THE EFFECTS OF EARLY CHILDHOOD EDUCATION ATTENDANCE ON COGNITIVE DEVELOPMENT: EVIDENCE FROM URBAN ETHIOPIA[1]

Tassew Woldehanna[2]

Abbreviation

PPVT	Peabody Picture Vocabulary Test
CDA-Q test	Cognitive Development Assessment – Quantity test
OLS	Ordinary least square
IV	Instrumental variables
UNESCO	The United Nations Educational, Scientific and Cultural Organization
UNICEF	The United Nations Children's Fund
MoFED	Ministry of Finance and Economic Development
EDRI	Ethiopian Development Research Institute
MoE	Ministry of Education
YL	Young Lives
YC	Younger cohort
OC	Older cohort
SNNP	Southern Nation and Nationalities People
NGO	**Non-Governmental Organizations**

[1] The final version of this article was submitted in August 2012.
[2] School of Economics, Addis Ababa University, Correspondence address: Tassew Woldehanna, Associate Professor, School of Economics, AAU. P. O. Box 150175, Addis Ababa, Ethiopia; email: tassew.woldehanna@gmail.com
[2] Young Lives is an international study of childhood poverty that follows 12 children in four countries namely Ethiopia, India, Vietnam and Peru (for details of the young Lives look at www.younglives.org.uk)

Abstract

Studies in developed countries indicate that preschool education can have strong impact on children's cognitive development, but there are no studies conducted in the context of developing countries including Ethiopia where pre-school education is left for private sector. To see if government investment in pre-school education is worth, we examined the effects of early childhood education attendance on cognitive development of preschool age children. Using data obtained from the Young Lives Longitudinal Survey in Ethiopia, we measured cognitive development of children by Peabody Picture Vocabulary Test (PPVT) and Cognitive Development Assessment – Quantity (CDA-Q) Test. We employed standard instrumental variable estimation, Wooldridge instrumental variable estimation and propensity score matching techniques to see the degree of association between preschool enrolment and cognitive outcome of five years old children. Across all these models, the results persistently show that early childhood education attendance is positively associated with children's cognitive development. More specifically, from propensity score matching result, children who have been attending preschool education have scored 31.2% higher in vocabulary test and 23.1 % in quantitative test than those of non-preschool attendees. The implications of these results are quite crucial and timely for the Ethiopian government. Despite the fact that preschool education increases children's cognitive development, public investment in this critical stage of education is currently very low in the country and left for the private sector. As a result, the subsector is dominated by fee charging kindergartens in which children from low socioeconomic background do have very little opportunity to attend this fist and critical stage of education implying that government has to do more in this first and essential stage of education.

1. Introduction

It is widely recognized that early childhood education is an integral part of basic education and it could be an essential step in achieving the goals of Education-for-All (UNESCO, 1996) in particular and human skill formation in general (see, among others, Cunha et al., 2006; Hackman, 1999; Currie 2001; Goodman and Sianesi, 2005). A good reason is that a well-conceived quality of early childhood education helps to meet the diverse needs of young children during the crucial early years of life, enhance their readiness for schooling, and have a positive and permanent influence on later schooling achievements (Carneiro, et al. 2006). For this reason, Woodhead (2009) states that early education may be the single most effective intervention for helping poor children, families, communities and nations break the intergenerational cycle of poverty.

The effect of early childhood education is not only limited to cognitive development of young children, but also to a number of non-cognitive skills such as motivation, self-discipline, socialization (see, Heckman, 1997; Cunha et al., 2006). More remarkably, these cognitive and non-cognitive skills interact and reinforce each other, characteristics termed as *self* and *cross- productivity* (Carneiro, et al., 2003; Helmers and Patnam, 2009). Nevertheless, the rationale here is that as cognitive ability is less malleable later in life than non-cognitive ability, intervening to improve cognitive functioning should be given more attention early in life. Additionally, cognitive skills are also much easier to measure at early ages compared to the non-cognitive skills (Connelly, 2008).

Nevertheless, in most developing countries a large share of children start education late in their ages and directly join primary schools skipping the nursery and kindergarten. As a result of this phenomenon, it is very common to see that high grade repetition and dropout rates are the main characteristics of their education sectors (UNESCO, 2005). The fact is that children with low levels of cognitive development before they enter school have lower school achievement and earn lower wages in their later lives (Currie and Thomas 2000; Case and Paxson 2006; Macours et al., 2008). Heckman (2006) also adds that low investment in childhood development in the first few years of life leads to lower cognitive development and reduces school performances, which have again a long-lasting adverse effect on human skill formation. Based on these grounds, low levels of cognitive development during early years of life have been tied to poor performance in school in a number of settings in those developing countries (Grantham-McGregor et al. 2007).

In the industrialized countries, to appreciate that early childhood education serves as a pathway for educational quality and in turn the acquired cognitive development is one of the basic predictors of success throughout life, many empirical studies are done on the link between early childhood education and cognitive development (see, Susanna et al., 2005; Magnuson, 2004; Campbell et al., 2002; Goodman, 2006). However, much less attention has been devoted to studying this kind of critical link in most developing countries, of which Ethiopia is one; and nor investment in early childhood development is seen as a critical foundation for school readiness and achieving success in school and life (Macours at al., 2008).

For instance, when we examine the nature of early childhood education in Ethiopia, It is predominantly provided by the private sector, non-government organizations, religious institutions, etc. Except in some technical support and quality monitoring, the government has very limited intervention in this critical stage of education. The point is that preschool education in Ethiopia is not compulsory and neither is any budget allotted by government for it. It is rather dominated by fee charging nursery schools and kindergarten which mainly supply to the needs of middle class parents living in urban and semi-urban areas of the country. Due to this nature of setting up, the enrolment rate of the subsector is very undersized. For example, in 2008/09, the gross enrolment rate of the country was only about 4.2% and concentrated in urban areas, mainly in Addis Ababa. This is a striking figure compared to the 94.2% (83%) national gross (net) enrolment of primary education for the same year. More specifically, at national level, 22.9% of pupils enrolled in grade one in 2008/09 had left school before reaching grade two. Likewise, the survival rates to grades five and eight were 78.9% and 43.6%, respectively. The primary education system also suffers from large numbers of out-of- school children and over-age children. Those entire situations put a logical question does preschool education help children to enter primary school at appropriate age, reduce their dropouts and repetition (UNICEF, 2007).

This paper aims to analyse the effects of early childhood education attendance on children's cognitive development using a unique Young Lives Longitudinal Data. To be specific, the objectives of this study are:

- To examine the pattern and trends of early childhood education enrolment;
- To analyse the effects of early childhood education on cognitive development of five years old children;
- To identify the main challenges that this subsector encounters; and
- To infer possible policy insights for the identified challenges in the subsector.

It is also equally important to note that the terms early childhood education and preschool education are used interchangeably in this study. Therefore, early childhood– or preschool – education in this study refers to educational efforts between three and six years of age that aim at fostering cognitive, social, motivational and emotional development of young children in order to provide them with a good start in formal primary education. A good start in primary school, in turn, increases the likelihood of favourable educational and social outcomes later in life.

Estimating the impact of preschool on cognitive development is problematic as there is selection problem in preschool enrolment, which biases the results. We tried to reduce the bias by using propensity score matching where children in preschools are matched to those who did not enrol in preschool using the characteristics of the parents, shocks household encountered and initial household wealth and composition before the children were born and reached the age of three and five. We also used instrumental variable estimation method in which the instruments are highly correlated with the preschool, but not with the outcome (cognitive score). We also used Wooldridge approach in which the propensity score (predicted probability from a probit model) is used as a main or excluded instrument (Wooldridge, 2002). Moreover, we have conducted relevance and validity test to see if the chosen instruments are correlated with the endogenous variables, but not with the error terms (Wooldridge 2002).

The finding of this study contributes to the existing literature in a number of ways. First, we document a clear association between early childhood education and cognitive development of children in Ethiopia that enable children to enter formal education ready to learn. Hence, this will cast an interesting implication for the expansion of public preschools, which are believed to be the foundation for sustainable educational quality (UNICEF, 2007). Second, beyond the supply side response, it may create a demand side response, in the sense that most people consider preschool as a luxury and can be substituted easily by the primary school. So, this study sheds some insight for creating communities' awareness to send their children to preschool education at their appropriate ages so that better educational performance can be achieved in the future. Moreover, this study also adds one developing country, Ethiopia, into the existing educational empirical literature.

The rest of the paper proceeds as follows. Section 2 presents related literature review together with assessment of the Ethiopian early childhood education trends. Section 3 lays the framework for the econometric analysis while section 4 presents the descriptive statistics and section 5 contains the empirical results. Concluding remarks are provided in section 6.

2. Literature Review

2.1 The Rationale for Early Childhood Education Investment

The early years of life are so critical for the acquisition of concepts, skills and attitudes that lay the foundation for lifelong learning (Cunha et al., 2006; Carneiro and Heckman, 2003). The fact is that early childhood is a sensitive period marked by rapid transformations in physical, cognitive, language, social and emotional development. It represents a window of opportunity for a lifetime development of a person (UNESCO, 2010).This is the time when children's brains development advances at a pace greater than any other stage in life. For example, childhood development researches indicate that by the age of 2½ years, a child's brain has achieved 50% of its adult weight, and by the age of 5, the brain has grown to 90% of its adult weight. In addition, many of the brain's structures and biochemical routes are developed in the first two years of life (see Bruner, 1999; Halfon et al., 2001). In view of that, Heckman (2009) states that the process of skill formation would be easier at early period as it is possible for this skill to grow along with the development of brain neurons. This implies that human skill formation is a multi-stage process (Figure 2.1) in which investment is done at preschool age, during school age and post school age -job training - (Cunha et al., 2006). For that reason, the theory of human capital emphasizes, among others, on the significance of early childhood education for its initial formation (Heckman and Klenow, 1997; Cunha and Heckman; 2003).

Over recent years, based up on the traditional theory of human capital (e.g., Becker 1962), a number of authors have developed a model of skill formation technology that allows to assess education and training policies over the life cycle of a person (see, among others, Heckman 2000; Carneiro and Heckman 2003; Cunha et al. 2006). The key insight of this skill formation model is that the formation of skills is a life cycle process that demonstrates both self- productivity and complementarity (Cunha et al., 2006). By self-productivity, we mean that education learned at one stage is an input into the learning process of the next stage, implying that skills are self-reinforcing. Similarly, by complementarity we mean that productivity with which investments at one stage of education are transformed into valuable skills is positively affected by the level of skills that a person has already obtained in the previous stages, implying that skills produced at one stage raise the productivity of investment at subsequent stages (Cunha et al., 2006; Woessmann, 2006).

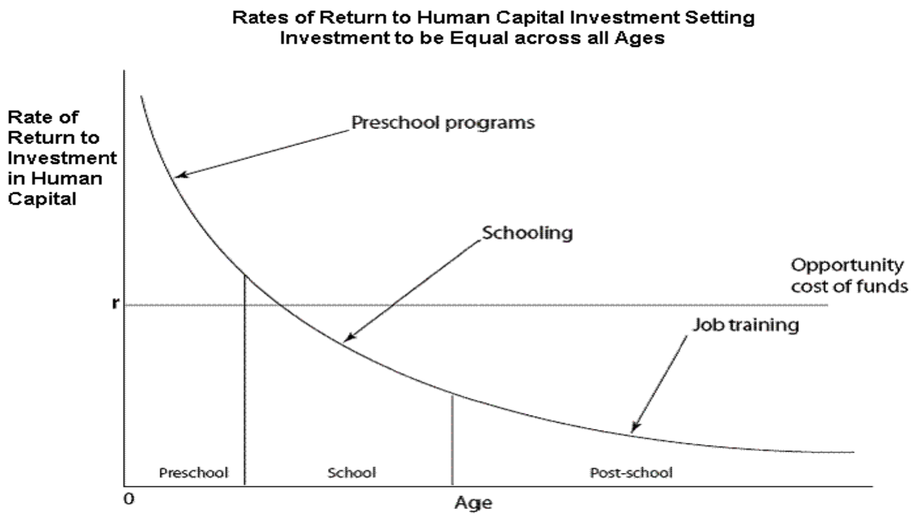
Jointly, these features of self-productivity and complementarities produce a skill multiplier whereby an investment in education at one stage raises not only directly the skills attained at that stage, but also indirectly the productivity with which educational investments at the next stage will be transformed into even further skills (Carneiro and Heckman 2003; Cunha et al., 2006; Woessmann, 2006). These multiplier effects explain why education is a dynamic synergistic process in which early learning begets later learning and the sooner it begins the greater the returns from it (Heckman and Klenow, 1997; Cunha et al., 2006).

The human capital argument also stresses that there are multiple important skills, both cognitive and non-cognitive, and that for some of these skills (most notably on the cognitive side), there are sensitive or even critical periods in a person's life cycle where investments are particularly effective or even crucial, and that inadequate early investments are difficult and costly to remedy later on (Cunha et al., 2006; Woessmann, 2006). This signifies that investment on education is better to those who start at their early stage of development than later years as young children's cognitive ability and behavior are more malleable compared to adults (Connelly, 2008). By the same token, Heckman and Masterov (2007) also provide productivity argument for investing over young children ascertaining the importance of early education by their maxim in the following ways:

> *Skill begets skill; learning begets learning. Early disadvantage, if left untreated, leads to academic and social difficulties in later years. Advantages accumulate; so does disadvantage. That is why, a large body of evidence shows that post-school remediation programs like public job training and General Educational Development (GED) certification cannot compensate for a childhood of neglect for most people.*

All these ideas underpin the case over the importance of early childhood education investment as relevant for the development of cognitive and non-cognitive ability and have profound implications for the efficiency of different policies that aim at fostering human capital. Taken from Cunha et al. (2006), the curve in Figure 2.1 summarizes the theoretical evidence on the rate of return to investment at different stages of the life cycle. The horizontal axis represents age, which is a proxy for an individual's stage in the life cycle of skill formation. The vertical axis represents the rate of return to investment assuming the same amount of investment is made at each age. The figure demonstrates that there is a higher rate of return at younger ages for equal amount of investment across the individual's years of life. The main idea of the figure is that learning is easier in early childhood than later in life, and cognitive stimulation early in life are critical for long-term skill development (Shonkoff et al., 2000; Cunha and Heckman, 2003). However, the economic argument for early investment does not preclude later investment; rather it argues that there are dynamic complementarities to be gained from investing at different stages of the life cycle that makes skill formation is a multi-stage process (Cunha et al., 2006; Orla Doyle et al., 2010).

Figure 2.1: Rate of return to human capital investment

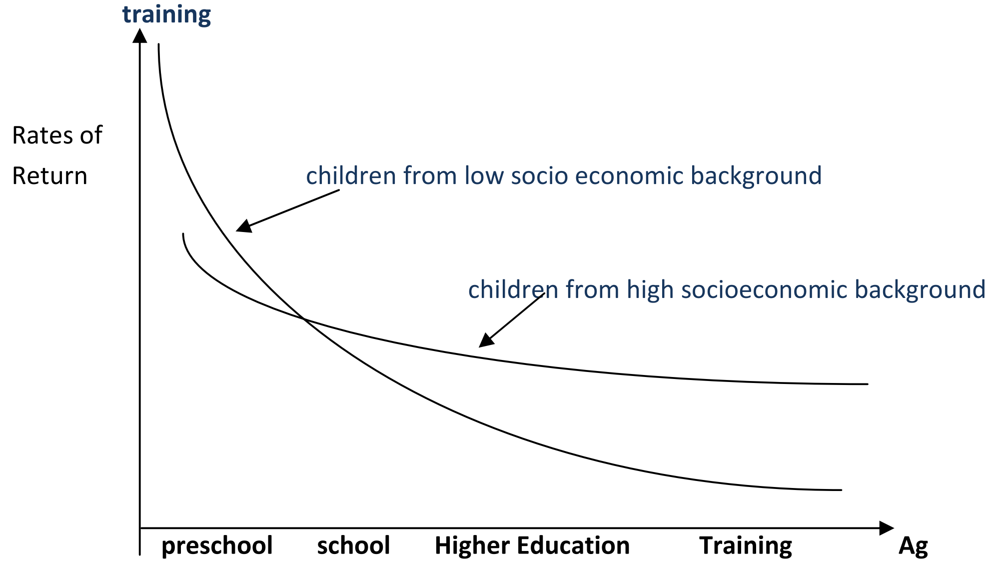


Rates of return to human capital investment setting investment to be equal across all ages

Source: Cunha et al. (2006)

Furthermore, in the education sector it is widely argued that there is always a trade-off between efficiency and equity objectives, and that one can only be achieved at the expense of the other. However, Eurydice network (2009) notes that, viewed in a wider perspective, equity and efficiency are in fact mutually reinforcing in early childhood education. It is both more efficient and more equitable to invest in education very early: correcting failure later on is not merely inequitable, but highly inefficient in comparison. This is so not only because early childhood education facilitates later learning, but also because it can produce large socio-economic returns, especially for disadvantaged children (Heckman, 1999, Eurydice network, 2009). Currie (2001) additionally suggests that it may be more effective for a government to equalize initial endowments through early childhood development programs than to compensate for differences in outcomes later in life. And she hypothesizes that families may under-invest in early childhood because of market failures such as liquidity constraints and information failures. Hence, public investments in quality early childhood education can produce important long-term improvements in the intellectual and social development of disadvantaged children. Supporting this concept, Cunha et al. (2008) and Woessmann (2006) state that return to educational investments tend to be highest for children from disadvantaged families at early stages and for the well-off at late stages of the life cycle. Taken from Woessmann (2006), the two curves, in Figure 2.2, illustrate that returns to early interventions are particularly high for children from disadvantaged backgrounds whose homes do not provide them with the foundation of skills necessary to prosper at later educational stages, but at older ages, they tend to be higher for children from well-off families. This implies that in the absence of public involvement, rates of return turn down more rapidly for children from low than from high socio-economic background. This in turn reveals complementarities between efficiency and equity at early stages and trade-offs at late stages

Figure 2.2: Stylized returns for a dollar spent at different stage of education and training



Source: Woessmann (2006), (originally adapted from Cunha et al. (2006))

In fact, beyond the human capital argument, there is growing recognition that all reasonable plans for human development begin early with measures to protect the rights of the child rather than wait for 18 years later. The UN Convention on the Rights of the Child (United Nations, 1989) has had a long history of supporting and encouraging interventions aimed at children and families. The human rights argument holds that children have a right to live and develop to their fullest potential. In line with this argument, organizations such as UNESCO and UNICEF strongly encourage investment in integrated early childhood programs as allowing children to live in environments without adequate health and well-being as well as the right to free education directed at the full development of the human person is a violation of the most basic human rights.

Moreover, in addition to the human capital argument, there is a strong justification from the human rights side, which remarkably argues that the responsibility to protect a child's human rights is the most fundamental reason to invest in programs that enhance early childhood development (Myer, 1995; Woodhead, 2009).

2.2 A review of Previous Studies: Experience of Other Countries

Huge body of empirical works that demonstrate the effects of early childhood education on a child's cognitive; language and social development are found mainly in North America and Western Europe, particularly in the United States. Still, only little has been learned about the relationship between early childhood education and child outcomes in settings outside of the developed countries although comparable research is being done in a growing number of developing countries (UNESCO, 2010). As the result, to reach in meaningful concluding remarks of the present study, it would be imperative to draw some lessons (experiences) from the developed; and if any from the developing countries as well.

In light of these facts, it would be sensible to start the review from the USA as it has rich experiences of preschool interventions. For instance, using data from the early childhood longitudinal study, Susanna et al. (2005) examines the influence of preschool centres on children's cognitive and social development. To identify the effects, the authors employed OLS, instrumental variables (IV) and matching estimates. The main findings indicate that preschool centres raise reading and math scores, but has a negative effect for socio-behavioural measures. More specifically, the duration of centre-based care matters: the greatest academic benefit is found for those children who start at ages 2–3 rather than at younger or older ages; negative behavioural effects are greater the younger the start age. These patterns are found across the distributions of family income. The intensity of centre-based care also matters: more hours per day lead to greater academic benefits, but increased behavioral consequences. Similarly, Magnuson (2004), using nationally representative longitudinal data of US children who joined pre-primary school in 1998-99 and finished first grade in 2000, analyses whether prekindergarten increases school readiness at kindergarten entry and whether any of the effects last long. In the analysis, controlled by various socioeconomic variables and state dummies, prekindergarten attendance is instrumented by state expenditures spent on early childhood education and care. The findings from IV estimator and matching regression indicate that prekindergarten increases math and reading skills at kindergarten entry, but are also associated with an increase in behavioural problems. Furthermore, while the cognitive gains largely fade out by the spring of first grade, the negative effects on class-room behaviour do not. In fact, the largest and most lasting academic gains were found for disadvantaged children.

Campbell et al. (2002) also examine the effect of Abecedarian Project, which is a scientific experimental centre, on children cognitive and non-cognitive behaviour by considering data for the years between 1972 and 1977. The result of OLS and matching regression shows that the treated children on average have relatively higher cognitive ability, which is captured by test scores, than untreated children. The treated children got higher results on reading and mathematical skills and attend more years of schooling compared to untreated and even they are more likely to join college or university. Additionally, the result of this study shows that the difference in cognitive ability between the treated and untreated is wider for girls than boys. Looking at non-cognitive skills, the result shows that teen pregnancy is greatly reduced though it was not possible to avoid it totally. Furthermore, using the Chicago Longitudinal Study (CLS), Reynolds et al (2000) study the impact of public preschools between 1983 and 1985 on children development outcomes. The study contained 1150 young children living in poverty. The centres provided services for children between the ages of 3 and 9, ensuring a stable transition from preschool to early elementary school. The children were enrolled for varying lengths of time, which allowed the researchers to examine the long-term effects of differing levels of participation beginning at different ages. The comparison group of 380 children was randomly selected from selected schools in poor neighbourhoods. The preschool program was a half-day experience and primary grades were full day. The program had no set curriculum, but included a structured set of activities that promote basic math, language and reading skills and encouraging psychological and social development. The study followed the participants through age 21. Results of the study demonstrate that children who were involved in the program had higher reading and math scores during adolescence than those who had not participated. Children who were involved in the program had experienced lower retention rates and lower special education placement by age 20. Children who participated in full-day kindergarten intervention had significantly lower rates of special education and grade retention. Cost-benefit analysis of the program indicates that every dollar invested in the preschool program returned $7.14 in education, social welfare, and socioeconomic benefits by reducing public expenditures for remedial education, criminal justice treatment, and crime victims.

Taking any early schooling (before the compulsory starting age of 5) and of preschool on a cohort of British children born in 1958, Goodman et al. (2005) assess whether any effects on cognition and socialization are long-lasting, as well as their net impact on subsequent educational attainment and labour market performance. Employing fully interacted linear model and matching method (the so-called average effect of treatment on the treated – ATT); and controlling for a particularly rich set of child, parental, family and neighbourhood characteristics, early education is found to have positive and long-lasting effects. Specifically, pre-compulsory education (preschool or school entry prior to age 5) was found to yield large improvements in cognitive tests at age 7and remained significant throughout the schooling years, up to age 16. The effects on socialization appear to be more mixed, with adverse behavioural effects from parental reports at age 7 persisting. In adulthood, pre-compulsory schooling was found to increase the probability of obtaining qualifications and to be employed at 33.

Based on a sample of 8,500 children with additional information merged in from a census of all preschool institutions, Cleveland and Krashinsky (1998) also separately assess the effects of preschool education on cognitive and behavioural development at age 5 and 10 of several different types of ordinary preschool programmes. Based on 'analyses of variance' controlling for a number of important socio-economic and family factors, they find that preschool generally boosts cognitive attainment at ages 5 and 10. In terms of problem behaviour, preschool attendance was found to have no effect at age 5 but to increase some types of behavioural problems at 10, in particular conduct disorder, although the latter associations were relatively weak. The study also found weak evidence for the benefits of nursery education being slightly greater for socially disadvantaged children, although this difference was small compared to the general benefit of preschool for all children.

In the same country, UK, Sammons et al. (2003) conducted a large-scale study following children (2-year-olds or more) attending preschool. At school entry they have better cognitive outcomes (pre-reading, early number, and language) and superior social and behavioural skills than their peers without preschool experiences; longer preschool attendance leads to higher cognitive gains when entering school; and the cognitive gains of attending preschool are larger for disadvantaged children. Additionally, Claessens et al. (2006) tested how school age skills are related to both the achievement and non-cognitive skills that children bring to kindergarten. The study used sample data from Early Childhood Longitudinal Study-Kindergarten Cohort (ECLS-K) for children who were in kindergarten in 1998 and 1999. From the OLS regression result, children's skill in math, reading and attention skills at entry of kindergarten are not predictive of subsequent math achievements. Similarly, socio-emotional skills except attention skill are not predictive of fifth grade math and reading skills. As the theoretical argument, it is found that the gains in reading and math scores in the years of pre-primary education are indicative of future reading and math achievements. In the same way, Carneiro, Crawford and Goodman (2006) investigated the determinants and consequences of cognitive skills and social adjustment of children in Britain, which is one part of non-cognitive skill. In the study, cognitive and non-cognitive skill is analyzed at age 7 and 11 of the children. The standardized test score in maths, reading, copying and drawing, are used to measure cognitive skill. The result showed that children from professional and managerial social classes have higher cognitive and non-cognitive skills at age 7. Moreover, the parents education, the interest of parents in their children education and how much the parents reads also have positive contribution to the development of cognitive skills.

However, unlike the above effects, research findings from Canada report negligible effects. Lefebvre et al. (2000) investigate the relationship between child care arrangements and developmental outcomes of young children using data from Cycle 1 of the Canadian National Longitudinal Survey of Children and Youth. Models of the determinants of Motor and Social Development (MSD) scores for children aged 0-47 months, and of the Peabody Picture Vocabulary Test assessment scores (PPVT-R) for children aged 4-5 years are estimated controlling for a variety of non-parental childcare and early education characteristics. The results suggest that infant-toddler non-parental care arrangements have insignificant or negligible impacts on developmental outcomes (MSD). For pre-schoolers, modes of care and early education do not, on average, influence cognitive development (PPVT). The results of fixed effect estimates for a sample of siblings aged 0-47 months confirm the preceding conclusion. The analysis is repeated to identify the determinants of the probability the child's MSD (PPVT) score is in the bottom part of the distribution of MSD (PPVT) scores and the conclusions are similar.

Passing to other country's experience, we get New Zealand. Based on longitudinal study of the Competent Children Project, Wylie et al. (2003) analyse the effect of early childhood educational experiences on cognitive, social, communication, and problem solving competencies. The study also examined the effect of family resources, home activities, and engagement in school on these same items. To do so, since 1993, the researchers have collected developmental data on over 500 children and the children are assessed when they are 12 and 14 years of age. Results from the study indicate that starting age of early childhood education was significant in affecting cognitive skills. Children who started before age 2 had higher scores in mathematics, curiosity and reading comprehension. Additionally, parental education levels and family income affected children's literacy and mathematic scores positively. A study in Nepal (2001) also showed that investing in early childhood resulted in halving the dropout rate and grade repetitions reduced to less than one fifth of the former rate. In Brazil, there were dramatic increases in grade completion rate, from 2% to 40% as a result of a community-based early childhood programme. A programme in Brazil that focuses on including children in good quality programmes points out that a child in preschool costs no more than $ 100, a child on the street costs $ 200 and a child in the penal system $ 1000. By improving children's chances of success in school, early childhood education, has the potential of addressing disadvantage squarely.

A well-known study from Jamaica shows that children randomly assigned to receive home-based early stimulation have substantial improvements in cognitive development and subsequent school performance (Powell et al. 2004). Similarly, in Argentina, Using methods of intention-to-treat and treatment-on-the-treated effect, Berlinski et al. (2009) found that expansion of supply of preschool education has increased the attendance of preschool education and showed positive effect on the cognitive scores and non-cognitive skills. Meaning, preschool education has increased the score of the students and improved classroom attention, effort, discipline and participation. In fact, for the same country, developing within household estimator and instrumenting preschool attendance by locality of residence and birth group, (Berlinski et al. (2008) also show that preschool attendance has positive impact on the children's completion years of primary and secondary education, accompanied by low dropout and repetition rate in each grade for the treated compared to untreated children.

When come to Africa, related research activities are very scant. Very few studies have been conducted to assess the effects of preschool attendance on cognitive development of children. For instance, Peter Glicket et al. (2007) estimate the determinants of cognitive ability among 14 to 17 year olds in Senegal. Unlike standard school-based samples, tests were administered to current students as well as to children no longer or never enrolled. Result of the study indicates that years of schooling strongly affects cognitive skills, but conditional on years of school, parental education and household wealth, as well as local public school quality, have surprisingly modest effects on test performance. Instead, family background primarily affects skills indirectly through its impacts on years of schooling. Therefore closing the schooling gaps between poor and wealthy children will also close most of the gap in cognitive skills between these groups.

Crossing from western Africa to the east of the continent, we find a study conducted by Malmberg et al. (2011), within the context of East-African preschool policy; they investigate the effects of the Madrasa Resource Centre (MRC), a child-cantered intervention program, on East-African (Kenya, Zanzibar, and Uganda) preschool children's cognitive development. Altogether 321 children (153 non-interventions and 168 interventions) participated in a cross-sequential study over three time-points during preschool (mean ages 4.3, 6.0, and 7.1 years). A multilevel model (MLM; time-points nested within children nested within schools), in which time was coded flexibly (i.e., child's age operationalized as months from start of the intervention), showed a beneficial curvilinear effect of the intervention program on children's cognitive gains. A moderation analysis suggested that the effect of observed preschool quality (ECERS) was stronger in the intervention program.

Coming to the target country, Ethiopia, as far as our knowledge, very few studies have been conducted related to the early childhood education. Using the Young Lives Longitudinal data, Woldehanna et al. (2008) estimate correlates of preschool enrolment for the 5-year-old children from a probit model. The findings of this study indicate that while parents' educational level, Muslim households and long-term health problems have positive and statistically significant association with the probability of a child being enrolled in preschool, sibling order is found to have negative association. Yet, no statistically significant association is found for two other independent variables, namely child's gender and wealth index of household in Round 1. Beyond this simple correlates, there are no studies that highlight the effects of early childhood education on cognitive development and subsequent school participation in Ethiopia As the result, our empirical review of the sector is limited only to this simple correlates estimation. Nevertheless, before we proceed to the empirics of question, we still find it imperative to examine the general trend and structure of the Ethiopian early childhood education sector for data acquired from the Ministry of Education (MoE).

In summary our review of literature reveals that early childhood education among other investments such as early childhood nutrition and care is an investment with a good return in the future. Early childhood education increases the cognitive outcome of children and makes the ready to receive well when they enter primary education. The return to such investment would be even higher for children from lower socio-economic background. Many of the Ethiopian children are disadvantaged in the sense that their socio-economic background is very low stunted (MoFED, 2008): more than 40% of the Ethiopian children are stunted, and 38% of the Ethiopian children are living with parents living in absolute poverty and unable to consume sufficient calories required to perform normal activity. Given this situation, preschool education could result a huge benefit to children in improving their cognitive development.

2.3 An Overview of Early Childhood Education in Ethiopia

Formally, compulsory education in Ethiopia starts at age seven in primary schools. Nevertheless, children can join pre-primary schools between age three to six depending on the availability of the program in their areas. As it is already explained in section one, early childhood education is structured in the form of kindergartens and predominantly provided by the private sector, Non-Governmental Organizations (NGO), communities and faith-based organizations. The government has very limited intervention on this regard. To be precise, in its 2007 report, the Ministry of Education (MOE) states that the government does not run preschool education program essentially for two main reasons. While one explanation, as stated in the document, is to enhance the involvement of the private sector in the education sector, the second justification is to maximize the government's effort in the other levels of the sector. As a result of this government's limited intervention, enrolment rate for preschool education has remained very low and especially absent in most rural areas of the country.

In light of this fact, this subsection provides an overview of the general trends of early childhood education in Ethiopia. To do so, Table2.1 depicts gross enrolment rates by gender, preschool population, enrolment by level, and number of kindergartens across time in the country. For instance, in 2008/09, out of the estimated 6.95 million children of the appropriate age group, only about 4.2 per cent of the children have been reported to have access to pre-primary education in 2904 kindergarten sites all over the country.

In fact, when we examine Table 2.1 critically, though enrolment is small when compared to the relevant age group, it has been growing since 2003/04 at an average of about17.5% per year for five years. For this reason, we see that Goss Enrolment Rate (defined as the percentage of total number of children in kindergarten, irrespective of age, out of total population of relevant age groups) for kindergarten in 2008/09 was 4.2% which is 0.3% higher than its previous year. Similarly, given the small number of kindergartens, the number of kindergartens has shown an increasing trend from year to year. For example in 2000/01, there were only 964 kindergartens. But, in 2008/09, this number has jumped to 2,904 with average annual growth rate of 23.02 %. This comparative rise in number of preschool indicates an ever increasing involvement of the private sector in this stage of education as the government has done nothing at this level of education.

Table 2.1: Trend of Early Childhood Education in Ethiopia

Year	Gross enrolment rate (%)			Number of kindergartens
	Boys	Girls	Total	
2000/1	2	2	2	964
2001/2	2.1	2	2.1	1189
2002/3	2	2	2	1067
2003/4	2.2	2.1	2.2	1244
2004/5	2.4	2.3	2.3	1497
2005/6	2.8	2.6	2.7	1794
2006/7	3.2	3.1	3.1	2313
2007/8	3.9	3.9	3.9	2740
2008/9	4.2	4.2	4.2	2904

Source: Author's compilation based on MOE 2001 to 2009 statistical abstract

Despite the increasing trend of the enrolment rate at the national level, regional and residence (urban versus rural) differences remain significant. In actual fact in a country as large and diverse as Ethiopia, differences among regions and residence are to be expected, but the difference is big enough to require special emphasis. Given such wide disparity, the national average of early childhood gross enrolment ratio is to some extent deceiving. As we can see from Table 2.2, with the exception of Addis Ababa, enrolment rate of children of the relevant age group has remained to be at its lowest stage. In Addis Ababa, for instance, enrolment rate has shown an impressive improvement from about 36% in 2001/02 to 47% in 2006/07 and further to 73.8% in 20008/09. Next to the capital city, enrolment rate in Harari and Dire Dawa is higher compared to the other areas. However, for all the remaining regions (Tigray, Afar, Amhara, Oromiya, Somali, Benishangul-Gumuz, SNNPR and Gambella) enrolment rates were less than the national average gross enrolment rate over the years 2000/01 to 2008/09. Especially, for the two previously most underserved regions, Afar and Somali, the preschool gross enrolment rate has never been above 0.6%. The basic reason for the low preschool attainments in these two regions is the fact that most of the people in those regions are pastoral and semi-pastoral communities in which community of the regions are highly movable from one place to another across seasons in a year.

Generally, despite the fact that the government has been doing little in this critical stage of education, the gross enrolment rate of the subsector has been growing since 2004/05 at approximately 17.5% per year. However, this enrolment growth is basically owing to the involvement of the private sector, mainly in urban parts of the country. For instance, it is apparent to see that large share of the gross enrolment rates are taken by Addis Ababa, Harari and Dire Dawa, which are dominantly urban areas. At the same time the Ministry of Education (MoE, 2008/09) reports that in urban areas of the country, there is a large increase in parents' desires to have their children attend Kindergartens. All those points, directly or indirectly, indicate that there is no as such expansion of early childhood education in rural areas of the country. That is why we have limited the scope of the study mainly to urban areas of the country. However, hope is given that the implication of the urban results will serve for relevant educational policy insight for the whole country with respect to this fundamental stage of education.

Table 2.2: Gross enrolment rate (%) across regions and overt time

Regions	2000/1	2001/2	2002/3	2003/4	2004/5	2005/6	2006/7	2007/8	2008/9
Tigray	2.2	2.1	2.0	2.1	2.2	2.1	2.3	1.0	1.0
Afar	0.6	0.6	0.6	0.6	0.5	0.5	0.5	0.5	0.6
Amhara	1.3	1.2	1.2	1.3	1.4	1.3	1.5	2. 0	22
Oromiya	1.4	1.4	1.4	1.7	1.6	1.8	2.3	3.0	3.4
Somalia	0.3	0.7	0.6	0.6	0.6	0.6	0.6	0.6	0.6
B.Gumuz	1.8	1.9	1.8	2.9	2.8	2.7	3.9	3.4	3.7
SNNPR	1.5	1.5	1.5	1.6	2.0	2.7	2.6	3.0	3.5
Gambella	0.8	1.1	1.1	1.0	3.7	1.4	0.7	3.1	2.4
Harari	12	8.6	7.9	7.2	6.6	6.1	11.0	12.6	11.4
Addis Ababa	35.7	36.4	33.1	31.6	32.9	40.3	46.7	74.9	73.8
Dire Dawa	9.7	10	9.1	11.6	10.6	3.5	7.9	10.7	19.7
Total	2.0	2.1	2.0	2.2	2.3	2.7	3.1	3.9	4.2

Source: Author's compilation based on MOE 2001 to 2008 statistical abstract

3. Methodology

3.1 Data Source and Measurements

The data used for the empirical analysis come from the Young Lives Longitudinal Survey in Ethiopia, which is part of the international project tracking the livelihoods of children in Peru, Vietnam, and the state of Andhra Pradesh (India). This longitudinal survey was conducted in two phases. While the first phase took place in 2002, the second phase was administered in 2006/07. In these two phases of data collection, the targeted children were six to 18 months old and 4.5 to 5.5 years old, respectively. These children are found in 20 sites (communities) of the five big regions of the country where more than 96% of the Ethiopian children live. In line with the focus on child poverty, there was an emphasis on poor rural and urban sites during the sampling. Between the two rounds, there was an attrition rate of 1.33%, which is low compared with other longitudinal studies (Woldehanna, 2008). Furthermore, the data covers a comprehensive range of topics including information on children’s and their families’ access to key services (education and health), work patterns and social relationships, as well as core economic indicators such as assets.

What we are looking for is to understand the degree of association between preschool enrolment (between ages three and five) and cognitive development test scores measured in round 2 (at the age of five). During the round-two survey, each parent was asked if the child was enrolled in preschool education when she or he was three years old. Then in round two cognitive development outcome was measured when children were at around the age of five (in round 2) using Peabody Picture Vocabulary Test (PPVT) and Cognitive Development Assessment – Quantity (CDA-Q) Test. While the former is a test of vocabulary recognition that has been widely used as a general measure of cognitive achievement, the latter is a common test used in assessing cognitive development of young children. More precisely, PPVT is a test of receptive vocabulary adaptable according to age. In PPVT test, a child hears a word (‘boat’, ‘lamp’, ‘cow’, ‘goat’ etc) and is then asked to identify which of four figures corresponds with the spoken word. After the test was administered, the PPVT score was built based on the datasets provided by Young Lives Study. Likewise, CDA-Q test which essentially measures children’s understanding of quantity-related concepts was built by Young Lives Study. (see Cueto et al., 2009).

Additionally, to achieve better measures of the cognitive development, the raw score tests were adjusted in to standardized scores. For instance, the PPVT raw scores were adjusted in order to corrected PPVT raw score, This is undertaken by correcting errors in test administration e.g. (1) eliminating items above the ceiling that could not be administered; (2) taking the highest item reached by the child; and (3) subtracting lower items that were answered incorrectly and also lower items with poor statistical behaviour. In the same manner, corrected raw score of Cognitive Developmental Assessment (CDA) test was developed by Young Lives Study (Cueto et al., 2009). Except items eliminated due to poor statistical behaviour, the score was built as the sum of all items (1 for correct, 0 for blank or incorrect). In fact, to make language-comparable, both PPVT and CDA-Q test raw scores were further standardized by language spoken as there was language diversity in undertaking the longitudinal survey across the country. In short, to materialize the objectives of this study, alternative forms of PPVT and CDA-Q tests were developed as proxy measures of cognitive development of preschool (five-years-old) children from the second round survey.

3.2 Econometric Model Estimation

We start by specifying cognitive development, measured by PPVT or CDA-Q test scores, to be a function of the child's preschool education attendance, and a set of child and household characteristics that mainly influence those test scores. In other words, although our main interest is to evaluate the effects of early childhood education on cognitive development outcomes, we must acknowledge the fact that those outcomes are greatly subjected to various factors beyond early childhood education. For this reason, we introduce a set of socioeconomic control variables to improve the robustness of the results. For instance, nutritional status of early childhood is one of the most important determinants of cognitive development and achievement in subsequent schools. To do so, we have included, Z-score of height-for-age, measured at round one of the longitudinal survey. The key insight is that a child's learning productivity at preschool seems to be partially determined by parental investments in health and nutrition during infancy. To fix the idea, Vector *X* in equation (1) includes: (a) nutritional status of children in round one; (b) wealth index of household in round 1; (c) household demographic composition and dependency ratio in round 1. Indeed, community-fixed effects are also included to reduce bias from characteristics common to children within the same community.

Accordingly, the basis of our estimation strategy can be summarized by the following education production function.

$$Log(Y_{ij}) = \beta_j + \beta X_{1ij} + \beta_{2i} Z_{ij} + U_{ij} \quad \textbf{(1)}$$

where Y_i is child i's cognitive development, measured by the test score of interest, living in community j; X_{ij} is a dummy variable whether child 'i' is enrolled in preschool or kindergarten school (KG) at age three or above, Z_i is other (confounding) factors affecting child 'i' cognitive development since birth, which basically include, as stated above, nutritional status, household composition, wealth and child characteristics; β_j is a community fixed effect; and U_i is a disturbance term

However, estimating equation (1) by ordinary least squares (OLS) is likely to lead to biased estimates of the parameters of interest as many of the inputs we consider are possibly correlated with household unobservable characteristics such as household or child specific effect and to other factors of early childhood experiences. Also, no matter how comprehensive our list of inputs, it is possible that there are omitted terms in equation (1). All of these considerations give rise to serious concerns about the exogeneity of the input measures; that is there are good reasons to suspect that some of the explanatory variables specifically preschool enrolment is correlated with the error term. For instance, given the above equation, preschool education attendance is highly influenced by parents' preference for child quality. Such kind of preference makes preschool education attendance to be interpreted as a parent's and/or community's characteristics. Hence preschool attendance should not be included as an exogenous variable, as it is clearly endogenous variable in many ways. This endogeneity, $E(x_{ij},U_ij) \neq 0$, makes OLS estimates biased, which indeed needs instrumental variables. The same is also true for nutritional status as it could be endogenous variable in the specified model. Hence, in order to deal with this problem we need to impose further assumptions to equation (1).In particular, we will assume that

$$U_{ij} = V_{ij} + \varepsilon_{ij} \quad \textbf{(2)}$$

where (V_{it}) captures child and household specific unobserved factors and ε_{ij} is white noise.

Equation (2) really indicates that we have to use instrumental variable (IV) estimation method where we have to obtain valid instruments for the suspected endogenous variables. Although it is hard to find relevant and valid instruments for a given endogenous variable, there are two conditions to be satisfied for the instruments to be appropriate. First, the instruments must be relevant (strongly correlated with the endogenous variable), which is testable via first-stage F tests. Second, instruments must be valid (uncorrelated with the error terms), conditional on other explanatory variables which could be verified by the specification test or test of over identification provided we have more than two instruments for one endogenous variable. Hence, based on these criteria, we have chosen community dummies for preschool enrolment and pre and post natal shocks that affects household welfare negatively for nutritional status. We tried to use distance to preschools as instrument for Preschool enrolment, but it is insignificantly correlated with the enrolment in preschool due to lack of variation within the sample.

Hence, for the IV or two-stage least squares (2SLS) estimation, the first-stage (reduced form) equations are stated as:

$$KG_{ij}= B_1CD_j + Shocks_{ij}+ B_3X_{ij} + \alpha_{2j} + \varepsilon_{ij}. \quad \mathbf{(3)}$$

where CD cluster dummies and Shocks are pre and post natal shocks that affects the household welfare negatively. Similarly, for other endogenous variable it becomes.

$$Z\text{-}HFAR1_{ij} = {}^{\gamma}{}_{1S}hocks_{ij}+ \gamma_2CD_j + \gamma_3X_{ij} +\alpha_{2j} + \varepsilon_{ij} \quad \mathbf{(4)},$$

where$Z\text{-}HFAR1_{ij}$ *is* z-score of height for age in round 1.

Following these estimations, two- stage-least squares (2SLS) estimates are estimated using generalized method of moment (GMM) with standard errors computed after adjusting for the general form of heteroscedasticity, or robust standard errors adjusted for the clustering of data at the community level and with additional correction for the two-stage estimation process.

Moreover, we have used alternative instrumental variable estimation method known as Wooldridge approach in which propensity score or predicted probability of preschool enrolment is used as an instrument for preschool enrolment (Wooldridge (2002, p.623). In this case we first run a probit model of preschool enrolment where cluster/site dummies and pre and post natal shocks as well as other factors that affect the cognitive outcome are included. Then we predicted the probability of preschool enrolment propensity score of preschool enrolment. In the second stage regression, we used the predicted probability of preschool enrolment as instrument for preschool while we exclude cluster/site dummies and pre and post natal shocks in order to identify the parameters. According to Wooldridge (2002, p.623), this estimator provides efficient estimates. In both instrumental variable estimations, we tested the relevance via first-stage F tests and validity of instruments via Hansen J test or over identification test.

In both the OLS and IV estimation methods, the explanatory variables used to explain cognitive outcome of children at the age of five are preschool enrolment, Z-score of height for age at the age of one, age of the child in months during round one survey, sex of the child (male dummy), wealth index in round one, household composition, highest grade completed by caregiver and father of the child, number of months the child breastfed, birth weight measured by five point Likert scale for the relative size of birth weight (-2 to 2), and dummy for child had health problems at the age of one year. Among these explanatory variables, as discussed above, preschool enrolment and Z-score of height for age at the age of one are the endogenous variables. In the standard IV estimation, urban site dummies and pre-natal and post-natal shocks that affect household welfare negatively are used as instrumental variables for preschool enrolment and Z-score of height for age. In the Wooldridge IV estimation, the predicted probability of preschool enrolment obtained from the propensity score estimation for the matching regression is used as instrument for preschool enrolment.

3.3 Propensity Score Matching

The accuracy of estimating impact of pre-school attendance on cognitive development of children from an econometric model largely depends on the availability of instruments that satisfy relevance and validity conditions. Since it is very difficult to test appropriate so as to ensure validity of instruments used for pre-school attendance, we also used propensity score matching techniques to assess the impact of pre-school on cognitive development and substantiate the OSL and IV estimation results.

If we want to infer causality, selecting an appropriate comparison group through propensity score matching also offers an alternative way to obtain comparable results and (compared to OLS and IV), which requires some assumptions about the "correct" functional form. The matching assumptions ensure that the only remaining relevant difference between the two groups is program participation (preschool education attendance, in this case) provided that the differences can be captured by the observables and there is no individual effect. However, since we have a lot of observables including what happened to mothers and the households before the birth of child and the child and household characteristics before the child was enrolled in preschool. Therefore, we can capture much of the differences among the children and hence we can match the treated (preschool attendee with the untreated (non-preschool children) in a better way than many studies relying on cross-section data sets. The propensity score analysis proceeds in two steps. First, we estimate a propensity score for each individual as the conditional probability (from a probit model) of attending kindergarten given the full set of covariates. The propensity score is next used to create a matched control group of children who did not attend preschool education. We use Kernel matching method and limit the sample to children for whom there is sufficient overlap in propensity scores between the kindergarten and comparison group (the area of common support). The robustness of the results was checked by using other matching techniques such as nearest neighbourhood matching, radius matching and one-to-one matching. If the matching process proceeds correctly, the treatment and control children will have similar measured characteristics and the effects of preschool education can be estimated by comparing the matched groups' means. That is, the average treatment effect on the treated (ATT), which is calculated as

$$ATT \equiv E(CD_{i1} - CD_{i0} \mid d_i = 1) \qquad \textbf{(6)}$$

Where d_i is preschool enrolment dummy which is 1 if enrolled in preschool education and 0 otherwise, CD_{i1} and CD_{i0} are cognitive outcomes, with CD_{i1} the score of outcome that would be observed if the child attended preschool education and CD_{i0} the outcome score observed on the same age if the child did not attend preschool education. Child's participation equation is specified as

$$d_i = \alpha_1 X + \varepsilon_i \quad \textbf{(7)}$$

Where d is dummy variable for a child's participation in preschool, while *X* are variables that affect both the participation in pre-school and cognitive development outcomes.

Implementing matching requires choosing a set of variables covariates, *x*, that credibly satisfy the condition that the outcome variable (cognitive development in our case) must be independent of treatment (pre-school enrolment) conditional on the propensity score. Heckman, Ichimura, and Todd (1997) also show that omitting important variables can seriously increase bias in resulting estimates. Moreover, there is no guideline on how to choose conditioning variables, *x*, (Smith and Todd, 2005), hence selection of *x* variables intuitively is very important and hence these covariate must include variables that affect both the outcome and the participation in programs. Smith and Todd (2005) or Sianesi (2004) advised to have a deeper knowledge of setup of the program (in our case) pre-school education is important in order to select variables to be included in the probit model of propensity score.

In the estimation of propensity score of preschool enrolment, the explanatory variables used are those that affect both the preschool enrolment (participation of children in preschool) and the cognitive outcomes. These variables include urban site dummies (one site in Addis Ababa a comparison group), pre-natal and post-natal economic shocks that affect household welfare negatively, Z-score of height for age at the age of one, age of the child in months during round one survey, sex of the child (male dummy), wealth index in round one, household composition, highest grade completed by caregiver and father of the child, number of months the child breastfed, birth weight measured by five point Likert scale for the relative size of birth weight (-2 to 2), and dummy for child had health problems at the age of one year. It is the predicted probability of preschool enrolment from this model that we use as instrumental variable for preschool in the estimation of Wooldridge IV.

4. Descriptive Statistics

Although we have presented the general trend of the of early childhood education in previous sections using secondary data acquired from the Ministry of Education, before we proceed to the estimation part, it is equally important to reconsider the pattern of preschool education directly from the Young Lives Longitudinal Data. Accordingly, Table 4.1 presents summary statistics of key variables. Out of 1912 preschool age children included in the survey, 762 are from urban areas while the rest 1150 are from rural part of the country. With respect to kindergarten enrolment, 57% of the urban preschool age children have been enrolled in kindergarten since the age of three or four. However, this participation rate is overestimated due to exclusion of the rural areas. If we have a close look at the rural children, only 3% of them have been attending preschool education. This implies that out of total sample of 1912 preschool age children, only 25% of them have been attending preschool education in the country. This obviously indicates preschool is an urban experience in Ethiopia. Off course, even in the urban areas, as it is predominantly provided by the private sector, it is unequally distributed. This means that only some privileged children have access to this fundamental stage of education. For instance, if we look for an issue -who runs the preschool, out of the total enrolled children, 71.10% of them have been attending their early childhood education in private kindergarten, whereas the rest 28.90% have been attending their education in kindergartens owned by community, public and others. This clearly indicates that the subsector is dominated by fee charging kindergartens in which children from socio economic disadvantaged background do have only little opportunity to attend this fist and critical stage of education. In light of this fact, since the opportunity to attend preschool is almost entirely restricted to urban children and privileged ones, the analyses of this study, mainly represent for the urban areas of the country.

Coming to the cognitive outcomes, Table 4.2 reports various outcomes: PPVT, CDA-Q test scores, these test scores adjusted for language difference among children, log of PPVT, log of PPVT standardized by language; log of CDA-Q test; log of CDA-Q test adjusted for language differences. In all measures children enrolled in preschool education scored higher than those who did not enrol in preschool. We will see later if this difference in cognitive development test score persists when we control of other factors and compare the preschool enrolled children with matched non-preschool children.

Table 4.1: .Summary statistics of test scores and household characteristics for urban children

	N	Mean	S.D.	Min	Max
Raw score of PPVT test in round 2	739	26.79	14.85	3.00	121.00
Logarithm of PPVT test score in round 2	739	3.17	0.47	1.10	4.80
Standardized core of PPVT test in round 2	741	76.25	21.14	40.00	160.00
Logarithm of Standardized score of PPVT test in round 2	741	4.30	0.26	3.69	5.08
Raw score in the CDA test in round 2	742	9.44	2.73	0.00	14.00
Logarithm of Raw score in the CDA test in round 2	740	2.20	0.35	0.69	2.64
% of math questions correctly answered in round 2	742	62.93	18.21	0.00	93.33
Logarithm of % of math questions correctly answered in round 2	740	4.09	0.35	2.59	4.54
Raw score in the PPVT test standardized by language in round 2	497	318.10	50.20	193.10	519.00
Raw score in the CDA test standardized by language in round 2	658	315.30	48.20	115.30	471.60
Logarithm of Raw score in the PPVT test in round 2 standardized by language	498	5.75	0.15	5.26	6.25
Logarithm of Raw score in the CDA test standardized by language in round 2	658	5.74	0.16	4.75	6.16
Dummy variable for child being enrolled in preschool	745	0.57	0.50	0.00	1.00
Z-score of height-for-age at the age of one year	725	-0.94	1.77	-6.04	8.17
age of child in month in round 1	745	12.33	3.58	6.10	18.23
Wealth index for 1-year-olds (Round 1)	745	0.33	0.15	0.01	0.76
Dummy for male	745	0.52	0.50	0.00	1.00
Number of children below 7 and above 65 years old	745	1.52	0.70	1.00	5.00
Number of children between 7 and 17 years old	745	1.30	1.34	0.00	8.00
Number of male family members > 17 and less than 65 years	745	1.17	0.82	0.00	5.00
Number of female family members > 17 and less than 65 years	745	1.41	0.82	0.00	6.00
What is the highest grade completed by primary caregiver?	745	4.98	4.30	0.00	14.00
What is the highest grade completed by father?	745	6.59	4.75	0.00	14.00
Number of months a child was breastfeed	745	28.94	13.52	0.00	36.00
Five point Likert scale for the relative size of child when	745	0.06	1.07	-2.00	2.00
Dummy for child had health problems at the age of one year	745	0.43	0.50	0.00	1.00

With respect to nutritional status, the mean values for rural and urban children are 1.46 and 0.95 below the standard deviation of the median's Z-score height –for age, respectively. Here, height for age represents the long term effect of malnutrition. This indicator is sometimes described as the nutrition poverty height. The standard for WHO (international standard) is that any child with -2Z (below) is stunted in terms of height and are said to be chronically malnourished. Accordingly, though the mean of Z scores are a bit above -2, this does not mean that most of the children are not stunted. For instance, the minimum combined average of the two residences is rather about 7.68 below the median's Z-score height –for age with that standard deviation of 1.96. This indicates that there is wide spread of severe malnutrition among the preschool age children in both rural and urban part of the country. Similarly, concerning the demographic composition of the sample, the sex ratio of the sample is a bit taken by male to the extent of 53%. It is also apparent to see that the visited communities are dominated by high dependency ratio. For instance, the mean values for the number of children below 7 and above 65 years old are 1.70 and 1.52 for rural and urban areas, respectively. This indicates that there is high ratio of dependents to the active force in each household of the visited communities.

Table 4.2. Test score for five years old children by preschool attendance

	Non preschool children	Preschool children
Raw score in the PPVT test	21.9	30.9
Raw score in the CDA test	8.3	10.4
standardize score in the CDA test by language spoken	297.6	326.5
standardize score in the PPVT test by language spoken	288.2	327.3

5. Estimation Results

5.1 OLS and IV Estimates

Because of the absence of preschools in many rural areas of Ethiopia, number of rural children enrolled in preschool are very few and hence we are obliged to estimate preschool enrolment model only for urban areas. Accordingly, Table 5.1 presents the estimated results from the basic OLS and instrumental variable estimation methods where the cognitive development score is unadjusted and adjusted for mother tong language differences among children. The explanatory variables used to explain cognitive outcome of children at the age of five are preschool enrolment, Z-score of height for age at the age of one, age of the child in months during round one survey, sex of the child (male dummy), wealth index in round one, household composition, highest grade completed by caregiver and father of the child, number of months the child breastfed, birth weight measured by five point Likert scale for the relative size of birth weight (-2 to 2), and dummy for child had health problems at the age of one year.

All estimates including OLS, standard IV estimates and Wooldridge IV estimates show that preschool attendance is positively and strongly associated with the cognitive scores though OLS provides lower estimates compared to the IV estimates. The IV estimates always have the same sign but are much larger than the corresponding OLS estimates. We know OLS is not an efficient estimator if we have endogenous variables in our model. While the Wooldridge IV estimator is a more efficient estimator, there is no way of testing for mis-specification or overidentification because we have only one instrument. The overidentification test suggests that the IV estimates where preschool is the only endogenous variable have specification problem hence we do not use these for interpretation of the results. We have reasonably better estimates in the standard IV estimates where both preschool enrolment and nutritional status of children are endogenous because it satisfies the relevance (based on first stage regression) and validity (based on overidentification test) tests (see Tables A5.4a, A5.4b).

More specifically in the Wooldridge IV estimates, *ceteris paribus,* five years-old-children who have attended preschool education scored 21.6% higher in PPVT and 15% higher in language adjusted PPVT than the non-preschool attendees of the same age. Similarly, five years-old-children who have attended preschool education scored 23.4% higher in cognitive development assessment - quantitative test- CDA-Q test and 6.5 % higher in language adjusted CDA-Q test than those who did not attend -preschool education of the same age.

The magnitude of the association is only slightly higher in the estimates of standard IV version where both preschool enrolment and nutritional status are endogenous. The other IV estimate where only preschool is endogenous is not much different either from other IV estimates indicating the robustness of the results. In all estimations, the association declines when the cognitive development test scores are adjusted for language differences perhaps because many of the differences in cognitive differences among children is due to relationship between test score and language used to administer the test.

It is important to mention the associations between cognitive outcomes of children and other confounding factors provided in Tables A5.4a and A5.4b in the appendix. In all the estimation methods discussed so far, we found household wealth and parents education (including the education level of both the father and the caregivers) to have strong association with cognitive development outcome measured by PPVT and CDA-Q test scores. With respect to nutritional status of children at age one, we found nutritional status is positively associated with test score measured by PPVT and CDA-Q test scores. This indicates that early nutrition status is an important correlate of subsequent cognitive development. However, such correlations become very weak and less statistically significant when the cognitive outcomes are adjusted for language. Tables A5.4a and A5.4b show the gender differential in cognitive development. In all estimation methods, we found that there is higher association between preschool attendance and cognitive development for girls than for boys. However, the association is only statistically significant in the IV estimates where both preschool and Z-score of height for age are endogenous indicating that unexplained gender differential cognitive outcome it is not robust. Moreover, it is apparent to see that there is no statistically significant association between demographic variables or household composition and cognitive development outcomes. This may indicate that number and composition of household does not matter for a child's cognitive outcomes in urban areas of the country.

Table 5.1: Summary of IV results on the relationship between preschool attendance and children's cognitive development at the age of five

	Ln (PPVT test score)	Ln (math test score)
Unadjusted for language difference among children		
OLS	0.172***	0.166***
IV - Wooldridge approach - only preschool as endogenous	0.216***	0.234***
IV – only preschool enrolment endogenous	0.318***	0.315***
IV- both preschool enrolment and z-score of height for age endogenous	0.279***	0.276***
Unadjusted for language difference among children		
OLS	0.071***	0.056***
IV - Wooldridge approach –only preschool as endogenous	0.149***	0.065***
IV – only preschool enrolment endogenous	0.212***	0.083***
IV- both preschool enrolment and z-score of height for age endogenous	0.207***	0.075**

note: *** p<0.01, ** p<0.05, * p<0.1; *ATT= Average treatment effect of the treated ; result of psmatch2 ; k=2 in the k-Nearest neighbourhood matching;* ln (PPVT test score) =Logarithm of PPVT test score in round 2; ln(math test score)=Logarithm of % of math questions correctly answered in round 2; Wooldridge approach – takes propensity score (predicted probability of enrolment in preschool education) as instrument for preschool enrolment

We are aware that finding good (relevant and valid) instrument variables is not that easy though the instruments we have chosen have passed the relevance and validity test (see Tables 5.3b, 5.4a and 5.4b). However, checking the robustness of the results is never an ended 'business'. That is why we still intend to examine the strengths of the discussed results using a propensity score matching estimation approach, which is basically built upon very certain assumptions in demonstrating the differential effects of preschool attendance on cognitive outcomes as compared to the previous models.

5.2 Propensity Score Matching Estimates

The goal of propensity score matching approach is to reproduce the treatment group among the non-treated ones, this way re-establishing the experimental conditions in a no experimental setting (Blundell et al, 2008). In the case of this study, while children who have been attending preschools belong to the treatment group, preschool age children who have not attended preschools are considered as non-treated group. Accordingly, to examine the differential effects of preschool attendance on the urban children's cognitive scores, we initially estimate propensity scores by running probit regressions of preschool attendance on a set of observed covariates (see Table A5.2 in the Appendix). We then estimate a regression based on Kernel matching technique and extracted only the ATT from the kernel regression results. To see the robustness of the results, we have also computed ATT using nearest neighbourhood, radius and calliper matching. In the kernel matching methods, all treated ones are matched with a weighted average of all controls with weights that are inversely proportional to the distance between propensity scores of treated and controls (Brand et al., 2003). In the analysis, the common support option has been selected; and the region of common support is given by [0.031, 0.998]. Though it is possible to estimate both the average treatment effect across the entire population (ATE) and the average treatment effect on the treated (ATT), we are only confined to the later one as our purpose of employing this technique is only to check the robustness of the previous OLS and IV estimates. The ATT here is the average difference of cognitive scores between the preschool attendees and the non-preschool attendees of urban preschool age children of the longitudinal data.

Accordingly, we obtain ATT estimates for both PPVT and CDA-Q test scores from the kernels and other matching regressions for all urban children of the sample. In fact, we also obtain cognitive scores separately for boys and girls. Average treatment effects on the treated (ATT) of cognitive scores are summarized in Table 5.2. Kernel matching estimates suggest that there is cognitive achievement advantage from attending early childhood education in the sense that preschool attendance is highly associated with cognitive outcome of children. To see the robustness of the results, we have obtained similar pattern when we use nearest neighbourhood and radius matching techniques (see A5.1 in the appendix). From the kernel matching estimates, we found that children who attended preschool education scored 31.2% higher in PPVT test and 23.1 in CDA-Q test when the test score is not adjusted for language differences. When the test scores are adjusted for language differences, preschool attendees scored 11.6% higher for PPVT test and 11.7% in CDA-Q test indicating gain that some of the association between preschool and cognitive outcome is due to the relationship of the tests to languages children speak. The result we obtain from the propensity score matching is much closer to IV estimates indicating the robustness of our results.

The sign or direction of association between preschool enrolment and cognitive outcomes is the same when we computed the estimate for girls and boys separately, but slight higher association between preschool and cognitive association for girls than for boys explaining the widely accepted result that preschool is more beneficial for girls (Campbell et al. 2002) because they are subject to discrimination by parents.

Table 5.2: Impact of preschool education on children's cognitive development at the age of five (kernel matching result)

	Girls + boys		Girls		Boys	
	ATTk	**T-stat**	**ATTk**	**T-stat**	**ATTk**	**T-stat**
Unadjusted for language						
Ln (PPVT test score)	0.312***	9.27	0.410***	6.24	0.307***	4.10
Ln(math test score)	0.231***	8.66	0.253***	4.54	0.233***	4.12
Adjusted for language						
Ln (PPVT test score)	0.116***	8.01	0.126***	6.59	0.100***	4.47
Ln(math test score)	0.117***	6.35	0.110***	5.00	0.103***	4.30

ATTk= Average treatment effect of the treated from kernel matching

If enrolment in pre-school education has a strong impact on cognitive development of children, why some parents fail to send their children to pre-school. The probit model we run to estimate the propensity score helps is to identify some of the factors correlated with parent's failure to send children to pre-school. The probit regression provided in Table A5.2 (in the appendix) indicated that in addition to the location dummies, household wealth and parents level of education are significantly correlated with the probability that a child in urban areas is enrolled in preschool education. This implies that enrolment in preschool is deterred by parents lack of awareness and money to send their children to preschool education. The results also indicate that one way of promoting pre-school education is establish government preschool so that children from poor families can be enrolled. Moreover, the result also imply that creation of awareness on the importance of preschool education to improve children's' cognitive development

6. Concluding Remarks

Using high quality data obtained from the Young Lives Longitudinal survey in Ethiopia, we have examined the associates of early childhood education attendance and cognitive development of preschool age children with particular emphasis on urban part the country. To check the robustness of the results, we employed several regression models including propensity score matching, OLS and instrumental variable estimation methods. Across all these models, though the point estimates change somewhat, the results persistently show that early childhood education attendance is associated with a substantial improvement in children's cognitive development. For instance, in the propensity score matching, children who have been attending preschool education have scored 31.2% higher in PVVT test and 23.1 % in CDA-Q test than those of non-preschool attendees. These figures are 11.6% and 11.7% when the cognitive outcomes are adjusted by language differences among the children. We have similar results in the instrumental variable estimation method: children who have been attending preschool education have scored 21.62% higher in PVVT and 23.4 % in CDA-Q test than those of non-preschool attendees, controlling for other factors. When cognitive outcomes are adjusted language differences among children, preschool education attendants scored 15% higher in PVVT and 6.5 % in CDA-Q test than those of non-preschool attendees.

Despite the evidence that preschool education is important for children's cognitive development, public investment in this critical stage of education is currently very small in the country. The majority of activities in this stage of education are rather left for private sector. As the result, it is obvious to see that the subsector is dominated by fee charging kindergartens in which children from low socioeconomic background do have very little opportunity to attend this fist and critical stage of education. For instance, in numerical terms, while the data from the Young Lives indicates that out of the 1912 total sample, 25% of them attend preschool education, we find a surprising figure for data acquired from the Ministry of Education for the whole country. Out of the 7 million preschool age children, only 4.2% of them attend preschool education. This clearly demonstrates that only some privileged children have an access to this fundamental stage education. Furthermore, though high enrolment rate in primary education is currently recording- to the extent of 94.3% gross enrolment, the sector is generally characterized by high dropout rate, repetition rate and overall quality deterioration.

Therefore, although early childhood education is not a panacea, the obtained results reveal that early education programs can substantially improve the cognitive development, academic success, and lives of children in poverty while benefiting the nation as a whole. Therefore, given this low participation rate and low quality of basic education, the government has to derive some lessons from the empirical analysis. Early childhood education is a key to later achievements in school and life. In view of that, government has to do more in this first and essential stage of education. It is paradoxical to expect that high quality basic education will be genuinely available for all children if the provision of early childhood schemes favours richer households in urban areas and exclude the poor and the marginalized. The government has to be aware that early childhood is a critical window of opportunity that helps break inter-generational transmission of poverty (Siraj-Blatchford, 2009). Investing in young children is one of the wisest investments a nation can make. The reviewed literatures indicate that countries that invest in early childhood education do so not because they have surplus resources but because they appreciate the advantages for children, families, communities and ultimately entire nations. So, the government needs to look at the existing preschool system with new eyes. Government instead of spending huge amount on capacity building of old people it is better to invest on kids who will be more able in the future. At least, by giving equal attention similar to that of primary, secondary and tertiary education, the government needs to establish public preschools for the poor in both rural and urban areas of the country.

Nevertheless, caution should be taken; public intervention in early childhood education does not necessarily imply supply by the public sector alone. In addition to wealth, we have seen parents' education is the key for children to attend in preschool education indicating the importance of awareness on the benefits of preschool school education. Helping stimulate demand by parents for such schemes through media campaigns could be one method of achieving this. Private provision should also be encouraged and incentives must be given to private providers who work mainly with disadvantaged groups.

References

Almond, Douglas and Currie, Janet (2010). Human Capital Development before age five. NBER Working Paper Series, Working Paper number 15827. NATIONAL BUREAU OF ECONOMIC RESEARCH 1050 Massachusetts Avenue, Cambridge, MA 02138, March 2010, at http://www.nber.org/papers/w15827, accessed on 13 December 2011.

Barnett, W. S. (1995). Long-Term Effects of Early Childhood Programs on Cognitive and School Outcome. *The Future of Children*, 5(3): 25-50.

Becker, G. S. (1962). *Human Capital.* New York: Columbia University Press.

Berlinksi S., Galiani, S., and Manacorda, M. (2006). Giving Children a Better Start: Preschool Attendance and School-Age Profiles. The Institute for Fiscal Studies Working Paper 06/18.

Berlinski, S., Galiani, S. and Gertler, P. (2009). The effect of Pre-primary Education on Primary School Performance, .*Journal of Public Economics 93, 219-234*

Blundell, R et al. (2008). Alternative Approaches to Evaluation in Empirical Microeconomics. ForschungsinstitutzurZukunft der Arbeit, Institute for the Study of Labor, IZA DP No. 3800.

Brand, E. (2003). Regression and Matching Estimates of the Effects of Elite College Attendance on Career Outcomes. University Of Wisconsin – Madison Annual Meeting of the American Sociological Association, Atlanta.

Bruner, J. T. (1999). *The Myth of the First Three Years: A New Understanding of Early Brain Development and Lifelong Learning.* New York. Free Press.

Campbell, et al. (2002). Early childhood education: Young adult outcomes from the Abecedarian Project. *Applied Developmental Science 6(1): 42-57.*

Carneiro P., Crawford and Goodman, (2006). Which Skills Matter? Center for the Economics of Education. London School of Economic.

Carneiro, P. and Heckman, J. (2003). *Human Capital Policy*. Cambridge, MA: The MIT Press.

Carneiro, P., Cunha, Flavio and Heckman , James (2003). Interpreting the Evidence of Family Influence on Child Development', paper presented at conference on The Economics of Early Child Development: Lessons for Economic Policy. Federal Reserve Bank, University of Minnesota, USA.

Case, A., and Paxson, C. (2006). Stature and Status: Height, Ability, and Labor Market Outcomes. *NBER Working Paper No.12466. Cambridge, MA*

Claessens, A., Duncan ,G. and Engle, M. (2008). Kindergarten skills and fifth grade achievement: Evidence from the ECLS-K" JEL Classification: I-Health, Education and Welfare, I20- General.

Cleveland, G. and Krashinsky, M. (1998). **The benefits and costs of good child care** – The Economic Rationale for Public Investment in Young Children. *A Policy Study,* Department of Economics, Childcare Resource and Research Unit, University of Toronto.

Connelly, C. (2008). Early Childhood Education in Buffalo. New York. *Found develop.wikispaces.com/.../Early+Childhood+Education+(2Revised), accessed on 24 August 2012.*

Cueto, Santiago, Juan Leon, Gabriela Guerrero and Ismael Muñoz (2009). Psychometric characteristics of cognitive development and achievement instruments in round two of Young Lives. Young Lives. *Technical Note* No. 15, January 2009. Found at http://www.younglives.org.uk/our-publications/technical-notes/psychometric-characteristics-of-cognitive-development-and-achievement-instruments-in-round-2-of-young-lives. Accessed on 15 December 2011.

Cunha, F., et al. (2006). Interpreting the evidence on life cycle skill formation. In *Handbook of the economics of education*, Volume 1. Amsterdam: North Holland

Currie, J. & Thomas, D. (2000). School quality and the longer-term effects of Head Start. *Journal of Human Resources*, 35(4): 755-774.

Currie, J. (2001). Early Childhood Education Programs, *Journal of Economic Perspectives*, 15(2):213-238.

Currie, J. and Duncan, T. (2001). Early Test Scores, School Quality and SES: Long run Effects on Wage and Employment Outcomes. *Research in Labor Economics* 20: 103-132.

Desai, S., P et al. (1989). Mother or Market? Effects of Maternal Employment on the Intellectual Ability of 4-Year-Old Children. *Demography,* 26(4), 545-561.

Doyle , O, Harmon, C. P., Heckman J. J. (2010). The Timing of Early Childhood Intervention: How early is Early? UCD Geary Institute & UCD School of Public Health and Population Science, University College Dublin.

Eurydice network. (2009). Early Childhood Education and Care in Europe: Tackling Social and Cultural Inequalities, Education, Audiovisual and Culture Executive Agency, P9 Eurydice

Glick, Peter. (2007). Cognitive Skills among Children in Senegal: Disentangling the Roles of Schooling and Family Background, *SAGA Working Paper, JEL: I21, J24*

Goodman, Alissa and Sianesi, Barbara. (2005). Early education and children's outcomes: How long do the impacts last? Institute for Fiscal Studies 26 (4): 513-548.

Grantham-McGregor, Sally ; Cheung , Yin Bun ; Cueto , Santiago ; Glewwe, Paul ; Richter , Linda ; Strupp, Barbara, and the International Child Development Steering Group (2007) "Developmental Potential in the First 5 Years for Children in Developing Countries", *Lancet* 369 (9555): 60-70.

Grover, D. (2005). Investing in Early childhood is a wise investment". *The CARK Education Forum, Dushanbe, Tajikistan. June 6-8*

Halfon N, Shulman E, Hochstein M. (2001). Brain development in early childhood: Building Community Systems for Young Children. UCLA Center for Healthier Children, Families and Communities, 2001.

Heckman, J. J., Ichimura, H. and Todd, P. E. (1997), "Matching As An Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme", Review of Economic Studies 64(4), 605-654.

Heckman, J. J., Rubinstein, Y. (2001). The importance of noncognitive skills: Lessons from the GED testing program. *American Economic Review* 91 (2): 145–149.

Heckman, James J. and Masterov, D. V. (2007). The Productivity Argument for Investing in Young Children. A paper presented at the Allied Social Sciences Association Annual Meeting. January 5-7

Heckman James J. and Klenow, P. J. (1997). *Human Capital Policy*. University of Chicago.

Helmers, C and Patnam, Manasa (2009). The Formation and Evaluation of Childhood Skill Acquisition: Evidence from India. *Young Lives, working Paper 38,* Oxford University, QEH.

Lefebvre, Pierre (2006). Impact of Early Childhood Care and Education on Children's Preschool Cognitive Development: Canadian Results from a Large Quasi-experiment. *Cahier de recherche/Working Paper 06-36*

Leuven, E. and Sianesi, B. (2003). "PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing". http://ideas.repec.org/c/boc/bocode/s432001.html

Macours, Karen, Schady, N. R., and Vakis, R. N. (2008). Cash Transfers, Behavioral Changes, and Cognitive Development in Early Childhood: Evidence from a Randomized Experiment. *Policy Research Working Paper 4759, Impact Evaluation Series No. 25.*

Magnuson, Katherine A., Ruhm, Christopher, Waldfogel, Jane (2004). Does prekindergarten improve school preparation and Performance? Technical Report 10452, NBER. Working Paper 10452, National Bureau of Economic Research

Malmberg, L. E., Mwaura, P , Sylva, K. (2011). Effects of a preschool intervention on cognitive development among East-African preschool children: A flexibly time-coded growth model. Early Childhood Research Quarterly 26(1):12-133.

Ministry of Education, (2001-2009). Educational statistics annual abstract", *Education Management Information Systems,* Addis Ababa, Ethiopia.

Myers, R. (1995). The twelve who survive: Strengthening programmes of early childhood Development in the third world, *Ypsilanti, MI: High/Scope Press.*

Paxson, C., and N. Schady (2007). Cognitive Development Among Young Children in Ecuador: The Role of Health, Wealth and Parenting. *Journal of Human Resources* 42(1): 49-84.

Powell, C., Walker, S. P., Himes, J. H., Fletcher, P. D. and Grantham-McGregor, S. (1995). Relationships between physical growth, mental development and nutritional supplementation in stunted children: the Jamaican study. *ActaPaedatrica* 84(1): 22-29.

Reynolds, A., Temple, J. A. (1998) "Extended early childhood intervention and school achievement: Age 13 findings from the Chicago Longitudinal Study." *Child Development,* 69(1), 231-246.

Sammons, P., Sylva, K., Melhuish, E., Siraj-Blatchford, I., Taggart, B., and Elliot, K. (2002) "Measuring the impact of preschool on children's cognitive progress Over the preschool period. Technical Paper 8a, London: Institute of Education.

Sanchez, Alan (2009). Early nutrition and later cognitive achievement in preschool children in Peru. Young Lives, working paper No. 57

Sanchez, Alan (2009). Early nutrition and later cognitive achievement in developing countries. *Paper commissioned for the EFA Global Monitoring Report 2010, reaching the marginalized* University of Oxford

Shonkoff, J. (2009). *Investment in Early Childhood Development Lays the Foundation for a Prosperous and Sustainable Society*. Center on the Developing Child at Harvard University, USA

Sianesi, B. (2004) An evaluation of the Swedish system of active labour market programmes in the 1990s. Review of Economics and Statistics ***86****(1): 133–155.*

Smith, J. A., and Todd, P. E. (2005). Does Matching Overcome LaLonde's Critiques of Non-experimental Estimators? *Journal of Econometrics*. 125(1-2): 305-353.

Susanna, L. (2005). How much is too much? The influence of preschool centers on children's social and cognitive development. *Economics of Education Review* 26(1):52–66.

UNESCO (2005). *Education for all: The quality imperative*. EFA Global Monitoring Report.

UNESCO (2006). EFA Global Monitoring Report 2007: Strong foundations – Early childhood Care and education. Oxford and Paris, Oxford University Press and UNESCO*.*

UNESCO (2010). The World Conference on Early Childhood Care and Education (ECCE): Building the Wealth of Nations. 2010*/ED/BAS/ECCE/ME/1 REV.2*

UNESCO (1996). Links between early childhood development and education and primary education. UNESCO Education Sector Monograph no. 6/1996.UNESCO, action research in family and early childhood, November 1996.

United Nations Children's Fund (UNICEF), (2007). A Human Rights-based Approach to Education for All: A framework for the realization of children's right to education and rights within education. United Nations Children's Fund, 3 United Nations Plaza, New York, NY 10017, USA, pubdoc@unicef.org, www.unicef.org,

Woessmann, L. (2006). Efficiency and Equity of European Education and Training Policies. CESIFO Working Paper No. 1779 Category 1: Public Finance.

Woldehanna, Tassew, Mekonnen, Alemu and Alemu, Tekie (2008). Young Lives: Ethiopia Round 2 Survey Report, *Young Lives Country Report*, Oxford: Young Lives

Woodhead, Martin (2009). Pathways through early childhood education in Ethiopia, India and Peru: Rights, diversity and equity. *Young Lives, working paper 54*

Woodhead, Martin et al. (2009). Equity and quality? Challenges for early childhood and primary education in Ethiopia, India and Peru. *Working Paper, No. 55, Studies in Early Childhood Transitions. The Hague, The Netherlands: Bernard van Leer Foundation.*

Woodhead, Martin, and Siraj-Blatchford (2009). Effective Early Childhood Programmes: Early Childhood. Focus 4, the Open University.

Wylie, C., et al, (2003). The long-term contribution of early childhood education to children's performance: Evidence from New Zealand. *International Journal of Early Years Education,* 11(1): 69-78.

Appendix

Table A2.1: Trend of Early childhood Education in Ethiopia

Year	Preschool Age Population			Enrolled pupils		
	Boys	Girls	Total	Boys	Girls	Total
2000/1	2781270	2709148	5490418	55616	53742	109358
2001/2	2934333	2864847	5799180	60809	58177	118986
2002/3	3067799	3000397	6068196	62574	60483	123057
2003/4	3207925	3142895	6350820	71435	67483	138918
2004/5	3355063	3292733	6647796	78884	74396	153280
2005/6	3509595	3450342	6959937	96604	90124	186728
2006/7	3546784	3484337	7031121	112400	106668	219068
2007/8	3439741	3322001	6761742	135122	128342	*263464*
2008/9	3538728	3418013	6956741	149988	142653	292641

Source: Author's compilation based on MoE 2001 to 2009 statistical abstract

Table A5.1. Impact of preschool education on children's cognitive development at the age of five for both unadjusted and adjusted for mother tong language differences among children

Method of matching	Type of statistics	Unadjusted for language		Adjusted for language	
		Ln (PPVT test score)	Ln (math test score)	Ln (PPVT test score)	Ln (math test score)
Kernel matching (default, k=5)	ATT	0.312***	0.231***	0.116***	0.117***
	T-ratio	9.27	8.66	8.01	6.35
Radius matching	ATT	0.307***	0.23***	0.114***	0.116***
	T-ratio	9.47	8.91	8.97	7.26
Local linear regression matching	ATT	0.25***	0.158***	0.038	0.045
	T-ratio	3.72	2.72	0.71	0.81
k-Nearest neighbourhood matching (k=2)	ATT	0.13	0.272***	0.092***	0.074***
	T-ratio	1.54	3.79	3.29	4.88
k-Nearest neighbourhood matching (k=1)	ATT	0.152**	0.269***	0.119***	0.080***
	T-ratio	2.63	5.63	3.58	4.52

note: *** p<0.01, ** p<0.05, * p<0.1; *ATT= Average treatment effect of the treated; result of psmatch2 ; k=2 in the k-Nearest neighbourhood matching;* ln (PPVT test score) =Logarithm of PPVT test score in round 2; ln(math test score)=Logarithm of % of math questions correctly answered in round 2

Table A5.2: First stage probit estimation of preschool

	Coefficient (t-ratio)
Dummy variable for site 2 in Addis Ababa	0.174
	(0.476)
Dummy variable for site 3 in Addis Ababa	-0.249
	(-0.685)
Dummy variable for urban site in Amhara region	-1.832***
	(-5.899)
Dummy variable for urban site in Oromia region	-2.159***
	(-6.871)
Dummy variable for urban site 1 in SNNP region	-3.273***
	(-8.670)
Dummy variable for urban site 2 in SNNP region	-0.640*
	(-1.946)
Dummy variable for urban site Tigray region	-2.757***
	(-7.661)
Dummy variable for other urban site s (where children moved out of their site)	-1.738***
	(-4.863)
Calculated age in months for deriving z scores in round 1	0.037*
	(1.890)
Wealth index for 1-year-olds (Round 1)	1.743***
	(2.828)
Dummy for male	0.030
	(0.217)
Number of children below 7 and above 65 years old	-0.155
	(-1.623)
Number of children between 7 and 17 years old	-0.051
	(-0.988)
Number of male family members > 17 and less than 65 years	-0.048
	(-0.510)
Number of female family members > 17 and less than 65 years	0.113
	(1.190)
Highest grade completed by primary caregiver	0.059***
	(2.912)
Highest grade completed by father	0.065***
	(3.717)
Dummy for a child had long term health problem?	-0.342

	Coefficient (t-ratio)
	(-1.207)
Dummy for illness	-0.025
	(-0.168)
Dummy for theft	0.215
	(0.931)
Dummy for increased input prices	0.193
	(1.030)
Dummy for divorce or separation of family	-0.732**
	(-2.408)
Dummy for place employment shutdown or job loss	0.004
	(0.019)
Dummy for decrease in food availability	0.208
	(1.205)
Dummy for job loss/source of income/family enterprise	-0.157
	(-0.814)
Dummy for divorce or separation	0.040
	(0.127)
Dummy for severe illness or injury	-0.052
	(-0.248)
Constant	0.078
	(0.157)
Number of observations	744
Log-Likelihood	-221.04
Adjusted R2	0.565
note: *** p<0.01, ** p<0.05, * p<0.1	

Table A5.3a. Wooldridge IV estimation of cognitive development on preschool enrolment (propensity score as instrument-Wooldridge (2002, P.623)

explanatory variables	Log (PPVT test score)	Log (Q-CDA test score)
	coefficient (t-ratio)	coefficient (t-ratio)
Dummy variable for child being enrolled in preschool	0.216***	0.234***
	(3.822)	(5.466)
Z-score of height-for-age at the age of one year	0.013	0.023***
	(1.372)	(3.282)
Calculated age in months for deriving z scores in round 1	0.029***	0.016***
	(6.134)	(4.625)
Wealth index for 1-year-olds (Round 1)	0.212	0.025
	(1.573)	(0.215)
Dummy for male	0.040	0.011
	(1.273)	(0.444)
Number of children below 7 and above 65 years old	-0.034*	-0.014
	(-1.696)	(-0.740)
Number of children between 7 and 17 years old	0.015	-0.001
	(1.336)	(-0.065)
Number of male family members > 17 and less than 65 years	-0.010	-0.004
	(-0.471)	(-0.259)
Number of female family members > 17 and less than 65 years	-0.009	-0.009
	(-0.399)	(-0.535)
Highest grade completed by primary caregiver?	0.018***	0.009**
	(3.761)	(2.555)
Highest grade completed by father?	0.002	-0.002
	(0.349)	(-0.508)
Number of months the child breastfed	-0.000	0.000
	(-0.181)	(0.488)
Five point Likert scale for the relative size of birth weight (-2 to 2)	-0.007	-0.006
	(-0.494)	(-0.559)
Dummy for child had health problems at the age of one year	-0.052*	-0.008
	(-1.665)	(-0.339)
Constant	2.603***	1.864***
	(28.609)	(22.543)
Number of observations	719	720
Adjusted R^2	0.199	0.139
Centered R^2 (r2c)	0.215	0.156
Uncentered R^2 (r2c)	0.983	0.979
Under identification test (Kleibergen-Paaprk LM statistic) (idstat)	452.038	453.068
P-value for under identification test (idp)	0.000	0.000

note: *** p<0.01, ** p<0.05, * p<0.1

Table. A5.3b. Wooldridge IV estimation of cognitive development adjusted for language on preschool enrolment (propensity score as instrument-Wooldridge (2002, P.623)

explanatory variables	Log(PPVT test score)	Log (Q-CDA test score)
	coefficient (t-ratio)	coefficient (t-ratio)
Dummy variable for child being enrolled in preschool	0.149***	0.065***
	(4.533)	(3.172)
Z-score of height-for-age at the age of one year	0.006	0.011***
	(1.600)	(3.500)
Calculated age in months for deriving z scores in round 1	0.008***	0.009***
	(4.124)	(5.179)
Wealth index for 1-year-olds (Round 1)	0.097*	0.072
	(1.725)	(1.338)
Dummy for male	0.021*	0.009
	(1.730)	(0.745)
Number of children below 7 and above 65 years old	-0.008	-0.010
	(-0.999)	(-1.103)
Number of children between 7 and 17 years old	0.002	-0.005
	(0.451)	(-1.301)
Number of male family members > 17 and less than 65 years	-0.006	-0.008
	(-0.793)	(-1.158)
Number of female family members > 17 and less than 65 years	-0.002	0.002
	(-0.278)	(0.362)
what is the highest grade completed by primary caregiver?	0.005***	0.005***
	(2.931)	(3.027)
Highest grade completed by father?	-0.002	-0.002
	(-1.287)	(-1.137)
Number of months the child breastfed	-0.000	-0.000
	(-0.466)	(-0.366)
Five point Likert scale for the relative size of birth weight (-2 to 2)	-0.001	-0.001
	(-0.148)	(-0.227)
Dummy for child had health problems at the age of one year	-0.026**	-0.008
	(-2.046)	(-0.702)
Constant	5.515***	5.594***
	(154.014)	(151.743)
Number of observations	486	642
Adjusted R^2	0.154	0.127
Centered R^2 (r2c)	0.178	0.146
Uncentered R^2 (r2c)	0.999	0.999
Under identification test (Kleibergen-Paaprk LM statistic) (idstat)	175.501	363.657
P-value for under identification test (idp)	0.000	0.000
note: *** p<0.01, ** p<0.05, * p<0.1		

Table A5.4a: Regression of log of PPVT test score and cognitive development assessments - quantitative (CDA-Q) test score

	Ln of PPVT test score			Ln of CDA_Q test score (% correctly answered)		
	OLS	**IV version 1**	**IV version 2**	**OLS**	**IV version 1**	**IV version 2**
	coef/t	**coef/t**	**coef/t**	**coef/t**	**coef/t**	**coef/t**
Dummy variable for child being enrolled in preschool	0.172***	0.318***	0.315***	0.166***	0.279***	0.276***
	(4.490)	(4.863)	(3.847)	(5.682)	(5.373)	(3.610)
Z-score of height-for-age at the age of one year	0.014	0.007	0.187***	0.024***	0.017**	0.232***
	(1.465)	(0.778)	(2.943)	(3.272)	(2.533)	(4.267)
Calculated age in months for deriving z scores in round 1	0.029***	0.027***	0.052***	0.017***	0.015***	0.045***
	(6.269)	(5.797)	(4.972)	(4.722)	(4.326)	(4.991)
Wealth index for 1-year-olds (Round 1)	0.246*	0.149	-0.050	0.079	-0.047	-0.214
	(1.832)	(1.131)	(-0.275)	(0.775)	(-0.406)	(-1.237)
Dummy for male	0.041	0.025	0.107**	0.013	0.007	0.090**
	(1.290)	(0.803)	(2.255)	(0.547)	(0.291)	(2.103)
Number of children below 7 and above 65 years old	-0.037	-0.036*	-0.023	-0.018	-0.022	0.002
	(-1.608)	(-1.790)	(-0.847)	(-0.999)	(-1.171)	(0.078)
Number of children between 7 and 17 years old	0.015	0.015	-0.003	-0.002	-0.004	-0.026*
	(1.211)	(1.304)	(-0.207)	(-0.166)	(-0.473)	(-1.821)
Number of male family members > 17 and less than 65 years	-0.010	-0.005	-0.048*	-0.004	-0.000	-0.047**
	(-0.495)	(-0.232)	(-1.793)	(-0.229)	(-0.009)	(-2.248)
Number of female family members > 17 and less than 65 years	-0.006	-0.008	-0.009	-0.003	-0.011	-0.010
	(-0.288)	(-0.332)	(-0.359)	(-0.215)	(-0.728)	(-0.447)
Highest grade completed by primary caregiver	0.019***	0.017***	0.013**	0.010***	0.009***	0.004
	(4.068)	(3.497)	(2.119)	(2.795)	(2.662)	(0.869)
Highest grade completed by father	0.002	-0.001	-0.002	-0.000	-0.001	-0.003
	(0.533)	(-0.137)	(-0.394)	(-0.123)	(-0.419)	(-0.728)
Number of months the child was breastfed	-0.000	0.000	0.000	0.000	0.001	0.001
	(-0.322)	(0.371)	(0.078)	(0.228)	(0.592)	(0.607)

	Ln of PPVT test score			Ln of CDA_Q test score (% correctly answered)		
	OLS coef/t	IV version 1 coef/t	IV version 2 coef/t	OLS coef/t	IV version 1 coef/t	IV version 2 coef/t
Five point Likert scale for the relative size birth weight (-2 to 2)	-0.008	-0.006	-0.038**	-0.008	-0.010	-0.033*
	(-0.537)	(-0.426)	(-1.962)	(-0.686)	(-0.907)	(-1.859)
Dummy for child had health problems at the age of one year	-0.050	-0.043	-0.025	-0.006	-0.017	0.037
	(-1.535)	(-1.377)	(-0.631)	(-0.234)	(-0.715)	(1.040)
Constant	2.607***	2.594***	2.560***	1.869***	1.909***	1.797***
	(26.289)	(28.399)	(22.229)	(24.653)	(23.604)	(15.796)
Number of observations (N)	719	719	719	720	720	720
Adjusted R^2	0.200	0.182	-0.196	0.146	0.122	-0.862
Centered R^2 (r2c)		0.198	-0.173		0.139	-0.826
Uncentered R^2 (r2c)		0.983	0.975		0.979	0.955
Under identification test (Kleibergen-Paaprk LM statistic) (idstat)		264.065	23.262		266.558	24.092
P-value for under identification test (idp)		0.000	0.026		0.000	0.020
Hansen J statistics (over identification test of all instruments) (j)		24.438	11.244		41.265	12.179
P-value for over identification test (jp)		0.018	0.423		0.000	0.350

Note: *** p<0.01, ** p<0.05, * p<0.1; The instruments are regional dummies and pre and post natal shocks. IV version1: only preschool enrolment is endogenous ; IV version 2: both preschool and z-score of height for age are endogenous variables

Table A5.4b: Regression of log of PPVT test score and cognitive development assessment - quantitative (CDA-Q) test score adjusted for language

	Ln of PPVT test score			Ln of CDA_Q test score (% correctly answered)		
	OLS	**IV version1**	**IV version 2**	**OLS**	**IV version1**	**IV version 2**
	coef/t	**coef/t**	**coef/t**	**coef/t**	**coef/t**	**coef/t**
Dummy variable for child being enrolled in preschool	0.071***	0.212***	0.207***	0.056***	0.083***	0.075**
	(4.225)	(5.309)	(4.554)	(4.035)	(3.738)	(2.099)
Z-score of height-for-age at the age of one year	0.008**	0.005	0.009	0.011***	0.010***	0.111***
	(2.162)	(1.274)	(0.414)	(3.098)	(3.328)	(4.223)
Calculated age in months for deriving z scores in round 1	0.009***	0.007***	0.008**	0.009***	0.009***	0.023***
	(4.808)	(3.827)	(2.030)	(5.079)	(5.775)	(5.268)
Wealth index for 1-year-olds (Round 1)	0.139**	0.057	0.058	0.078	0.058	-0.016
	(2.558)	(0.996)	(1.015)	(1.536)	(1.105)	(-0.188)
Dummy for male	0.026**	0.017	0.019	0.009	0.012	0.045**
	(2.109)	(1.359)	(1.259)	(0.748)	(1.052)	(2.226)
Number of children below 7 and above 65 years old	-0.010	-0.004	-0.004	-0.011	-0.005	-0.001
	(-1.143)	(-0.533)	(-0.539)	(-1.258)	(-0.516)	(-0.084)
Number of children between 7 and 17 years old	0.001	0.003	0.002	-0.006	-0.006	-0.015**
	(0.221)	(0.488)	(0.351)	(-1.244)	(-1.551)	(-2.147)
Number of male family members > 17 and less than 65 years	-0.007	-0.005	-0.006	-0.007	-0.006	-0.030***
	(-0.956)	(-0.581)	(-0.610)	(-1.016)	(-0.885)	(-2.850)
Number of female family members > 17 and less than 65 years	0.003	-0.007	-0.007	0.003	0.000	0.004
	(0.396)	(-0.825)	(-0.826)	(0.419)	(0.039)	(0.408)
Highest grade completed by primary caregiver	0.006***	0.005***	0.005***	0.005***	0.004***	0.003
	(3.392)	(2.738)	(2.651)	(3.061)	(2.577)	(1.112)
Highest grade completed by father	-0.001	-0.003*	-0.003*	-0.002	-0.002	-0.003
	(-0.742)	(-1.874)	(-1.833)	(-0.995)	(-1.220)	(-1.174)
Number of months the child was breastfed	-0.000	-0.000	-0.000	-0.000	-0.000	0.000
	(-0.633)	(-0.558)	(-0.582)	(-0.429)	(-0.284)	(0.072)

	Ln of PPVT test score			Ln of CDA_Q test score (% correctly answered)		
	OLS	**IV version1**	**IV version 2**	**OLS**	**IV version1**	**IV version 2**
	coef/t	**coef/t**	**coef/t**	**coef/t**	**coef/t**	**coef/t**
Five point Likert scale for the relative size birth weight (-2 to 2)	-0.002	0.001	-0.000	-0.001	0.002	-0.022**
	(-0.262)	(0.168)	(-0.050)	(-0.269)	(0.454)	(-2.112)
Dummy for child had health problems at the age of one year	-0.024*	-0.025**	-0.025*	-0.008	-0.005	0.003
	(-1.858)	(-1.969)	(-1.922)	(-0.672)	(-0.470)	(0.158)
Constant	5.538***	5.496***	5.496***	5.595***	5.576***	5.550***
	(143.235)	(144.007)	(144.278)	(154.694)	(154.944)	(100.082)
Number of observations (N)	486	486	486	642	642	642
Adjusted R^2	0.191	0.068	0.073	0.127	0.119	-0.971
Centered R^2 (r2c)		0.095	0.100		0.138	-0.928
Uncentered R^2 (r2c)		0.999	0.999		0.999	0.999
Under identification test (Kleibergen-Paaprk LM statistic) (idstat)		135.917	17.546		303.735	24.258
P-value for under identification test (idp)		0.000	0.093		0.000	0.019
Hansen J statistics (over identification test of all instruments) (j)		6.910	6.666		41.516	10.540
P-value for over identification test (jp)		0.806	0.757		0.000	0.483

Note: *** p<0.01, ** p<0.05, * p<0.1; The instruments are regional dummies and pre and post natal shocks. IV version1: only preschool enrolment is endogenous ; IV version 2: both preschool and z-score of height for age are endogenous variables

Table A5.5: Descriptive statistics of variables used the regressions above

	Mean	Std. Dev.	Min	Max
Z-score of height-for-age at the age of one year	-0.943	1.774	-6.040	8.170
Z-score of height for age at the age of eight years	-1.193	1.119	-6.570	4.590
Dummy variable for a child begun formal school	0.872	0.334	0.000	1.000
Calculated age in months for deriving z scores in round 1	12.329	3.582	6.100	18.233
Calculated age in months for deriving z scores in round 2	62.451	3.779	55.294	74.743
Wealth index for 1-year-olds (Round 1)	0.330	0.152	0.006	0.757
Wealth index for 5-year-olds (Round 2)	0.373	0.153	0.008	0.881
Dummy for male	0.522	0.500	0.000	1.000
Number of children below 7 and above 65 years old	1.524	0.701	1.000	5.000
Number of children between 7 and 17 years old	1.304	1.341	0.000	8.000
Number of male family members > 17 and less than 65 years	1.165	0.821	0.000	5.000
Number of female family members > 17 and less than 65 years	1.410	0.824	0.000	6.000
what is the highest grade completed by primary caregiver?	4.984	4.302	0.000	14.000
what is the highest grade completed by father?	6.590	4.749	0.000	14.000
3.12.1 how many months did you breastfeed name?	28.930	13.523	0.000	36.000
Five point Likert scale for the relative size of child when born (-2 to 2)	0.056	1.069	-2.000	2.000
Dummy for child had health problems at the age of one year	0.430	0.495	0.000	1.000
Site dummies				
Dummy for site 2	0.121	0.326	0.000	1.000
Dummy for site 3	0.117	0.322	0.000	1.000
Dummy for site 4	0.116	0.320	0.000	1.000
Dummy for site 5	0.117	0.322	0.000	1.000
Dummy for site 6	0.122	0.328	0.000	1.000
Dummy for site 7	0.125	0.331	0.000	1.000
Dummy for site 8	0.117	0.322	0.000	1.000
Dummy for site 8	0.048	0.215	0.000	1.000
Dummy for illness	0.349	0.477	0.000	1.000
Dummy for theft	0.097	0.296	0.000	1.000
Dummy for increased input prices	0.191	0.393	0.000	1.000
Dummy for divorce or separation of family	0.058	0.234	0.000	1.000
Dummy for place employment shutdown or job loss	0.172	0.378	0.000	1.000
Dummy for decrease in food availability	0.340	0.474	0.000	1.000
Dummy for job loss/source of income/family enterprise	0.183	0.387	0.000	1.000
Dummy for divorce or separation	0.059	0.236	0.000	1.000
Dummy for severe illness or injury	0.138	0.346	0.000	1.000

N=745