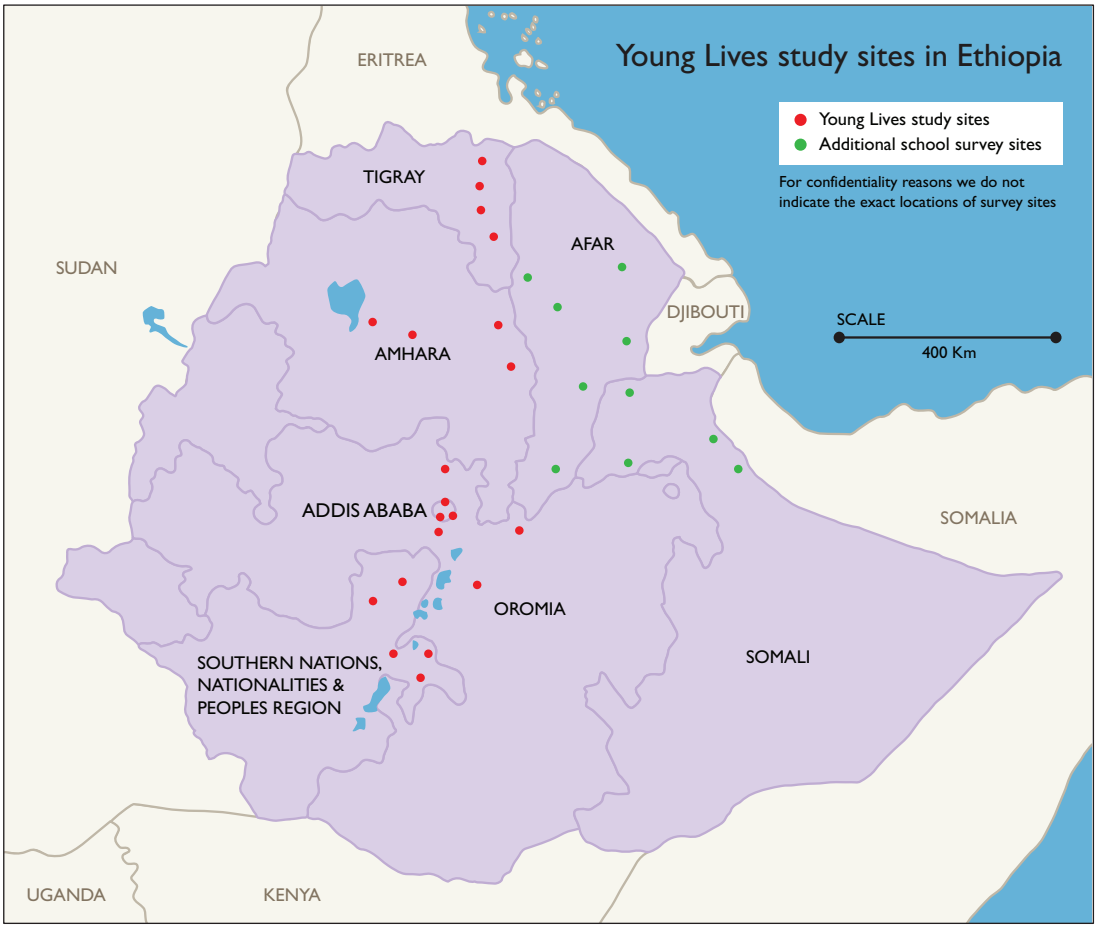
Appendix 4. Community-level information

Table A4.1. *Preschool attendance by community ID, with community description*

Comm. ID		Sentinel site	None	Preschool	Total	Description
1011	1	Addis Ababa	4	87	91	An overcrowded area in the centre of the capital city, Addis Ababa
1021	2	Addis Ababa	3	53	56	An industrial area in the southern part of the capital city, Addis Ababa
1022	2	Addis Ababa	2	37	39	An industrial area in the southern part of the capital city, Addis Ababa
1031	3	Addis Ababa	2	46	48	A slum area in the capital city, Addis Ababa
1032	3	Addis Ababa	4	35	39	A slum area in the capital city, Addis Ababa
2041	4	North Wello	51	35	86	A tourist town in Amhara region, with some extremely poor neighbourhoods
2051	5	North Wello	88	2	90	A poor rural community in Amhara region
2061	6	South Gondar	92	0	92	A rural area near Lake Tana in Amhara region
2071	7	South Gondar	98	0	98	A rural food-insecure area in Amhara region
3081	8	East Shewa	92	1	93	A rural area near Lake Ziway in Oromia region
3091	9	Arsi	88	2	90	A drought-prone rural area in Oromia region
3101	10	North Shewa	60	29	89	A fast-growing town in Oromia region
3111	11	East Shewa	78	14	92	A relatively rich rural area on the outskirts of Debrezeit town in Oromia region
4121	12	Gurage	93	3	96	A densely populated rural area growing enset (‘false banana’) in SNNP region
4131	13	Wolayita	86	5	91	A densely populated town in SNNP region
4141	14	Hawassa City	9	46	55	A fast-growing business and tourist town in SNNP region
4143	14	Hawassa City	2	14	16	A fast-growing business and tourist town in SNNP region
4144	14	Hawassa City	1	23	24	A fast-growing business and tourist town in SNNP region
4151	15	Sidama	86	8	94	A coffee-growing rural area in SNNP region
4161	16	Hadiya	91	8	99	A poor and densely populated rural community in SNNP region
5171	17	Southern Tigray	96	0	96	A drought-prone rural area highly dependent on government support in Tigray region
5181	18	Eastern Tigray	96	0	96	An extremely poor rural area dependent on the Productive Safety Net Programme and other government support in Tigray region
5191	19	Eastern Tigray	82	7	89	A small, very poor urban town in Tigray region
5201	20	Eastern Tigray	95	0	95	A model rural area in Tigray region known for its success in soil and water conservation
n/a		Migrated	37	21	58	Children who migrated between Rounds 1 and 2
Total			**1,436**	**476**	**1,912**	

Source: Adapted from Alemu et al. (2003: 13).

Figure A4.1. *Young Lives study sites in Ethiopia*



Note: Map was compiled prior to the formation of South Sudan.

Inequality in Attainment From Early Childhood to Adolescence: Longitudinal Evidence From Ethiopia

As the Ethiopian government moves to increase enrolment in preschool education from 34 per cent in 2013/14 to 80 per cent by 2020, this working paper provides some country-specific evidence on the predictors of preschool attendance and its association to attainment in Ethiopia. Using four rounds of Young Lives data, it examines: who typically attends preschool education; whether children who attend preschool education have higher attainment than children who do not; and how the association between preschool education and attainment evolves over time. Findings indicate that the key predictors of preschool attendance are being a first-born child, male, better nourished, speaking Amharic as a first language, having a more educated caregiver, belonging to a household with higher wealth, living in an urban area, and having a mother who is not employed.

The paper also examines the association between preschool education and children's attainment, while controlling for a rich array of child-, household- and primary school-level characteristics and variations between communities. Analysis finds that preschool education has a direct and substantial association with attainment, and points to an indirect one through on-time enrolment, grade progression and retention rates (captured in highest grade completed). It also shows that preschool education has a positive and statistically significant role at 5 and 8 years old. Further analysis shows that preschool education has a direct and substantial association with school progression by age 12. These findings indicate that investments in preschool education may doubly pay off – both directly towards attainment and indirectly through exposure to schooling.

Young Lives

An International Study of Childhood Poverty

About Young Lives

Young Lives is an international study of childhood poverty, involving 12,000 children in four countries over 15 years. It is led by a team in the Department of International Development at the University of Oxford in association with research and policy partners in the four study countries: Ethiopia, India, Peru and Vietnam.

Through researching different aspects of children's lives, we seek to improve policies and programmes for children.

Young Lives Partners

Young Lives is coordinated by a small team based at the University of Oxford, led by Professor Jo Boyden.

- *Ethiopian Development Research Institute, Ethiopia*
- *Pankhurst Development Research and Consulting plc, Ethiopia*
- *Centre for Economic and Social Studies, Hyderabad, India*
- *Save the Children India*
- *Sri Padmavathi Mahila Visvavidyalayam (Women's University), Andhra Pradesh, India*
- *Grupo de Análisis para el Desarollo (GRADE), Peru*
- *Instituto de Investigación Nutricional, Peru*
- *Centre for Analysis and Forecasting, Vietnamese Academy of Social Sciences, Vietnam*
- *General Statistics Office, Vietnam*
- *Oxford Department of International Development, University of Oxford, UK*

Contact:
Young Lives
Oxford Department of International Development,
University of Oxford,
3 Mansfield Road,
Oxford OX1 3TB, UK
Tel: +44 (0)1865 281751
Email: younglives@younglives.org.uk
Website: www.younglives.org.uk

Young Lives