Student Paper

Young Lives
An International Study of Childhood Poverty

Horizontal Inequalities in Children's Educational Outcomes in Ethiopia

Nardos K. Tesfay

Paper submitted in partial fulfillment of the requirements for the Degree of Master of Science in Educational Research Methodology at the University of Oxford, Department of Education.

The data used come from Young Lives, a longitudinal study of childhood poverty that is tracking the lives of 12,000 children in Ethiopia, India (Andhra Pradesh), Peru and Vietnam over a 15-year period. **www.younglives.org.uk**

Young Lives is core-funded from 2001 to 2017 by UK aid from the Department for International Development (DFID) and co-funded by the Netherlands Ministry of Foreign Affairs from 2010 to 2014.

The views expressed here are those of the author. They are not necessarily those of the Young Lives project, the University of Oxford, DFID or other funders.

HORIZONTAL INEQUALITIES IN CHILDREN'S EDUCATIONAL OUTCOMES IN ETHIOPIA

A *dissertation submitted in partial fulfillment of the requirements for the Degree of Master of Science in Educational Research Methodology at the University of Oxford, Department of Education*

Nardos K. Tesfay

Wolfson College

ABSTRACT

This paper examines the relationship between children's educational outcomes and ethnic group status Ethiopia. It builds on the main theories of education in developing countries – human capital theory and educational opportunity theory – and further draws on the theory of horizontal inequality as a lens for examining educational inequality. Using longitudinal data from the Young Lives study, information collected from 1,000 older cohort children who were followed up at ages 8, 12 and 15 were analysed to investigate the effects of ethnicity controlling for individual, home, community and school characteristics.

The findings showed that while school enrolment appears to have improved in the study sites, with most of the children enrolled in school at age 15, slow grade progression and low achievement remain challenges, particularly for certain minority ethnic groups. We further demonstrate significant regional in intra-regional differences, as well as differences between the centre (Addis Ababa) and the periphery. As regards the factors that predict schooling outcomes, we find that disparities in income level and poverty status cannot fully explain the variation in educational outcomes between groups. Contextual factors at the individual level, particularly schooling history, and school level variables also play a significant role.

The paper concludes by arguing that research and policy that emphasizes educational access, or enrolment, may understate how children differently experience schooling and may downplay the role of education in compensating for or reproducing inequalities. The paper further suggests that employing a group-based approach to inequality analysis, as opposed to an individual measure, can help to better understand the mechanisms through which inequalities take shape and may illuminate on role of schooling in aggravating or mitigating against individual and family level disadvantage.

ACKNOWLEDGEMENTS

The inspiration for this project originated in several years spent working in practice as a social policy officer with UNICEF in Africa. However, the initial ideas have been shaped by a remarkable group of motivated and creative individuals in the Department of Education at the University of Oxford. I am grateful to have been part of the class of 2012.

I thank, in particular, Dr. Lars-Erik Malmberg, who brought his extensive knowledge to this project, not only to technically guide the current research, but also to help strategically direct the focus of future work.

Many thanks to the team at Young Lives for their encouragement and continued support, beyond expectations.

And most of all, to my devoted family who have sacrificed abundantly to support my goals.

MAP OF ETHIOPIA

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Ethiopia has made enormous progress towards expanding access to primary education. Net enrolment has nearly doubled since 2002 (from 46 per cent to 81 per cent) and girls are as likely to be enrolled as boys, at 79 per cent and 84 per cent net enrolment, respectively (World Bank, 2011). However, two and a half million children remain out of school (World Bank, 2011), as retention and grade progression continue to threaten educational completion and achievement. The national dropout rate has actually risen from 11.8 per cent in 2005 to 18.6 per cent in 2009 (Overseas Development Institute, 2011: 7). Moreover, while the net attendance ratio for primary school is increasing, attendance remains generally low at 64 per cent (Demographic and Health Survey, 2011: 11), and disparities persist between geographic regions, income quintiles and urban-rural residence (Demographic and Health Survey, 2011: 11).

Poverty levels in Ethiopia are among the highest in the world. Nearly 80 per cent of the population live on less than $2 a day (World Bank, 2011). According to the multidimensional poverty index, which measures deprivation across a range of dimensions in addition to income poverty, the proportion of the poor is estimated at a staggering 89 per cent of the population (United Nations Development Programme, 2011: 144). Moreover, poverty in Ethiopia is persistent; the table below illustrates the extent to which poverty levels have stagnated over the last 30 years.

Poverty headcount in Ethiopia, 1982-2005

	Year			
	1982	1995	2002	2005
Poverty headcount (%)	90	85	86	78

Source: World Bank World Development Indicators, 2011.

In addition to the high levels of poverty and deprivation, violent conflicts are widespread in Ethiopia. In the past decades, armed conflicts have taken place over international boundaries (with Eritrea and Somalia), internal armed conflicts between government forces and insurgency groups are ongoing (in Somali

region), and local conflicts over resources and administrative boundaries often break out in various parts of the country (Norwegian Refugee Council, 2011).

Ethiopia therefore presents an interesting context for examining educational inequality. The high level and persistence of poverty points to the need to better understand its causes and consequences. Moreover, the incidence of vertical inequality in Ethiopia, or inequality between individuals, is not particularly high in relation to other African countries; with a Gini coefficient of 0.30 per cent (WIDER, 2007). Thus, an analysis of horizontal inequality, or cultural group-based inequality, could provide insights into how such disparities relate to the persistence of poverty and conflict.

The aim of the current research is to examine the extent and significance of horizontal inequalities in children's educational outcomes. The project is guided by the following research questions:

1. To what extent does *children's educational achievement* differ by ethnicity at age 15?
2. To what extent does *children's schooling participation* differ by ethnicity at age 15?
3. If differential patterns exist, what factors might help to *explain inter-group differences* in achievement and participation at the child, home, community and school level?
4. How does schooling participation *impact on children's cognitive development* between age 12 and 15 across ethnic groups?

The study draws on quantitative data collected as part of the Young Lives longitudinal study of childhood poverty and is organised into six parts. Chapter 1 introduces the theoretical and empirical literature in the area of education and development. The chapter goes on to highlight national trends in educational enrolment, progression and achievement in Ethiopia, as well as the key contextual variables identified as significant for children's schooling participation and performance in Ethiopia. Chapter 2 presents the conceptual framework for the current research, drawing together the literature presented in the Chapter 1.

Chapter 3 describes the methodological aspects of the research with a view to justifying the procedures used at each stage of the process. Chapter 4 highlights differences in schooling outcomes across ethnic groups in the Young Lives sample of older cohort children, and further demonstrates the variation between ethnic groups in a range of contextual factors. Chapter 5 discusses the contextual models of ethnicity and educational outcomes with a view to examining the significance of mean-level ethnic group differences in schooling outcomes and the extent to which this gap can be explained statistically by child, home, community and school factors. Finally, Chapter 6 reflects on the outcomes of the research with reference to the initial research aims and questions.

1. LITERATURE REVIEW

The purpose of this chapter is to introduce the theoretical and empirical literature in the area of education and development, as it relates to the current research. The chapter begins with a presentation of relevant theories of education in developing countries with a view to situating the research within the broader theoretical context. Section 2 discusses the theory of horizontal educational inequality and its relation to Ethiopia. Section 3 provides an overview of recent trends in educational enrolment, progression and achievement in Ethiopia and further identifies some of the challenges to achieving universal primary education. This is followed by a discussion of the empirical evidence on factors influencing schooling outcomes in section 4.

1.1. Theories of education in developing countries

This section provides a discussion of prominent theories of education, applied in developing countries, in two relevant disciplines – development economics and the sociology of education. The section further endeavours to illustrate the implications of each theoretical framework for educational inequality analysis.

1.1.1. Human capital theory

The most conventional theory of education in developing countries relates to the theory of human capital, which has a deeply rooted economics foundation. Human capital refers to the skills, knowledge and other competencies embodied in the ability to perform labour (Becker, 1995: 2). While the concept of human capital dates back to Adam Smith, the theory was more recently formalized by Gary Becker (1964). Becker argued that the objective of an educational system is to equip individuals with the skills and abilities required to become members of a productive workforce that can generate economic growth (Becker, 1964).

Human capital theory, has been highly influential in developing/labour surplus economies, and is the foundation of economic development theory, in which investment in human capital is seen as the “engine of economic growth” (Becker, 1995: 4). Household welfare is considered as both a predictor and an outcome of

human capital investments (Mincer, 1970; Becker, 1994; Card, 1999; Heckman et al., 1999; and Hanushek, 2005). Schooling decisions are therefore explained by rational choices made by individuals with a view to maximising household welfare, or utility, subject to resource constraints (Becker, 1964).

The limitation of human capital theory lies primarily in its conceptualization of the role of education. Schooling is not recognized as having value beyond its ability to produce economic growth. That is, the benefit of schooling is mainly its potential to improve peoples' access to employment opportunities. The benefits of schooling beyond productive purposes are not recognised. This has an important implication for inequality analysis. As Becker argues (1995:9):

> Education and training not only promote growth and efficiency, but they can reduce inequality and the impact of disadvantaged backgrounds. Education is the most effective way for able young people of poor backgrounds to rise in the economic hierarchy.

Thus, within the human capital framework, inequality is understood in terms of income and economic mobility. Education can play a role in reducing the gap between the highest and lowest income quintile. However, the role of the educational system in creating or perpetuating inequality is overlooked.

1.1.2. Educational opportunity theory

Jean Dreze and Amartya Sen's (1995) theory of educational opportunity goes beyond the human capital approach, arguing that education has the potential not only to improve peoples' access to productive work, but also to a number of social goods. They argue that primary education is valuable to individuals in at least five distinct ways (Dreze and Sen, 1995: 14-15). First, education has *intrinsic importance* in that being educated, or having knowledge, has value in itself. Second, education plays an *instrumental personal role*, for instance, by facilitating employment and other economic opportunities. Third, education plays an *instrumental social role* such that it facilitates public discussion of social needs and encourages the organisation of collective demands on duty bearers (e.g. basic social services). Fourth, education plays an *instrumental process role* by providing benefits beyond formal schooling objectives; for instance, by promoting

the development of life skills, and broadening aspirations. Finally, education plays an *empowerment and distribute role*. As argued by Dreze and Sen, "greater literacy and educational achievements of disadvantaged groups can increase their ability to resist oppression, to organize politically, and to get a fairer deal" (Dreze and Sen 1995: 14).

The theory of educational opportunity has been instrumental in the shaping human development theory. This is reflected in the Human Development Index, a composite score used to rank countries according to their level of human development. Education is a core component of the measure, which is derived from life expectancy at birth, mean years of schooling, expected years of schooling and per capita income (United Nations Development Programme, 2011: 127). Thus, human development theory suggests that education is considered as a both a means towards achieving developmental goals as well as an end in itself.

The main criticism of educational opportunity theory, relevant for the current research, is that by emphasizing the importance of educational opportunities – access to schooling – they downplay how schooling is differently experienced by individuals (Jeffrey, 2008: 6). Therefore, like human capital theory, the framework put forward by Dreze and Sen (1995) may also be limited in its ability to account for the transmission of societal inequalities – at the individual, home, community and school level – through educational outcomes.

1.1.3. Education and cultural identity

The focus on the importance of schooling seems a necessary strategy when considering 61 million primary school-aged children are still out of school (UNESCO, 2010). Despite global efforts to expand educational opportunities to the poor, according to UNESCO, "progress towards universal primary education is slowing and that the goal will not be met by 2015 if current trends continue" (UNESCO, 2012). Therefore, the notion of schooling as intrinsically and instrumentally beneficial is indeed a relevant one. However, it has been criticised for understating the role of schooling in the transmission of societal inequalities.

In their review of the history of educational research, Levinson and Holland (1996) discuss the emergence, in the 1970s, of a new era in the sociology of education. Social reproduction theory sought to demonstrate that schools were not simply a forum for achieving formal schooling objectives or a means through which social mobility could be attained, rather they served as mechanisms for the creation and perpetuation of social inequalities (see Althusser, 1971; Bowles and Gintis, 1976; Bourdieu and Passeron, 1977; and Giroux, 1983).

> As summarised by Levinson and Holland (1996: 5):
> … these scholars endeavoured to show that schools were not "innocent" sites of cultural transmission, or places for the inculcation of consensual values. Nor could schools be understood as meritocratic springboards for upward mobility – the great levelling mechanism… Rather, critical scholars argued that schools actually served to exacerbate or perpetuate social inequalities… schools served to reproduce rather than transform existing structural inequalities.

The work of Pierre Bourdieu further endeavoured to show the cultural reproduction of "privilege" in education and the continued subordination of non-dominant groups (Bourdieu, 1974; Bourdieu and Passeron, 1977). Bourdieu showed that in French schools dominant groups maintained their economic advantage by arbitrarily reproducing their "cultural capital", (Bourdieu, 1974: 42); where cultural capital refers to the "symbolic credit which one acquires through learning to embody and enact signs of social standing" (Levinson and Holland, 1996: 6). Bourdieu argued that this particular form of education, the transmission of the values and content of cultural capital, is a type of "symbolic violence" on non-dominant groups (Bourdieu and Passeron, 1977: x-xii). Implicit repression of non-dominant cultural groups is therefore used to explain differences in educational outcomes and life outcomes beyond schooling.

1.2. Horizontal inequality and education

This section introduces the theoretical and empirical literature on horizontal inequality, as it relates to education, within the context of the broader educational

inequality literature. It concludes with a discussion of the significance of the notion of horizontal inequality for Ethiopia's educational system.

1.2.1. Educational inequality

There is a voluminous empirical literature on educational inequality in developing countries. In particular, the extent of educational inequality by wealth quintile is extensively documented due, in part, to the addition of education modules in household sample surveys – such as the Welfare Monitoring Survey and Living Standards Measurement Survey. In 2010, UNESCO launched the 'Education for All (EFA)' Inequality Index for Income Groups (EIIIG) as part of the EFA Global Monitoring Report (UNESCO, 2010). The purpose of the index was to illuminate on disparities in EFA targets by wealth quintile (UNESCO, 2010: 279). The results showed that such disparities persist, and are particularly problematic in countries where progress towards meeting EFA goals is lagging.

> As noted by UNESCO (2010: 279):
> "The disparities are greatest in countries where overall EFA achievement is still low, such as Burkina Faso, Ethiopia, Mali, Mozambique and the Niger. These countries face the double challenge… [of developing] their educational systems while making them more inclusive by reaching and teaching the most marginalized."

The vast majority of educational inequality research focuses on vertical inequalities, defined as inequalities between individuals (Stewart, 2000), most notably, by income (see, for example, Engle and Black, 2008; UNESCO, 2010). However, with the adoption of the Millennium Development Goals, gender inequality has emerged as a major focus of analysis, as evidenced by the United Nations Gender Inequality index (UNDP, 2011: 140). In Ethiopia, gender gaps in education have also been well-researched (see Rose et al., 1997; Colclough et al., 2000; Rose and Al-Samarrai, 2001 and Camfield, 2011). Rose (2003) draws on official statistics to show that the issue of under-enrolment of girls is mainly a rural phenomenon in both primary and secondary school, and a particular concern in pastoralist areas (Rose, 2003: 4-5).

However, since the launch of the EFA campaign (1995) and the World Education Forum (2000), which forged a global consensus on key education priorities, there has been a growing interest in moving beyond addressing inequalities between individual towards understanding group-based inequalities. Analyses of group-based differential patterns often focus on geographic stratum and urban-rural residence (see UNESCO, 2010). In addition, there is a growing literature on the educational outcomes of children belonging to specific vulnerable groups – such as children with special needs and children affected by HIV and AIDS (see, for example, Case, Paxson and Abledinger, 2004; Ainsworth and Filmer, 2006; UNESCO, 2010).

1.2.2. Horizontal educational inequality

The current research adopts the theory of horizontal inequality as a lens for examining differential patterns of schooling outcomes across ethnic groups in Ethiopia. Horizontal inequality (HI) is defined as inequalities in economic, social or political dimensions or cultural status between culturally defined groups (Stewart, 2008a). The current study employs the definition of culturally defined groups proposed by Stewart (2008a) – groups encompassing common cultural identities such as ethnicity, language, race, religion or class (Stewart, 2008a).

HI is considered as both a cause of conflict and, as noted by Stewart (2000), "understanding the causes of conflict should include measures of horizontal inequality of each of the elements relevant in that society..." (Stewart, 2000: 253). More recent literature in this field goes beyond HI-based explanations of conflict in order to explore the role of persistent HI in perpetuating poverty traps.

> Stewart (2008a: 1) argues:
> In many cases Horizontal Inequalities or inequalities among groups [.] persist over long periods. Where they persist they are particularly deleterious as they trap people, generation after generation, in a situation of poverty.

There has been good progress towards improving knowledge on the forms of HIs in education and the role of educational inequality in perpetuating both conflict

and persistent poverty, across a range of poverty dimensions. Most studies deal with different elements of the educational system – its delivery, curriculum content and institutional processes. Educational access, or 'delivery,' is the most widely researched aspect of the educational system – particularly national patterns of enrolment and completion between groups (Johnson, 2007; Uwiko, 2007; Stewart, 2008b; also see Ostby, 2011). Moreover, numerous studies examine the diffusion of divisive curriculum and their effects on schooling outcomes (see Bush and Saltarelli, 2000; Salomon and Nevo, 2002; Davies, 2004; Harber, 2004; Johnson, 2007; Watson, 2007; and King, 2008). To a much lesser extent, educational processes and the transmission of societal inequalities through school policies, organisation and management have been examined (see Stephan, 1999; Salomon and Nevo, 2002; Johnson, 2007; and Brown 2007).

1.2.3. Significance of horizontal inequality for Ethiopia

Ethiopia presents an interesting context for examining differential patterns in schooling outcomes by cultural status. Overall, vertical inequality is moderate by African standards; with a Gini coefficient of 0.30 (WIDER, 2007). However, poverty levels are among the highest in the world (World Bank, 2011), and more than half of school aged children are out of school (Sedel, 2008). A recent study commissioned by the Ministry of Education identified two inequity issues in the education sector that are particularly relevant for the current study: (i) the persistence of significant regional disparities in terms of enrolment, with pastoral regions appearing to be most disadvantaged; and (ii) inequalities in enrolment patterns between urban and rural areas, with urban gross enrolment rates double that of rural areas (Jennings et al., 2010):

In addition to regional/geographic disparities, language and ethnicity is particularly significant for potential inequities in Ethiopia's educational system. Ethnicity in Ethiopia is commensurate with language.

> As noted by Keller and Smith (2005: chapter 11):
> Language moved to the centre of the political agenda [in 1994] in Ethiopia as a crucial element of the [national] education plan, which gave support to the development and use of nationality languages for all ethnic groups. This support was in part attributable to the fact that

> language represents a proxy for ethnic identity under ethnic federalism and therefore, linguistic identity has assumed an elevated level of importance.

With over 70 linguistic identities, in 1994 the Ministry of Education introduced a mother tongue education policy in grades 1-4 (Heugh et al., 2007). The Government recommends that students study English and Amharic (the official language) only as subjects, and switch to English instruction between grades 5 and 9. The language of instruction policy has had a significant impact. Results from the National Learning Assessments show that students with eight years of mother tongue education perform better across the curriculum in the Grade 8 tests than do those with fewer years of mother tongue instruction (Heugh et al., 2007: 104).

1.3. The problem of schooling in Ethiopia

This section briefly highlights patterns and trends in educational achievement, enrolment and progression in Ethiopia, at the national level. It seeks to situate the current research within the broader Ethiopian context with a view to demonstrating the relevance of the research aims questions presented in the subsequent chapter.

1.3.1. National patterns of educational achievement

The National Learning Assessments (NLA) were launched in 2000 by the Government of Ethiopia and followed-up in 2004 and 2007 with the aim of understanding student achievement and the factors that influence achievement (World Bank, 2008b). The assessments were administered to Grade 4 and Grade 8 students. Schools in Southern Nations Nationalities and Peoples' (SNNP) Region were slightly overrepresented to cover the diversity of that region (National Organisation for Examinations, 2007). In 2007, over 20,000 students were assessed across 10 regions – 13,300 Grade 4 students and 8,100 Grade 8 students.

In addition to reading, students were assessed in a number of other core subjects: maths, English and environmental sciences. Grade 8 students were

assessed in two additional sciences. Using quantitative questionnaires and qualitative interviews, background information on the students and the schools was gathered to correlate student/school characteristics with learning outcomes. Parents were also interviewed to capture information on parent expectations and on the responsiveness of the educational system to household needs.

The most significant finding relates to the generally low level of reading comprehension, with boys performing slightly better than girls (USAID, 2010). Statistically significant regional and language-group disparities were also observed. Table 1.1 and 1.2 present the mean and standard deviation of the test scores by subject and grade. The composite scores for both Grade 4 and Grade 8 reveal that mean student achievement was below the minimum requirement of 50 per cent for grade promotion, as benchmarked in the Ethiopian Education and Training Policy (1994).

Table 1.1 Descriptive statistics for NLA test scores for Grade 4 students, 2004

Subjects	N	Mean	Std. Error	Std. Dev
Reading	12886	64.49	.19	21.97
English	13028	38.68	.15	16.88
Mathematics	13143	39.70	.15	17.51
Env. Science	13005	51.74	.15	17.42
Composite	12507	48.48	.13	15.07

Source: National Organisation for Examinations, 2004.

Table 1.2 Descriptive statistics for NLA test scores for Grade 8 students, 2004

Subjects	N	Mean	SE	SD
Biology	7982	41.34	.15	13.42
Chemistry	7995	40.10	.16	14.74
English	8026	41.07	.16	14.68
Mathematics	8017	40.93	.18	16.15
Physics	8024	35.32	.14	12.14
Composite	**7898**	**39.74**	**.13**	**11.36**

Source: National Organisation for Examinations, 2004.

In addition, student performance was observed to have declined during the period of the educational expansion, from 48 per cent in 1999/2000 to 41 per cent in 2007 for Grade 4s and from 43 per cent to 40 per cent for Grade 8s (World

Bank, 2008). Improvement in student performance was found only in one subject among Grade 4 pupils (environmental sciences) and in two subjects among Grade 8 pupils – English and mathematics (National Organisation of Examinations, 2004: 73).

1.3.2. National patterns of access and progression

Ethiopia has made great advances in primary school enrolment in recent years. Net enrolment has increased from 46 per cent in 2002 to 81 per cent in 2010 (World Bank, 2011). However, retention and grade progression remain critical challenges for the educational system. Official statistics indicate a high primary school dropout rate, averaging 18.6 per cent in 2008/09 overall for grades 1-8; with 28 per cent of children dropping out in Grade 1 (ODI, 201: 7). Furthermore, between 2004/05 and 2008/09, the national dropout rate actually increased from 11.8 per cent to 18.6 per cent (OD!, 2011: 7).

In terms of attendance, the overall net attendance ratio for primary school increased from 30 per cent in 2000 to 42 per cent in 2005 and 64 per cent in 2011 (Demographic and Health Survey, 2000; 2005; 2011). However, there are huge gaps between geographic stratum and by poverty status. Figure 1.1 illustrates the large variation in the net attendance ratio – measured as the percentage of the primary school-aged (7-12 years) population that is attending primary school – by urban/rural residence, by region and by wealth quintile.

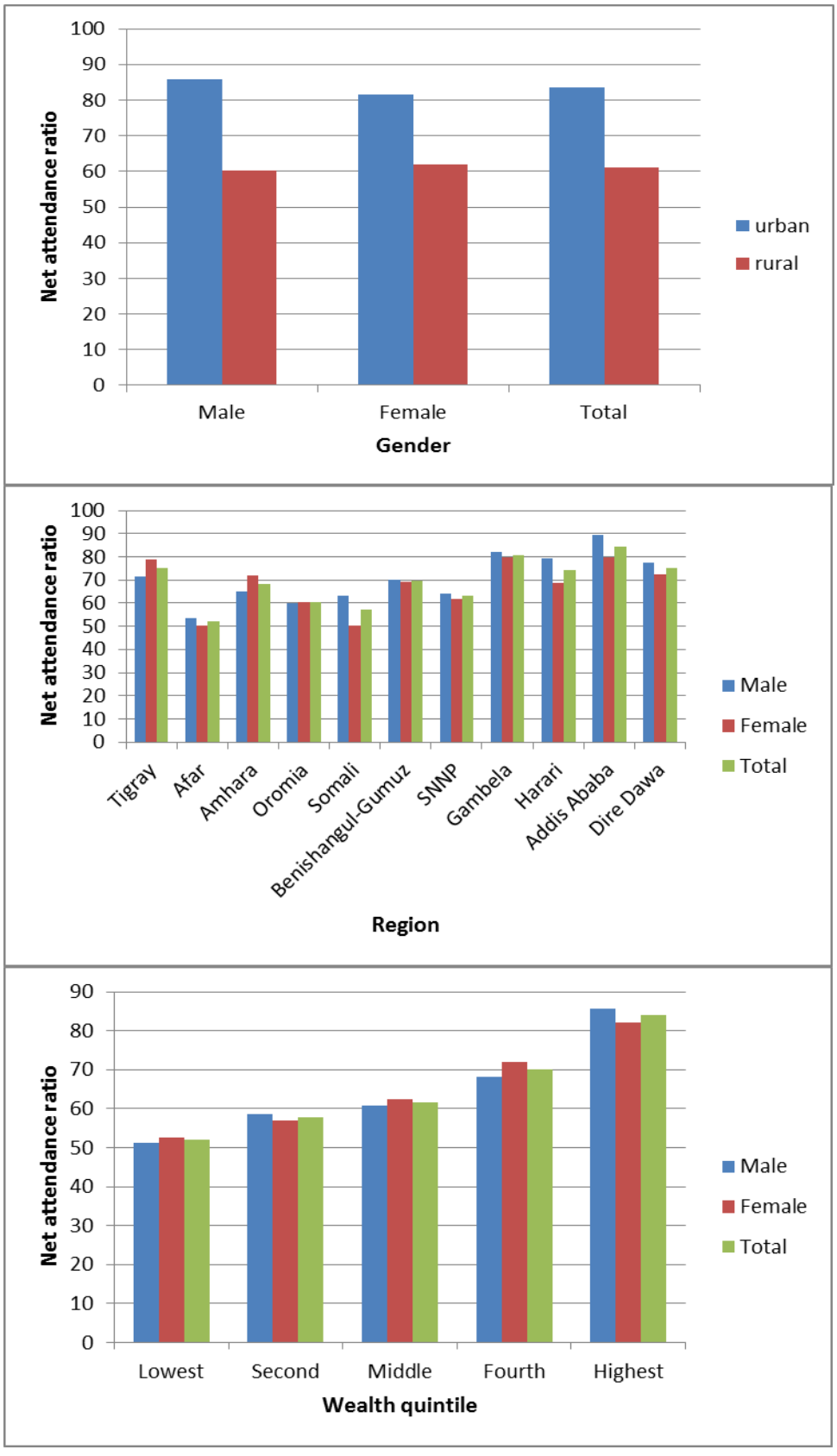


Figure 1.1 Net attendance ratio for primary school by residence, region and wealth quintile (Source: Demographic and Health Survey data tables, 2011: 11)

With regard to attendance patterns by wealth, only half of primary school-aged children were attending school in 2011 in the lowest quintile compared with 84 per cent in the highest wealth quintile. There are also significant differences between regions; with children in Addis Ababa significantly more likely to attend primary school (84 per cent) compared with children in Afar (52 per cent), Oromia (60 per cent), Somali (57 per cent) and SNNP (63 per cent). As regards residential differences, 84 per cent of children in urban areas were attending school in 2011 compared to 61 per cent of children in rural areas.

1.4. Factors associated with educational outcomes in Ethiopia

Factors predicting patterns of enrolment and drop-out have been widely studied in Ethiopia. However, there are significantly fewer studies that examine how children progress through the primary school grades (e.g. the extent/determinants of late enrolment, grade repetition and grade completion). Until recently, predictors of learning outcomes had been particularly under researched due to the absence of achievement data. However, with the introduction of the National Learning Assessments in 2000, good progress has been made towards understanding children's educational outcomes in relation to wider socio-economic and demographic trends.

This section seeks to present the main influences on educational outcomes in Ethiopia, as identified in the quantitative empirical literature. It is organised around three layers of environmental factors, identified in child developmental theory as significant for children's schooling experiences: (i) child level characteristics; (ii) family and community background factors; and (iii) influences at the school level (the rationale for which is provided in Chapter 2).

1.4.1. Child characteristics

There is a vast literature on the role of learning difficulties and nutrition in shaping cognitive development in Ethiopia. The role of child hunger, health and nutrition on schooling participation and performance has been well documented (USAID, 1994; Culclough, 2000; Dercon and Krishnan, 2000; Sharp et al., 2003; Dercon,

2011). Culclough et al. (2000) find that poor health is not only a barrier to school attendance but also affects enrolment as parents may be reluctant to send their children to school due to the presence of disease within the school (Culclough et al., 2000). As regards learning difficulties, Zehle (2009) uses official statistics as well as qualitative research in two regions of Ethiopia, Southern Nations Nationalities and People's (SNNP) Region and Addis Ababa (the capital city) to investigate the extent to which children's learning difficulties might help to explain the high levels of early primary school dropout.

In addition, a number of smaller studies have been undertaken to investigate the motivational predictors of achievement and their implications for a wide range of schooling and life outcomes. Most notably, Raju and Asfaw (2009) look at the predictive nature of test anxiety on student achievement in Ethiopia, taking into account several other factors: academic self-concept, socio-economic status, study habits and parental involvement. A sample of 497 students was randomly selected from seven schools within a suburb of Addis Ababa. The results indicated a weakly negative correlation between test anxiety and achievement in the presence of other variables. A strong positive relation was found between the remaining predictors and achievement, with the exception of socio-economic status, which was found not to be associated with achievement.

1.4.2. Family and community background

There is substantial empirical evidence to support various home disadvantage-based explanations of school enrolment. Most studies draw on large-scale national surveys and employ multivariate regression analysis techniques to examine proxy measures of socio-economic status. Home factors that appear to have to most significant effect on children's enrolment include parental educational level, household wealth, direct and indirect costs of schooling, mother's literacy, as well as having educated women in the community. Relevant studies include Krishnan, 1996; Weir and Knight, 1996; Weir, 1999; Admassie et al., 2003; and Weir, 2011. Each of these studies draws on the Ethiopia Rural Household Survey.

Additional studies include Woldehanna et al. (2005), who examined the relationship between enrolment and a range of contextual variables, drawing on the Young Lives longitudinal study of childhood poverty. Himaz (2009) and Orkin (2011) also use the Young Lives datasets to investigate the relationship between parental illnesses and death on children's schooling. Himaz (2009) in particular found that maternal orphans aged 8-12 are less likely to read than non-orphans and paternal orphans, while the death of the mother reduces enrolment by 20 per cent (Himaz, 2009). Other relevant studies include Schaffner (2003), who uses the Welfare Monitoring Survey (1999/2000), the Ethiopian Labour Force Survey (1999) and the Demographic and Health Survey (2000) to examine household level determinants and some school quality variables. Alvi and Dendir (2011) draw on the Ethiopian Labour Force Survey to study at sibling effects on school attendance. They find that earlier-born males in rural areas are more likely to attend school than their younger siblings.

There is a growing body of literature on the opportunity costs of schooling in Ethiopia. Examples include World Bank, 1998 cited in Andvig et al., n.d.; Colclough et al., 2000; Sharp et al., 2004; Rose and Dyer, 2008; Admassu, 2008 and Orkin, 2011. Admassu (2008) suggests that as Ethiopian children grow older they may contribute more time to household work and less time to school, particularly in rural areas. The argument is based on the finding that "as a child's age increases by one year, the expected number of grades lagged increases by a factor of 2.0 and 1.6 in rural and urban areas respectively" (Admassu, 2008). However, it has also been shown that when the household does not own land, children are actually more likely to be in school as the opportunity cost for their time is reduced (Sharp and Devereux, 2004 cited in Rose and Dyer, 2008).

1.4.3. School factors

The National Learning Assessment (2007) identifies a number of school level variables affecting learning outcomes. The most significant factors cited by children, parents and teachers relate to the availability of school supplies (e.g. teaching and learning materials); the availability and condition of schools and school facilities; the distance/time to get to school; and teaching experience and

ability, all of which had a positive effect on NLA test scores (NLA, 2007). In addition to the NLA, there are few nationally representative quantitative studies that link school quality variables to children's schooling outcomes. One exception is Chaudhury et al. (2006), who link Welfare Monitoring Survey (1996, 2000) data with official school records to analyse various factors related to children's schooling. They find that investment in child's schooling is significantly influenced by positive education externalities at the household and community level, the availability and distance to schools as well as the quality of school infrastructure (as measured by the pupil-teacher ratio, the proportion of qualified teachers and the proportion of qualified female teachers).

Additional studies that focus on school characteristics include Colclough et al. (2000) and Rose and Al Samarrai (2001). The studies use mixed methods designs and draw on the same sample to identify household, community and school level factors affecting girls' enrolment and progression in rural areas. Approximately 240 grade six students within mainly rural schools were selected from Jimma and Gojjam zones (Rose and Al-Samarrai, 2001). The cost of schooling and school materials was identified as one of the most significant determinants of children's schooling, particularly in the early years. Opportunity costs of schooling and demand for children's time were also found to be relevant for enrolment and the duration of enrolment. Finally, repetition was identified as a key deterrent to completion.

> As stated by Rose and Al-Samarrai (2001: 55-56):
> If children have to repeat a grade they will be older before they reach the last grade of primary school, which again increases the opportunity cost of their time and increases the chances of girls withdrawing when they reach puberty. Furthermore, a large proportion of children repeat in early grades, which causes them to lose interest in school. Of the dropouts who had repeated a grade, two-thirds repeated the first grade.

While these studies were carried out at the school level, they reveal more about child and family factors in comparison to school level predictors.

Verwimp (1999) takes into account the fact that most schools in Ethiopia are flexible and apply a multiple shift system; around 44 per cent of the primary government schools and 80 per cent of secondary government schools (World Bank, 2005: 145-146). Building on this, Verwimp constructs two measures of educational quality to examine their association with enrolment. The first is an institutional (school level) measure and the second, an individual (teacher level) measure. Verwimp carried out interviews with 35 teachers and seven school principals, randomly selected from seven rural primary schools. The school level measure was derived from items relating to teacher effort, school flexibility and infrastructure. The teaching quality measure was derived from the information generated from teacher attitude and behavior scales (Verwimp, 1999). Based on the constructed measures, Verwimp found school quality to be positively correlated with increased enrolment. He further showed that teaching quality is negatively correlated with increased pupil-teacher ratios (pupils per teacher in the whole school). The evidence further suggests that class size has less of an effect on enrolment than teacher time allocation.

2. THEORETICAL FRAMEWORK

This chapter draws together the literature discussed in the previous chapter with a view to illustrating the relevance and significance of the aims and questions of the current research. A conceptual model is then proposed for understanding children's educational trajectories and the environmental influences on educational outcomes in Ethiopia. The model draws on the theory of horizontal inequality as well as the Bronfenbrenner bioecologial model of human development.

2.1. Justification for the research

This research aims to contribute to two complementary aspects of the literature on education and development – horizontal inequality and child developmental theory – in three ways. First, the research examines ethnic group differences in children's educational outcomes at key stages during the school cycle. This component of the research draws on the theory of horizontal inequality, a new and emerging area of research with implications for education planning in conflict-prone contexts. Horizontal inequality is defined as inequality in economic, social or political dimensions or cultural status between culturally defined groups.[1] While there is a voluminous literature on vertical inequality between individuals, cultural group-based, or horizontal, inequality is a significantly under researched phenomenon, particularly in the Ethiopian context. The current study proposes to examine differences in schooling outcomes by ethnic group, a key dimension of cultural identity in Ethiopia. The Young Lives study, on which the current study draws, asks about self-perceived ethnicity (Young Lives, 2002: 8), which will be retained for the current research.

Ethiopia provides a unique case for examining patterns of cultural inequality. Overall, vertical inequality is moderate by African standards. However, poverty levels are among the highest in the world (World Bank, 2011), and more than half

[1] Culturally defined groups refer to groups encompassing common cultural identities such as ethnicity, race, religion or class. Refer to chapter 2 for a more complete discussion of the horizontal inequality literature.

of school aged children are out of school (Sedel, 2008). Thus, an analysis of horizontal inequalities and education could provide insights into how group disparities in education threaten human development and exacerbate poverty traps. Moreover, while patterns of educational access and achievement in Ethiopia have been widely researched, few studies have attempted to examine children's educational experiences or patterns of change over time, due to the absence of appropriate data. Thus, the longitudinal design of the current research (discussed in the subsequent chapter) provides a unique opportunity for understanding child development in Ethiopia in the context of wider socio-economic and demographic trends.

The second objective of the study is to investigate the factors that predict inter-group differences in a systematic way. This aspect of the research will build on the growing body of literature that aims to explore the role of environmental factors – at the individual, home, school and community level – in shaping children's educational trajectories, drawing on the Bronfenbrenner model of child development. Developmental theory suggests that the child's immediate environment influences his/her developmental processes (Bronfenbrenner, 1977; Bronfenbrenner and Ceci, 1994; Bronfenbrenner and Morris, 1998). This bioecological model of human development is depicted in Figure 2.1. Each layer of environmental influences is "nested" within the next (Bronfenbrenner and Morris, 1998; 995); with interactions occurring between them. The Bronfenbrenner model therefore provides an appropriate framework for simultaneously examining the impact of multiple layers of deprivations on children's developmental processes.

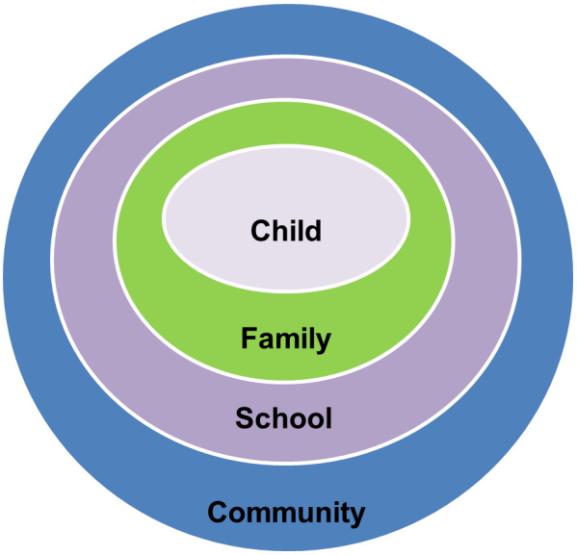


Figure 2.1 Bronfenbrenner's model of human development

Much of the quantitative research in this area has focused on the home environment as the most important factor influencing child development (see, for example, Bronfenbrenner and Ceci, 1994; Brooks-Gunn, et al., 1997; Bradley et al., 2001; Melhuish et al., 2008). However, other research has emphasised the importance of child and school level factors for learning outcomes (refer to Creemers, 1994; Glewwe and Jacoby, 1994; Blatchford, 2003; Wößmann, 2003; Sylva et al., 2004; Lewin, 2007; Strand, 2011a; Strand, 2011b). In the Ethiopian context, there is limited empirical research that examines multiple layers of predictors simultaneously, while at the same time linking these factors directly to learning outcomes. The current research would be the only known study to apply this approach in this context.

Thirdly, the study investigates the effect of schooling participation on children's cognitive development over time, and the extent to which ethnic group differences persist at this level. This component of the research draws together the previous analysis of children's educational achievement and their progression through the school system.

2.2. Specific research questions

Based on the existing literature and the identified gaps in the literature, the aim of the research is to better understand how children's educational experiences are shaped by horizontal inequalities in Ethiopia – with a focus on ethnic group differences – and the extent to which environmental factors can help to explain those differences. The study will therefore attempt to answer the following four questions:

1. To what extent does *children's educational achievement* differ by ethnicity at age 15?

2. To what extent does *children's schooling participation* differ by ethnicity at age 15?

3. If differential patterns exist, what factors might help to *explain inter-group differences* in achievement and participation at the child, home, community and school level?

4. How does schooling participation *impact on children's cognitive development* between age 12 and 15 across ethnic groups?

2.3. Conceptual model

Figure 2.2 presents the general conceptual model developed to address the research aims and questions. The model is based on the academic literature, the Ethiopian education context, the available data, and the hypothesised relationship between achievement, schooling progression and cultural status. As illustrated in the model, the research focuses on one form of cultural identity – ethnic group status.

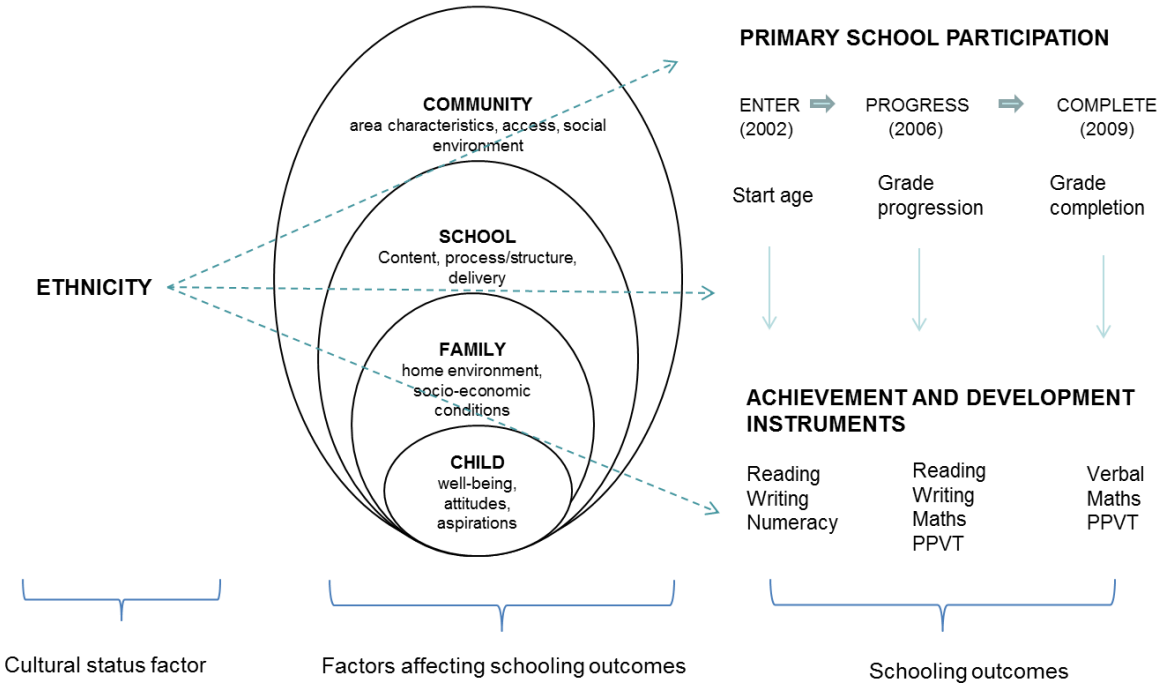


Figure 2.2 Conceptual model of factors predicting children's educational outcomes

It is hypothesised that differences in educational outcomes between ethnic groups are explained by child, home, school and community factors. The research suggests that home disadvantage or income poverty-based explanations – which are the most commonly cited determinants of educational deprivation – are insufficient to account for ethnic gaps in schooling outcomes. It is argued that a range of other factors, at the individual, school and community level, also need to be considered.

We anticipate that much of the influence of ethnicity on achievement and progression will be accounted for by individual child factors, including the child's life chances at birth, physical development and schooling history. However, the research further explores the possibility that the school may play a mediating role against individual and family level disadvantage.

3. METHODOLOGY

The current chapter describes the methodological aspects of the research. The section is organised into five parts which seek to describe and justify the research procedures used. The first discusses the research strategy, and further provides an overview of the Young Lives study, on which the current research is based. The second part details the sampling strategy used for the Young Lives study, emphasising the limitations of the sample, particularly with regard to the ethnic composition of the sub-sample of interest. The third describes the instruments, measures and study variables, and further endeavours to link each measure and variable to the relevant theoretical and empirical literature with a view to establishing validity. Section 4 proposes an analytical strategy for addressing the research questions. Section 5 outlines the ethical considerations made at different stages of the research.

3.1. Research strategy

The current research employs secondary data analysis drawing on the Young Lives longitudinal study of childhood poverty, a project led by the University of Oxford's Department for International Development. The study provides rich and extensive data on children's educational experiences in Ethiopia, and on the specific factors that shape them. It therefore provides an appropriate basis for addressing the research aims and questions.

3.1.1. Secondary data analysis

Secondary data analysis refers to the analysis of data by researchers who were not involved in the actual data collection process and for purposes that were not envisaged by those who collected the data (Bryman, 2008: 296). The benefits to using secondary data lie primarily in their ability to provide extensive, high quality data quickly and without a financial cost implication. Large-scale secondary data sets further offer the possibility of studying sub-groups of the original target population and afford social researchers the opportunity to "[discover] new relationships, new theories… [and to] make fuller use of the data" (Bryman, 2008: 298-99).

However, a number of limitations need to be considered when using secondary data, which have implications for how the data can be analysed. First, a secondary researcher may not be familiar with the original data collection process (Bryman, 2008: 300), and may therefore lack information on how the instruments were actually designed/piloted or how the data were recorded and coded. Thus, research based on secondary data may be limited in its ability to account for data quality vis-à-vis the reliability and validity of the instruments and measures. A second issue relates to data relevance. Whilst the availability of wide-ranging data permits the examination of many new and different relationships, key variables may be missing (Bryman, 2008: 300). Finally, large-scale secondary data sets can be complex, often requiring a significant time investment to clean and merge, and to construct the study variables of interest.

3.1.2. Benefits of longitudinal research designs

By extension, longitudinal designs involve the collection of data over time and are therefore particularly well-suited for analysing patterns of change, and with significantly more precision than cross-sectional data – collected at one point in time. With regard to the Young Lives data, the use of repeated measures of a range of educational outcomes, on the same sample of children, enable us to measure cognitive development and schooling participation/progression with greater accuracy than would be the case with cross-sectional survey data.

Longitudinal surveys also afford researchers the opportunity to identify the factors that drive change or the 'causal factors', as we are able to observe the actual order of events (e.g. which of the independent variables precede the dependent one). The additional advantage to using longitudinal data is therefore in their ability to provide information about the direction of causal relationships, with significantly less ambiguity than a cross-sectional design. As regards the Young Lives data, in addition to portraying how a child develops over time, the breadth of data collected further helps to illuminate on the factors that might explain changes in a child's developmental outcomes.

3.1.3. Objectives and design of the Young Lives longitudinal study

The Young Lives data have been collected to investigate topics that relate to the causes and consequences of childhood poverty (Young Lives, 2002). Household and community data collection began in 2002 in 20 purposively selected sentinel sites in each of the four study countries – Ethiopia, India (Andhra Pradesh), Peru and Vietnam. In each of the countries, the study follows 3,000 index children in two age cohorts, 1,000 children born in 1994/1995 and 2,000 children born in 2001/2002, over a fifteen year period.

The research design takes the form of a longitudinal survey of children and their parents/caregivers, and involves both quantitative and qualitative components. The Young Lives surveys therefore provide a unique opportunity to analyse children's development over time in a developing country context, and to examine the socio-economic and demographic factors that might help to explain their developmental processes. Moreover, the Young Lives data sets have never been used to examine cultural group-based inequality in Ethiopia. The current research is therefore unique both in its purpose and design.

A major component of the research involves understanding the role of children's educational experiences as both a cause and consequence of childhood poverty. Thus, as a complement to the household survey, in 2009 a comprehensive school survey was introduced. The school survey was designed to capture more detailed information about children's schooling with a view to linking the school data with the main survey components in order to analyse the relationship between schooling and child development (Young Lives, 2011). Taken together, the Young Lives data therefore offer a good basis for addressing the research aims and questions of the current study.

3.2. The Young Lives sample

The Young Lives target population in Ethiopia included two cohorts of children born in 2001/2002 and 1994/1995 in 20 sentinel sites. In 2002, a younger cohort of 2,000 children between the ages of 6 and 18 months and an older cohort of 1,000 children between the ages of 7.5 and 8.5 years were selected for the study.

The children are followed over a fifteen year period, which corresponds to 1-16 years of age for the younger cohort and 8-23 years for the older cohort.

The current study focuses on the older cohort children, who were school-aged in 2002. Extensive educational data have been collected for these children over the survey period, 2002-2009. The older cohort was therefore chosen purposively with a view to examining children's educational trajectories over time.

3.2.1. Young Lives sampling strategy

The sampling strategy followed the sentinel site surveillance methodology, which involved a two stage sampling procedure: (i) purposively selecting sites based on pre-determined criteria; and (ii) applying a random sampling of households at the site level. For certain practical reasons, exclusions were made. For instance, among those excluded from the sample were children living in pastoral areas in order to keep the tracking costs manageable (Woldehanna et al, 2011: 16).

During the first stage of sampling, five regions were selected with a view to ensuring national coverage – 96 per cent of the Ethiopian population live in the five selected regions (Outes-Leon and Sanchez, 2008: 6). Figure 3.1 shows the location of each of the sentinel sites. Within each of the regions three to five districts (*woredas*) were selected (20 in total), and within each district a sentinel site was selected corresponding to the lowest level administrative unit in the district (*kebele*). The criteria for the selection of districts and sites included achieving adequate variation in poverty status, ensuring cultural/geographic diversity and a balance between urban and rural areas (Outes-Leon and Sanchez, 2008: 6). At the second stage of sampling, a village within each sentinel site was randomly selected and all households within the village and its surrounding areas were interviewed until 150 households were reached (Outes-Leon and Sanchez, 2008: 7). This corresponds to 100 younger cohort household and 50 older cohort households.

In order to fully exploit the range of educational data available, the current study also uses subsets of the Young Lives school survey. In 2010, school data were collected in all 20 sites covering 2,132 children in 120 schools – 692 children in

the older cohort, 952 younger cohort and 487 older cohort siblings (Young Lives, 2011: 6-7). Different sampling procedures were used in urban and rural areas. In rural areas, all schools attended by a Young Lives child/older cohort sibling were surveyed (Young Lives, 2011: 6-7). In urban areas, 40 per cent of the younger and older cohorts were selected randomly. Within the selected schools, all Young Lives children were surveyed (Young Lives, 2011: 6-7). Thus, in both urban and rural areas, only the Young Lives children were selected for the school survey, class peers were not included.

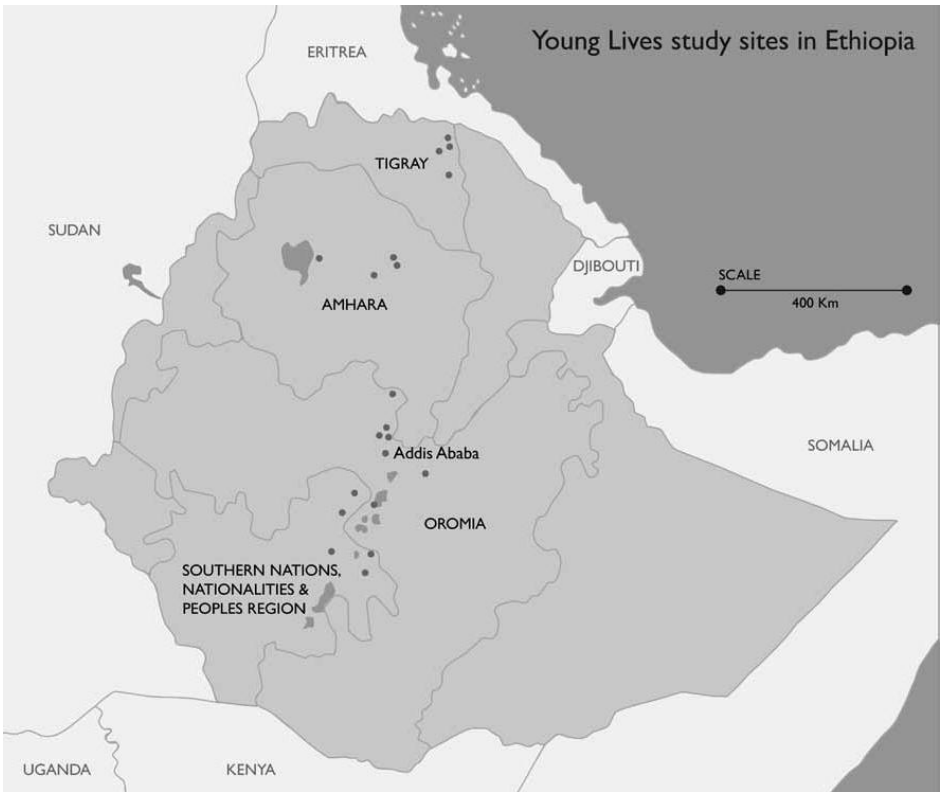


Figure 3.1 Young Lives study sites in Ethiopia (Source: Woldehanna, 2011: 19)

3.2.2. Addressing the limitations of the sampling strategy

Certain measures have been taken to address issues in the design of the school survey. First, only general indicators of school quality were used with a view to better understanding the broader characteristics of the schools in which the Young Lives index children are studying. Data on teaching and learning practices were not considered. This was mainly due to the relatively small size of the older cohort school sample (n=692), the large number of missing observations, the fact that class peers were not selected for the survey, and the

lack of curriculum-based educational achievement data which limited the scope for an analysis of school and classroom level effectiveness.

Biases in the household sample which had implications for the current study, have also been accounted for. First, households in the sample appear to be slightly better-off than the average Ethiopian household in terms of income and access to services. However, they are less likely to own assets (Outes-Leon and Sanchez, 2008: 22). Income, access to services and asset ownership are all likely to affect children's/household schooling decisions. Therefore, the current study examines each socio-economic variable separately, rather than examining a composite measure of household wealth or poverty status. The current research also takes into account the limitations of a non-representative data set. We focus on patterns of change between children in the study regions, rather than national patterns of schooling outcomes. Using the data in this way reduces the potential for non-random bias resulting from a non-representative sample.

3.2.3. Ethnic composition of the Young Lives older cohort sample

The most critical sampling issue for the current study relates to the ethnic composition of the older cohort children. Whilst the sample includes a range of ethnic groups in each of the study regions, many are represented in relatively low numbers. Figure 3.2 shows the ethnic composition of the older cohort children. The figure illustrates the small size of many of the sampled ethnic groups covered by the study, with some having only a few observations.

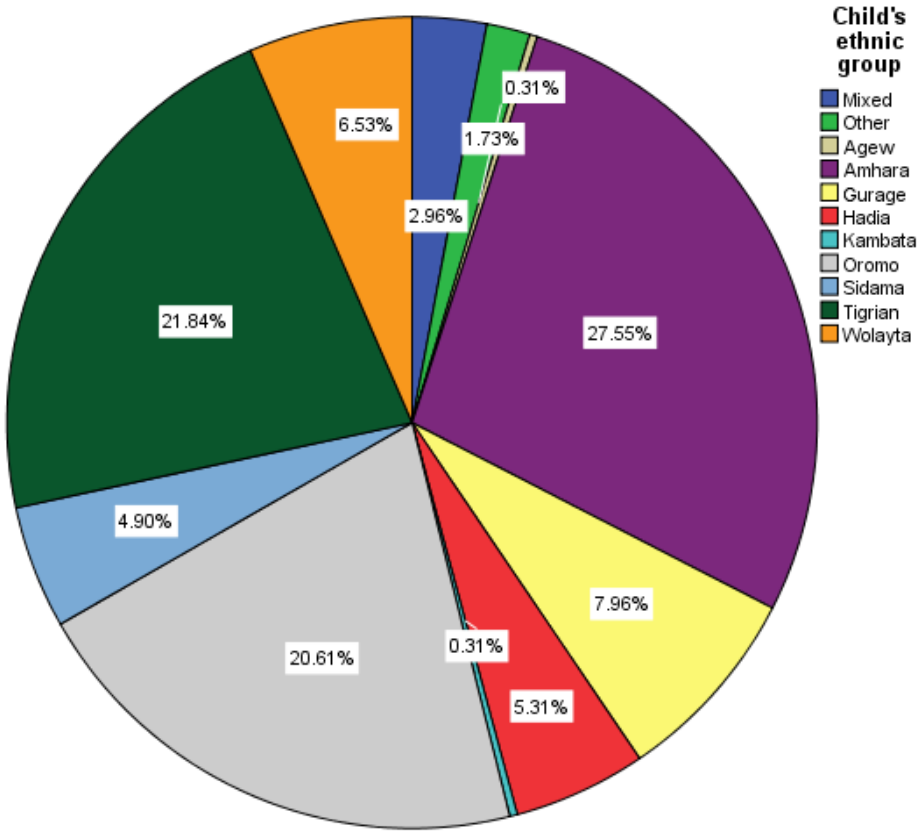


Figure 3.2 Ethnic composition of the Young Lives sample of older cohort children

To address this limitation, two measures were taken which reduced the number of ethnic group categories from eleven to eight. Children belonging to ethnic groups that are significantly underrepresented in the sample ($n < 10$) were excluded from the study in order to minimise the effect of data scarcity. This affected the samples of children of Agew and Kambata ethnic origins. Children who reported 'mixed' as their ethnic status were designated their mother's ethnicity. Mother's ethnicity was used instead of father's as there were fewer missing cases for mothers ($n = 25$) than for fathers ($n = 55$).

Figure 3.3 shows the ethnic composition of the older cohort children by region. The figure reflects the high level of ethnic homogeneity in three of the five study regions: Amhara, Oromia and Tigray. This is consistent with the ethnic composition of the regional populations. According to the most recent Ethiopian

national census, 92 per cent of the local Amhara population is inhabited by the Amhara ethnic group, 97 per cent of the population of Tigray region is of Tigrian origin and 88 per cent of the local Oromo population belong to the Oromo ethnic group (National Population Census, 2007: Table 3.1).

The Addis Ababa sample is significantly more heterogeneous in ethnicity. However, the Amhara ethnic group dominate the sample at 34 per cent. This is also reflective of national figures which suggest that although all ethnic groups are represented in the Capital, the Amhara are the largest group composing nearly half of Addis Ababa's population (National Population Census, 2007; Table 3.1). The SNNP sample is also notably heterogeneous and reflective of the actual diversity within that region, which is composed of over 45 indigenous ethnic groups (National Population Census, 2007: Table 3.1).

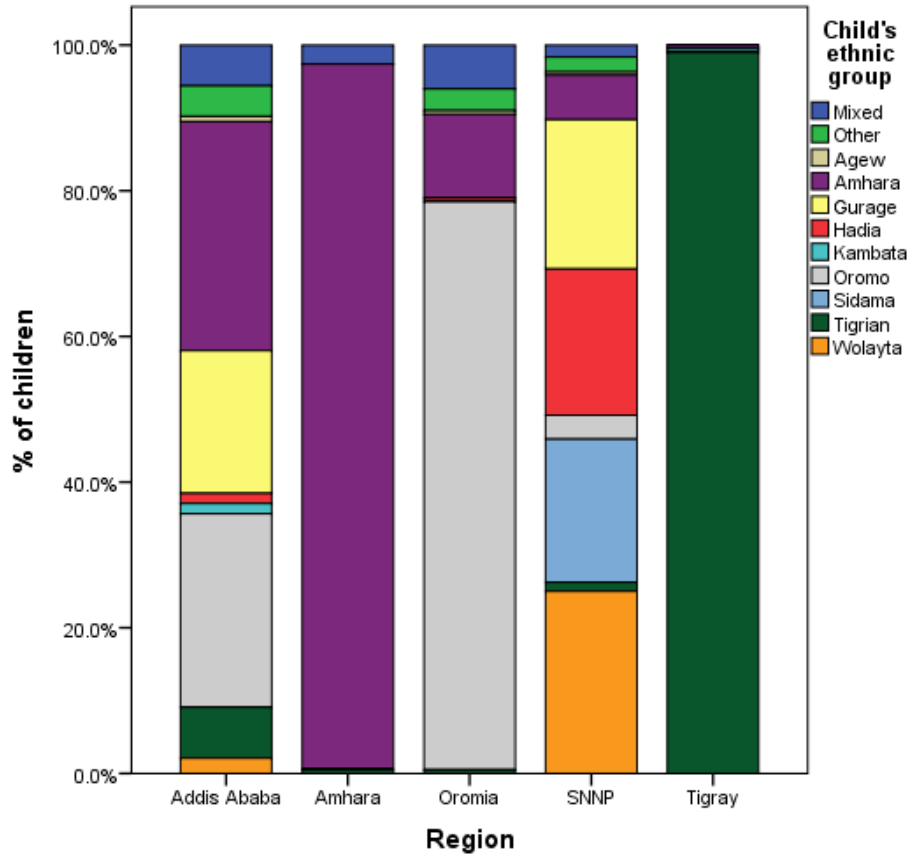


Figure 3.3 Ethnic composition of the Young Lives sample of older cohort children, by region

3.3. The Young Lives data

The Young Lives children and their parents/caregivers were followed up on two occasions, in 2006 and 2009, with a fourth round of data collection planned for 2012-13. The research instruments include both quantitative and qualitative tools. Quantitative data were collected through structured questionnaires. Closed items were used to gather more factual information about the child, home and community characteristics, while multiple-scale items were used to explore perceptions, attitudes and behaviour.

3.3.1. The instruments used for the current study

Table 3.1 provides a snapshot of the Young Lives research process in terms of the instruments used. The current study proposes to use the main quantitative household (caregiver/child) data as a primary source of information. As a secondary source, the school structured observation data and principal questionnaire will be used for general information about the quality of the schools in which the Young Lives older cohort children study. These data sets were chosen for their relevance to addressing the current research questions.

The instruments were developed through an iterative process that involved working with national and regional teams to translate and adapt international items to local contexts, as well as to pilot and refine the instruments accordingly (Morrow, 2009). Therefore, each measure has been carefully designed with a view to maximising inter-country comparability while at the same time ensuring country level relevance and appropriateness.

Table 3.1 Young Lives longitudinal data collection framework

HOUSEHOLD QUESTIONNAIRE*					
	Round 1	**Round 2**	**Round 3**	**Round 4**	**Round 5**
	2002	2006-2007	2009	2013	2016
Younger cohort	6 to 18 months	4 to 5 years	7 to 8 years	11 to 12 years	14 to 15 years
Older cohort	7 to 8 years	11 to 12 years	14 to 15 years	18 to 19 years	21 to 22 years
QUALITATIVE INTERVIEWS AND CASE STUDIES					
	Round 1	**Round 2**	**Round 3**	**Round 4**	
	2007	2008	2011	2014	
Younger cohort	5 to 6 years	6 to 7 years	9 to 10 years	12 to 13 years	-
Older cohort	12 to 13 years	13 to 14 years	16 to 17 years	19 to 20 years	-
SCHOOL SURVEY					
	Round 1	**Round 2**			
	2010	2012			
Younger cohort	8 to 9 years	10 to 11 years	-	-	-
Older cohort	15 to 16 years	17 to 18 years	-	-	-

* The household questionnaire is complemented with a community level questionnaire targeting community leaders. The community protocol is implemented with each round of quantitative data collection.

The quantitative questionnaires were implemented through structured interview-based techniques with a view to minimising the risk of response inaccuracy and misunderstanding; however, self-administered questionnaires were also used for older children to cover more sensitive topics (Boyden, 2009). As regards the school component, the structured observation data were collected by field workers to record specific characteristics of schools and classrooms in a standardised, comparable format. The main benefit to using structured questionnaires and observations for the current research lies in their ability to:

- provide a systematic way of studying children's educational experiences;
- provide extensive standardised data on children's schooling, which facilitates comparisons of schooling experiences between children and between study sites; and
- enable the precise measurement of aspects related to children's learning.

3.3.2. Measures of educational outcomes

Three measures of cognitive ability were used to investigate educational achievement. The measures are based on assessments of literacy, numeracy and general cognitive skill. The outcome variables are introduced in this section and further described in Appendix I.

At age 8 and 12, tests were carried out to measure basic reading and writing ability. The tests involved reading a sentence from a card and writing a sentence that was read aloud. Children's reading levels were assessed on a 4-point scale according to whether they could not read, read letters, read words, or read sentences. Writing ability was measured on a 3-point scale – not able to write, able to write with difficulty/errors and able to write without difficult.

For the current research, reading and writing ability scales were recoded in order to produce a comprehensive measure of literacy corresponding to the expected literacy levels for Ethiopian children at age 8 and 12. In line with the Ministry of Education's Minimum Learning Competencies framework, by age 8 Ethiopian children should be able to read at least words and to write, even with errors. By age 12, the children should be able to read sentences and write with ease (Ethiopian Ministry of Education, 2008). Thus, literacy at age 8 was computed on a 2-point scale: (i) not able to read or write; and (ii) able to read at least words and write, even with difficulty. Literacy at age 12 was computed on a 3-point scale so as to assess how the children have progressed over the 4-year period as follows: (i) not able to read or write; (ii) able to read at least words and write even with difficulty; and (iii) able to read at least sentences and write without difficulty. At age 15, a set of 24 cloze (or fill in the blank items) items were used to measure verbal ability on a 2-point scale (correct/incorrect).

Basic numeracy at age 8 was assessed by a single arithmetic calculation. At age 12, the children were tested with a 10-item maths achievement test scored 1 for correct and 0 for blank or incorrect. At age 15 a more complex 30-item test was administered. Most of the items included in the test were selected from the Trends in International Mathematics and Science Study, 2003 (Cueto et al., 2009: 14). A range of items were selected to ensure sufficient variation in the level of

difficulty. In addition, age 8 basic numeracy items were retained with a view to ensuring comparability over time.

To measure general cognitive ability at age 12 and 15, the Peabody Picture Vocabulary Test (PPVT) was used. The PPVT is a widely-used test of receptive vocabulary and is internationally used for measuring vocabulary acquisition (Cueto et al., 2009: 12). The test is highly correlated with some commonly-used measures of intelligence (Cueto et al., 2009: 12). It is therefore considered a useful measure of assessing general intelligence and scholastic aptitude.

The age 12 and 15 mathematics and PPVT scores, as well as the age 15 verbal test scores, were then converted into Rasch scores in order to allow for measurement on a uni-dimensional interval scale. However, there are significantly more missing observations when the Rasch scores are employed (54 per cent of the data are missing for PPVT scores and 17 per cent for both math and verbal). As cognitive outcomes are the main concepts of interest, loss of power due to missing data was considered of primary concern. Thus, we used instead the corrected raw scores – scores corrected for poor statistical behaviour.

Validity and reliability tests were carried out by Young Lives researchers on each of the cognitive development and achievement instruments. In order to assess the consistency/stability of the test scores, Classical Test Theory and Item Response Theory were used to estimate reliability indicators (Cueto et al., 2009: 3). The validity analysis sought to establish the extent to which test scores (of various proposed instruments) were supported by theory and reported evidence – for instance, whether test scores correlate with parental educational level as expected, and whether children in higher grades perform better on the tests than children in lower grades as expected (Cueto et al., 2009).

Schooling participation was measured progressively by three different variables selected/constructed to measure how well the children progressed through the school system. First, the age at which the children started formal school was computed to assess the extent of late enrolment. This variable demonstrated significant variation by a range of contextual variables and was therefore deemed

most appropriate for addressing aims and questions of the current research. Age of entry into formal school was not immediately available in the data set. The variable was computed by subtracting the number of years of schooling children received from their current age (age 15), which, in turn, was computed as the year in which they started school subtracted from the current year. Grade repetition was computed to assess how well the children progressed through the grades. The official age of entry into formal school in Ethiopia is 7. Therefore, assuming normal age-for-grade progression, the children would be expected to be studying in Grade 8 in 2009. Grade repetition was therefore computed using the following formula (equation 4.1):

Number of grades repeated = (8 – highest grade completed) – (start age – 7) … 3.3.1

The final measure of schooling participation, school completion, was measured by the highest grade completed by the child at age 15. This variable was readily available in the data set.

It is worth noting that a range of non-cognitive measures were also available in the data set. More dynamic attitudinal aspects of the child – such as the child's sense of agency, self-esteem, school well-being and educational aspirations – were analysed using exploratory factor analysis. These constructs provided high reliability scores using to Cronbach's alpha (0.71) – a measure of internal consistency. Moreover, the addition of these variables to these models, particularly pride and educational aspirations, helped to explain some of the variance in test scores (e.g. motivational factors explained 9 per cent of the variation in PPVT scores over and above 'structural' child characteristics such as health and nutrition). However, despite providing a good model fit, the variables were not retained for the current study, as they were not considered as directly relevant for addressing the research questions.

3.3.3 Predictors of educational outcomes

Influences on cognitive achievement and schooling participation were examined at three levels – at the individual child level, within the home and community, and at school. Table 3.2 provides a summary of how each of the study variables

relate back to the academic literature with a view to establishing face validity of each variable and measure used. The variables are discussed descriptively in Chapter 4; more technical details (including scales and scale type) are provided in Appendix I.

Table 3.2 Sources for measures and explanatory variables used in the current study

Level	Measures	Variables	Sources
Child	Gender	Sex	USAID, 1994; Culclough, et al., 2000; WFP, 2003; Sharp et al., 2003; National Learning Assessment, 2007; Admassu, 2008; Dercon and Sanchez, 2009; WHO, 2011.
	Hunger and nutrition	Age-for-height	
	Early cognitive support	Years of early childhood care	
	School participation	Age of entry into formal school	
	Grade progression	Number of grades repeated	
	Child work	Hrs.of household work per day Hrs. of productive work per day	
Home/ community	Mother's education	Mother's years of schooling	Krishnan, 1996; Weir and Knight, 1996; World Bank, 1998; Weir, 1999; Colclough et al., 2000; Sharp and Devereux, 2004; Woldehanna et al., 2005; Rose and Dyer, 2008.
	Household wealth/ ability to pay for school	Housing quality index score Services quality index score Asset index score	
	Sibling age order	Number of older siblings Number of younger siblings	
	Geographic residence	Urban or rural residence	
School	Principal's education	Principal's highest qualification	Colclough et al., 2000; Rose and Al Samarrai, 2001; World Bank, 2005; Chaudhury et al., 2006; Heugh et al, 2006; Verwimp (1999)
	School flexibility	Full-day, half-day or shift school	
	Quality/condition of school facilities	Security of school Maintenance of school Availability of toilet facilities Availability of library Availability of play area	

Six variables were used to assess child well-being in relation to schooling at the individual child level: gender; height-for-age at age 8, 12 and 15; the number of years of early childhood care; school entry age; the number of grades repeated;

and the number of hours of work per day (both household work and productive work outside of the home).

At the household level, three proxy variables were used to examine family socio-economic status – a housing quality index, an asset index and a services quality index. The housing quality index is composed of indicators of housing conditions, including the number of rooms and the type of material used to construct floors, and walls (Boyden, 2009). The asset index relates to the ownership of consumer durables – such as a radio, television, fridge, bicycle, mobile phone, etc. (Boyden, 2009). The services quality index measures of the availability of services at the household level, including electricity, water, sanitation facilities and cooking fuels (Boyden, 2009). Mother's highest educational qualification was available for 966 sample members. Father's highest educational qualification was available for significantly fewer participants at 913. Taking into account the high number of missing observations for fathers but also the importance of mother's literacy for children's education in Ethiopia, mother's educational level was retained for the study. The number of older and younger siblings was also used to examine the effect of sibling age order.

At the school and community level, urban vs. rural residence was used to examine its effect on educational outcomes. The school principal's highest educational qualification was used as a proxy for the overall school quality. Drawing on the literature and taking into account data availability, additional variables were selected to measure school quality: the security of the school, the cleanliness or maintenance of the school, and whether the children have access to certain facilities – sanitation facilities, a library and a play area. The demographic data were available for all of the older children, whereas the school level information was available only for the sub-sample of the children selected for the school survey.

3.4. Analytical strategy

3.4.1. Differential patterns in educational outcomes

The first educational outcome measured in this analysis is educational achievement at age 15; the second is schooling participation. The data are analysed in a multiple regression framework, which identifies the contribution of each explanatory variable to explaining the variation in the dependent (outcome) variable. For each measure of educational achievement (literacy, numeracy and general scholastic aptitude) and for each measure of schooling participation (age of entry into formal school, number of grades repeated by the age of 15 and the highest grade completed by age 15) groups of explanatory variables are entered hierarchically, in logical blocks which are specified in the following steps.

1. *Base model*: The first model establishes the association between educational outcomes and ethnic group membership after controlling for region.
2. *Child characteristics*: Individual child level variables that are relatively stable over time are then entered. These include influences related to the child's physical health and well-being.
3. *Family background*: This block takes into account structural features of the home that remain relatively fixed over time – such as family assets, access to services, housing quality and parental educational qualifications.
4. *School and community*: The last block includes contextual variables related to the school and community.

The four single-equation models are represented as follows:

$$Y_i = \alpha_1 + \sum_j \beta_j \, eth_j + e_i \quad \text{… 3.4.1.1}$$

$$Y_i = \alpha_1 + \sum_j \beta_j \, eth_j + \sum_k \gamma_k \, C_{ki} + e_i \quad \text{… 3.4.1.2}$$

$$Y_i = \alpha_1 + \sum_j \beta_j \, eth_j + \ + \sum_k \gamma_k \, C_{ki} + \sum_h \theta_h \, F_{hi} + \ e_i \quad \text{… 3.4.1.3}$$

$$Y_i = \alpha_1 + \sum_j \beta_j \, eth_j + \ + \sum_k \gamma_k \, C_{ki} + \sum_h \theta_h \, F_{hi} + \sum_m \varphi_m \, S_{mi} + \ e_i \quad \text{… 3.4.1.4}$$

where the βs are the estimated coefficients and the e_s capture the unobservables. The educational outcome variable, Y_i, is defined as the corrected raw test score in a particular cognitive domain, school start age, number of grades repeated or highest grade completed. The ethnicity dummy variable, *eth* (introduced in equation 3.4.1) takes on the value of '0' or '1' corresponding to the ethnic group in observation. The variable C_i (equation 3.4.2) represents the child's structural characteristic (for all *k* characteristics). Family background and school/community variables are denoted by F_i and S_i, respectively (equations 3.4.3 and 3.4.4). The models are estimated for each cognitive domain / schooling participation variable separately in order to examine the effects of contextual variables on each educational outcome.

3.4.2. Differential patterns in educational progression, age 12-15

The third outcome modelled in this analysis is educational progress over the period 2006-2009. Educational progress is a dynamic concept that encompasses both children's schooling participation and their cognitive development over time. In order to examine children's cognitive development between 12 and 15, we consider both achievement at age 15 and prior achievement at age 12. We further take into account how children have progressed through the school system from the time they start school until the age of 15. Schooling progression is measured by the three variables discussed in section 3.3 – the age at which the children started school, the number of grades they repeat by age 15 and the highest grade they complete by age 15. Simultaneously examining cognitive development and schooling progression enables us to analyse how grade progression impacts on cognitive outcomes.

However, a single-equation model would be inappropriate for analysing the effects of schooling progression on cognitive development due to the endogenous relationship between the measures. That is, estimating the parameters of a single equation without taking into account the parameter estimates of the other equations in the system would produce biased results (Gujarati, 1995: 636).

Therefore, for each cognitive domain (literacy, numeracy and general scholastic ability), we estimate the following simultaneous-equation, or path, model:

$$Y_{i15} = \beta_1 + \beta_2 Y_{i12} + \beta_3 grade_{i15} + \beta_4 grade_{i12} + e_1 \quad \ldots 3.4.2.1$$

$$grade_{i15} = \beta_5 + \beta_6 grade_{i12} + \beta_7 Y_{i12} + e_2 \quad \ldots 3.4.2.2$$

$$Y_{i12} = \beta_8 + \beta_9 grade_{i12} + \sum_j \beta_j eth_j + e_3 \quad \ldots 3.4.2.3$$

$$grade_{i12} = \beta_{10} + \sum_j \beta_j eth_j + e_4 \quad \ldots 3.4.2.4$$

where the βs are the estimated coefficients and the e_s are the error terms. The ethnicity dummy variable, *eth* (equation 3.4.3.3) takes on the value of '1' for the observed ethnic group. The outcome variables are cognitive ability (Y_i) at age 12 and 15 (equations 3.4.2.1 and 3.4.2.3) and the highest grade completed ($Grade_i$) at age 12 and 15 (equations 3.4.2.2 and 3.4.2.4). Both $Grade_{i12}$ and Y_{i12} are also explanatory variables for the other equations defined in the system.

Figure 3.4 illustrates the model as a path diagram which provides a visualisation of the child's educational progression from age 12 to 15.

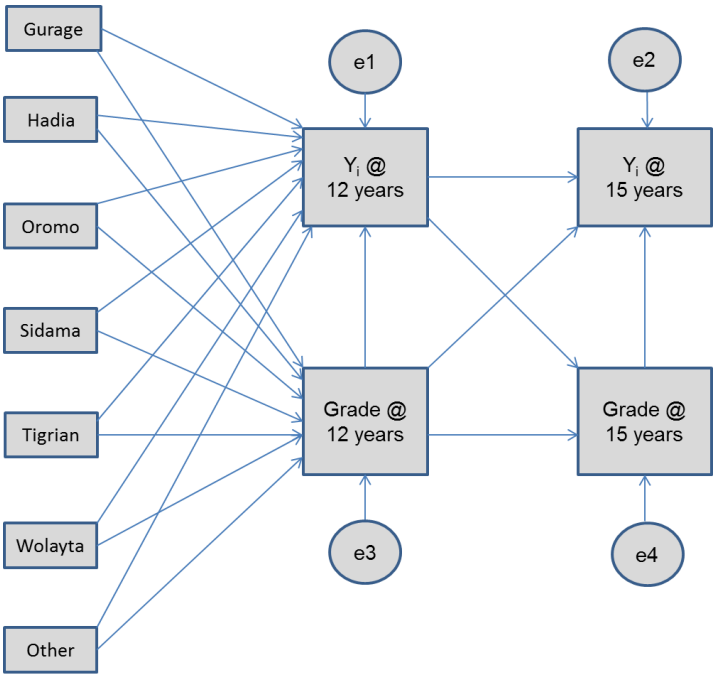


Figure 3.4 : Educational progression path model

3.4.2. Attrition and non-response

Attrition refers to the loss of participants between data collection periods, and is a common problem in longitudinal designs. If attrition is non-random, it may lead to biased inference and undermine the precision of research. However, in the case of the Young Lives Ethiopia sample, the attrition rate among the older cohort children has been low relative to other longitudinal surveys, at 2.1 per cent between 2002 and 2009 (Woldehanna, 2011: 13). Table 3.3 shows that by the end of round three (2009), six of the children in the older cohort sample had died, seven refused to participate in the study and 14 were untraceable.

Table 3.3 Attrition among the sample of older cohort children, 2002-2009

	Year		
	2002	2006	2009
Attrition due to death		6	6
Attrition due to refusal		6	7
Number of untraceable children		8	14
Cumulative attrition rate (%)		1.4	2.1
Sample size	1,000	980	973

Source: Woldehanna, 2011: 13

Attrition rates in the Young Lives study are considered low relative to other longitudinal studies (Outes-Leon and Dercon, 2008: 14). The low level of attrition may be related to the rigorous tracking mechanism set up by the project, effective training of enumerators and follow-up by field researchers (Outes-Leon and Dercon, 2008: 14). The provision of incentives to participants – such as compensation for lost time – may also be a factor (Morrow, 2009).

At the item level, missing data appears to be a problem in relation to the some of the study variables as shown in Table 3.4. Of the six dependent variables, only verbal test scores have a significant amount of missing data, at 17 per cent. As regards the independent variables, only the school quality variables had more than 5 per cent missing data. This is due related to the sampling strategy used for the school survey; a sub-sample of 692 out of a total of 973 older cohort children were selected for the school study.

Table 3.4 Missing Value Analysis for each of the study variables

Variable	N	Missing	
		Count	Percent
Dependent:			
PPVT score (2009)	962	38	3.8
Verbal test score (2009)	832	168	16.8
Math test score (2009)	973	27	2.7
Highest grade completed (2009)	965	35	3.5
School start age (2009)	965	35	3.5
Grades repeated (2009)	965	35	3.5
Independent:			
PPVT score (2006)	953	47	4.7
Math test score (2006)	950	50	5.0
Literacy level (2006)	956	44	4.4
Region (2002)	1000	0	.0
Ethnic group (2002)	974	26	2.6
Sex (2002)	980	20	2.0
Health status (2009)	973	27	2.7
Early childhood care	968	32	3.2
Household work (2009)	973	27	2.7
Productive work (2009)	973	27	2.7
Housing quality (2009)	974	26	2.6

Variable	N	Missing	
		Count	%
Services quality (2009)	974	26	2.6
Consumer durables (2009)	974	26	2.6
Mother's education level (2006)	966	34	3.4
Older siblings (2006)	977	23	2.3
Younger siblings (2006)	976	24	2.4
Rural residence (2002)	980	20	2.0
School principal qualification (2009)	1861*	251	11.9
Type of school (2009)	1862*	250	11.8
Security of school (2009)	1840*	272	12.9
Maintenance of school (2009)	1851*	261	12.4
Availability of toilets (2009)	1869*	243	11.5
Availability of library (2009)	1853*	259	12.3
Availability of play area (2009)	1867*	245	11.6

*Sample includes younger cohort children and older cohort siblings

Missing data have been treated through listwise deletion of missing observations. It is recognised that this approach is considered a less robust way of handling missing data in comparison to single or multiple imputation techniques, particularly when more than five per cent of the data are missing (Little and Rubin, 2002; Schafer and Graham, 2002). By controlling for the uncertainty of *missingness* – particularly with regard to the explanatory variables – data imputation can yield more precise estimates of the parameters, and is therefore preferred to more naïve methods such as listwise/pairwise deletion and mean substitution. However, most of the data scarcity in the current sample was due to the sampling strategy employed for the school survey, which resulted in a sub-sample of the older cohort children rather than the full sample. Moreover, variables for which the extent of missingness was approaching the five per cent level related to the outcome variables. For these reasons, it was decided that listwise deletion would be an appropriate approach.

3.5. Ethical considerations

In addition to completing the Central University Research Ethics Committee (CUREC) protocol, the research proposal for the current study was submitted to the Young Lives Research Coordinator in order to gain access the non-public

school survey data; to verify the study's relevance and feasibility, given the available data; and to obtain additional information about data collection procedures. The proposal was submitted on 16 November 2011 and approved in January, 2012. The data were used solely for the purposes outlined in the research proposal which helped to ensure respect for data protection – anonymity and confidentiality of the participants (Bryman, 2008: 124).

Moreover, the Young Lives project obtained ethics approval from the University of Oxford in 2006. Ethical questions have been addressed through the following provisions: (i) the research fieldworkers underwent an extensive training in which ethical research was discussed; (ii) the fieldworker manuals were designed with explicit guidance on ethical research procedures; (iii) fieldworkers were required to record any ethical questions or emerging dilemmas in their field notes which they discussed with team leaders; (iv) ethics have been discussed with the research and policy teams involving differing academic traditions and disciplines. Further details are provided in Morrow (2009) "*The Ethics of Social Research with Children and Families in Young Lives: Practical Experiences*." Young Lives Working Paper No. 53, Department of International Development, University of Oxford: Oxford.

4. ETHNIC GROUP DIFFERENCES IN EDUCATIONAL OUTCOMES

This analysis focuses on the Young Lives sample of 1,000 older cohort children. Educational data have been collected for these children over the period 2002-2009. Given the longitudinal aspect of the data, and the breadth of data collection, several patterns related to children's schooling can be highlighted. Section 4.1 examines patterns of achievement, overall and by region and ethnicity. This is followed by an analysis of inter-group differences in schooling participation in section 4.2. Section 4.3 presents the key contextual variables that might help to explain the extent of ethnic group differences in the sample.

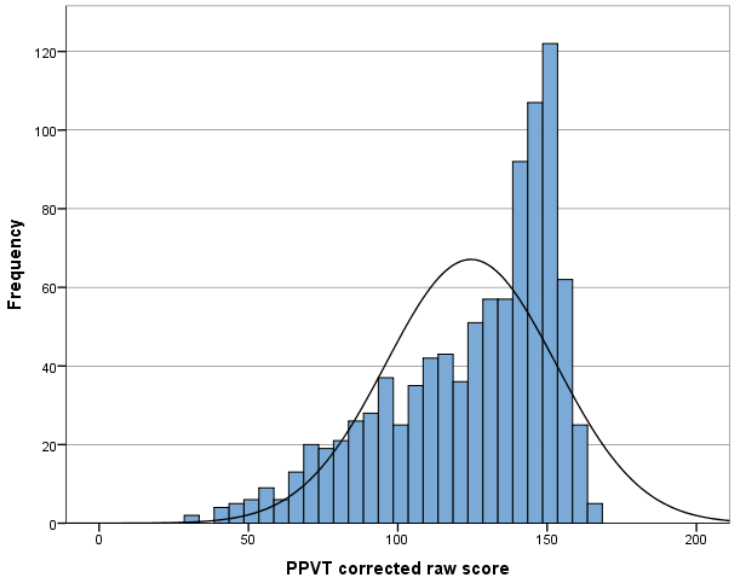
4.1. Differential patterns of achievement

An important starting point to this analysis is to examine achievement at age 15, which is measured by ability test scores in three cognitive domains – general scholastic aptitude, mathematics ability and verbal ability. The analysis begins with an overview of the children's performance which is followed by an examination of the extent of regional and ethnic group differences.

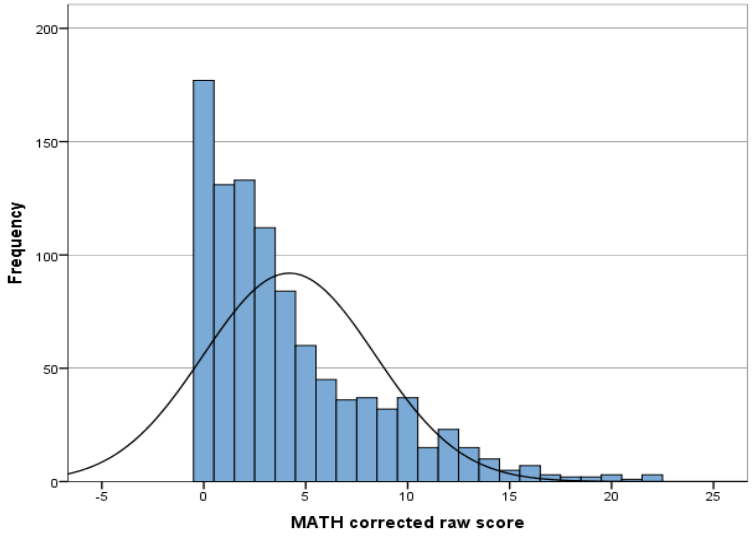
4.1.1. Overall patterns of achievement

Figure 4.1 shows the distribution of the test scores at age 15. Maths and verbal scores appear to have positively skewed distributions with most of the values appearing in the left tail, indicating that the children performed below the expected level for their age.[2] Moreover, both distributions are truncated; with approximately 15 per cent of the children scoring zero points on each test, suggesting that the test items may have been too difficult to capture the full range of the children's abilities. Contrastingly, the negatively skewed distribution of the PPVT scores suggests that test items may not have been difficult enough.

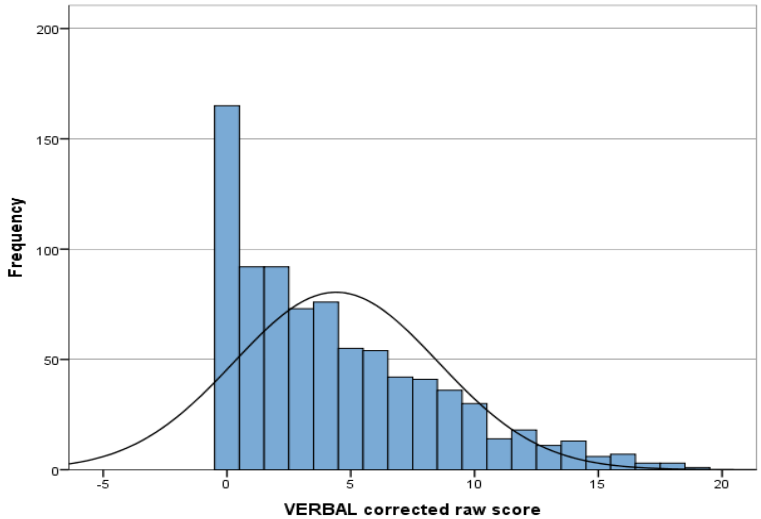
[2] Test items were benchmarked against international standards, as discussed in chapter 4.



Skewness = -0.97
Kurtosis = 0.60



Skewness = 1.28
Kurtosis = 1.46



Skewness = 0.98
Kurtosis = 0.38

Figure 4.1 Distribution of test scores of the older cohort children at age 15

It is also interesting to consider how the children have progressed in terms of their cognitive ability since 2002 (age 8). Figures 4.2 and 4.3 compare reading and writing skills between age 8 and 12 by region, using the constructed measures of literacy discussed in the previous chapter. As figure 4.2 indicates, literacy skills start off lower than the expected level for 8-year olds. At age 8, the majority of the children could not read anything at all (74 per cent). While the children do improve in their literacy skills by age 12, Figure 4.3 shows that a large proportion are still below the expected level for their age; with 33 per cent of the children only able to read words (rather than sentences) and to write with errors. Moreover, 10 per cent of the children remain neither able to read nor write.

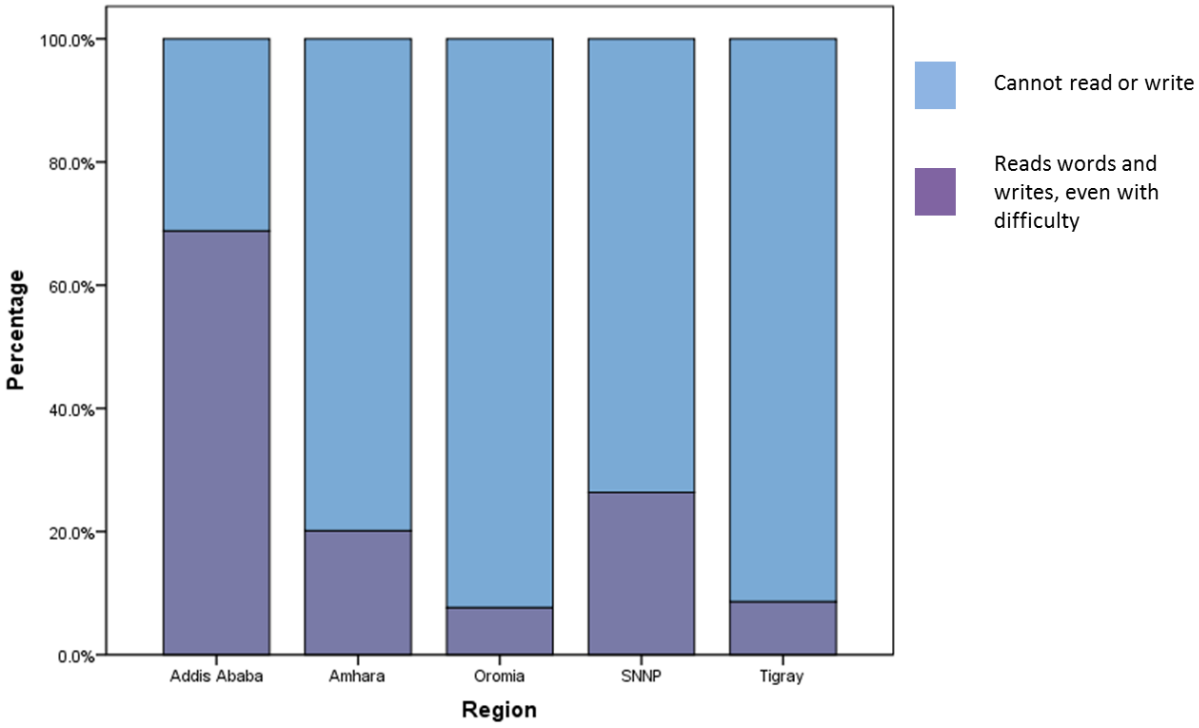


Figure 4.2 Literacy at age 8 for the Young Lives older cohort, 2002

Multivariate tests (Turkey's t-test) were used to examine the significance of the mean-level difference in literacy skills at age 12 between the regions. The results confirmed that the difference is not statistically significant for Addis Ababa and Amhara; where literacy levels appear to be higher in comparison with the other

three regions (p < .05). The difference between Oromia, SNNP and Tigray also appears to be insignificant (p < .05).

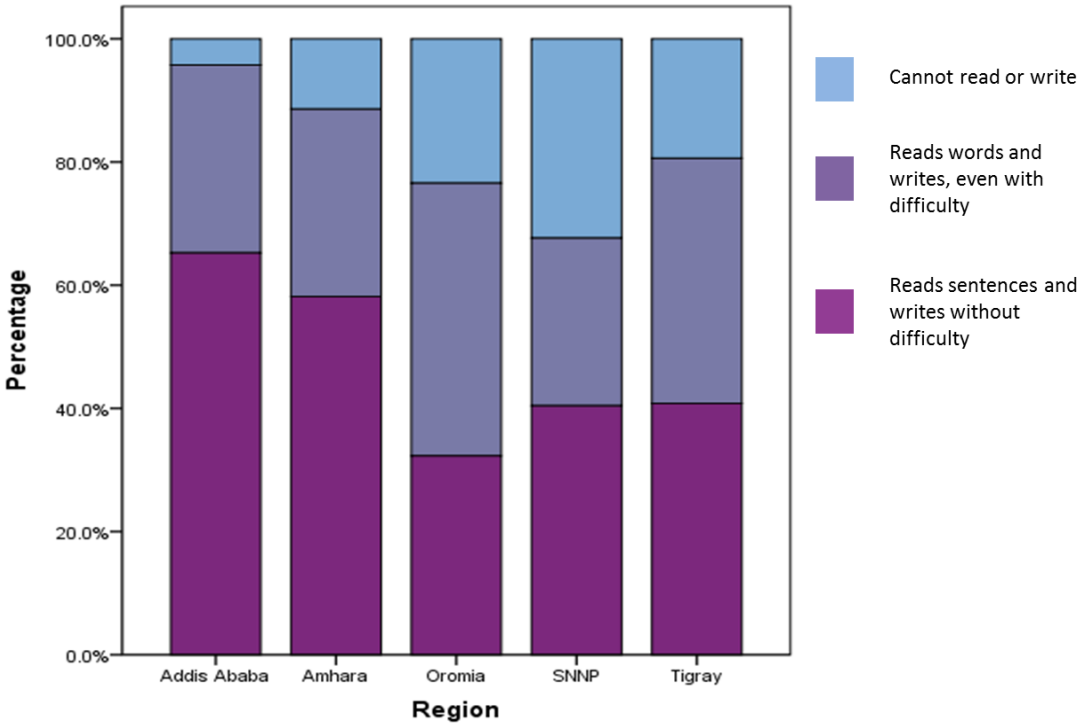


Figure 4.3 Literacy at age 12 for the Young Lives older cohort, 2006

Interestingly, the main spoken and written language in both Amhara Region and Addis Ababa (the capital city) is Amharic, which is the language of instruction in these regions, as well as the official national language (Federal Democratic Republic of Ethiopia, 1994: Article 5). Until 1994, Amharic was also the official language of instruction in formal schools, at which time the Ministry of Education introduced a mother tongue education policy which recommends eight years of mother tongue instruction (Heugh et al., 2007).

With regard to numeracy, only one third of the children were able to compute a simple addition calculation at age 8, with the difference between regions not statistically significant, suggesting a low starting level of basic numeracy across the sample. Figure 4.4 shows the distribution of math test scores at age 12 by region. While the overall mean score is low (4 out of 10), numeracy levels had improved in Addis Ababa (the capital) and Tigray Region, while the others

regions lagged behind. Multivariate tests confirmed the statistical significance of the difference between Addis Ababa/Tigray and the remaining regions ($p < .05$); with no difference between Amhara, Oromia and SNNP regions.[3]

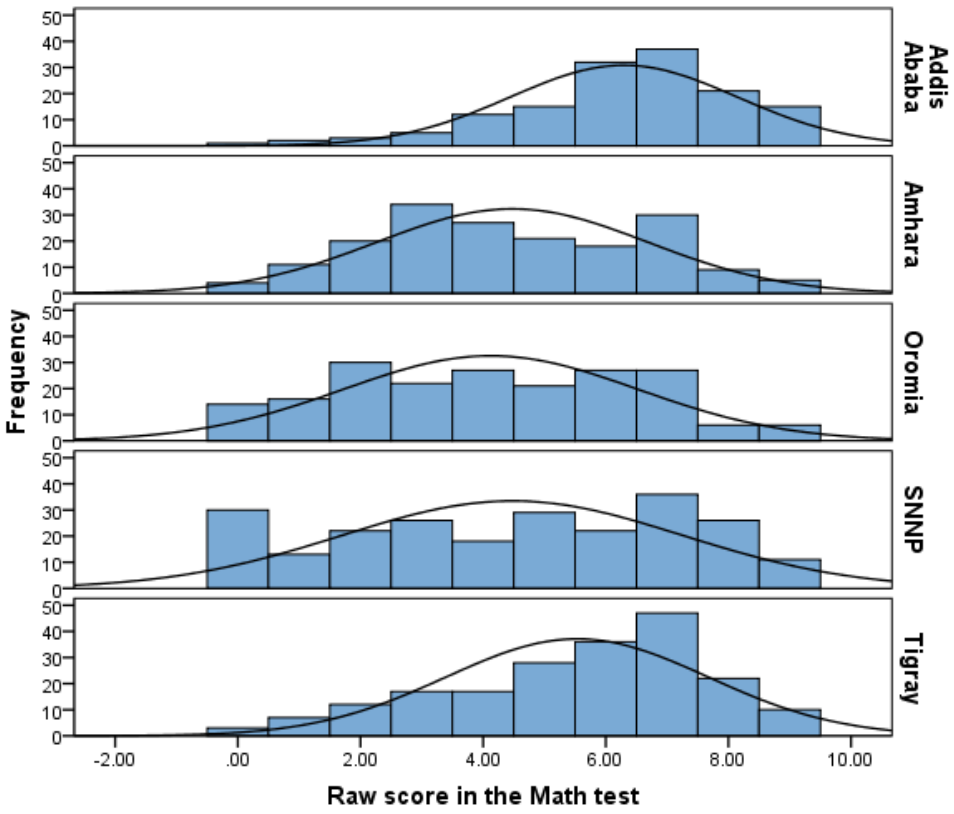


Figure 4.4 Distribution of math scores for the older cohort at age 12, by region

4.1.2. Patterns of achievement by region and ethnicity

Table 4.1 presents the mean and standard deviation of all age 15 test scores by region and ethnicity; with subscripts indicating which region and ethnic groups are statistically similar in terms of performance. Turkey's t-tests were used to examine the significance of the mean-level differences.

Beginning with an analysis of PPVT scores at age 15, the results indicate that children in the capital, Addis Ababa, do better than children in the periphery; with Amhara and Oromia regions having the lowest mean scores ($p < .05$). As regards math test scores at age 15, the results confirm a significant difference

[3] Turkey's t-tests were carried out to examine all pairwise comparisons.

between Addis Ababa/Tigray and the remaining three regions ($p < .05$); which is consistent with the above analysis of age 12 math test results discussed above. For verbal scores, the mean-level difference is significant for all regional pairwise comparisons, with the exception of SNNP and Tigray ($p < .05$). The highest verbal tests scores emerge in Addis Ababa, with the lowest mean scores occurring in Amhara and Oromia.

Table 4.1 Mean and standard deviation of test scores of the older cohort at age 15, by region and ethnicity

	PPVT		Math		Verbal	
	Mean	(SD)	Mean	(SD)	Mean	(SD)
Region						
Addis Ababa	144.34	(12.96)	6.02_{a}	(4.64)	7.70	(4.57)
Amhara	115.03_{a}	(27.84)	3.53_{b}	(3.90)	3.40	(3.30)
Oromia	$117.65_{a,b}$	(33.30)	3.38_{b}	(3.88)	2.37	(3.47)
SNNP	$123.06_{b,c}$	(27.23)	3.76_{b}	(4.03)	4.72_{a}	(4.09)
Tigray	127.88_{c}	(25.42)	4.90_{a}	(4.29)	4.87_{a}	(5.13)
Total	124.51	(28.36)	4.20	(4.22)	4.41	(4.13)
Ethnicity						
Amhara	123.27_{a}	(27.88)	4.37_{a}	(4.10)	4.35_{a}	(4.39)
Gurage	121.89_{a}	(30.12)	5.15_{a}	(3.96)	5.41_{a}	(4.72)
Hadia	107.41_{b}	(20.88)	1.00_{b}	(1.55)	$1.43_{a.b}$	(1.78)
Oromo	122.48_{a}	(33.48)	$3.20_{a,c}$	(3.62)	3.82_{b}	(4.17)
Sidama	$118.71_{a,b}$	(19.75)	$2.83_{b,c}$	(2.56)	$1.96_{a,b}$	(2.72)
Tigrian	$128.87_{a,c}$	(25.23)	5.16_{a}	(4.38)	4.97_{a}	(4.23)
Wolayta	138.40_{c}	(19.41)	6.17_{a}	(4.46)	4.62_{a}	(3.77)
Other	$124.16_{a,b,c}$	(28.35)	$5.06_{a,b}$	(3.99)	$4.26_{a,b}$	(3.69)
Total	*124.16*	*(28.35)*	*4.38*	*(4.12)*	*4.19*	*(4.21)*

Notable mean-level ethnic differences in achievement are also observed at age 15. Children of Gurage and Wolayta origin are among the groups achieving the highest mean scores, while children of Hadia and Sidama origin attain the lowest scores. This points to the significance of intra-regional variation in outcomes as all four of the aforementioned ethnic groups are indigenous to SNNP region.

4.2. Differential patterns in schooling participation

4.2.1. Overall patterns of school enrolment

We continue with an analysis of how the Young Lives older cohort children enter and progress through the school system. Assuming the children start primary school at the official age of entry (which is 7 years); by age 15 the majority of the children would have been expected to be studying in grade 8 or 9. However, while most of the children are indeed enrolled in formal school at age 15, very few have progressed through the grades at the expected rate.

Figure 4.5 illustrates the large spread of the older cohort children across the primary school grades; with a median grade of six. Overall, only 18 per cent of enrolled children are actually in grade 8 or above. Thus, 78 per cent are overage for their grade. A striking 20 per cent of 15-year olds are still studying in the early primary school grades 1-4. This high level of variation in grade enrolment suggests that many children have either started school after the official age of entry and/or repeated grades.

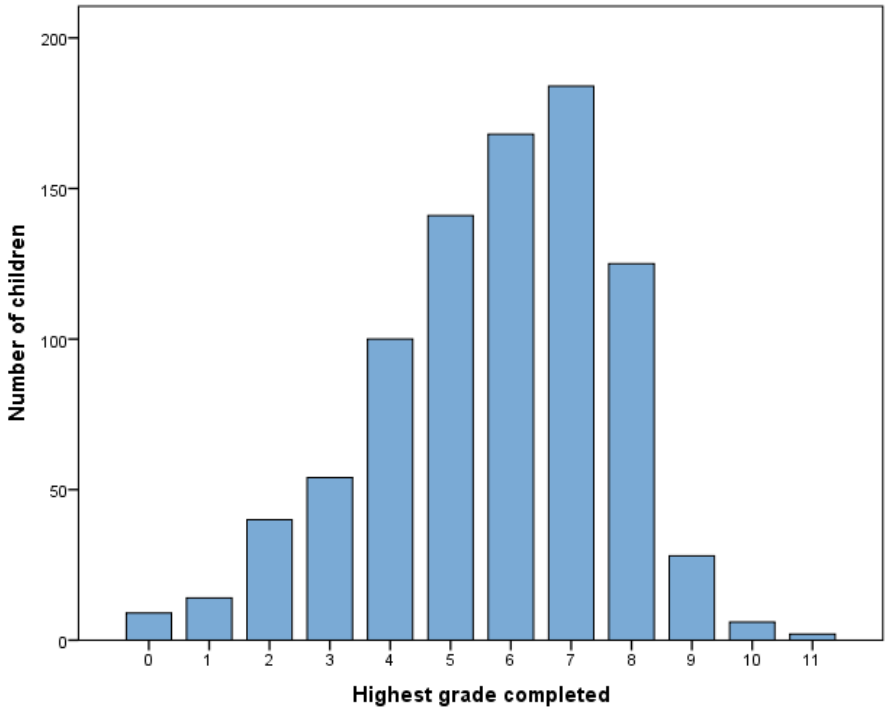


Figure 4.5: Grade enrolment of the older cohort at age 15, 2009

4.2.2. Patterns of enrolment by region and ethnicity

Table 4.2 shows the mean and standard deviation of relevant schooling participation variables; with subscripts indicating which region/ethnic groups are statistically similar. The variables are as defined as follows: (i) *school start age* – measured by the age at which the child entered primary school; (ii) the overall *number of grades repeated* by age 15; (iii) the *highest grade completed* by the child by age 15; and (iv) the *dropout rate* – measured as the proportion sampled children that were not in school at age 15. All values are reported by region and ethnic group.

Beginning with an examination of school start age, the table shows that while the mean age of entry is 8.7 years overall, we observe a large variation across the cohort. Late enrolment appears to be a major issue in each of the sampled regions outside of the capital city (Addis Ababa), and particularly relevant in SNNP where 47 per cent of children started school at age 10 or later. The table shows that there are also substantial differences in start age by ethnicity. Within SNNP region, late enrolment is most important for the Hadia and Sidama ethnic groups; with mean start ages of 11.06 and 9.75, respectively. Grade repetition is common to all regions and ethnic groups. Children are most likely to repeat grades in Oromia Region; with children repeating .84 grades on average by age 15. With regard to ethnicity, among the Sidama ethnic group, on average children repeat one grade. However, the mean-level difference is not significant ($p < .05$).

Table 4.2 Educational progression for the older cohort children, by region and ethnicity

	School start age		No. of grades repeated (2009)		Grade completed (2009)		Dropout (2009)
	Mean	(SD)	Mean	(SD)	Mean	(SD)	(%)
Region							
Addis Ababa	7.40	(1.28)	0.71_{a}	(.89)	6.89	(1.55)	1.40
Amhara	8.50_{a}	(1.97)	$0.67_{a,b}$	(.85)	5.70_{a}	(2.09)	13.22_{a}
Oromia	8.83_{a}	(2.01)	0.84_{b}	(1.20)	5.10_{b}	(2.12)	12.10_{a}
SNNP	9.47	(2.07)	0.75_{a}	(1.41)	4.79_{b}	(2.07)	10.41_{a}
Tigray	8.70_{a}	(1.58)	0.35_{a}	(.68)	5.82_{a}	(1.66)	12.40_{a}
Ethnicity							
Amhara	8.39_{a}	(1.95)	0.78_{a}	(1.04)	5.83_{a}	(2.11)	10.0_{a}

	School start age		No. of grades repeated (2009)		Grades completed (2009)		Dropout (2009)
	Mean	(SD)	Mean	(SD)	Mean	(SD)	(%)
Gurage	8.29_{a}	(2.01)	0.59_{a}	(1.11)	6.12_{a}	(2.21)	8.8_{a}
Hadia	11.06_{b}	(1.74)	0.65_{a}	(0.95)	3.29_{b}	(1.70)	18.0_{a}
Oromo	8.62_{a}	(1.91)	0.88_{a}	(1.28)	5.50_{a}	(2.07)	8.2_{a}
Sidama	9.75_{c}	(2.10)	1.06_{a}	(1.79)	$4.19_{b,c}$	(1.86)	9.1_{a}
Tigrian	8.62_{a}	(1.56)	0.49_{a}	(0.85)	5.89_{a}	(1.64)	11.6_{a}
Wolayta	$8.78_{a, c}$	(1.70)	0.75_{a}	(1.11)	5.48_{a}	(1.64)	8.6_{a}
Other	7.53_{a}	(1.58)	2.00	(2.26)	$5.47_{a,c}$	(2.12)	2.7_{a}
Total	8.70	(1.95)	0.75	(1.17)	5.54	(2.06)	10.8

With regard to grade completion, Hadia and Sidama children are also less likely to reach the expected grade for their age (Grade 8 or 9); with a mean grade level of 3.29 or 4.19. It is not surprising therefore that in SNNP, where indigenous children start school later and are more likely to repeat grades, that the mean grade level is 4.79.

Dropout is a factor in each of the study regions, although less prominent in Addis Ababa. Children of Hadia and Sidama origin have higher dropout rates at 18 per cent and 11.8 per cent, respectively. However, the results are not statistically significant ($p < .05$). It is also important to note that the dropout rate increase as the grades progress. In 2006, when the children were aged 12, around 3 per cent were not in school compared with 10 per cent at age 15.

4.3. Predictors of inter-group differences in schooling participation and performance

The next step towards answering our research questions is to examine the factors that might help to explain ethnic group differences in children's schooling participation and performance. The variables used for this analysis have been selected on the basis of the reviewed quantitative, empirical literature on education in Ethiopia as well as the broader literature on predictors of educational outcomes in developing country contexts. The section also draws on the analysis presented in sections 4.1 and 4.2. Detailed information about the constructed measures and the variables is provided in chapter 3.

Appendix II reports the means and standard deviations of a wide range of socio-economic and demographic factors identified as relevant for children's schooling in Ethiopia by each of the three measures of cognitive ability at age 15 – general scholastic aptitude, math achievement and verbal ability – and by the following educational progression variables – the age of entry into formal school, the number of grades repeated and the highest grade completed. The purpose of Appendix II is to illustrate the difference in educational outcomes between the children according to the deprivation they face at the individual, family, community and school level.

The subsequent discussion presents the contextual variables identified as most relevant for ethnicity. For many of the characteristics, the results show a large regional and ethnic gap.

4.3.1. Child characteristics

Beginning with an examination of individual child factors, we find that in general children living in Addis Ababa appear to be significantly better off across a range of variables in comparison to children in the periphery. Figure 4.6 shows the means and 95 per cent confidence bands for the children's height-for-age z-scores at age 8 (an indicator of early health and nutrition). It is important to note that a child is considered stunted and hence chronically undernourished if his/her z-score is below -2.0, or 2 standard deviations below the standardised mean (WHO, 2011). The mean values appear to be significantly higher in Addis Ababa (-1.00) and lowest in Amhara (-1.81) and Tigray (-1.74) regions.

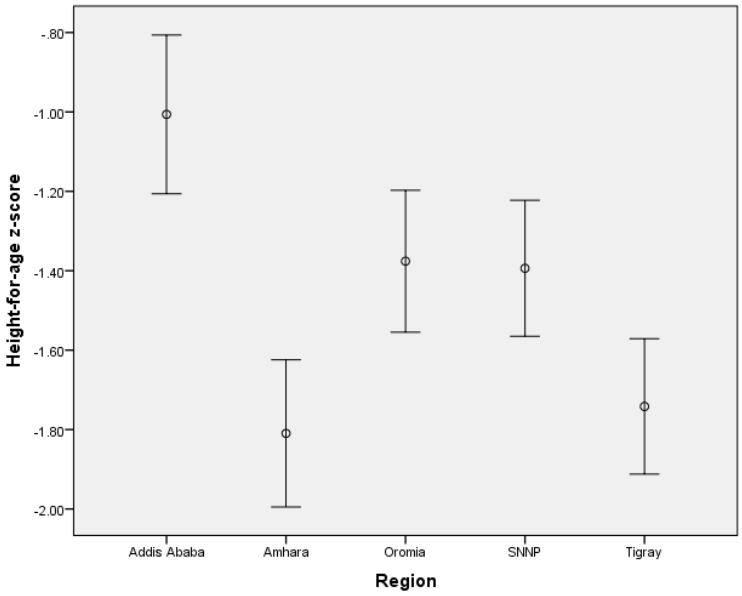


Figure 4.6: 95 per cent confidence interval for height-for-age at age 8, by region

Disaggregating by ethnicity reveals notable inter-group disparities within SNNP region. Figure 4.7 shows the means and 95 per cent confidence bands for age-for-height by ethnicity. Among the four ethnic groups sampled in the SNNP region, children of Sidama (-1.97) origin have the lowest mean-level index scores; with a high level of within group variation. We also observe within group variation among Hadia children (mean of -1.67). Children of Gurage and Wolayta origin appear to be much better of in terms of general health and nutrition; with mean height-for-age z-scores of -1.03 and -1.13, respectively.

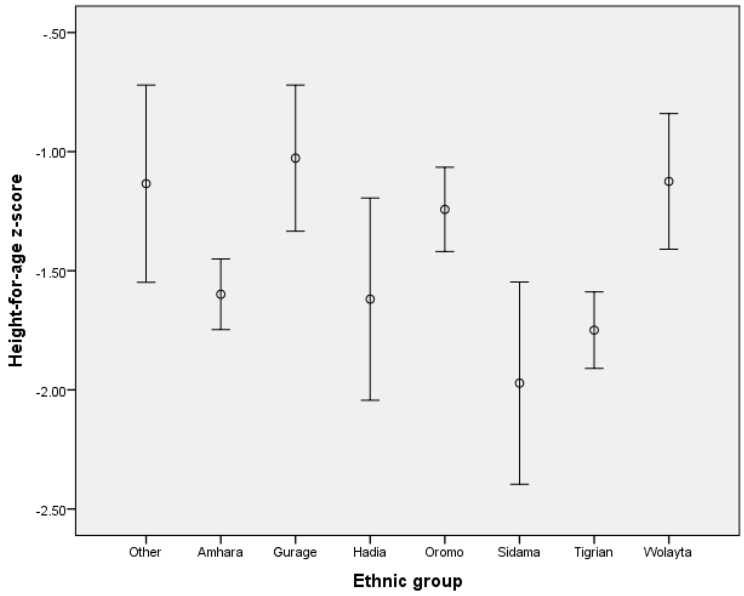


Figure 4.7 : 95 per cent confidence interval for height-for-age at age 8, by ethnicity

Turning to access to early childhood care, recognised as important for children's future cognitive development, Figures 4.8 and 4.9 show that children in Addis Ababa spend significantly more years in early care compared with children living outside of the capital. Comparisons between ethnic groups reveal further intra-group disparities within SNNP Region. Most notably, children of Gurage origin spend an average of 2.34 years in care compared with only .53 years for Hadia children.

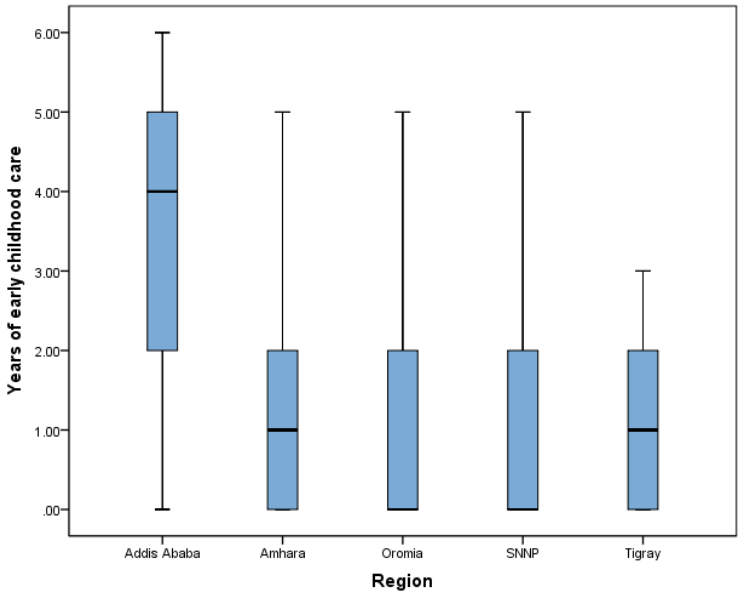


Figure 4.8 : Boxplot for years of early childhood care for the older cohort, by region

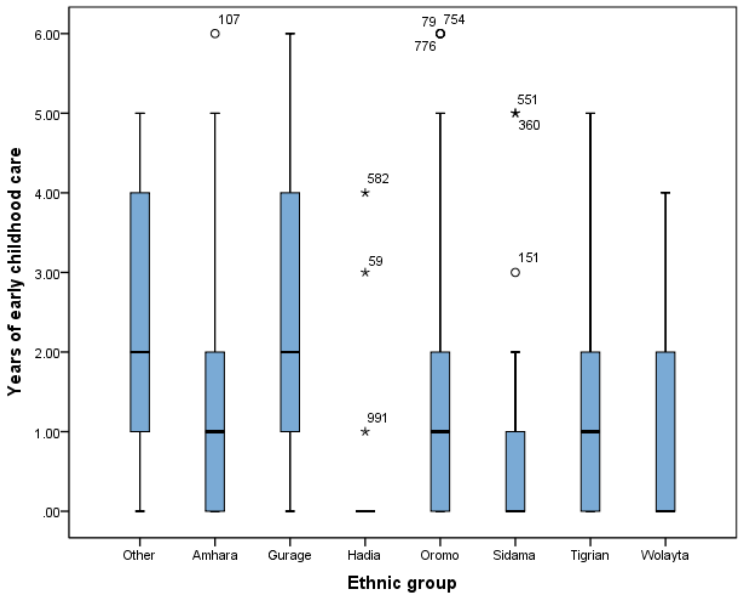


Figure 4.9: Boxplot for years of early childhood care for the older cohort, by ethnic group

Table 4.3 shows the mean and standard deviation of the hours children spend working for the household or for pay/profit; with subscripts denoting the statistical significance of between group comparisons. The regional difference in both household and productive working hours is statistically significant for Addis Ababa only; with children in Addis Ababa working substantially fewer hours per

day than their peers in other regions. With regard to ethnic groups, the most interesting finding relates to Hadia and Sidama children providing more household support – in terms of caring for others or carrying out domestic tasks – than children of other ethnic groups.

Table 4.3 Hours of work per day for the older cohort at age 15 (2009)

	Household work			Productive work			Total work		
	Mean	SD	N	Mean	SD	N	Mean	SD	N
Region									
Addis Ababa	2.21	1.34	141	.46	1.39	141	1.29	.51	141
Amhara	3.42_{a}	2.16	190	$1.85_{a,b}$	2.56	190	2.14_{a}	.79	190
Oromia	3.65_{a}	6.32	198	2.33_{a}	2.46	198	2.27_{a}	.70	198
SNNP	4.06_{a}	2.02	241	1.48_{b}	2.34	241	2.11_{a}	.71	241
Tigray	3.57_{a}	6.44	201	2.30_{a}	2.90	201	2.26_{a}	.73	201
Ethnicity									
Amhara	3.26_{a}	2.03	279	1.49_{a}	2.35	279	$2.02_{a,b}$	.81	279
Gurage	3.10_{a}	2.44	76	1.39_{a}	2.60	76	$1.82_{b,c}$	.75	76
Hadia	$4.58_{a,b}$	1.96	53	1.94_{a}	2.24	53	2.30_{a}	.66	53
Oromo	3.02_{a}	1.62	211	1.91_{a}	2.23	211	$2.07_{a,b}$	.74	211
Sidama	$4.24_{a,b}$	1.49	49	1.59_{a}	2.22	49	2.36_{a}	.66	49
Tigrian	3.46_{a}	6.25	215	2.17_{a}	2.88	215	2.20_{a}	.76	215
Wolayta	3.84_{a}	1.63	63	1.11_{a}	2.32	63	1.68_{c}	.66	63
Other	7.26_{b}	19.88	19	2.68_{a}	3.66	19	$1.89_{a,b,c}$	.87	19

4.3.2. Family background

For descriptive purposes, we consider three proxy variables to examine family socio-economic status – a housing quality index, an asset index and a services quality index. Table 4.4 shows the mean and standard deviation of the variables at age 8 as a predictor of educational progression and achievement. Comparing first across regions, the mean-level differences are significant for few pairwise comparisons ($p < .05$); with the most improved housing conditions found in Addis Ababa and Tigray; services most accessible in Addis Ababa; and asset ownership significantly lower in Amhara and Tigray regions. In terms of ethnicity, Housing quality appears to be lowest among Hadia and Sidama households. Children of Hadia and Sidama origin are also less likely to have access to basic services and own assets.

Table 4.4 Family socio-economic status at age 8 (2002)

	Housing quality index			Services quality index			Asset ownership index		
	Mean	SD	N	Mean	SD	N	Mean	SD	N
Region									
Addis Ababa	0.41 $_{a}$	0.14	141	0.56	0.20	141	0.48	0.15	141
Amhara	0.30$_{b}$	0.16	190	0.37$_{a}$	0.25	190	0.18$_{a}$	0.14	190
Oromia	0.29$_{b}$	0.15	199	0.52$_{a}$	0.18	199	0.34	0.17	199
SNNP	0.24$_{b}$	0.20	241	0.43$_{a}$	0.22	241	0.30	0.19	241
Tigray	0.43$_{a}$	0.19	201	0.32$_{a}$	0.19	201	0.15$_{a}$	0.17	201
Ethnicity									
Amhara	0.33$_{a}$	0.16	279	0.44$_{a,c}$	0.25	279	0.27$_{a}$	0.19	279
Gurage	0.27$_{a}$	0.21	76	0.36$_{a,c}$	0.26	76	0.30$_{a}$	0.20	76
Hadia	0.08$_{b}$	0.11	53	0.39$_{b}$	0.16	53	0.19$_{b}$	0.11	53
Oromo	0.30$_{a}$	0.15	212	0.53$_{c}$	0.17	212	0.37$_{a}$	0.19	212
Sidama	0.22$_{b}$	0.22	49	0.36 $_{b}$	0.16	49	0.27$_{a}$	0.12	49
Tigrian	0.43$_{c}$	0.19	215	0.32 $_{a, c}$	0.19	215	0.18$_{b}$	0.19	215
Wolayta	0.37$_{c}$	0.16	63	0.60 $_{d}$	0.19	63	0.39$_{a}$	0.22	63
Other	0.33$_{a,b,c}$	0.17	19	0.50$_{a,c,d}$	0.20	19	0.39$_{a}$	0.16	19

In terms of mother's educational level by region, the most notable finding relates to the number of mother's that have no educational experience. Figure 4.10 shows similar proportions of mothers in Amhara and SNNP have no education; with a slightly lower proportion of mothers in Oromia having never been to school. A striking 80 per cent of mothers in Tigray Region have no educational experience, compared with less than 20 per cent in Addis Ababa. Looking at educational level by ethnic group (Figure 4.11) shows similarly large proportions of Tigrian mother's having never been to school.

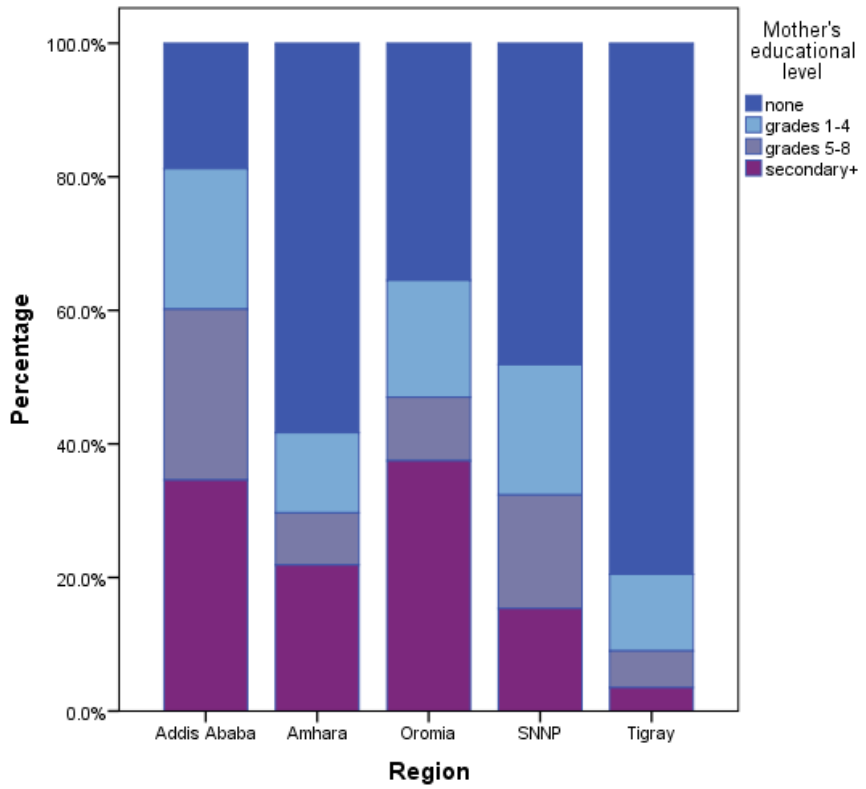


Figure 4.10 Mother's educational level for the older cohort, by region

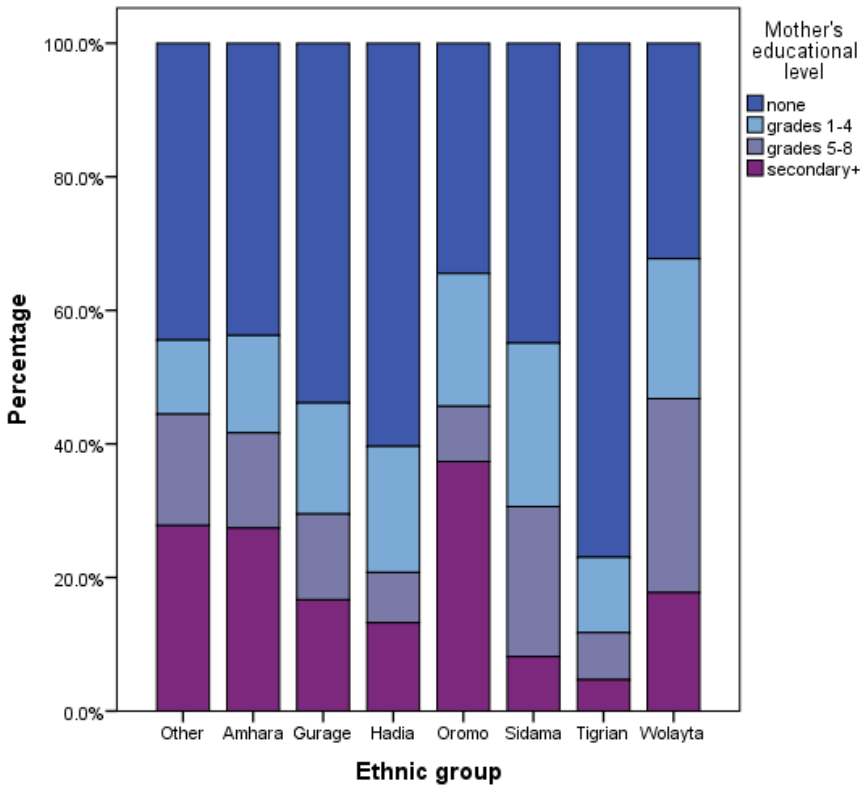


Figure 4.11 Mother's educational level for the older cohort, by ethnic group

4.3.3. School and community

Looking first at the urban-rural distribution of the older cohort sample, Table 4.5 shows that the majority of selected children live in rural areas. We further observe that nearly all of the children of Hadia and Sidama origin live in rural dwellings of SNNP region. In comparison, all of the children of Wolayta origin, which is another ethnic group indigenous to SNNP Region, live in urban parts of the country.

Table 4.5 Urban-rural residence by region and ethnic group

	Urban		Rural	
	N	% of urban	N	% of rural
Region				
Addis Ababa	143	36.11	0	0
Amhara	49	12.37	143	24.49
Oromia	56	14.14	144	24.66
SNNP	97	24.49	147	25.17
Tigray	51	13.79	150	25.68
Total	*396*	*100*	*584*	*100*
Ethnicity				
Amhara	119	30.51	164	28.08
Gurage	33	8.46	45	7.71
Hadia	3	0.77	50	8.56
Oromo	93	23.84	119	20.38
Sidama	1	0.26	48	8.22
Tigrian	66	16.92	150	25.68
Wolayta	64	16.41	0	0
Other	11	2.82	8	1.37
Total	*390*	*100*	*584*	*100*

Tables 4.6 and 4.7 shows the self-reported qualification of principals in the schools in which the older cohort children study. We find that children that attended schools in Tigray Region, and children of Tigrian origin, were more likely to have highly qualified school principals. Interestingly, none of the children of Sidama origin had principals with a high educational level. Similarly, only one child of Hadia origin had a highly educated school principal.

Table 4.6 Principal qualification for the older cohort children, by region

	Region				
	Addis Ababa	Amhara	Oromia	SNNP	Tigray
Number of children	*72*	*165*	*171*	*187*	*165*
Low qualification (%)	1.4	1.8	4.1	4.3	0.0
Medium qualification (%)	47.2	62.4	80.1	59.9	34.5
High qualification (%)	51.4	35.8	15.8	65.6	39.2

Table 4.7 Principal qualification for the older cohort children, by ethnic group

	Ethnic group							
	Amhara	Gurage	Hadia	Oromo	Sidama	Tigrian	Wolayta	Other
Number of children	*216*	*54*	*43*	*173*	*40*	*175*	*45*	*11*
Low qualification (%)	1.9	14.8	0.0	3.5	0.0	0.0	2.2	0.0
Medium qualification (%)	58.8	59.3	97.7	73.4	100.0	36.0	8.9	54.5
High qualification (%)	39.4	25.9	2.3	23.1	0.0	64.0	88.9	45.5

Tables 4.8 and 4.9 describe the older cohort sample by school type. Looking at enrolment by region, we find that with the exception of Addis Ababa, most of the children are enrolled in shift (or flexible) schools. However, a substantial proportion of the children in SNNP (36 per cent) attended full-day schools. Looking at ethnic groups within SNNP region, the table shows that the children of Gurage and Hadia origin are more likely to attend full-day schools. All of the Sidama and 93 per cent of Wolayta children are enrolled in shift schools.

Table 4.8 Type of school for the older cohort children, by region

	Region				
	Addis Ababa	Amhara	Oromia	SNNP	Tigray
Number of children	*73*	*165*	*170*	*187*	*165*
Full-day (%)	100.0	0.0	0.0	36.4	1.2
Half-day (%)	0.0	0.0	20.6	0.0	13.9
Shift (%)	0.0	100.0	79.4	63.6	84.9

Table 4.9 Type of school for the older cohort children, by ethnic group

	Ethnic group							
	Amhara	Gurage	Hadia	Oromo	Sidama	Tigrian	Wolayta	Other
Number of children	*216*	*55*	*43*	*172*	*40*	*175*	*45*	*11*
Full-day (%)	8.8	81.8	88.4	12.8	0.0	5.1	6.7	54.5
Half-day (%)	0.5	0.0	0.0	19.2	0.0	13.1	0.5	9.1
Shift (%)	90.7	18.2	11.6	68.0	100.0	81.7	93.3	36.4

In terms of the availability and quality of school facilities, Tables 4.10 and 4.11 describe the Young Lives sample of schools by region and ethnicity. Children in Addis Ababa were found to be most likely to be enrolled in schools that are secure and well-maintained. However, they were less likely to have access to working toilets and a play area. Libraries were significantly more prevalent in Amhara and Tigray Regions. Children in school in Amhara Region were generally better equipped in terms of these particular facilities compared with children in other regions.

Table 4.10 School facilities by region

	Region				
	Addis Ababa	Amhara	Oromia	SNNP	Tigray
N	*73*	*165*	*169*	*177*	*165*
Boundary is secure (%)	84.9	49.4	41.4	68.4	64.8
Well-maintained (%)	100.0	85.5	99.4	73.4	97.6
Working toilets (%)	80.6	99.4	100.0	95.2	94.5
Library available (%)	79.5	98.2	74.9	73.8	94.5
Play area available (%)	31.5	98.2	65.5	98.9	75.8

Table 4.11 School facilities ethnic group

	Ethnic group							
	Amhara	Gurage	Hadia	Oromo	Sidama	Tigrian	Wolayta	Other
N	*210*	*55*	*43*	*169*	*40*	*174*	*43*	*11*
Boundary is secure (%)	51.9	81.8	11.6	46.7	100.0	65.5	88.4	72.7
Well- maintained (%)	88.6	80.0	76.7	98.2	100.0	97.7	44.2	100.0
Working toilets (%)	97.7	94.4	79.1	98.8	100.0	93.7	97.8	81.8
Library available (%)	95.8	76.4	11.6	75.7	100.0	93.7	97.8	54.5
Play area available (%)	87.0	78.2	100.0	63.6	100.0	74.3	92.7	80.0

5. CONTEXTUAL MODELS OF ETHNICITY AND EDUCATIONAL OUTCOMES

The previous section showed descriptively that there are notable differences in schooling outcomes across ethnic groups in the Young Lives sample of older cohort children in Ethiopia. We further demonstrated that there is a large variation between groups in a range of contextual factors identified as important for children's schooling participation and performance in Ethiopia. In this section we set out to examine the significance of mean-level ethnic group differences in educational outcomes at age 15 and the extent to which this gap can be explained statistically by child, home, community and school factors.

5.1. Modelling the effect of ethnicity on achievement

In the first part of this analysis we investigate the difference between ethnic groups in cognitive achievement across three cognitive domains – general scholastic aptitude, mathematics and verbal ability – and the extent to which contextual factors account for such differences. Separate regression models are developed with a view to: (i) examining differential patterns by each cognitive domain; and (ii) isolating the significance and size of the effect of contextual variables on achievement. For each cognitive outcome, groups of explanatory variables are entered in logical blocks (as discussed in Chapter 4). The results are discussed below.

5.1.1. Contextual models of general cognitive ability at age 15

Table 5.1 presents the results of the multiple regression analysis for general cognitive ability, measured by PPVT test scores at age 15.

Model 1: Base model

Before taking into account the contextual variables, we observe a strong association between PPVT test scores at age 15 and ethnic group status after controlling for geographic region; with children of Gurage, Hadia, Oromo and Sidama origin performing less well than children in the base group (the Amhara ethnic group). Most notably, children of Hadia origin score 29 points lower than ethnic Amhara children. It is important to recall that the three ethnic groups that

achieve the lowest results – the Gurage, Hadia and Sidama – are ethnic groups indigenous to SNNP Region, an ethnically and linguistically diverse region in the southern part of the country.

Model 2: Child characteristics

Taking into account the characteristics of the child, ethnicity remains significant for each of the aforementioned ethnic groups who continue to achieve lower results. We also find most of the variables to be strongly and significantly related to general cognitive ability. The explanatory variables in this block explained 38 per cent of the variation in PPVT scores. The most significant variables related to gender, schooling history, and time spent engaged in productive work.

Model 3: Family background

Model four includes key home factors related to children's schooling in Ethiopia. The variables include index scores for housing quality, services quality and asset ownership. Maternal education and the presence of older and younger siblings are also included. These variables in the model explained 39.6 per cent of the variation in test scores. Housing quality and assets appear to be most strongly and significantly associated with cognitive ability. What is noteworthy in this model is that even after accounting for family factors, children of Gurage, Hadia, Oromo and Sidama origin continue to below Amhara children on the PPVT suggesting that community and school level influences may play a stronger role than home disadvantage.

Model 4: School and community

School and community factors explained an additional 1.3 per cent of the variation in PPVT test scores. The difference is mainly attributed to the urban-rural gap; with children living in rural areas scoring 10 points below urban children. School principal qualification is positively associated with cognitive ability at age 15; with the highest qualification related to a higher score of 7 points. Most of the variables used to measure the quality of school facilities and infrastructure are insignificant, with the exception of the presence of toilets and a play area. These variables are negatively correlated with test scores.

Furthermore, once community and school level influences are controlled for, the effect of ethnicity is no longer significant for Hadia and Sidama children, suggesting that what distinguishes children from these particular ethnic groups is that they may live in communities with poorer quality schools.

5.1.2. Contextual models of math ability at age 15

Table 5.2 presents the results of the multiple regression analysis for maths ability measured in terms of math test scores at age 15.

Model 1: Base model

We observe a strong association between math test scores at age 15 and ethnic status. With the exception of Tigrian children, most of the children perform significantly less well than the Amhara children. Similar to the PPVT results, children of Hadia and Sidama origin perform particularly less well.

Model 2: Child characteristics

After taking into account the characteristics of the child, ethnicity remains an important predictor of maths ability. The inclusion of these variables, particularly gender, years of early childhood care, school start age and grade repetition, improves the fit of the model by 33 per cent.

Model 3: Family background

The inclusion of the family background variables improved the model fit marginally, by one per cent. In addition to ethnicity, the most significant variables are housing quality access to services (e.g. electricity and water), which appear to be more important for maths performance than family asset ownership.

Model 4: School and community

The school and community variables included explained an additional one per cent of the variation in math test scores. The most notable school/community influence is geographic stratum ($p < 0.1$); with children living in rural areas scoring one point below their urban peers.

5.1.3. Contextual models of verbal ability at age 15

Table 5.3 presents the results of the multiple regression analysis for literacy measured by verbal test scores at age 15.

Model 1: Base model

The base model for verbal ability demonstrates a strong positive association between verbal ability at age 15 and ethnicity; with children of Gurage, Hadia, Sidama and Oromo backgrounds achieving up to five points lower than children from the Amhara ethnic group.

Model 2: Child characteristics

After taking into account the characteristics of the child, which explain 41 per cent of the variation in test scores, ethnic status remains significant. Children of ethnic Tigrian origin appear to perform better on the verbal test than all children in the sample. The most significant variables are gender, years of early childhood care, school start age and grade repetition.

Model 3: Family background

The inclusion of family background variables improved the model fit by only 1.4 per cent. Services quality appears to be positive and significant for verbal ability. The presence of older siblings appears to be negatively associated with test scores. The results suggest that after controlling for gender and the child's schooling history, home socio-economic factors are not as strong a determinant of verbal achievement as for maths.

Model 4: School and community

The addition of the school and community variables improved the model fit by 4 per cent, suggesting that school and community factors are more important for verbal ability than family background. In other words, the school may play a compensatory role for individual and family level deprivation, or it may perpetuate deprivation. We observe a strong negative relationship between geographic stratum and verbal scores, and a strong positive association with principal qualification.

Table 5.1 Contextualised models for PPVT test scores at age 15 (2009)

		Model 1 (n=955)		Model 2 (n=945)		Model 3 (n=929)		Model 4 (n=713)	
Variable	Value	β	S.E.	β	S.E.	β	S.E.	β	S.E.
Base model									
Intercept	Intercept	151.054***	3.027	203.160***	8.017	182.408***	8.856	178.405***	13.041
Region	Dummy: 1 if Amhara	-36.068***	3.561	-24.581***	3.598	-15.963***	3.827	1.926	6.113
(Base group: Addis Ababa)	Dummy: 1 Oromia	-27.228***	3.400	-12.607***	3.454	-10.032***	3.533	2.879	4.985
	Dummy: 1 SNNP	-15.033***	3.741	-5.124*	3.454	-1.223	3.490	5.546	5.434
	Dummy: 1 Tigray	-27.452***	6.651	-14.090**	6.002	-5.724	6.253	6.157	7.941
Ethnic group	Dummy: 1 if Gurage	-19.664***	4.388	-18.629***	3.801	-11.339***	3.905	-5.877	5.714
(Base group: Amhara)	Dummy: 1 if Hadia	-28.720***	5.277	-10.614**	4.675	-3.747	4.711	6.414	6.486
	Dummy: 1 if Oromo	-7.079**	3.605	-6.645**	3.100	-3.380	3.173	-1.545	4.036
	Dummy: 1 if Sidama	-17.306***	5.477	-7.866*	4.791	-3.070	4.767	9.248	6.936
	Dummy: 1 if Tigrian	3.888	6.740	2.142	5.787	2.988	5.859	3.609	6.989
	Dummy: 1 if Wolayta	1.661	5.094	3.752	4.430	4.371	4.440	2.561	6.426
	Dummy: 1 if Other	-14.287**	6.590	-12.665**	5.772	-9.598*	5.837	-3.481	8.278
Child characteristics									
Gender	Dummy: 1 if female			-8.868***	1.620	-8.046***	1.607	-8.522***	1.947
Health status (2009)	Index score (0-4)			-2.796*	1.509	-3.129**	1.493	-3.021*	1.664
Early childhood care (2009)	Number of years of care			.657	.921	.026	.926	-.389	1.184
School starting age	Age in years			-6.193***	.634	-5.671***	.634	-4.664***	.807
Grades repeated (2009)	Number of grades			-4.197***	.693	-4.176***	.684	-3.609***	.922
Household work (2009)	Number of hours per day			.041	.173	.135	.170	.138	.242
Productive work (2009)	Number of hours per day			-1.601***	.353	-1.200***	.354	-1.948***	.569
Family background									
Housing quality (2009)	Index score					12.320**	5.183	5.161	6.174
Services quality (2009)	Index score					5.109	4.112	1.902	4.981

Asset index (2009)	Index score			13.976**	5.540	12.683*	6.908
Mother's education (2006)	Highest qualification			.444**	.207	.517**	.239
Older siblings (2006)	Number of older siblings			-.205	.321	.238	.380
Younger siblings (2006)	Number of younger siblings			-1.500**	.599	-.759	.687
School and community							
Geographic stratum (2002)	Dummy: 1 if rural					-10.055***	2.992
Principal qualification (2009)	Dummy: 1 if high level					7.266***	2.488
Type of school (2009)	Dummy: 1 if shift school					.306	.228
School is secure (2009)	Dummy: 1 if yes					-1.887	2.620
School is well-maintained (2009)	Dummy: 1 if yes					2.347	3.524
School has a library (2009)	Dummy: 1 if yes					4.463	3.791
School has toilets (2009)	Dummy: 1 if yes					-8.918*	5.247
School has a play area (2009)	Dummy: 1 if yes					-11.302***	3.257
Percentage of the variance in ppvt test scores explained (R^2)		*.164*		*.381*		*.396*	

*p < .10; **p < .05, ***p < .01

Table 5.2 Contextualised models for math test scores at age 15 (2009)

		Model 1 (N=965)		Model 2 (N=954)		Model 3 (N=938)		Model 4 (N=722)	
Variable	Value	β	S.E.	β	S.E.	β	S.E.	β	S.E.
Base model									
Intercept	Intercept	6.827***	.465	11.377***	1.235	9.092***	1.383	10.527***	1.980
Region	Dummy: 1 if Amhara	-3.287***	.546	-.965*	.556	-.235	.600	.377	.937
(Base group: Addis Ababa)	Dummy: 1 Oromia	-2.524***	.520	.356	.532	.559	.551	1.298*	.759
	Dummy: 1 SNNP	-.285	.575	1.547***	.535	1.894***	.548	2.049**	.835
	Dummy: 1 Tigray	-1.422	1.023	.976	.931	1.662	.983	1.649	1.221
Ethnic group	Dummy: 1 if Gurage	-1.239*	.675	-1.101*	.590	-.702	.614	.180	.878
(Base group: Amhara)	Dummy: 1 if Hadia	-5.034***	.806	-2.282***	.720	-1.695**	.736	-.958	.993
	Dummy: 1 if Oromo	-1.070**	.553	-1.030**	.480	-.836*	.498	-.596	.619
	Dummy: 1 if Sidama	-4.582***	.843	-2.769***	.743	-2.537***	.750	-1.599	1.066
	Dummy: 1 if Tigrian	-.492	1.037	-.534	.898	-.582	.922	.204	1.075
	Dummy: 1 if Wolayta	-1.936**	.784	-1.211*	.687	-1.234*	.698	-1.679*	.988
	Dummy: 1 if Other	-1.676*	1.014	-1.221	.896	-1.004	.918	.158	1.274
Child characteristics									
Gender	Dummy: 1 if female			-1.445***	.250	-1.309***	.252	-1.204***	.298
Health status (2009)	Index score (0-4)			-.158	.231	-.136	.232	-.227	.252
Early childhood care (2009)	Number of years of care			.591***	.142	.547***	.145	.446**	.181
School starting age	Age in years			-.690***	.097	-.622***	.099	-.696***	.122
Grades repeated (2009)	Number of grades			-.975***	.107	-.949***	.107	-1.094***	.141
Household work (2009)	Number of hours per day			.009	.027	.018	.027	.024	.037
Productive work (2009)	Number of hours per day			-.071	.054	-.025	.055	.055	.087
Family background									
Housing quality (2009)	Index score					1.703**	.814	1.468	.948

Services quality (2009)	Index score			-1.545**	.644	-1.629**	.763
Asset index (2009)	Index score			2.370***	.868	1.853*	1.057
Mother's education (2006)	Highest qualification			.038	.032	.032	.036
Older siblings (2006)	Number of older siblings			-.006	.050	-.004	.058
Younger siblings (2006)	Number of younger siblings			-.030	.094	-.012	.106
School and community							
Geographic stratum (2002)	Dummy: 1 if rural					-.943**	.460
Principal qualification (2009)	Dummy: 1 if high level					.432	.382
Type of school (2009)	Dummy: 1 if shift school					-.034	.035
School is secure (2009)	Dummy: 1 if yes					-.379	.399
School is well-maintained (2009)	Dummy: 1 if yes					-.406	.537
School has a library (2009)	Dummy: 1 if yes					.116	.580
School has toilets (2009)	Dummy: 1 if yes					-.257	.796
School has a play area (2009)	Dummy: 1 if yes					.343	.498
Percentage of the variance in math test scores explained (R^2)		*.101*	*.335*	*.355*		*.363*	

*p < .10; **p < .05, ***p < .01

Table 5.3 Contextualised models for verbal test scores at age 15 (2009)

		Model 1 (N=826)		Model 2 (N=819)		Model 3 (N=805)		Model 4 (N=617)	
Variable	Value	β	S.E.	β	S.E.	β	S.E.	β	S.E.
Base model									
Intercept	Intercept	8.626***	0.445	14.644***	1.235	12.663***	1.355	16.932***	1.997
Region	Dummy: 1 if Amhara	-4.912***	0.506	-3.383***	0.516	-2.776***	0.555	-3.498***	.883
(Base group: Addis Ababa)	Dummy: 1 Oromia	-4.736***	0.482	-2.645***	0.494	-2.499***	0.508	-2.353***	.712
	Dummy: 1 SNNP	-2.560***	0.560	-1.291***	0.510	-1.122***	0.516	-1.917**	.804
	Dummy: 1 Tigray	-5.127***	0.938	-3.626***	0.850	-3.508***	0.893	-4.225***	1.131
Ethnic group	Dummy: 1 if Gurage	-1.572**	0.636	-1.646***	0.552	-1.026*	0.572	-1.166	.860
(Base group: Amhara)	Dummy: 1 if Hadia	-4.877***	1.548	-2.552**	1.340	-2.440*	1.320	-2.666	1.885
	Dummy: 1 if Oromo	-1.382***	0.512	-1.439***	0.441	-1.326***	0.453	-2.074***	.574
	Dummy: 1 if Sidama	-2.887***	1.610	-1.532*	1.395	-1.835*	1.377	-2.168	2.382
	Dummy: 1 if Tigrian	1.716*	0.951	1.600**	0.818	2.224**	0.830	2.294**	.982
	Dummy: 1 if Wolayta	0.030	0.933	0.109	0.807	-0.007	0.799	-.512	1.063
	Dummy: 1 if Other	-0.989	0.950	-0.782	0.837	-0.716	0.850	-2.985**	1.216
Child characteristics									
Gender	Dummy: 1 if female			-0.360	0.249	-0.212	0.248	-.533*	.297
Health status (2009)	Index score (0-4)			-0.235	0.231	-0.200	0.230	-.189	.249
Early childhood care (2009)	Number of years of care			0.304**	0.137	0.258	0.139	-.013	.179
School starting age	Age in years			-0.831***	0.098	-0.764***	0.098	-.821***	.128
Grades repeated (2009)	Number of grades			-1.038***	0.111	-1.004***	0.110	-1.101***	.152
Household work (2009)	Number of hours per day			-0.004	0.024	0.005	0.024	.030	.034
Productive work (2009)	Number of hours per day			0.108**	0.055	0.164**	0.055	.178**	.085
Family background									
Housing quality (2009)	Index score					0.539	0.796	-.350	.957

Services quality (2009)	Index score			1.241**	0.624	.126	.749
Asset index (2009)	Index score			1.336*	0.847	1.568	1.036
Mother's education (2006)	Highest qualification			0.008	0.031	.029	.035
Older siblings (2006)	Number of older siblings			-0.085**	0.048	-.101*	.056
Younger siblings (2006)	Number of younger siblings			-0.088	0.090	-.109	.103
School and community							
Geographic stratum (2002)	Dummy: 1 if rural					-1.072**	.436
Principal qualification (2009)	Dummy: 1 if high level					1.113***	.363
Type of school (2009)	Dummy: 1 if shift school					-.021	.032
School is secure (2009)	Dummy: 1 if yes					-.161	.377
School is well-maintained (2009)	Dummy: 1 if yes					-.717	.598
School has a library (2009)	Dummy: 1 if yes					-.885	.575
School has toilets (2009)	Dummy: 1 if yes					-.656	.867
School has a play area (2009)	Dummy: 1 if yes					.354	.456
Percentage of the variance in verbal test scores explained (R^2)		*.208*	*.414*	*.430*		*.469*	

*p < .10; **p < .05, ***p < .01

5.2. Modelling the effect of ethnicity on schooling progression

In the first part of this analysis we investigated the relationship between ethnicity and achievement. The next step is to examine differences in schooling participation between ethnic groups. Separate regression models are developed with a view to examining differential patterns in: (i) schooling access – measured by the age at which enter formal school; (ii) schooling progression – measured by the number of grades repeated by age 15; and (iii) schooling completion – measured as the highest grade completed by the child by age 15. The models further aim to examine the extent to which contextual factors account for inter-group differences. Groups of explanatory variables are therefore entered in identical logical blocks as used for the previous analysis in section 5.1.

5.2.1. Contextual models of grade completion by age 15

Table 5.4 presents the results for the highest grade completed by age 15.

Model 1: Base model

The results confirm a strong positive association between grade completion at age 15 and ethnicity, and suggest that children of Hadia and Sidama ethnic origin are significantly behind the official grade for their age. Compared with Amhara children, Hadia and Sidama children appear to be 1.3 and 2.3 grades behind.

Model 2: Child characteristics

Ethnicity is no longer significant once we take into account child factors. Schooling history and time spent engaged in productive work are the most important factors, and, overall, explained 73 per cent of the variance. Gender was not a strong predictor.

Model 3: Family background

Once we controlled for child factors, the family background variables were not significant, suggesting that schooling history and individual level deprivations are more important for schooling completion than family disadvantage.

Model 4: School and community

Contrastingly, school and community variables were very significant suggesting that community and school factors may be more important predictors of the grade attainment/progression than the home environment. Strong positive associations were found between grade completion and principal qualification, whether the school attended was a shift/flexible school and the availability of a school library. Ethnicity does remain significant for children of Sidama origin, suggesting that we have not identified all of the determinants of the difference in grade levels.

5.2.2. Contextual models of grade repetition by age 15

Table 5.5 presents the results for the number of grades repeated by age 15.

Model 1: Base model

Ethnicity is significant only for the Sidama ethnic group; with children of Sidama origin more likely to repeat than Amhara children. It is important to note that region is not a significant predictor of grade repetition; with the exception of Oromia where grade repetition is more likely.

Model 2: Child characteristics

This model explains 19.6 per cent of the variance in repetition; with years of early childhood care, age of entry into school showing a strong negative correlation (e.g. earlier entry into school and more years of early care results in fewer grades repeated). Moreover, the more hours worked the more grades repeated.

Model 3: Family background

The family background variables explain seven per cent of the variation. However, child characteristics remain the most important factors.

Model 4: School and community

School and community factors explain 3 per cent of the variation in repetition over and above the family background model. Taken together, these variables are more significant for grade repetition than family background, which suggests that school factors may be at least as important for whether a child repeats a grade as individual and family level deprivation.

5.2.3. Contextual models of school entry age

Table 5.6 presents the results for the age at which the child starts schooling.

Model 1: Base model

The results confirm that children of Hadia and Sidama ethnic origin are significantly less like to start school in line with the official age of entry. Specifically, the Sidama children start school 0.87 years later than Amhara children, whereas Hadia children start school 2.3 years later.

Model 2: Child characteristics

Ethnicity remains significant after controlling for child factors. However, we observe that Sidama children start school earlier than would be expected given their individual level of disadvantage; 0.46 years before Amhara children. Tigrian and Wolayta also appear to start earlier than Amhara children – by 0.73 and 0.75 years. The variables in this model explain 64 per cent of the variation in school entry age. Significant child level predictors include gender, general health and nutrition and the number of years of early childhood care – which are all strongly and negatively associated with school entry age.

Model 3: Family background

The family background variables explained an additional 1.4 per cent of the variation. Only service quality shows a strong negative association with start age, suggesting that children with better access to services (including schools) start school earlier. However, we observe that ethnicity remains significant for suggesting that the variables in this model have not captured the main predictors of school start age for these groups.

Model 4: School and community

This block does not include school quality factors, as these data were collected in 2009, after the children started school. Thus the only variable added in this model is rural residence, which appears to be uncorrelated with school entry age.

Table 5.4 Contextualised models for grade completion at age 15 (2009)

		Model 1 (N=958)		Model 2 (N=905)		Model 3 (N=888)		Model 4 (N=689)	
Variable	Value	β	S.E.	β	S.E.	β	S.E.	β	S.E.
Base model									
Intercept	Intercept	6.896***	.219	12.065***	.382	11.764***	.425	10.444***	.511
Region	Dummy: 1 if Amhara	-1.202***	.257	.070	.168	.076	.187	-.301	.247
(Base group: Addis Ababa)	Dummy: 1 Oromia	-1.792***	.245	-.255*	.162	-.240	.176	-.230	.201
	Dummy: 1 SNNP	-1.363***	.271	.026	.159	.015	.165	-.187	.217
	Dummy: 1 Tigray	-1.352***	.482	.440*	.281	.381	.306	-.312	.324
Ethnic group	Dummy: 1 if Gurage	.084	.318	.088	.175	.070	.186	.261	.229
(Base group: Amhara)	Dummy: 1 if Hadia	-2.280***	.381	-.367*	.218	-.362*	.230	-.107	.268
	Dummy: 1 if Oromo	.028	.260	.102	.145	.051	.153	-.197	.166
	Dummy: 1 if Sidama	-1.345***	.398	-.551*	.225	-.506*	.235	-.639**	.285
	Dummy: 1 if Tigrian	.281	.488	-.053	.272	.009	.282	.025	.289
	Dummy: 1 if Wolayta	-.121	.369	-.005	.205	-.119	.214	-.403	.260
	Dummy: 1 if Other	-.425	.477	-.522*	.278	-.583*	.289	-.052	.328
Child characteristics									
Gender of child (2002)	Dummy: 1 if female			-.079	.075	-.081	.077	-.044	.078
Health status (2009)	Index score (0-4)			-.072	.069	-.055	.070	-.021	.066
Early childhood care (2009)	Number of years of care			.111***	.044	.116***	.046	.063	.049
School starting age (2006)	Age in years			-.715***	.031	-.703***	.032	-.642***	.034
Grades repeated	Number of grades			-.284***	.038	-.284***	.038	-.288***	.040
Household work (2009)	Number of hours per day			-.006	.010	-.005	.011	.010	.010
Productive work (2009)	Number of hours per day			-.130***	.016	-.129***	.017	-.053**	.023
Family background									
Housing quality (2002)	Index score					-.018	.270	.225	.270
Services quality (2002)	Index score					.306*	.186	.723***	.203

Asset index (2002)	Index score			-.270	.365	-.214	.384
Mother's education (2006)	Highest qualification			.010	.010	.002	.009
Older siblings (2006)	Number of older siblings			.025*	.015	.015	.015
Younger siblings (2006)	Number of younger siblings			.027	.029	.007	.028
School and community							
Geographic stratum (2002)	Dummy: 1 if rural					.250*	.134
Principal qualification (2009)	Dummy: 1 if high level					.490***	.098
Type of school (2009)	Dummy: 1 if shift school					.031***	.009
School is secure (2009)	Dummy: 1 if yes					.124	.105
School is well-maintained (2009)	Dummy: 1 if yes					.171	.139
School has a library (2009)	Dummy: 1 if yes					.464***	.158
School has toilets (2009)	Dummy: 1 if yes					.101	.215
School has a play area (2009)	Dummy: 1 if yes					-.093	.133
Percentage of the variance in grade completion explained (R^2)		*.167*	*.727*	*.731*		*.793*	

*p < .10; **p < .05, ***p < .01

Table 5.5 Contextualised models for grade repetition at age 15 (2009)

		Model 1 (N=958)		Model 2 (N=955)		Model 3 (N=938)		Model 4 (N=721)	
Variable	Value	β	S.E.	β	S.E.	β	S.E.	β	S.E.
Base model									
Intercept	Intercept	.678***	.132	2.812***	.366	3.030***	.409	4.053***	.506
Region	Dummy: 1 if Amhara	.087	.155	-.241	.169	-.220	.189	.074	.246
(Base group: Addis Ababa)	Dummy: 1 Oromia	.509***	.148	.192	.162	.206	.177	.128	.203
	Dummy: 1 SNNP	-.093	.163	-.093	.163	-.072	.170	.130	.224
	Dummy: 1 Tigray	-.284	.291	-.557**	.283	-.499*	.308	.232	.326
Ethnic group	Dummy: 1 if Gurage	-.027	.192	-.168	.180	-.125	.191	-.388*	.233
(Base group: Amhara)	Dummy: 1 if Hadia	.042	.230	.176	.220	.198	.232	-.120	.270
	Dummy: 1 if Oromo	-.181	.157	-.210	.146	-.159	.155	.080	.167
	Dummy: 1 if Sidama	.477**	.241	.419*	.226	.423*	.237	.488*	.283
	Dummy: 1 if Tigrian	.071	.295	-.037	.274	-.065	.284	-.169	.287
	Dummy: 1 if Wolayta	.156	.223	-.027	.210	.074	.220	.383	.267
	Dummy: 1 if Other	1.191***	.288	.700***	.272	.776***	.283	-.069	.342
Child characteristics									
Gender of child (2002)	Dummy: 1 if female			.085	.076	.094	.077	.083	.080
Health status (2009)	Index score (0-4)			.067	.071	.053	.072	.042	.068
Early childhood care (2009)	Number of years of care			-.129***	.043	-.133***	.045	-.071	.049
School starting age (2006)	Age in years			-.242***	.029	-.248***	.029	-.288***	.031
Grades repeated	Number of grades			-	-			-	-
Household work (2009)	Number of hours per day			.019**	.008	.019**	.008	.000	.010
Productive work (2009)	Number of hours per day			.148***	.016	.150***	.016	.061***	.023
Family background									
Housing quality (2002)	Index score					-.004	.278	-.234	.279

Services quality (2002)	Index score			-.195	.191	-.695***	.210
Asset index (2002)	Index score			.200	.375	.260	.393
Mother's education (2006)	Highest qualification			-.009	.010	-.005	.010
Older siblings (2006)	Number of older siblings			-.028*	.016	-.010	.016
Younger siblings (2006)	Number of younger siblings			-.031	.029	-.017	.028
School and community							
Geographic stratum (2002)	Dummy: 1 if rural					-.306**	.137
Principal qualification (2009)	Dummy: 1 if high level					-.543***	.100
Type of school (2009)	Dummy: 1 if shift school					-.032***	.009
School is secure (2009)	Dummy: 1 if yes					-.138	.106
School is well-maintained (2009)	Dummy: 1 if yes					-.134	.143
School has a library (2009)	Dummy: 1 if yes					-.411***	.154
School has toilets (2009)	Dummy: 1 if yes					.034	.213
School has a play area (2009)	Dummy: 1 if yes					.147	.133
Percentage of the variance in grade repetition explained (R^2)		*.058*	*.196*	*.203*		*.236*	

*p < .10; **p < .05, ***p < .01

Table 5.6 Contextualised models for age of entry into formal school

		Model 1 (N=958)		Model 2 (N=923)		Model 3 (N=906)		Model 4 (N=906)	
Variable	Value	β	S.E.	β	S.E.	β	S.E.	β	S.E.
Base model									
Intercept	Intercept	7.426***	.207***	10.885***	.196	11.454***	.243	11.339***	.263
Region	Dummy: 1 if Amhara	1.115***	.244***	-1.639***	.182	-1.933***	.200	-1.981***	.204
(Base group: Addis Ababa)	Dummy: 1 Oromia	1.283***	.232***	-1.322***	.174	-1.502***	.188	-1.553***	.193
	Dummy: 1 SNNP	1.456***	.256***	-.550***	.183	-.675***	.188	-.699***	.189
	Dummy: 1 Tigray	1.637***	.456***	-.976***	.314	-1.137***	.338	-1.213***	.344
Ethnic group	Dummy: 1 if Gurage	-.057	.301	.021	.203	-.175	.214	-.190	.214
(Base group: Amhara)	Dummy: 1 if Hadia	2.238***	.360	.541**	.246	.239	.256	.229	.257
	Dummy: 1 if Oromo	.153	.246	-.044	.164	-.110	.172	-.106	.172
	Dummy: 1 if Sidama	.867**	.377	-.446**	.254	-.694***	.262	-.720***	.263
	Dummy: 1 if Tigrian	-.353	.462	-.728**	.306	-.809***	.313	-.793**	.313
	Dummy: 1 if Wolayta	-.036	.349	-.750***	.234	-.623***	.243	-.584**	.245
	Dummy: 1 if Other	-.765*	.452	-.567*	.306	-.741*	.315	-.743**	.315
Child characteristics									
Gender of child (2002)	Dummy: 1 if female			-.182**	.078	-.193**	.078	-.193**	.078
Health status (2002)	Height-for age index score			-.260***	.033	-.257***	.033	-.257***	.033
Early childhood care (2009)	Number of years of care			-1.033***	.034	-1.010***	.036	-1.006***	.037
School starting age (2006)	Age in years			-	-	-	-	-	-
Grades repeated	Number of grades			-	-	-	-	-	-
Household work (2009)	Number of hours per day			-	-	-	-	-	-
Productive work (2009)	Number of hours per day			-	-	-	-	-	-
Family background									
Housing quality (2002)	Index score					-.304	.315	-.202	.327

Services quality (2002)	Index score			-.579***	.212	-.457**	.237
Asset index (2002)	Index score			-.130	.417	-.128	.417
Mother's education (2006)	Highest qualification			-.012	.011	-.012	.011
Older siblings (2006)	Number of older siblings			-.019	.018	-.020	.018
Younger siblings (2006)	Number of younger siblings			-.004	.032	-.008	.033
School and community							
Geographic stratum (2002)	Dummy: 1 if rural					.162	.140
Principal qualification (2009)	Dummy: 1 if high level					-	-
Type of school (2009)	Dummy: 1 if shift school					-	-
School is secure (2009)	Dummy: 1 if yes					-	-
School is well-maintained (2009)	Dummy: 1 if yes					-	-
School has a library (2009)	Dummy: 1 if yes					-	-
School has toilets (2009)	Dummy: 1 if yes					-	-
School has a play area (2009)	Dummy: 1 if yes					-	-
Percentage of the variance in start age explained (R^2)		*.166*	*.636*	*.650*		*.651*	

*p < .10; **p < .05, ***p < .01

5.3. Modelling the effect of ethnicity on educational progress

The third outcome modelled in this analysis is educational progress over the period 2002-2009. Educational progress is a dynamic concept that encompasses both children's grade progression and their cognitive development over time. The conceptual framework for this model is discussed in Chapter 2, and the analytical strategy in Chapter 4.

As an intermediary step, we returned to four single-equation models of achievement estimated in section 5.1. To model the relationship between ethnicity and cognitive progression over time, rather than achievement at a point in time, we include a prior achievement model. Prior achievement is entered before the remaining contextual variables with a view to examining whether key contextual variables – namely school start age and grade repetition – remain significant once prior achievement is taken into account. The results, which are included in Appendix III, show that these variables are particularly important for cognitive development across all cognitive domains – general scholastic aptitude, mathematics progression and literacy development. The rest of this section discusses the main findings of the analysis.

5.3.1. Model of cognitive development and grade progression

Table 5.7 presents the results for general cognitive development and grade progression. We find that both grade level at age 12 and how well children progress through the educational system are strong determinants of their general intellectual development between ages 12 and 15. Section 5.1 showed that slow grade progression is largely predicted by schooling history, namely, enrolment age and the incidence of grade repetition. Thus, the current model illustrates the longer term effects of these phenomena on cognitive development over time.

We further find that some ethnic groups are more likely than others to be affected by late enrolment, and therefore more likely to progress slowly and achieve lower results. Table 5.7 shows that children of Hadia and Sidama origin are significantly more likely to be in the lower primary school grades in comparison to other children, with the exception of Amhara children. They start off with lower

results at age 12; and by age 15, they continue to perform less well (as shown in section 5.1).

Table 5.7 Model of cognitive development and grade progression, age 12-15

	PPVT (age 15)		Grade (age 15)		PPVT (age 12)		Grade (age 12)	
Predictor	β	S.E.	β	S.E.	β	S.E.	β	S.E.
Grade (15 yrs)	2.416***	.916	-	-	-	-	-	-
PPVT (12 yrs)	.309***	.033	.003**	.001	-	-	-	-
Grade (12 yrs)	3.538***	1.061	1.020***	.019	6.577***	.467	-	-
Base (Amhara)								
Gurage	-	-	-	-	7.320***	2.744	.146	.192
Hadia	-	-	-	-	-5.764*	3.472	-2.422***	.229
Oromo	-	-	-	-	7.524***	1.812	-.369***	.126
Sidama	-	-	-	-	-2.593	3.470	-1.418***	.238
Tigrian	-	-	-	-	4.337**	1.797	-.275**	.125
Wolayta	-	-	-	-	11.639***	3.012	-.444**	.210
Other	-	-	-	-	-2.419	5.384	.161	.376
N = 951								
X^2 (df=35)=518.71, p=000								
CFI = .811								
RMSEA = .118								

5.3.2. Model of mathematics development and grade progression

Table 5.8 presents the parameter estimates for maths progression between the ages of 12 and 15. Contrary to the previous model, grade progression does not appear to be as important for mathematics development as prior attainment. Moreover, ethnicity is not an important predictor of maths achievement at age 12 for the lowest performing groups – the Hadia and the Sidama. This may reflect problems of sample size or the fact that all of the children performed quite poorly on the maths tests. However, similar to the previous model, ethnicity is significant for grade level at age 12.

Table 5.8 Model of maths development and grade progression, age 12-15

	Math (age 15)		Grade (age 15)		Math (age 12)		Grade (age 12)	
Predictor	β	S.E.	β	S.E.	β	S.E.	β	S.E.
Grade (15 yrs)	.498***	.137	-	-	-	-	-	-
Math (12 yrs)	.691***	.056	.120***	.014	-	-	-	-
Grade (12 yrs)	.036	.152	.944***	.020	.797***	.041	-	-
Base (Amhara)	-	-	-	-	-	-	-	-
Gurage	-	-	-	-	.583**	.240	-2.424***	.229
Hadia	-	-	-	-	-.147	.303	-.300**	.125
Oromo	-	-	-	-	-.059	.158	.137	.192
Sidama	-	-	-	-	-.413	.303	-.378***	.126
Tigrian	-	-	-	-	.955***	.157	-1.445***	.238
Wolayta	-	-	-	-	.627**	.263	-.483**	.210
Other	-	-	-	-	.298	.471	.327	.376

N = 947
X^2 (df=35)=468.51, p=000
CFI = .836
RMSEA = .111

5.3.3. Model of literacy development and grade progression

Table 5.9 presents the estimation results for literacy development between the ages of 12 and 15. Grade level at age 12 and grade progression emerge as strong predictors of literacy development. Similar to the previous models, we observe a high level of variation in grade level at age 12 by ethnic group. Prior attainment appears to be uncorrelated with literacy skills at age 15. However, this may be due to the difference in the instruments used at age 12 and 15.

Table 5.9 Model of literacy development and grade progression, age 12-15

	Y1 = Literacy (age 15)		Y2 = Grade (age 15)		Y3 = Literacy (age 12)		Y4 = Grade (age 12)	
Predictor	β	S.E.	β	S.E.	β	S.E.	β	S.E.
Grade (15 yrs)	.827***	.145	-	-	-	-	-	-
Literacy (12 yrs)	-.020	.022	.004	.006	-	-	-	-
Grade (12 yrs)	.426**	.167	1.038***	.018	.083	.113	-	-
Base (Amhara)								
Gurage	-	-	-	-	-.811	.660	.132	.191
Hadia	-	-	-	-	.133	.835	-2.398***	.229
Oromo	-	-	-	-	-.307	.436	-.379***	.126
Sidama	-	-	-	-	.091	.835	-1.428***	.238
Tigrian	-	-	-	-	.247	.432	-.281**	.125
Wolayta	-	-	-	-	.284	.725	-.456**	.210
Other	-	-	-	-	.243	1.295	.177	.375

N = 830
X^2 (df=35)=508.65, p=000
CFI = .775
RMSEA = .080

6. DISCUSSION

Ethiopia presents a particularly interesting context for examining the different ways in which horizontal inequalities manifest through educational outcomes. On the one hand, vertical inequality is considered moderate in relation to other African countries, with a Gini coefficient of 0.30 (WIDER, 2007). However, at least 80 per cent of the population lives in poverty (UNDP, 2011); a reduction of only 10 per cent in the last 30 years (UNDP, 2011), suggesting that the poverty situation in Ethiopia is persistent and complex. Moreover, huge disparities persist in education, particularly with regard to achievement, retention and grade progression. These issues are critical challenges for Ethiopia's educational system as they affect not only children's schooling outcomes but also their life chances beyond schooling. The significance of the current research therefore relates to improving our understanding of how horizontal educational inequalities contribute to the persistence of poverty.

6.1. Summary of main findings and policy implications

The objective of the current research was to examine differential patterns in children's educational achievement (at age 15) and schooling participation by ethnic group, and to identify the factors that might help to explain inter-group differences. The research further explored the effect of slow grade progression on cognitive development between age 12 and 15.

6.1.1 Contextualising ethnic gaps in attainment at age 15

Our findings suggest that, even after controlling for geographic region, there is substantial variation in achievement between different ethnic groups in the sample. Achievement levels are generally low in each of the three cognitive domains – general intelligence, mathematics ability and verbal ability – although consistently higher in Addis Ababa. This is consistent with national patterns of achievement which show a large gap in children's performance between the centre, Addis Ababa, and the rest of the country (National Organisation for

Examinations, 2004, 2007; USAID, 2010; Demographic and Health Survey, 2011).

However, children from certain ethnic groups appear to perform significantly less well than others. We find that within the most ethnically and linguistically diverse region in the sample, Southern Nations, Nationalities and People's (SNNP) Region, intra-regional disparities persist between indigenous groups. Children of Hadia and Sidama ethnic origin score, on average, 23 points below Wolayta children on the PPVT, three points below on the math test and four points below on the verbal test. This finding is consistent with the notion of horizontal inequality, discussed in Chapter 1; however, there are no known studies that attempt to compare achievement outcomes between ethnic groups in Ethiopia.

Table 6.1 shows the ethnic composition of the older cohort children by each of the study sites in SNNP, the region in which there is the highest variation in performance. The purpose of the table is to illustrate the extent of ethnic homogeneity within each of the study sites. Most of the children from the Sidama ethnic group are located in Site 4, and the majority of Hadia children reside in Site 5. Moreover, a large proportion of children from the Gurage and Wolayta groups live in the urban centre, Site 3. The table therefore demonstrates the importance of geographic location and urban vs. rural residence for learning outcomes. In other words, ethnic gaps in attainment further reflect differences between sites.

Table 6.1 Ethnic composition of the Young Lives sample of older cohort children in SNNP, by site

	SNNP study sites				
	Site 1	Site 2	Site 3	Site 4	Site 5
Number of children	*46*	*49*	*47*	*49*	*49*
Amhara ethnic group (%)	0.0	4.0	26.0	4.0	0.0
Gurage ethnic group (%)	98.0	0.0	12.0	0.0	0.0
Hadia ethnic group (%)	0.0	2.0	0.0	0.0	98.0
Oromo ethnic group (%)	0.0	4.0	15.0	0.0	0.0
Sidama ethnic group (%)	0.0	0.0	1.0	96.0	0.0
Tigrian ethnic group (%)	0.0	0.0	6.0	0.0	0.0
Wolayta ethnic group (%)	0.0	88.0	38.0	0.0	0.0
Other ethnic group (%)	2.0	2.0	2.0	0.0	2.0
Total (%)	*100.0*	*100.0*	*100.0*	*100.0*	*100.0*

With regard to the factors that predict achievement at age 15, for each of the cognitive outcomes, the models were able to account for the poor attainment of the Hadia and Sidama children in the sample. Children's schooling participation – measured by three variables selected to capture the extent of late enrolment and slow grade progression – is found to be particularly significant for cognitive ability among Hadia and Sidama children. After controlling for schooling history, home environmental factors – such as household wealth, mother's educational level and the number of siblings – explained only 1-2 per cent of the variance in test scores over and above child factors. This suggests that differences in income and poverty status cannot fully explain the variation in achievement outcomes between the ethnic groups. Rather, when a child enters formal school and how well they progress through the school system may be more important for cognitive outcomes than the deprivation they face at the family level. While we recognise that family background would indeed influence children's schooling experiences, we argue that school level policies that seek to address late enrolment and slow grade progression may compensate for such disadvantages and improve children's cognitive achievement and development over time. This is supported by additional evidence that children of Hadia and Sidama origin not only start school later and progress through the grades more slowly than other children in the sample; they may also be enrolled in poorer quality schools.

6.1.2 Contextualising ethnic gaps in schooling participation and progression

The findings show that Hadia and Sidama children are significantly behind the official grade for their age in comparison to their peers. The mean grade level for the older cohort children at age 15 is 5.54 years. This figure is low, given that at age 15 the children would be expected to be studying in Grades 8 or 9. Hadia and Sidama children are most likely to be in the early primary school grades (Grades 1-4). The ethnic gap in grade completion is largely attributable to late enrolment, which not only affects the rate of grade progression – as children who start school late are less likely to progress through the grades at the expected rate than children who start school on time – but also impacts upon children's cognitive development over time (as shown in 6.1.1).

Children may enrol late and repeat grades for a wide range of reasons. Data from the Young Lives survey suggests that the cost of schooling is an important determinant for many children (Young Lives data, 2009). Families that are not able to pay for school and essential school supplies cannot send their children to school on time or ensure their participation throughout the schooling cycle. The children also cite illness and the need to work as factors (Young Lives data, 2009). Parents report disability and illness of a family member as key determinants of whether children attend school (Young Lives data, 2009). The literature also suggests that poor health and chronic undernourishment may contribute to slow progression and underachievement (Dercon and Krishnan, 2009). Other factors cited by Young Lives children and their parents include transportation problems, low motivation on the part of children and poor performance in school as factors affecting schooling participation and progression (Young Lives data, 2009).

The findings show that once we control for the children's schooling history, family background does not account for much of the variation in the rate of grade progression of Hadia and Sidama children in the sample. Moreover, school and community factors were found to be more significant for grade attainment/progression than the children's home environment. Thus, similar to the findings for achievement, while we recognise that the child's home environment would influence his/her access to and progression through school, we argue that school quality factors may be as important for their ability to progress through the grades.

Despite the significance of the mean-level difference in a range of individual and family background variables, we were not able identify the most important predictors of school entry age for Hadia and Sidama children. None of the schooling participation models were able to account for the late enrolment among these groups. This may be attributable to the limited set of variables used in this component of the research. A more extensive list of variables would be needed to explain the late enrolment of the Hadia and Sidama group.

6.2. Limitations and areas for future research

The research questions of the current study focused on demonstrating the significance of horizontal educational inequalities in Ethiopia. Drawing on the data collected for the Young Lives project, the results showed the low achievement and poor grade progression of Hadia and Sidama children in the sample. The study further showed the importance of late enrolment for their performance and participation in school.

The contextual models have not accounted for the difference in enrolment age. Despite the use of a wide range variables identified in the literature as significant for children's schooling in Ethiopia, the reasons for which Hadia and Sidama children start late and progress through the grades more slowly remain unclear. This result helps to illustrate that explanations based on standard indicators of household poverty and socio-economic status (e.g. family wealth and parental educational level) are insufficient for explaining differences in grade progression, completion and achievement.

While the study does indeed demonstrate the importance of individual and community/school factors for children's participation and performance, only a limited number of individual and school effects have been examined here. Future research should explore complementary data collected by the Young Lives project – on student and parent attitudes towards school, the role of non-cognitive skills, teacher and learning practices , etc. – in shaping child/household decisions regarding when children start school and how well they progress through the school system.

6.3. Conclusion

The theory of horizontal inequality provides a unique lens for examining differences in educational outcomes and the role of such differences in perpetuating the cycle of poverty. Research and policy that emphasizes educational access – as a means of growth promotion, socio-economic mobility and the empowerment of individuals – may downplay the role of education in

compensating for or reproducing societal inequalities. Adopting a group-based approach to inequality analysis, as opposed to an individual measure, can help to illuminate on the actual mechanisms through which inequalities are formed – socio-economic, political or cultural. For example, the current study showed that a regional analysis of educational deprivation would have failed to account for (intra-regional) variation between indigenous ethnic groups, and might have missed the opportunity to account for more persistent inequalities that trap individuals, generation after generation.

BIBLIOGRAPHY

Admassie, A. and Bedi, A.S. (2003) Attending school: two'Rs' and child work in rural Ethiopia. *ISS Working Paper Series/General Series,* 387, 1-44.

Admassu, K. (2008), Primary School Enrolment and Progression in Ethiopia: Family and School Factors, *Paper presented at the annual meeting of the American Sociological Association Annual Meeting, Boston, MA.* http://www.allacademic.com/meta/p241783_index.html [accessed 12 February 2012].

Ainsworth, A. and Filmer, D. "Inequalities in Children's Schooling: AIDS, Orphanhood, Poverty and Gender." World Development, 34 (2006), 1099-1128.

Althusser. Louis (1971) *Lenin and Philosophy and Other Essays*. New York: New Left Books.

Alvi, E. and Dendir, S. (2011) Sibling Differences in School Attendance and Child Labour in Ethiopia. *Oxford Development Studies,* 39 (3), 285-313.

Andvig, J.C., Canagarajah, S. and Kielland, A. (n.d.) *Child Labour in Africa: The Issues*. Washington, D.C: World Bank. http://info.worldbank.org/etools/docs/library/74184/winter2002/proceedings/pdfpapers/mod10ja.pdf [Accessed 12 August 2011].

Becker, G.S. (1964) *Human capital: a theoretical and empirical analysis, with special reference to education.* New York: National Bureau of Economic Research.

Becker, G.S. (1995) *Human capital and poverty alleviation.* Washington, D.C.: World Bank.

Blatchford, P. (2003), A systematic observational study of teachers' and pupils' behaviour in large and small classes, *Learning and Instruction,* 13 (6), 569-95

Bourdieu, Pierre. 1974. The School as a Conservative Force: Scholastic and cultural inequalities. In *Contemporary Research in the Sociology of Education*. John Eggleston (ed.) London: Methuen.

Bourdieu, P. and Passeron, J.C. (1977) *Reproduction in education, society and culture.* London: Sage.

Boyden, J. (2009) Young Lives: an International Study of Childhood Poverty: Round 3. Colchester, Essex: UK Data Archive [distributor].

Bowles, S. and Gintis, H. (1976) *Schooling in capitalist America: educational reform and the contradictions of economic life.* London: Routledge & Kegan Paul.

Bradley, R. H., Corwyn, R. F., Burchinal, M., McAdoo, H. P., & Garcia-Coll, C. (2001). The Home Environments of Children in the United States Part I: Variations by Age, Ethnicity, and Poverty Status. *Child Development, 72*, 1844-1867.

Bronfenbrenner, U. (1977) Toward an experimental ecology of human development. *American psychologist,* 32 (7), 513.

Bronfenbrenner, U., and Ceci, S. J. (1994). Nature-Nurture Reconceptualized in Developmental Perspective: A Bioecological Model. *Psychological Review*, 101 (4), 568-586.

Bronfenbrenner, U., & Morris, P. A. (1998). The Ecology of Developmental Processes. In W. Damon (Ed.), *Handbook of Child Psychology: Theoretical Models of Human Development (5th ed.)*. New York: John Wiley & Sons.

Brooks-Gunn, J., Duncan, G. J., & Maritato, N. (1997). Poor Families, Poor Outcomes: The Well-being of Children and Youth. In G. J. Duncan & J. Brooks-Gunn (Eds.), *Consequences of Growing Up Poor.* New York: Russell Sage Foundation.

Brown, G.K. (2007) Making ethnic citizens: The politics and practice of education in Malaysia. International Journal of Educational Development, 27 (3), 318-330.

Bryman, A. (2008) *Social research methods.* Oxford: Oxford University Press.

Bush, K. and Saltarelli, D. (2000) *The Two Faces of Education in Ethnic Conflict: Towards a Peacebuilding Education for Children*. Florence: UNICEF Innocenti Research Centre.

Camfield, L. (2011) A girl never finishes her journey: mixing methods to understand female experiences of education in contemporary Ethiopia. *Research Papers in Education,* 26 (4), 393-412.

Card, D. (1999) Causal effect of education on earnings, in *Handbook of labor economics*, Orley Ashenfelter and David Card (Eds). Amsterdam: North-Holland, 1801–1863

Case A, Paxson C, and Abledinger J.(2004) Orphans in Africa: parental death, poverty and school enrolment. *Demography,* 41, 483-508.

Chaudhury, N., Christiaensen, L. and Asadullah, M.N. (2006) Schools, Household, Risk, and Gender: Determinants of Child Schooling in Ethiopia. *CSAE Working Paper Series*.

Colclough, C., Rose, P. and Tembon, M. (2000) Gender inequalities in primary schooling: the roles of poverty and adverse cultural practice1. *International Journal of Educational Development,* 20 (1), 5-27.

Creemers, B.P.M. (1994), The history, value and purpose of school effectiveness studies, in D. Reynolds (ed.), *Advances in Schools Effectiveness Research and Practice* Oxford, UK: Pergamon.

Cueto, S., Leon, J., Guerrero, G. and Munoz, I. (2009) Psychometric characteristics of cognitive development and achievement instruments in Round 2 of Young Lives. Oxford, UK: Young Lives.

Davies, L. *Education and Conflict: Complexity and Chaos*. London: Routledge Falmer, 2004.

Dercon, S., Krishnan, P., Dercon, S. and Krishnan, P. (2000) In Sickness and in Health: Risk Sharing within Households in Rural Ethiopia. *The Journal of Political Economy,* 108 (4), 688-727.

Dercon, S. (2011) Long-Term Implications of Under-Nutrition on Psychosocial Competencies: Evidence from Four Developing Countries. Oxford, UK: Young Lives.

DeStefano, J. and Elaheebocus, N. (2009) School Effectiveness in Woliso, Ethiopia: Measuring Opportunity to Learn and Early Grade Reading Fluency. Washington, DC: USAID.

Drèze, J. and Sen, A. (1995) India: economic development and social opportunity. Delhi: Oxford University Press.

Engle, P. and Black, M. (2008) The Effect of Poverty on Child Development and Educational Outcomes. *Annals of the New York Academy of Sciences,* 1136, 243-256.

Ethiopian Central Statistical Agency and ICF International (2000) *Ethiopia Demographic and Health Survey 2011*. Addis Ababa, Ethiopia and Calverton, Maryland, USA: Central Statistical Agency and ICF International.

Ethiopian Central Statistical Agency and ICF International (2005) *Ethiopia Demographic and Health Survey 2011*. Addis Ababa, Ethiopia: Central Statistical Agency.

Ethiopian Central Statistical Agency and ICF International (2011) *Ethiopia Demographic and Health Survey 2011*. Addis Ababa, Ethiopia: Central Statistical Agency.

Ethiopian Central Statistical Agency (2007) National Population Census Statistical Tables. Addis Ababa, Ethiopia: Population Census Commission.

Ethiopian Ministry of Education (2008) *Minimum Learning Competencies*. Addis Ababa: Ministry of Education.

Federal Democratic Republic of Ethiopia (1994) *Constitution of Ethiopia* http://www.apapeth.org/Documents/EthiopianLaws.html. Retrieved 3 December 2011. [Accessed on 03 August, 2012].

Giroux, H., Roger, S., Smith, P., and Ellsworth, E. (1989) *Schooling Popular Culture, and Everyday Life*. New York: Bergin and Garvey.

Glewwe, P. and Jacoby, H. (1994), Student achievement and schooling choice in low-income countries: Evidence from Ghana, *Journal of Human Resources,* 29 (3), 843-64.

Gujarati, D.N. (1995) *Basic Econometrics, 3rd*. New York: McGraw-Hills Inc.

Hanushek, E (2005). *Economic outcomes and school quality*. International Institute for Educational Planning. http://unesdoc.unesco.org/images/0014/001410/141027e.pdf [Accessed on 12 July, 2012].

Hanushek, E. and Woessmann, L. (2008). The role of cognitive skills in economic development, *Journal of Economic Literature* 46 (3): 607–608.

Heckman, J., Lochner, L., and Todd, P. (2006) Earnings functions, rates of return and treatment effects: The Mincer equation and beyond, in *Handbook of the Economics of Education*, Eric A. Hanushek and Finis Welch (Eds). Amsterdam: North Holland, 307–458.

Harber, C. (2004) *Schooling as Violence: How Schools Harm Pupils and Societies*. New York: Routledge.

Heugh, K., Benson, C., Bogale, B. and Yohannes, M.A.G. (2007) Final report: Study on medium of instruction in primary schools in Ethiopia. Addis Ababa, Ministry of Education.

Himaz, R. (2009) The impact of parental death on schooling and subjective well-being: evidence from Ethiopia using longitudinal data. Oxford, UK: Young Lives.

Jennings, M., Mekonnen, E. and Feye, B. (2010) Social Assessment for the Education Sector. Addis Ababa, Ethiopia: Ministry of Education.

Jeffrey, C.(2008) Degrees without freedom?: education, masculinities and unemployment in north India. Stanford, Calif.: Stanford University Press.

Johnson, D. Building citizenship in fragmented societies: the challenges of deracialising and integrating schools in post-Apartheid South Africa. *International Journal of Educational Development,* 27 (2007), 306-317.

Keller, Edmond J. and Lahra Smith (2005) Obstacles to implementing territorial decentralisation: the first decade of Ethiopian federalism. In Philip Roeder and Donald Rothchild (eds.) *Sustainable Peace: Democracy and Power-Dividing Institutions After Civil Wars.* Ithaca, NY: Cornell University Press.

King, E. (2008)*The Role of Education in Violent Conflict and Peacebuilding in Rwanda* Toronto [unpublished PhD thesis].

Krishnan, P. (1996) Family Background, Education and Employment In Urban Ethiopia. *Oxford Bulletin of Economics and Statistics,* 58 (1), 167-183.

Levinson, B.A., Holland, D.C. and Foley, D.E. (1996) The cultural production of the educated person: critical ethnographies of schooling and local practice. Albany: State University of New York Press.

Lewin, Keith (2007), Improving Access, Equity and Transitions in Education: Creating a Research Agenda, *CREATE Pathways to Access Monographs 1.* Brighton: Consortium for Research on Educational Access Transitions and Equity, University of Sussex), 1-50.

Melhuish, E., Phan, M.B., Sylva, K., Sammons, P., Siraj-Blatchford, I. and Taggart, B. (2008) Effects of the home learning environment and preschool center experience upon literacy and numeracy development in early primary school. *Journal of Social Issues,* 64 (1), 95-114.

Mincer, J. (1970). The distribution of labor incomes: a survey with special reference to the human capital approach. *Journal of Economic Literature* 8 (1): 1–26.

Morrow, V. (2009) *The Ethics of Social Research with Children and Families in Young Lives: Practical Experiences*. Young Lives Working Paper No. 53. Oxford: Young Lives.

National Organisation for Examinations (2004) *Ethiopian Second National Learning Assessment of Grade 4 students.* Addis Ababa: National Organisation for Examinations.

National Organisation for Examinations (2004) *Ethiopian Second National Learning Assessment of Grade 8 students.* Addis Ababa: National Organisation for Examinations.

National Organisation for Examinations (2007) *Ethiopian Second National Learning Assessment of Grade 4 students.* Addis Ababa: National Organisation for Examinations.

National Organisation for Examinations (2007) *Ethiopian Second National Learning Assessment of Grade 8 students.* Addis Ababa: National Organisation for Examinations.

Norwegian Refugee Council (2011) Ethiopia: Monitoring of conflict, human rights violations and resulting displacements still problematic. Geneva: Internal Displacement Monitoring Centre.

Orkin, K. (2011) If God wills... next year I will send her back to school: The Effects of Child and Parental Illness on School Participation in Rural Ethiopia. CREATE Pathways to Access. Research Monograph No. 60.

Østby, G. and Urdal, H. (2011) Education and Civil Conflict: A Review of the Quantitative, Empirical Literature. Background paper prepared for the Education for All Global Monitoring Report. Paris: UNESCO.

Outes-Leon, I. and Dercon, S. (2008) Survey attrition and attrition bias in Young Lives. Young Lives Technical Note No, 5. Oxford, UK: Young Lives.

Outes-Leon, I. and Sanchez, A. (2008) An assessment of the Young Lives sampling approach in Ethiopia: Young Lives Technical Note No. 1. Oxford, UK: Young Lives.

Overseas Development Institute (2011) *Ethiopia's progress in education: A rapid and equitable expansion of access.* London: ODI.

Raju, P.M. and Asfaw, A. (2009) Recalled test anxiety in relation to achievement, in the context of general academic self-concept, study habits, parental involvement and socio-economic status among Grade 6 Ethiopian students. *Education 3–13,* 37 (3), 269-285.

Rose, P., Yoseph, G., Berihun, A. and Nuresu, T., 1997, Gender and Primary Schooling in Ethiopia, *IDS Research Report* No 31, Institute of Development Studies, Brighton: IDS.

Rose, P. and Al-Samarrai, S. (2001) Household constraints on schooling by gender: Empirical evidence from Ethiopia. *Comparative Education Review,* 45 (1), pp. 36-63.

Rose, P. (2003) Can Gender Equality in Education be attained?: Evidence from Ethiopia. *Background paper for 2003 UNESCO Global Monitoring Report*. Paris: UNESCO.

Rose, P. and Dyer, C. (2008) Chronic Poverty and Education: A review of the literature. *Background paper for The Chronic Poverty Report,* 9.

Salomon, G. and Nevo, B. (2002) *Peace Education: The Concept, Principles, and Practices around the World*. London: Lawrence Erlbaum Associates.

Sanchez, A (2009) *Birth weight, nutrition and the acquisition of cognitive skills: evidence from four developing countries*. Initial findings. Background paper for EFA Global Monitoring Report 2010. Paris: UNESCO.

Sedel, S. (2008) Setting the scene, in *The Challenge of Learning: Improving the Quality of Basic Education in Sub-Saharan Africa,* Adriaan M. Verspoor (ed.) Paris: Association for the Development of Education in Africa.

Schaffner, J. (2003) *The Determinants of Schooling Investments Among Primary School Aged Children in Ethiopia*. Washington, D.C.: World Bank.

Schutt, R. K. (2006). *Investigating the Social World: the Process and Practice of Research*. USA: Pine Forge Press.

Sharp, K., Devereux, S. and Amare, Y. (2003). Destitution in Ethiopia's Northeastern Highlands (Amhara National regional State). *IDS and SC-UK Ethiopia policy research project*. London: DFID.

Stephan, W. (1999) *Reducing Prejudice and Stereotyping in Schools.* New York: Teachers College.

Stewart, F. (2000) Crisis prevention: tackling horizontal inequalities. *Oxford Development Studies* 28, 245–262.

Stewart, F. (2008a) Horizontal Inequalities: Explaining Persistence and Change, in *Horizontal Inequalities and Conflict: Understanding Group Violence in Multiethnic Societies.* Frances Stewart (ed.) New York: Palgrave MacMillan.

Stewart, F. (2008b) Horizontal Inequalities and Conflict: an introduction and some hypotheses, in *Horizontal Inequalities and Conflict: Understanding Group Violence in Multiethnic Societies.* Frances Stewart (ed.) New York: Palgrave MacMillan.

Strand, S. (2011a) The limits of social class in explaining ethnic gaps in educational attainment. *British Educational Research Journal*, 37 (2),197-229.

Strand, S. (2011b) The White British-Black Caribbean achievement gap: tests, tiers and teacher expectations. *British Educational Research Journal,* Forthcoming., 1-27.

Sylva, K., Melhuish, E. C., Sammons, P., Siraj-Blatchford, I. and Taggart, B. (2004),The Effective Provision of Pre-School Education (EPPE) Project: *Technical Paper 12 - The Final Report: Effective Pre-School Education.* London: DfES / Institute of Education, University of London.

UNESCO Institute for Statistics (2012) *Statistical Tables*. http://www.uis.unesco.org/Education/Pages/out-of-school-children.aspx [Accessed on 2 August, 2012]

UNESCO (2010) *Education for All Global Monitoring Report: reaching the marginalized.* Oxford, UK: Oxford University press.

USAID (2010) Ethiopia early grade reading assessment: data analytic report. Addis Ababa: USAID.

United Nations Development Programme (2011) *Human Development Report 2011: Sustainability and equity a better future for all*. New York, N.Y: Palgrave MacMillan.

US Agency for International Development (1994). The Demand for Primary Schooling in Rural Ethiopia. USAID research study, Addis Ababa.

Uwiko, U. Education, horizontal inequalities and ethnic relations in Nigeria. *International Journal of Educational Development,* 27 (2007), 266-281.

Verwimp, P. (1999) Measuring the quality of education at two levels: A case study of primary schools in rural Ethiopia. *International review of education,* 45 (2), 165-194.

Watson, K. Language, education and ethnicity: whose rights will prevail in an age of globalisation? *International Journal of Educational Development,* 27 (2007), 252-265.

Weir, S. (1999) The effects of education on farmer productivity in rural Ethiopia. *CSAE Working Paper Series*.

Weir, S. (2011) Parental Attitudes and Demand for Schooling in Ethiopia. *Journal of African Economies,* 20 (1), 90.

Weir, S. and Knight, J. (2007) Production externalities of education: evidence from rural Ethiopia. *Journal of African Economies,* 16 (1), 134.

WIDER (2007) *World International Inequality Database (WIID2).* Helsinki, Finland: World Institute for Development Economics Research.

Wößmann, L. (2003), Schooling resources, educational institutions and student performance: the international evidence, *Oxford Bulletin of Economics and Statistics,* 65, 117-70.

Woldehanna, T., Jones, N. and Tefera, B. (2005) *Children's educational completion rates and achievement: implications for Ethiopia's second poverty reduction strategy (2006-10).* Oxford, UK: Young Lives.

Woldehanna, T., Gudisa, R., Tafere, Y. and Pankhurst, A. (2011) *Understanding changes in the lives of poor children: initial findings from Ethiopia round 3 survey*. Oxford, UK: Young Lives.

World Bank (2005) *Education in Ethiopia: strengthening the foundation for sustainable progress.* Washington, D.C.: World Bank.

World Bank (2008) *Assessing achievement levels in education*. Washington, D.C: World Bank.

World Health Organisation (2011) *Global database on child growth and malnutrition*. http://www.who.int/nutgrowthdb/about/introduction/en/index5.html [Accessed on 10 August, 2012].

Young Lives (2002) *Round 1 survey justification document*. [unpublished].

Young Lives (2011) *School survey handbook: data user guide*. [unpublished].

Zehle, J. (2009) Dropout and Learning Difficulties in 1st Grade Government Primary Schools in Ethiopia. Proceedings of the 16th International Conference of Ethiopian Studies, pp. 1117-1124.

APPENDIX I: DESCRIPTION OF THE VARIABLES

Variable name	Variable description
PPVT ability test score at age 15 (2009)	Dependent, ratio scale, continuous variable; Verified normal distribution; (min = 31, max = 166)
PPVT ability test score at age 12 (2009)	Dependent, ratio scale, continuous variable; Verified normal distribution; (min = 11, max = 127)
Maths achievement test score at age 15 (2009)	Dependent, ratio-scale continuous variable; Verified (truncated) normal distribution; (min = 0, max = 22)
Maths achievement test score at age 12 (2006)	Dependent, ratio-scale continuous variable; Verified (truncated) normal distribution; (min = 0, max = 9)
Verbal ability test score at age 15 (2009)	Dependent, ratio-scale, continuous variable; Verified (truncated) normal distribution; (min = 0, max = 19)
Literacy skill at age 12 (2006)	Dependent, ordinal-categorical variable; Codes are: 0 = cannot read or write; 1 = reads words and writes with errors; 2 = reads sentences and writes easily
Region (2002)	Independent, nominal-categorical variable; Codes are: 1 = Addis Ababa; 2 = Amhara; 3 = Oromia; 4 = SNNP; 5 = Tigray
Ethnic group (2002)	Independent, nominal-categorical variable; Codes are: 1 = Amhara; 2 = Gurage; 3 = Hadia; 4 = Oromo; 5 = Sidama 6 = Tigrian; 7 = Wolayta; 8 = Other
Gender (2002)	Independent, nominal-categorical variable; Codes are: 1 = Male; 2 = Female
Height-for-age age 8, 12 and 15 (2002, 2006, 2009)	Independent, ratio-scale, continuous variable; (min = -6.0, max = 6.0)
Years of early childhood care (computed using 2009 data)	Independent, interval-continuous variable; (min = 0, max = 6)
Age of entry into formal school (computed using 2009 data)	Independent, interval-continuous variable; (min = 1, max = 15)
Number of grades repeated (computed using 2009 data)	Independent, interval-continuous variable; (min = 0, max =6)
Hours of household work per day (computed using 2009 data)	Independent, ratio-scale, continuous variable; (min = 0, max = 13)
Hours of productive work per day (computed using 2009 data)	Independent, ratio-scale, continuous variable; (min = 0, max = 13)

Housing quality index score at age 8, 12 and 15	Independent, ratio-scale, continuous variable; (min = 0, max = 1)
Services quality index score at age 8, 12 and 15	Independent, ratio-scale, continuous variable; (min = 0, max = 1)
Asset index score at age 8, 12 and 15	Independent, ratio-scale, continuous variable; (min = 0, max = 1)
Mother's educational level at age 12 (recoded using 2006 data)	Independent, interval-continuous variable; (min = 0, max = 11)
Number of older siblings at age 12 (2006)	Independent, interval-continuous variable; (min = 0, max = 14)
Number of younger siblings at age 12 (2006)	Independent, interval-continuous variable; (min = 0, max = 6)
Geographic stratum at age 8 (2002)	Independent, nominal-categorical variable; Codes are: 1 = Urban; 2 = Rural
School principal's highest qualification (2009)	Independent, ordinal-categorical variable; Codes are: 1 = Low; 2 = Medium; 3 = High
Type of school (2009)	Independent, ordinal-categorical variable; Codes are: 1 = Full-day; 2 = Half-day; 3 = Shift
School is secure (2009)	Independent, ordinal-categorical variable; Codes are: 0 = No; 1 = Yes
School is well-maintained (2009)	Independent, ordinal-categorical variable; Codes are: 0 = No; 1 = Yes
Availability of school toilets (2009)	Independent, ordinal-categorical variable; Codes are: 0 = No; 1 = Yes
Availability of school library (2009)	Independent, ordinal-categorical variable; Codes are: 0 = No; 1 = Yes
Availability of school play area (2009)	Independent, ordinal-categorical variable; Codes are: 0 = No; 1 = Yes

APPENDIX II: EDUCATIONAL OUTCOMES BY KEY CONTEXTUAL VARIABLES

Table i: Mean and SD of PPVT, maths and verbal test scores at age 15 by key contextual variables

		PPVT			Math			Verbal		
Variable	Value	Mean	SD	N	Mean	SD	N	Mean	SD	N
Individual factors:										
Gender	Male	126.21	28.30	488	4.69	4.522	496	4.49	3.942	419
	Female	122.72	28.34	466	3.69	3.820	475	4.32	4.306	413
	Gap									
Early nutrition (2002)	Severely undernourished	111.29	30.37	112	2.52	3.280	115	2.38	2.761	89
	Not undernourished	127.55	27.16	624	4.66	4.409	633	4.95	4.257	542
	Gap									
Early childhood care	No early childhood care	111.42	28.99	424	2.46	2.786	438	2.27	2.742	325
	At least one year of care	135.06	22.90	525	5.64	4.647	529	5.79	4.272	501
	Gap									
School start age	7 years (official age) or younger	139.25	20.44	288	6.53	4.936	288	6.87	4.518	276
	10 years or older	107.93	29.11	286	2.14	2.854	297	2.11	2.995	209
	Gap									
Grade repetition	No grades repeated	128.13	26.40	483	5.27	4.624	495	5.27	4.376	423
	At least 2 grades repeated	122.49	29.66	127	2.89	3.170	129	2.65	2.800	108
	Gap									
Household work (2006)	No work	127.64	27.42	94	5.47	4.973	98	5.52	4.544	92
	At least 3 hours per day	119.39	28.34	330	3.32	3.536	338	3.17	3.286	278
	Gap									
Productive work (2009)	No work	131.44	25.34	510	4.66	4.401	518	5.02	4.315	455
	At least 3 hours per day	111.27	30.75	211	3.19	3.749	217	3.32	3.745	189
	Gap									
Household factors:										
Household wealth (2009)	Bottom quintile	109.29	27.55	182	3.13	3.696	191	3.31	3.326	147
	Fourth quintile	145.87	13.60	270	7.12	4.673	77	7.70	4.428	71

	Gap									
Mother's educational level	No education	117.84	29.48	470	3.61	3.915	479	3.76	3.736	411
	Secondary or above	144.57	14.98	74	7.14	4.683	74	6.44	4.564	70
	Gap									
Father's educational level	No education	115.33	29.25	248	3.33	3.826	251	3.70	3.902	212
	Secondary or above	139.48	21.73	138	6.02	4.843	138	6.36	4.425	121
	Gap									
Older siblings (2006)	No siblings	127.74	27.93	194	4.92	4.661	197	4.99	4.254	175
	At least 4 siblings	123.93	27.66	457	4.23	4.210	465	4.37	4.096	386
	Gap									
Younger siblings (2006)	No siblings	132.11	25.61	198	4.74	4.605	199	5.68	4.598	180
	At least 4 siblings	117.48	28.90	280	3.74	4.028	286	3.72	3.781	236
	Gap									
Community and school:										
Geographic stratum	Urban	139.80	19.50	388	5.76	4.636	390	6.42	4.373	345
	Rural	114.02	28.72	566	3.15	3.555	581	2.98	3.263	487
	Gap									
Distance to school (2009)	15 minutes or less	129.35	26.15	394	4.81	4.411	400	4.80	4.149	347
	30 minutes or more	117.82	27.20	132	4.18	3.982	137	4.53	4.062	129
	Gap									
Teacher qualification (2009)	No qualification	117.33	26.28	169	2.71	3.229	172	2.92	2.785	106
	Degree or above	138.43	21.28	198	6.77	4.850	199	7.26	4.466	190
	Gap									
School electricity (2009)	No electricity	119.08	25.72	108	2.93	3.508	111	3.99	3.914	70
	Electricity	130.38	25.13	482	4.77	4.304	487	4.53	3.963	454
	Gap									
School water (2009)	No running water	120.03	26.93	157	3.54	3.464	162	3.49	3.519	136
	Running water	128.26	26.24	454	4.75	4.465	461	4.68	3.966	421
	Gap									

Table ii: Mean and SD of school start age, grade repetition and completion at age 15 by key contextual variables

		School start age			Number of repeated grades			Highest grade completed		
Variable	Value	Mean	SD	N	Mean	SD	N	Mean	SD	N
Individual factors:										
Gender	Male	8.800	2.023	490	.632	1.028	435	5.65	2.041	435
	Female	8.586	1.858	474	.698	1.127	435	5.82	1.878	435
	Gap									
Early nutrition (2002)	Severely undernourished	10.162	2.121	111	.676	.834	102	4.25	2.018	102
	Not undernourished	8.337	1.757	631	.686	1.156	577	6.07	1.790	577
	Gap									
Early childhood care	No early childhood care	10.384	1.478	432	.524	.966	372	4.17	1.549	372
	At least one year of care	7.332	.965	530	.772	1.147	496	6.90	1.331	496
	Gap									
School start age	7 years (official age) or younger	-	-	-	.933	1.391	270	7.47	1.318	270
	10 years or older	-	-	-	.439	.809	246	3.51	1.405	246
	Gap									
Grade repetition	No grades repeated	-	-	-	-	-	-	6.15	1.909	495
	At least 2 grades repeated	-	-	-	-	-	-	4.40	1.603	129
	Gap									
Chores (2002)	No chores	8.609	1.862	323	-	-	-	-	-	-
	Chores	8.743	1.995	640	-	-	-	-	-	-
Household work (2006)	No work	-	-	-	.627	.797	86	5.86	2.001	86
	At least 3 hours per day	-	-	-	.755	1.206	302	5.36	1.995	302
	Gap									
Productive work (2009)	No work	-	-	-	.690	1.112	497	5.02	4.315	455
	At least 3 hours per day	-	-	-	.780	1.079	146	3.32	3.745	189
	Gap									
Household factors:										
Household wealth (2002)	Bottom quintile	9.372	2.061	513	.624	1.119	439	5.07	2.056	439
	Fourth quintile	7.142	1.014	21	.571	.597	21	7.29	1.102	21

	Gap									
Household wealth (2009)	Bottom quintile	-	-	-	.405	.847	148	4.93	2.134	148
	Fourth quintile	-	-	-	.826	1.189	75	6.88	1.602	75
	Gap									
Caregiver literacy (2002)	Caregiver cannot read	8.995	1.988	608	-	-	-	-	-	-
	Caregiver reads easily	7.871	1.576	210	-	-	-	-	-	-
Mother's education (2006)	No education	-	-	-	.548	.869	410	5.38	1.986	410
	Secondary or above	-	-	-	.611	.942	72	6.89	1.338	72
	Gap	-	-	-						
Father's education (2006)	No education	-	-	-	.549	.895	211	5.30	2.036	211
	Secondary or above	-	-	-	.619	1.017	134	6.66	1.773	134
	Gap									
School-aged children (2002)	None	8.638	1.998	144	-	-	-	-	-	-
	At least 3 children	9.045	1.868	287	-	-	-	-	-	-
	Gap									
Older siblings (2006)	No siblings	-	-	-	.661	1.076	183	5.98	1.978	183
	At least 4 siblings	-	-	-	.580	1.003	410	5.71	1.972	410
	Gap									
Younger siblings (2006)	No siblings	-	-	-	.652	.910	184	6.27	1.908	184
	At least 4 siblings	-	-	-	.567	.940	252	5.52	1.967	252
	Gap									
Community and school:										
Geographic stratum	Urban	7.845	1.503	389	.718	.971	376	6.46	1.647	376
	Rural	9.269	2.001	575	.625	1.153	494	5.18	2.004	494
	Gap									
Distance to school (2009)	15 minutes or less	-	-	-	.662	.954	400	5.86	1.885	400
	30 minutes or more	-	-	-	.489	1.284	137	5.80	1.798	137
	Gap									
Teacher qualification (2009)	No qualification	-	-	-	.735	1.252	170	4.94	1.897	170
	Degree or above	-	-	-	.356	.898	199	7.37	1.544	199

	Gap	-	-	-						
School electricity (2009)	No electricity	-	-	-	.590	1.429	110	4.84	2.228	110
	Electricity	-	-	-	.612	.978	482	6.24	1.640	482
	Gap	-	-	-						
School water (2009)	No running water	-	-	-	.592	1.006	157	5.25	1.670	157
	Running water	-	-	-	.621	1.117	457	6.11	1.891	457
	Gap									

APPENDIX III: CONTEXTUALISED MODELS OF PROGRESSION

Table i: Contextualised models of general cognitive development between age 12 and 15 (2006-2009)

		Model 1 (N=955)		Model 2 (N=929)		Model 3 (N=916)		Model 4 (N=900)		Model 5 (N=691)	
Variable	Value	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.
Base model											
Intercept	Intercept	151.054***	3.027	111.174***	4.601	176.288***	9.030	164.592***	9.629	166.972***	13.768
Region	Dummy: 1 if Amhara	-36.068***	3.561	-22.076***	3.597	-17.781***	3.768	-12.506***	3.956	2.643	6.196
(Base group: Addis Ababa)	Dummy: 1 Oromia	-27.228***	3.400	-16.493***	3.338	-8.965***	3.511	-7.980**	3.609	2.866	5.050
	Dummy: 1 SNNP	-15.033***	3.741	-5.822*	3.621	-.946	3.486	1.723	3.550	6.732	5.527
	Dummy: 1 Tigray	-27.452***	6.651	-10.534*	6.405	-5.827	6.036	-1.136	6.296	8.022	8.053
Ethnic group	Dummy: 1 if Gurage	-19.664***	4.388	-19.118***	4.149	-17.877***	3.807	-12.268***	3.941	-6.437	5.741
(Base group: Amhara)	Dummy: 1 if Hadia	-28.720***	5.277	-18.742***	5.059	-8.817*	4.711	-4.066	4.780	5.980	6.571
	Dummy: 1 if Oromo	-7.079**	3.605	-7.058**	3.383	-6.521**	3.053	-4.004	3.158	-1.934	4.035
	Dummy: 1 if Sidama	-17.306***	5.477	-10.430**	5.274	-4.719	4.855	-1.896	4.864	9.623	7.048
	Dummy: 1 if Tigrian	3.888	6.740	-2.045	6.333	-.961	5.725	.790	5.864	3.038	7.081
	Dummy: 1 if Wolayta	1.661	5.094	1.035	4.842	2.361	4.447	2.382	4.497	2.919	6.480
	Dummy: 1 if Other	-14.287**	6.590	-9.385*	6.191	-11.069*	5.787	-10.119*	5.907	-3.096	8.294
Prior attainment											
PPVT raw test score (2006)	Number of points			.392***	.035	.199***	.035	.164***	.036	.131***	.042
Child characteristics											
Gender of child (2002)	Dummy: 1 if female					-9.277***	1.799	-9.130***	1.801	-9.717***	2.210
Health status (2006)	Height-for-age index score					-.598	.891	-.411	.889	-.350	1.041
Health status (2009)	Height-for-age index score					.376	.908	.283	.905	.513	1.045
Early childhood care (2009)	Number of years of care					.281	.915	-.197	.929	-.639	1.203
School starting age (2006)	Age in years					-5.259***	.677	-5.093***	.681	-4.348***	.872

Grades repeated	Number of grades			-3.387***	.709	-3.551***	.708	-3.066***	.952
Household work (2006)	Number of hours per day			-.073	.470	.319	.480	.702	.556
Productive work (2006)	Number of hours per day			-1.220***	.443	-.772*	.447	.137***	.562
Household work (2009)	Number of hours per day			.036	.172	.104	.171	.051	.243
Productive work (2009)	Number of hours per day			-1.238***	.369	-1.012***	.371	-2.149	.602
Family background									
Housing quality (2009)	Index score					10.181**	5.250	5.328	6.310
Services quality (2009)	Index score					4.825	4.152	3.052	5.083
Asset index (2009)	Index score					10.381*	5.589	10.822	7.031
Mother's education (2006)	Highest qualification					.325*	.209	.432*	.244
Older siblings (2006)	Number of older siblings					-.124	.324	.149	.386
Younger siblings (2006)	Number of younger siblings					-1.486***	.608	-.912	.702
School and community									
Geographic stratum (2002)	Dummy: 1 if rural							-8.347***	3.122
Principal qualification (2009)	Dummy: 1 if high level							6.758***	2.541
Type of school (2009)	Dummy: 1 if shift school							.302	.229
School is secure (2009)	Dummy: 1 if yes							-2.372	2.660
School is well-maintained (2009)	Dummy: 1 if yes							.812	3.611
School has a library (2009)	Dummy: 1 if yes							4.806	3.841
School has toilets (2009)	Dummy: 1 if yes							-11.509**	5.383
School has a play area (2009)	Dummy: 1 if yes							-10.595***	3.281
Percentage of the variance in general cognitive development explained (R^2)		*.164*		*.269*		*.411*		*.420*	

*p < .10; **p < .05, ***p < .01

Table ii: Contextualised models of mathematics progression between age 12 and 15 (2006-2009)

		Model 1 (N=955)		Model 2 (N=929)		Model 3 (N=916)		Model 4 (N=907)		Model 5 (N=703)	
Variable	Value	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.
Base model											
Intercept	Intercept	6.827***	.465	1.116**	0.525	4.068***	1.451	2.980**	1.569	4.156**	2.117
Region	Dummy: 1 if Amhara	-3.287***	.546	-1.399***	0.492	0.111	0.562	0.481	0.598	1.251	.918
(Base group: Addis Ababa)	Dummy: 1 Oromia	-2.524***	.520	-0.640*	0.466	1.210**	0.525	1.325**	0.546	1.957***	.736
	Dummy: 1 SNNP	-.285	.575	0.427	0.504	1.520***	0.520	1.797***	0.535	2.289***	.806
	Dummy: 1 Tigray	-1.422	1.023	-0.308	0.892	1.510*	0.897	1.846**	0.951	1.651	1.171
Ethnic group	Dummy: 1 if Gurage	-1.239*	.675	-0.774	0.590	-0.574	0.571	-0.329	0.595	.358	.837
(Base group: Amhara)	Dummy: 1 if Hadia	-5.034***	.806	-2.483***	0.743	-1.326*	0.729	-0.916	0.747	-.525	.975
	Dummy: 1 if Oromo	-1.070**	.553	-0.709	0.485	-0.741*	0.463	-0.621	0.483	-.252	.593
	Dummy: 1 if Sidama	-4.582***	.843	-2.441***	0.746	-1.485**	0.728	-1.437**	0.737	-.802	1.024
	Dummy: 1 if Tigrian	-.492	1.037	-0.665	0.902	-0.589	0.861	-0.617	0.890	.331	1.028
	Dummy: 1 if Wolayta	-1.936**	.784	-1.359**	0.686	-0.824	0.666	-0.911	0.680	-1.531*	.941
	Dummy: 1 if Other	-1.676*	1.014	-1.520*	0.902	-0.838	0.893	-0.712	0.922	.313	1.262
Prior attainment											
Maths test score (2006)	Number of points			0.870***	0.051	0.590***	0.058	0.561***	0.059	.578***	.066
Child characteristics											
Gender of child (2002)	Dummy: 1 if female					-1.104***	0.269	-1.006***	0.273	-.924***	.320
Health status (2006)	Height-for-age index score					0.137	0.134	0.146	0.135	.232	.151
Health status (2009)	Height-for-age index score					-0.076	0.136	-0.100	0.136	-.057	.152
Early childhood care (2009)	Number of years of care					0.559***	0.140	0.522***	0.144	.363**	.178
School starting age (2006)	Age in years					-0.232**	0.113	-0.219**	0.115	-.248*	.140
Grades repeated	Number of grades					-0.567***	0.112	-0.580***	0.113	-.681***	.144
Household work (2006)	Number of hours per day					-0.121*	0.072	-0.094	0.074	-.066	.082
Productive work (2006)	Number of hours per day					-0.190***	0.069	-0.153**	0.071	-.145*	.085

Household work (2009)	Number of hours per day			0.002	0.026	0.006	0.026	.013	.035
Productive work (2009)	Number of hours per day			-0.069	0.056	-0.033	0.057	.059	.087
Family background									
Housing quality (2009)	Index score					1.555**	0.800	1.530*	.917
Services quality (2009)	Index score					-0.476	0.627	-1.625**	.733
Asset index (2009)	Index score					1.664**	0.850	1.276	1.020
Mother's education (2006)	Highest qualification					0.009	0.032	.003	.035
Older siblings (2006)	Number of older siblings					-0.020	0.050	-.016	.056
Younger siblings (2006)	Number of younger siblings					-.077	.093	-.058	.102
School and community									
Geographic stratum (2002)	Dummy: 1 if rural							-.949**	.449
Principal qualification (2009)	Dummy: 1 if high level							.522	.369
Type of school (2009)	Dummy: 1 if shift school							-.030	.033
School is secure (2009)	Dummy: 1 if yes							-.353	.385
School is well-maintained (2009)	Dummy: 1 if yes							-.142	.520
School has a library (2009)	Dummy: 1 if yes							.276	.578
School has toilets (2009)	Dummy: 1 if yes							-.844	.813
School has a play area (2009)	Dummy: 1 if yes							.244	.490
Percentage of the variance in maths progression explained (R^2)		*.164*		*.269*		*.411*		*.408*	*.426*

*p < .10; **p < .05, ***p < .01

Table iii: Contextualised models of literacy progression between age 12 and 15 (2006-2009)

		Model 1 (N=826)		Model 2 (N=812)		Model 3 (N=806)		Model 4 (N=792)		Model 5 (N=608)	
Variable	Value	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.
Base model											
Intercept	Intercept	8.626***	0.445	8.679***	0.447	14.984***	1.267	12.917***	1.391	17.210***	2.047
Region	Dummy: 1 if Amhara	-4.912***	0.506	-4.936***	0.512	-3.433***	0.524	-2.829***	0.562	-3.562***	.894
(Base group: Addis Ababa)	Dummy: 1 Oromia	-4.736***	0.482	-4.636***	0.487	-2.706***	0.502	-2.591***	0.515	-2.394***	.717
	Dummy: 1 SNNP	-2.560***	0.560	-2.554***	0.559	-1.299***	0.512	-1.115**	0.517	-1.854**	.808
	Dummy: 1 Tigray	-5.127***	0.938	-5.112***	0.938	-3.658***	0.853	-3.509***	0.895	-4.222***	1.134
Ethnic group	Dummy: 1 if Gurage	-1.572**	0.636	-1.600***	0.637	-1.641***	0.555	-0.974*	0.574	-1.106	.862
(Base group: Amhara)	Dummy: 1 if Hadia	-4.877***	1.548	-4.940***	1.547	-2.508*	1.345	-2.428*	1.323	-2.698	1.885
	Dummy: 1 if Oromo	-1.382***	0.512	-1.441***	0.520	-1.434***	0.449	-1.283**	0.462	-2.077***	.581
	Dummy: 1 if Sidama	-2.887***	1.610	-2.918*	1.609	-1.517	1.399	-1.852	1.380	-2.289	2.383
	Dummy: 1 if Tigrian	1.716*	0.951	1.677*	0.952	1.605*	0.821	2.241**	0.832	2.294**	.982
	Dummy: 1 if Wolayta	0.030	0.933	-0.015	0.934	0.101	0.810	-0.032	0.801	-.601	1.064
	Dummy: 1 if Other	-0.989	0.950	-1.056	0.951	-0.770	0.842	-0.688	0.853	-2.967**	1.216
Prior attainment											
Literacy test score (2006)	1: if reads and writes			0.005	0.022	-0.009	0.019	0.011	0.022	.015	.031
Child characteristics											
Gender of child (2002)	Dummy: 1 if female					-0.325	0.252	-0.158	0.250	-.450	.300
Health status (2006)	Height-for-age index score					-0.231	0.233	-0.196	0.232	-.183	.251
Health status (2009)	Height-for-age index score					0.272**	0.140	0.222*	0.141	-.066	.182
Early childhood care (2009)	Number of years of care					-0.865***	0.102	-0.789***	0.103	-.841	.133
School starting age (2006)	Age in years					-1.046***	0.112	-1.011***	0.112	-1.086	.153
Grades repeated	Number of grades					-0.005	0.025	0.005	0.024	.029***	.034
Household work (2009)	Number of hours per day					0.114**	0.056	0.174***	0.057	.195***	.087
Productive work (2009)	Number of hours per day					-0.325	0.252	-0.158	0.250	-.450	.300

Family background									
Housing quality (2009)	Index score					.491	.815	-.406**	.965
Services quality (2009)	Index score					1.234**	.633	.168	.754
Asset index (2009)	Index score					1.366	.867	1.623	1.044
Mother's education (2006)	Highest qualification					.021	.032	.037	.035
Older siblings (2006)	Number of older siblings					-.090*	.050	-.117**	.057
Younger siblings (2006)	Number of younger siblings					-.102	.093	-.132	.104
School and community									
Geographic stratum (2002)	Dummy: 1 if rural							-1.077**	.438
Principal qualification (2009)	Dummy: 1 if high level							1.151***	.364
Type of school (2009)	Dummy: 1 if shift school							-.021	.032
School is secure (2009)	Dummy: 1 if yes							-.146	.382
School is well-maintained (2009)	Dummy: 1 if yes							-.731	.601
School has a library (2009)	Dummy: 1 if yes							-.845	.597
School has toilets (2009)	Dummy: 1 if yes							-.732	.877
School has a play area (2009)	Dummy: 1 if yes							.310	.465
Percentage of the variance in literacy progression explained (R^2)		*.208*		*.206*		*.410*		*.426*	*.467*

*p < .10; **p < .05, ***p < .01