YOUNG LIVES STUDENT PAPER

Title: Early childhood education in India: A possible investment in better outcomes? A quantitative analysis using Young Lives India

Authors: Sabilah Eboo Alwani

Date: 19 October, 2018

Thesis submitted in part fulfilment of the requirements for the degree of MPhil etc

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

Young Lives was core-funded by UK aid from the Department for International Development (DFID) until June 2018.

UNIVERSITY OF CAMBRIDGE
Faculty of Education

Faculty of Education

MPhil in Education

Thesis

Early childhood education in India: A possible investment in better outcomes? A quantitative analysis using Young Lives India

Name: Sabilah Eboo Alwani
Advisor: Dr Ricardo Sabates
Submission Date: 19 October, 2018

Abstract

This paper explores the relationship between early childhood education and academic outcomes for children in India by estimating the ability of preschool participation at age 5 to predict results on major cognitive assessments at age 12. Initially looking at differences in means, it moves on to utilise regression analysis first in an uncontrolled model, and then in a model which controls for both gender and maternal education, as these have been deemed important inputs for academic attainment in the wider literature on human capital development. The sample used for this research is constructed from Young Lives (India), which from 2002 to 2017 surveyed two cohorts of children across Andhra Pradesh and Telangana, with a pro-poor sampling strategy. Surprisingly, the results of the analysis find that participation in early childhood education had a negligible effect on test scores, even when controlling for gender and maternal education. Meanwhile, maternal education emerged as a strong predictor of test results. These findings contradict much of the existing evidence that demonstrates associations between early childhood education and cognitive development, and, in turn, improved economic outcomes. Accordingly, it raises questions about the generalisability of the existing evidence and the quality of India's ECE offering.

Keywords:

Early Childhood Education, India, Longitudinal Research, Academic Outcomes, Poverty Alleviation

Declaration of Originality

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except where specifically indicated in the text. This thesis does not exceed 20,000 words in length.

Table of Contents

Abbreviations and definitions

Abbreviations

ECCE: Early Childhood Care and Education

ECE: Early Childhood Education

ECD: Early Childhood Development

HCT: Human Capital Theory

ICDS: Integrated Child Development Services

IECEI: Indian Early Childhood Education Impact study

LAMI: Low-and-middle income (referring to countries)

MDG: Millennium Development Goals

NHST: Null Hypothesis Significance Testing

PIAT: Peabody Individual Achievement Test

PPVT: Peabody Picture Vocabulary Test

SDG: Sustainable Development Goals

SES: Socioeconomic status

YL: Young Lives

Definitions

ECCE: ECCE summarises the set of inputs and processes that enables children 0-8 to flourish throughout the life course by creating the building blocks of later social, cognitive, emotional and economic development (Kaul *et al*., 2017).

ECE: Used within this paper, ECE refers to regularly attended education in a setting or establishment outside of a child's home for children approximately aged 3-5, during the period immediately preceding primary school (Woldehanna and Gebremedhin, 2012).

ECD: "A multifaceted concept from an ecological framework that focuses on the child's outcome (development), which depends on characteristics of the child and the context, such as health, nutrition, protection, care and/or education". Covers age range from pre-birth to 8 years, as per Convention on the Rights of the Child and Education for All declaration (Britto, Engle and Super, 2013).

Preschool: Used within this paper interchangeably with ECE.

List of tables and figures

Figures

Tables

1. Introduction

1.1 Research context and foundation

Research over the past half-century has shown conclusively that the period of early childhood is the single most critical phase in human development (Shonkoff and Phillips, 2000). The foundational capacities established during this time support the acquisition of key skills for the life course and beget improved outcomes which extend to future generations (Black *et al*., 2016). Access to early childhood care, health and education (ECCE) has been shown to be fundamental to nurturing the growth of these skills, enabling children to achieve cognitive outcomes that support more secure futures (Attanasio, Meghir and Nix, 2015 and Engle *et al*., 2007). This paper explores the relationship between early childhood education (ECE) and later cognitive outcomes for a sample of children in India, to determine whether India's ECE programmes are capitalising on this critical period by helping children to attain academic outcomes as consistent with the literature on ECE.

Undertaking such evaluation is becoming increasingly salient for policies relating to economic growth and to international development. For example, numerous studies have linked preschool participation to increased individual earnings, and more widely to progress toward poverty alleviation (see Heckman, 2006; Engle *et al*., 2007; Shonkoff and Richter, 2013 and Black *et al*., 2016). Quantifying returns to ECE for policy purposes has also been increasingly convincing: for every dollar invested in ECE, returns have been shown to be as high as 7-10% per annum (Heckman, 2011). The subsequent growth in literature on ECE, marrying evidence from child development, public health, economics and education, has catalysed widespread belief in early childhood education as a critical lever for economic development outcomes.

But the evidence has also shown that this is the period in which individual disparity begins to take effect, and that early education has a role to play as an equalizer (Singh and Mukherjee, 2016). Accordingly, early education is now viewed globally as offering possibilities for improved individual opportunities, particularly for those who are socioeconomically or otherwise disadvantaged (Streuli, Vennam and Woodhead, 2011). Evidence of the hopes placed upon it are enshrined in Goal 1 of the Dakar Framework for Action (Education for

All), to expand and improve ECCE around the world, and in Sustainable Development Goal 4.2, which stipulates early childhood development, care, and education for all children by 2030 (UNESCO, 2000). Accordingly, both research and international policy now make the case for ECE as a cost-effective ways to break intergenerational poverty cycles (Indian Ministry of Women and Child Development, 2013).

Consequently, improving ECE interventions is a priority for rapidly growing economies, because the evidence supports the notion that they will have much to gain (Attanasio, Meghir and Nix, 2015). Governments of many LAMI (low-and-middle income) countries are progressively instituting national policies for ECCE, and international development agencies are inscribing them as standard features of poverty reduction strategies (UNICEF, 2007 and Becker, Murphy and Tamura, 1990). But implementation of these policies which are fairly new and untested has been uneven and problematic, as service provision is usually honed for primary rather than pre-primary education, and coordination across various ministries which house the related services for ECCE has proved challenging.

Therefore, making the right improvements can prove difficult, as to date the focus has been more on developing practice rather than research and evaluation. Much of the evidence comes from developed countries, who have very different economic, cultural, and political contexts (Levin and Schwartz, 2006). Moreover, factors which are negligible in these contexts, such as gender or geography, might play a significant role in outcomes elsewhere. In order for improvements to be effective, more contextual evidence is needed (Woodhead, 2009 and Yoshikawa and Nieto, 2013). This raises the question of how applicable the existing evidence is to the contexts of LAMI countries, and whether the thinking applied to policymaking on ECE fails to take into account local particulars.

This paper explores whether this is the case for India. The goal is to determine whether patterns on ECE from the global evidence body are exportable; that is, whether early childhood education is strongly associated with academic attainment in this context, as this would be the expectation based on existing evidence. To date, research on early childhood education in India has been limited, and the Indian government has acknowledged that its ECE services are not up to standard (Alcott *et al.*, 2018 and Government of India, 2013). But with India's rapidly growing economy and extensive early years infrastructure, it offers interesting parameters for the assessment of ECE's role. This paper's findings are therefore

intended to contribute to the limited evidence and to offer analysis that could be useful in the improvement of India's early years educational provision. In particular, it hopes to shed light on whether India's preschool performance is linked to the kinds of academic outcomes that promise economic returns.

The primary aim of this paper is to estimate the ability of participation in preschool to predict cognitive outcomes for children, as they have been widely linked to improved economic opportunities. However, key researchers on early education in India have pointed out the pressing need for studies to go beyond establishing the simple effects of ECE and understand other dynamics at play (Kaul *et al*., 2017). In order to respond to this suggestion, I also consider two factors which have been identified in the literature as important to India's educational experience: gender and maternal education.

Gender, in particular, has been a noted axis of disparity in access, enrolment, completion and attainment. It has sparked a sub-theme, that of 'gender and education', which comprises its own body of literature with input from various perspectives, be they human capital or social justice (Unterhalter, 2007). The evidence from these disparate perspectives agrees that there has historically been a gender gap in educational experiences and outcomes, and that this remains a problem in much of the developing world. Underpinning disparate studies on gender and education is an abiding concern with inequality. In India particularly, statistics show that 25% of girls are out of school, and 17% of girls have never been to school at all. There are knock-on effects of these statistics as well: only 2% of girls in India are in higher education (Education-inequalities.org, 2018). As these girls grow into motherhood, there are further intergenerational consequences. Accordingly, both gender and maternal education represent important points for inquiry in wider research on ECE and outcomes in India.

Drawing from the body of evidence on ECE and on these other relevant variables, I test the hypothesis that preschool participation should emerge as a good predictor of cognitive outcomes. This hypothesis is tested using secondary longitudinal data on Indian students who participated in early childhood education, from the Young Lives (YL) study. Young Lives is a renowned study on childhood poverty consisting of multiple quantitative and qualitative survey layers, undertaken to inform effective policy and interventions for children (Vennam *et al*., 2009). Accordingly, it contains specific, relevant data on early childhood education experiences. (Early Childhood Development: informing policy and making it a priority,

2018). Additionally, as it is concerned with childhood poverty, it is informed and inspired by the same concerns and themes that underpin this paper.

This paper therefore follows the same line of motivation, but focuses specifically on understanding how educational opportunities in early childhood are linked with academic achievement. The related policy question of interest is how such evidence, emanating from local rather than global data, might help to reduce "the intergenerational transmission of limited educational attainment, poor personal and social adjustment, unemployment or low paid work, impaired health, and decreased longevity;" in short, to improve quality of life with ECE as a foundational start (Shonkoff and Richter, 2013).

The indicators selected to represent the concepts 'early childhood education' and 'cognitive development' are preschool participation at approximately age 5 and test score results on the Peabody Picture Vocabulary Test (PPVT) and mathematics assessment (taken from the Trends in International Mathematics and Science Study or TIMSS) at age 12. These types of tests are in line with widely used indicators of cognitive development from related studies.

The research questions addressed in this paper are as follows:

- Was early childhood education a good predictor of test scores on the PPVT and maths tests?
- Did boys who attended preschool fare better than girls on these tests?
- Did children with educated mothers fare better on these tests?
- Once gender and maternal education were controlled, was preschool a useful predictor of test scores?

These research questions called for the use of empirical data recorded as categorical and continuous variables. Accordingly, my methodological approach necessitated quantitative analysis, for which I employed statistical modelling to estimate the relationships between these variables. Whilst the original YL research conducted also included qualitative data, its use for a mixed methods approach was not possible due to access constraints, but could have added a valuable layer of nuance to this paper.

This research intends to fill a gap in the literature on India firstly because there is very little, and secondly because more evidence on ECE's impact is needed in order to mitigate the risks of generalising from other, very different, contexts. However, the generalisability of these results themselves are limited due to geographical and sampling constraints (Singh and Mukherjee, 2017). But as one of the implications informing these research questions is whether existing evidence has over-generalised the expectation of ECE's impact, the limited generalisability of evidence from this paper has a unique context: ultimately, the transformative power of ECE may prove universal, but only localised evidence may show how, for whom, and to what extent this is the case.

In the remainder of this paper, I will delineate the background, execution, and conclusions related to my research inquiry. In Section 2, I will offer a conceptual framework to situate the paper's research questions. In Section 3, I will review the literature on ECE to date. In Section 4, I discuss the context of ECE in India. Section 5 gives an overview of the various data and variables used, while Section 6 outlines the paper's methodology, including limits and problems with the models. Section 7 covers the paper's findings, and Section 8 discusses them in more detail. Section 9 offers concluding thoughts and opportunities for future research.

1.2 Using Young Lives data

This paper's sample is drawn from Young Lives, a cross-sectional, longitudinal, international study on childhood poverty. The project surveyed children from four LAMI countries (India, Ethiopia, Peru, and Vietnam) every four years from 2002 - 2017. Each country's sample was comprised of a younger cohort, aged 1 at the inception of the study, and an older cohort, aged 5 at the inception of the study. The project's major contribution has been to provide evidence on how children change over time, as impacted by the causes and consequences of poverty. Three types of questionnaires at different levels of experience were utilised as data gathering tools: a child questionnaire, a household questionnaire, and a community questionnaire. The data from these surveys includes detailed child-specific and household information, including information related to education participation and outcomes, which therefore makes it a rich resource for this paper's research questions (Singh and Mukherjee, 2016). This paper focuses specifically on YL data from India, drawing from the younger cohort of children (originally

n=2,000). I make use of data comprised of information from Rounds 2 and 4, which took place when these children were approximately aged 5 and 12.

There are two important limits to generalisability introduced to this paper due to usage of YL's data. The first is geographic: participants from YL India were only from one (later two) states within India: Andhra Pradesh and Telangana[1]. Together these areas comprise only 7% of the population of India (Younglives-india.org, 2018). Whilst Andhra Pradesh and Telangana offer a range of attributes common to other parts of India, such as large areas of rural population coupled with a few major urban areas, they are unique in other ways which bar resulting evidence from generalisability, but still allow it to indicate potential patterns.

The second limitation is due to YL's sampling strategy. In order to construct a sample to study childhood poverty, the team over-sampled from poor households using purposive sampling[2] (Attanasio, Meghir and Nix, 2015). This was coupled with sentinel site surveillance, with sites corresponding to the interests of the YL research programme (Kumra, 2008). However, the YL team has discussed its endeavours to redress this limitation, showing that despite the sampling bias, the sample included households across a wide range of socioeconomic characteristics[3] (Kumra, 2008). Additionally, they have noted that the data was considered to be "sufficiently large for statistical analysis in general, allowing for the detection of moderate-sized differences between sub-groups of children" (Young Lives Survey Design and Sampling in India, 2014). However, they confirm that it is not suitable for generalisation to the wider population of India due to its purposive sampling and geographical specificities (Kumra, 2008).

Despite threats to generalisability arising from geography or sampling constraints, the YL research team has noted that its sample is an 'appropriate and valuable instrument for analysing causal relations, modelling child welfare, and its longitudinal dynamics in Andhra Pradesh' (Kumra, 2008). It is also a robust, academically accepted, large-scale, and geographically relevant survey which includes data on early childhood education experiences

[1] Telangana became its own state in June 2014, as the fourth round of the survey was being finalised (Young Lives Survey Design and Sampling in India, 2014).
[2] Families sampled by YL India had a mean income of approximately 850 USD per family per annum in 2005, lower than the average mean income for Andhra Pradesh as a whole.
[3] As cited by Attanasio, Meghir, and Nix (2015), "evidence on the standard deviation indicates that within [the] sample there is a considerable degree of heterogeneity in socio-economic background".

and academic outcomes. Accordingly, it provided an applicable sample for this paper's inquiry.

2. Human Capital as a conceptual framework

2.1 Early investments and future returns

Policies and practices relating to early childhood are characterised by a number of competing discourses. Of greatest interest for this paper is the political and economic perspective, which is informed by principles of child development and underpinned by economic models of human capital theory (HCT), a dominant paradigm in the space of education and development (Woodhead, 2006).

The value of the human capital perspective on early childhood is that it offers useful theoretical parameters for exploring the economic justification for investment in early childhood education. It does this by helping to outline a two-part continuum: the first part is the idea that *early investments in education boost academic attainment.* The second part is that *high academic attainment generates economic returns*. In fact, much of human capital literature on education can be divided into these two groups: early investments, and future returns. The latter idea is well established in the literature and is what links this paper's inquiry to overarching concerns with education's role in poverty alleviation and economic development; however, it is the former on which this paper focuses.

The notion that *education generates future returns* represents the initial conception of education by human capital theorists. A human capital perspective posits that educational attainment is a measure of cognitive development, which is ultimately of interest because of its link to economic returns. Accordingly, early human capital research provided evidence of the role of education as an economic determinant. Specifically, education economists showed that cognitive ability was an important determinant of labour market outcomes (Heckman, 1995). These outcomes were calculated via their rate of return (RoR), and education economists demonstrated that the RoR to schooling was a critical factor in individual income as well as GDP (Harmon and Walker, 1995). Evidence that estimated labour-market returns to education often used measures of cognitive ability such as test scores or years of schoolings in earnings regressions (Rosenzweig and Wolpin, 1994)[4]. For example, researchers such as Denison (1985), Barro (1989), and Becker (1964) used years of schooling

[4] For this paper, I have also therefore selected to use test scores as a measure of cognitive ability, following the precedent set by the wider literature.

to 'explain' variance in per capita earnings distributions in a variety of education contexts. This body of literature demonstrated that rates of return to education rose as the quantum of education increased (Ben-Porath, 1967).

Accordingly, the application of evidence on *education generating future returns* to decisions on improving earning capacity is often how investments in education have been considered (Ben-Porath, 1967). Logically, questions were therefore asked about how to best maximise those investments. The second notion, that *early investments maximise academic attainment*, stems from this inquiry. This has been explored through the human capital framework in a number of contexts, particularly in the research of James Heckman and Flavio Cunha, who argue strongly in favour of early investments in human capital.

The original notion of early investments was predicated upon the idea of a 'time profile of investment' in which early investments offered a longer period to realise future returns (Mincer, 1958; Becker, 1962 and 1964). Furthering this idea, Cunha *et al.* demonstrated the value of looking at childhood as two distinct time periods, showing that the RoR to a dollar of investment made in early childhood was higher than for the same dollar invested later (2006). This dichotomisation was valuable because it showed that children's outcomes were built upon earlier acquisition of skills. Cunha specified this as a *production function for skills*, in which skills beget skills through self-productivity and complementarity: self-productivity meaning that early skill attainment raises later skill attainment, and that early investments (in education) facilitate the productivity of later investments (Cunha *et al*., 2006). With this lens, pre-literacy, pre-numeracy and other skills developed specifically during the preschool years were shown to lay the groundwork for higher academic attainment later on. Positioning the case for ECE within an economic framework was therefore progressively compelling.

In addition to offering a useful continuum linking early educational investments to economic outcomes, the value of human capital's approach to ECE comes also from its endorsement amongst actors in international ECCE policy, because this endorsement means that research undertaken from this perspective is easily linked to a wider group of studies concerned with the improvement of international and national ECE policy.[5] In this way, the human capital

[5] Endorsement has reached levels as high as the World Bank, which depicts a linear relationship between investments in pre-primary education and improved rates of return across a range of socioeconomic outcomes (Woodhead 2006). Such institutions and their partners have positioned education as the main vehicle for

framework helps to position this paper because it also informs the policy narratives around poverty alleviation in today's context. Accordingly, adopting this perspective helps to clarify the motives of this research, and also to imbue these research questions with both meaning and purpose.

Adopting this framework as a conceptual foundation for this paper does, however, constrict the interpretation of 'valuable outcomes' for education, as this framework has been criticised for casting an instrumental view of a child as a resource belonging to the state, whose educational value is merely his or her potential economic contribution (Woodhead, 2006). However it should be noted that I have made use of this framework to direct the most appropriate line of inquiry, methods, and comparison to related literature, rather than to exclude all other views of the child and of childhood, which add much value to wider discussions on the range of social benefits of education.

2.2 Human capital and women's education

Whilst human capital theory does not offer much on gender and ECE specifically, its approach to gender and wider education is strong. The writings of Paul Schultz in particular advocate for the importance of educating women, especially as a strategy for poverty alleviation (1993). His arguments served as the foundation of other work which endorsed this perspective, and which took an instrumental view of women's education (see King and Hill, 1993 from the World Bank, which has often been the bastion of applying human capital theory to development practice, as well as Herz and Sperling, 2016, who compiled much of the evidence on gender and schooling from this viewpoint). Other major thinkers on human capital and education have also endorsed this perspective; Erik Hanushek (2008) highlighted that gender equality in education was a human capital investment with important economic outcomes, and Harry Patrinos (2008) undertook specific econometric methodologies to calculate the rates of return to women's education.

The implications from the related body of literature are that women's education is instrumental to a range of social benefits which extend over the course of generations (Unterhalter, 2007). Though human capital's instrumentalist perspective of women's

development, and are increasingly highlighting early education as the prize lever (Becker, Murphy and Tamura, 1990).

education obscures rights-based and social justice arguments, it supports the relationship between women's education and economic growth. It therefore helps to frame gender's role within this paper, as gendered inequalities during the preschool years could precipitate inequalities with major socioeconomic implications.

2.3 A meaningful framework

As a framework, HCT posits that *early investments in education boost overall educational attainment*, and that ultimately *education generates future returns.* Therefore, whilst my research approach did not involve the econometric methodologies which have characterised the development of these theories, the conceptual framework of HCT informs and underpins this paper with these two notions, and helps to connect it to the vast literature on early investments in human capital. It also provides the foundation for this paper's hypothesis by theorising that ECE should 'work'. Therefore, any study which does not yield such results ought to raise questions as to why. This is an important point within the context of this paper, which queries whether ECE practice in LAMI contexts might be informed by irrelevant evidence, thereby impeding improvements in practice and hindering countries back from achieving important development goals.

In addition to offering theoretical parameters, the related human capital literature also offers useful methodological parameters. Nearly all of the research informed by this theory explores the relationship between ECE and its outcomes as an empirical problem, which necessitates statistical methods to solve. Following on from the literature, I adopted a similar methodological approach, which will be detailed in Section 6.

3. Literature review

A review of the existing literature on early years education can be enumerative, and the previous section has covered evidence on the economic value of early investments in education. The aim of this section is therefore to highlight other trajectories of literature on ECE which frame the research question at hand.[6] This includes the ECE policy context, gaps within the existing evidence body, an overview of research from India, and highlights of related literature on gender and maternal education.

3.1 ECE and its related policy context

ECE's role in development policy initially stemmed from the wider framework of ECD. The origination of ECD as an area of study was neurological research highlighting the rapid pace of brain development specifically during the first years of life (Karoly *et al.*, 1998 and Young, 2007). The amalgamation of other important multidisciplinary evidence on child development from neuroscience, psychology, sociology and health collectively highlighted the importance of the early years for cognitive development, and shortlisted a number of critical inputs including education[7] (Campbell *et al*., 2001; Shonkoff and Phillips, 2000; Heckman, 2011 and Kohlberg, 1968). A landmark example which in particular helped to catalyse global policy engagement was an ECD-focused series in *The Lancet*, which quantified the loss of development potential and impact on long-term outcomes for children who lacked strong starts in education, nutrition and material stability. Consequently, ECD became tied to international agendas proposing human capital approaches to development. Many studies on ECD captured 'development' indicators through measures related to schooling, such as test scores or grade completion. Accordingly, policies and frameworks came to include ECE as an important lever for development (Un.org, 2018).

Studies of ECE found consistently that early education helped to shape opportunities across the life course, including improving educational attainment, earnings and market

[6] This review does not cover the human rights, developmental, or social and cultural perspectives on early childhood, though acknowledges that they have contributed significantly to the development of a multifaceted understanding of early childhood, and to important research on this topic.

[7] The effects of ECE are not limited only to cognitive outcomes, nor is its purpose solely instrumental; however, these are the relationships this paper explores. For further reading on ECE's impact on non-cognitive skills, see Cunha *et al*., 2006, or Gormley Jr. e*t al*., 2011.

competitiveness (Carneiro and Heckman, 2003; Cunha *et al.*, 2006; Heckman, 2011; Becker, 1993 and Cunha and Heckman, 2007)[8]. These studies showed evidence of returns as high as 7-13%, or $6-$17 per dollar invested, outstripping investments made in any other period of childhood (see Garcia *et al*., 2016; Singh and Mukherjee, 2017 and Britto, 2018). In this way, ECE came to have its own significance for development agendas and came to be linked to poverty reduction and socioeconomic mobility (Nadeau *et al*., 2011). Summarised neatly by Woodhead (2009), ECE came to be considered the single most effective intervention for correcting the intergenerational poverty cycle.

By the year 2005, ECCE's role in offsetting social and economic disadvantages had led the World Bank to finance loans to over 50 LAMI countries specifically for ECE and ECD programmes (Vennam *et al*., 2009). Meanwhile, the Education for All initiative inscribed ECCE as its 'bedrock' (Strong Foundations: Early Childhood Care and Education, 2006). Additionally, under the Sustainable Development Goals (SDGs), investment in ECE was seen not only as a social aim tackling inequality but also a means to achieving other critical goals on poverty alleviation, economic development and equalising opportunities (Richter *et al*., 2016 and Morabito, Vandenbroeck and Roose, 2013). These highly prominent initiatives and policies spurred a number of LAMI countries to initiate their own provisions early childhood education.

Therefore, the literature relating ECE to economic development policy frames this paper's research by offering evidence linking ECE to outcomes essential for progress toward benchmarks encompassed by the MDGs and SDGs. Specifically, it underscores the role of early education in facilitating the levels of cognitive attainment which are linked to important development goals.

3.2 Gaps in evidence

The literature on ECE is vast and substantive, though primarily from developed country contexts. Longitudinal studies from the US have found ECE to be a successful predictor of both academic attainment and economic outcomes (Currie and Thomas, 1995 and Campbell and Ramey, 1994). Two American studies in particular that are widely cited as foundational

[8] A number of studies on the effects of early intervention programs on economic outcomes is summarised neatly by Cunha (Cunha *et al*., 2006).

evidence are the Abecedarian Study (Campbell and Ramey, 1994) and the HighScope Perry Pre-School Project (Currie, 2001 and Heckman, 2011). A similar study in the UK, the Effective Provision of Preschool Education Study, affirmed the positive effects of preschool on children's cognitive development amongst 3,000 3-7 year-olds (Sylva *et al*., 2011). The evidence drawn from these studies is widely cited because the studies involved experimental research; therefore, their results are interpreted as 'causal'. Grouped with this evidence were the ongoing results of two landmark ECCE programmes, HeadStart, launched in the USA in the 1960s, and Sure Start, its British counterpart (Woodhead, 2006). This group of studies confirmed that ECE improved cognitive abilities and served as a strong foundation for academic success later on (Black *et al*., 2016). Such research was marked by the use of test scores to measure children's intellectual ability, such as well-established local examinations or more internationally used tests such as the PIAT (Peabody Individual Achievement Test) (Rosenzweig and Wolpin, 1994 and Harmon and Walker, 1995). For example, Attanasio, Meghir and Nix used PPVT scores and maths test scores and found them to have signal-to-noise ratios of 40% and 68% respectively, indicating that they were useful and informative proxies for measuring cognitive ability (2015). Such studies also concluded that the most effective time to invest in education in order to equalise initial differences in endowments was in the early years (Currie, 2001). In fact, a large portion of the research on ECE underpinned by the human capital framework tested the hypothesis that interventions in early education could remediate circumstantial disadvantages and serve as an equaliser of opportunities, particularly for children in poverty (Woodhead, 2006).

However, though comprehensive, the evidence being primarily from developed countries is potentially problematic (see Sylva *et al*., 2011; DeCicca and Smith, 2013; Heckman and Masterov, 2007; Engle *et al*., 2007; Belfield, Nores and Schweinhart, 2006; Magnuson *et al*., 2004 and Goodman and Sianesi, 2005 for further examples). In fact, despite policies tying ECE to development goals, there has been a noted lack of evaluation of preschool programs in developing countries (Currie, 2001 and Woldehanna and Gebremedhin, 2012)[9]. This represents a gap in the literature on this topic and points to the existing hegemony of

[9] The noted paucity of global evidence on ECD outcomes has given rise to an important World Bank initiative entitled SABER-ECD, which now collects, analyses, and disseminates global ECD related information for the use of policy-makers and other parties.

evidence.[10] Accordingly, today's literature landscape for ECE can be characterised as complex and substantive, but still immature in evidence emanating from LAMI countries (Woodhead, 2006). The implications of this evidence gap are that policies for international development using ECCE as a lever are being informed by findings from contexts which are potentially very different from those for which the policies are intended.

3.3 ECE literature from India

This evidence gap extends to ECE research on India, which is extremely limited (Alcott *et al*., 2018; Kaul and Sankar, 2009; Chopra, 2012; Singh and Mukherjee, 2017 and Kaul *et al*., 2017 have all documented this scarcity).[11] Only a small body of evidence on the association between preschool participation and later developmental outcomes on ECE exists (see Arora *et al*., 2006; Datta *et al.,* 2010; Nagajara and Anil, 2014 and Shabana *et al.,* 2013 for examples), but it is cross-sectional rather than experimental, and limited in its examination of longitudinal effects. Most of it also covers only particular regions or states, and is therefore limited in generalisability. For example, Arora, Bharti and Mahajan (2006) linked preschool participation to higher cognitive development, but only for children in Jammu and Kashmir. A few cover wider geographies but are stock-taking studies, such as CECED's 2013 review (Indian Ministry of Women and Child Development, 2013). A small number show the effect of preschool on primary school outcomes such as retention or school readiness (Kaul *et al.*, 1993 and NCERT, 1996). Others examine the comparative effects of government versus private pre-primary, and have shown that private school students outperform (Kingdon, 1996; Tooley and Dixon, 2003; Singh and Mukherjee, 2017 and Pratham, 2010). But overall, the research is still limited, and longitudinal evidence is particularly scarce. Moreover, several researchers have highlighted that India's existing ECE evidence does not delve into more nuanced issues such as ECE quality (Kaul *et al*., 2017). Marope and Kaga (2015), for example, found that the quality of the educational intervention impacts how beneficial it is, but not much research has taken this further.

[10] One of the few exceptions includes the compendium published by Engle and colleagues highlighting the status of ECD globally, though this covers not only education but moreso and more often related issues such as health and social welfare (Britto, Engle and Super, 2013).

[11] There is however a small but growing body of empirical evidence from LAMI countries showing high quality ECE generates improved labour market outcomes and long term cognitive abilities; see Grantham-McGregor *et al*., 2007; Rao *et al*., 2013 and the recent 2016 Lancet series on evidence of economic and health outcomes linked to early years investments from a range of LAMI countries (Engle *et al*., 2011).

Aside from the comprehensive (but geographically limited) data from YL India, the only other sizeable examples of evidence on ECE are from the ASER Centre, which publishes an 'Annual Status of Education' report (Early Childhood Development: informing policy and making it a priority, 2018) and, most recently, the India Early Childhood Education Impact (IECEI) study on the impact of preschool on school readiness in Assam, Rajasthan and Telangana. However, even the IECEI is limited in its ability to offer long-term evidence as it only covers a four-year period.

Perhaps closest to my research is the very recent study by Singh and Mukherjee (2018), which takes YL India data to examine the effect of private preschool on both cognitive achievement and subjective wellbeing at age 12, but uses propensity score matching where this paper utilises OLS regression. While some of the findings overlap, the conclusions and their implications vary slightly. Additionally, as this paper looks only at private provision of ECE, it is not able to comment on ECE in India more widely. In total, therefore, the literature on India is characterised by scarcity and requires further substantiation of the evidence base.

3.4 Evidence on gender and maternal education

The literature on gender in India shows that disparity begins early and persists throughout life, on indicators ranging from infant survival to labour market participation. Additionally, whilst India used to rank well on the World Economic Forum's Gender Parity Index, in the last year or two it has slipped over twenty places to well below the global average, indicating a persistence of this issue (The Global Gender Gap Report, 2017).

Evidence of gender bias at all levels schooling, including preschool, has been cited in numerous papers (Kaul *et al.*, 2017). But other evidence on India suggests that this disparity is regularising; official statistics show that primary enrolment has a 1.167-to-1 ratio of boys to girls, and almost identical enrolment for boys and girls in ECE (Data.worldbank.org, 2018). Historically, gender has been a regular predictor of disparity in educational and economic outcomes in India (Asadullah and Yalonetzky, 2012 and Boserup, 1970). However, in some studies, including some using YL data, gender has not been found a significant contributor to academic outcomes (Streuli, Vennam and Woodhead, 2011 and Vennam *et al*., 2009). Meanwhile looking specifically at early education, the evidence argues that gender may affect the extent to which children benefit from ECE (Magnuson *et al*., 2016 and Garcia,

Heckman and Ziff, 2018). The results pertaining to gender's role in academic outcomes, and to disparity in education, are therefore mixed. And whilst India has made progress toward gender parity in education, the evidence is still conflicting as to whether or not it has been achieved[12]. However, despite the debate as to whether or not gender parity in education has been reached, almost every available study on Indian education disaggregates information by gender. Moreover, the juxtaposition between gender as a historical axis of disparity and contemporary evidence showing the gap is closing highlight the need to continue to monitor the situation of gender in Indian education closely; accordingly, I have included gender as a control variable within this study.

Tied to the notion of gender is that of maternal education. Rightly or wrongly, women's education is often discussed in the literature for its instrumental value in intergenerational outcomes. There is a significant body of literature stemming from both sociology and economics that supports evidence of the intergenerational persistence of economic status (see Solon, 1999; Hauser and Logan, 1992 and Asadullah and Yalonetzky, 2012 for examples). In tandem, it is now widely recognised that parental education has a role to play in breaking intergenerational poverty cycles. The literature shows that more educated parents have children with higher educational attainment and ultimately improved life outcomes, even when other factors such as socioeconomic status are controlled for (Black, Devereux and Salvanes, 2005). Mothers' education in particular has been shown to be associated with higher earnings and better educational outcomes for children (Aakvik *et al*., 2003 and Rosenzweig and Wolpin, 1994). For example, longitudinal data from the National Longitudinal Survey of Youth uses PPVT and PIAT test scores to show that maternal education and children's academic outcomes are significantly related (Rosenzweig and Wolpin, 1994 and Leibowitz, 1974). There are even studies using data from YL India which have shown significant association between maternal education and the completion of secondary school, which itself is linked to improved economic opportunities in other research (Singh and Mukherjee, 2015). As with research on ECE, a number of these studies on the intergenerational impact of education come from from US and UK contexts (see Dearden *et al*., 1997 and Mulligan, 1999), but there is also research from LAMI countries on the strength

[12] The international importance of measuring gender's impact is also evidenced by the regular disaggregation of school outcomes such as enrolment, grade completion, and test scores by gender at the highest levels of global research in education (Handbook on Measuring Equity in Education, 2018).

of this link, as measured by impact of maternal education on school readiness (Kaul *et al*., 2017).

This body of evidence supports the idea that maternal education has a role to play in the intergenerational transmission of human capital (Galab, Reddy and Himaz, 2008 and Richter *et al*., 2016). Evidence on maternal education's relation to cognitive outcomes therefore has implications for policies relating to human capital growth (Rosenzweig and Wolpin, 1994 and Black, Devereux and Salvanes, 2005). It has therefore been included as a variable of interest within this paper.

In summary, there is a well-established multidisciplinary body of literature on the early years as a formative period for cognition and learning and as a key input for economic outcomes. However, these outcomes have been cited in development discourse and policy as particularly desirable for developing economies, but the evidence may not be pertinent to those contexts. Therefore, evidence that more closely examines the role that ECE plays in such contexts would be valuable.

4. Context of ECCE in India

4.1 Overview of policies and infrastructure

The provision of early childhood services in India has long been conceptualised as an investment in human capital (Streuli, Vennam and Woodhead, 2011 and Mohite and Bhatt, 2008). This notion of ECCE informed India's National Policy for Children and its launch in 1974 of the Integrated Child Development Services (ICDS), today the world's largest publicly funded early childhood system. Through the ICDS, India provides almost universal access to health, nutrition and education services (Alcott *et al.*, forthcoming and Indian Ministry of Women and Child Development, 2013). The operating infrastructure for these services includes 1.3 million *anganwadi* centres (preschools). Nearly every village has at least one *anganwadi* (Kaul *et al*., 2017), and together the system served over 104.5 million beneficiaries in 2014, ranging from expectant mothers to children aged 6 (Richter *et al*., 2016).

The system is appropriately large, as India hosts the largest population of school-going children globally (Streuli, Vennam and Woodhead, 2011). But pre-primary enrolment in India is still only at 12.9% despite the sizeable infrastructure (Data.worldbank.org, 2018). Moreover, evidence shows that the *anganwadis* are poor providers of early childhood education, due to insufficient teacher training, substandard facilities and lack of regulation[13] (Rao and Kaul, 2017 and Singh and Mukherjee, 2018). Poor quality pre-primary provision has precipitated widespread lack of school readiness: evidence indicates that pre-literacy and pre-numeracy skills at age 5 are vastly below expected levels (Save the Children, 2009 and Beyond Basics, 2018). Meanwhile, educational demands in India's burgeoning economy are quickly outpacing the services *anganwadis* can offer (Streuli, Vennam and Woodhead, 2011). As a result, disillusioned parents are seeking other options such as low-cost private preschools (Kaul *et al*., 2017 and Alcott *et al.*, forthcoming).

Private preschools range from low-cost to highly-priced, but almost all promise English medium instruction, which parents see as a path to upward socioeconomic mobility and better

[13] *Anganwadis* are noted to have poorly trained workers, with the second-in-commands having either no education or only very basic. Moreover, qualitative research from YL shows that teachers generally seem disinterested and disengaged.

school preparedness (Indian Ministry of Women and Child Development, 2013). This has attracted even socioeconomically disadvantaged families[14], though research shows more boys than girls are sent to private preschool when resources are scarce (Streuli, Vennam and Woodhead, 2011).

However, despite widespread uptake, private schools have been found to further social inequities and have been deemed developmentally inappropriate for children, as their pedagogy and curricula tend to mirror what is used in primary school (Singh and Bangay, 2014 and Indian Ministry of Women and Child Development, 2013). Parents are therefore limited in options, as private provisions offer better-equipped schools in English medium but are often unregulated and developmentally inappropriate; yet government-backed public preschools often lack suitable infrastructure, regularity of supervision and quality programming (Streuli, Vennam and Woodhead, 2011 and Kaul *et al*., 2017). These challenges present major problems with quality that transverse ECE providers across India.

The Indian government has recognised shortcomings of the ICDS and has made improving ECE a priority (Streuli, Vennam and Woodhead, 2011). Progress has been made by strengthening the policy framework; India's twelfth Five-Year Plan (2012-2017) has shifted attention toward early education (Singh and Mukherjee, 2017). Additionally, the government launched the National Early Childhood Care and Education (ECCE) Policy in 2013, specifically to expand the educational component of the ICDS, followed by the National Curriculum Framework and Quality Standards, which together are meant to serve as a suite of policies and guidelines for the implementation of quality ECE nationwide (Kaul *et al*., 2017). However, while these demonstrate the political will to restructure and enhance the ICDS, the system still suffers from lack of resources, direction, and governance (Richter *et al*., 2016). Consequently, there remains a large gap between policy and practice (Alcott *et al*., forthcoming).

[14] For example, double (67%) as many of the younger cohort from YL India attended government ECE as attended private ECE (34%), even amongst this more socioeconomically disadvantaged group (Streuli, Vennam and Woodhead, 2011).

Alongside quality issues, gender, caste and class represent three of the major 'axes of stratification' plaguing India's school system from the preschool years[15] (Kaul *et al*., 2017). In particular, gender disparity has been shown to be significant at all levels of schooling and is linked to a number of wider social issues. It has also been identified as an important factor in educational decision-making, with parents often investing more in sons' education, particularly for private schooling (Woodhead *et al*., 2009).

Another problematic phenomenon is noncompliance with the official ages of entry to primary school. Evidence has shown that children from ages 4-8 rarely progress on time to the correct grade, whether due to changing from public to private and being forced to repeat a grade, or being sent to school early by parents keen to earn their children a competitive edge (Alcott *et al*., forthcoming). This is further complicated by incoherence between national and state policies on age of entry to primary: national education policy states that children should be in pre-primary provision at age 5, yet many state policies stipulate age 5 as the minimum age of entry to primary (Government of India, 2009). The resulting nonlinear trajectories experienced by many children lead to exposure to developmentally inappropriate settings, including both curriculum and pedagogy (Indian Ministry of Women and Child Development, 2013).

On the whole, the Indian government has made important strides in education policy for children under 14, but has left out children under 6: Article 21A, passed in 1950, stipulated free and compulsory education for children aged 6-14. Sarva Shiksha Abhiyan (SSA) also prioritised universalising primary learning, but not pre-primary (Vennam *et al*., 2009 and Singh and Mukherjee, 2016). Children under the age of 6 are therefore not covered by a rights-based approach to education, which some have cited as rationale for India's lack of focus on the educational element of the ICDS provision to date.

Therefore, outstanding challenges range from quality to disparity to coordination of services, but lack of research on the specificities impedes the ability to optimise the status quo. However, the enabling policy trajectory to date shows that the government is a willing participant in improvements and investments. This research seeks to respond to this issue by

[15] Combating gender disparity has however been on the government's agenda; it launched the National Programme of Education for Girls at Elementary Level, which helps to free girls from sibling care so that they can attend school (Streuli, Vennam and Woodhead, 2011).

evaluating India's ECE provision and its impact within a particular state context, as well as exploring the roles of two highly important factors determined by long-standing lines of disparity.

4.2 The context of Andhra Pradesh

The state of Andhra Pradesh, the home of this paper's sample, has a complex set of challenges. It has a long-established government-run ECCE system but also a growing trend of private schooling at all levels, and across all socioeconomic backgrounds (Asadullah and Yalonetzky, 2012). Yet students fall vastly below the expected levels of attainment (Singh and Mukherjee, 2016).[16] It also has extreme geographic and socioeconomic dichotomies; whilst Hyderabad is India's IT capital with a wealthy middle class and high-technology job sector, rural Andhra Pradesh is often plagued by agrarian crises (Vennam *et al*., 2009).

In many cases, children in Andhra Pradesh are first generation learners; this was the case for more children than not in the YL India sample (Streuli, Vennam and Woodhead, 2011). This has an impact on maternal education's ability to improve outcomes for this generation. Secondly, the rapid growth of private schools is attracting many parents with English-medium teaching, resonating with wider national patterns. The knock-on effects in Andhra Pradesh are interestingly that rather than increasing total enrolment, children who were already attending *anganwadis* often transition to private ECE (Alcott *et al*., 2018). Meanwhile, whilst poverty levels are usually indicative of preschool type, the rise of private options has actually created intra-class stratification, with some poor families completely unable to send children to private pre-primary, and others just about managing to do so. Where this is possible, it is usually for sons. The government of Andhra Pradesh is therefore under pressure to compete, and is now moving toward English-medium instruction in some secondary schools, though primary and ECE remain Telugu-medium (Streuli, Vennam and Woodhead, 2011).

Aside from English instruction, many parents in Andhra Pradesh are attracted to private pre-primary because of disappointment with *anganwadi* provisions, a dissatisfaction that appears

[16] Comprising just under 10% of India's population, Andhra Pradesh is largely rural, but importantly is home to the capital of India's IT sector, Hyderabad. Telugu is its major language, spoken by 85% of the population (Vennam *et al*., 2009).

characteristic of tendencies elsewhere in India (Streuli, Vennam and Woodhead, 2011). This dissatisfaction is even apparent amongst the poorest quintile of families, with less than 50% of parents in this quintile believing that the *anganwadi* offering was good or excellent (Streuli, Vennam and Woodhead, 2011). Consequently, this attraction toward private preschools means that many children end up attending more than one preschool, with the majority going in the direction of government to private.

Overall, educational disparity in Andhra Pradesh is decreasing (Asadullah and Yalonetzky, 2012). However, the trends discussed above show that education in Andhra Pradesh is characterised by differentiation of preschool experiences by wealth status, urban versus rural location, gender, and parental education (Streuli, Vennam and Woodhead, 2011). These provide an important backdrop for this paper and justification for the inclusion of gender and maternal education as explanatory variables alongside preschool participation.

5. Data

5.1 Construction of the sample

This paper draws data from the Young Lives (YL) study in India, which surveyed families from undivided Andhra Pradesh (later Andhra Pradesh and Telangana). The sample used for the purpose of this study is the younger cohort, which was comprised of 2,000 children aged 1 at the start of the survey in 2002[17]. The younger cohort was selected because information for this group on both preschool participation at age 5 and test score results at age 12 was collected. Within this sample, only children who participated in both Rounds 2 and 4 were included, as critical data was needed from both rounds (gender, preschool participation and mother's level of education from Round 2, and test results from Round 4). Finally, my sample included only children for whom responses to the key predictor variable of preschool participation were recorded, as this allowed the analysis to reflect more accurately on relationships between the variables in question. Though constraining the sample in this way could have introduced some bias, the non-response count on this question was low, and it was deemed important for the accuracy of the statistical models. These constraints left a final sample size of 1,665 children. Table 5.1 below offers descriptive statistics on all variables.

Table 5.1: Frequencies and percent of sample for each input variable

Variable	*n*	%
Gender		
Male	890	53.5
Female	775	46.5
Total	1665	100
Preschool		
Participated	946	56.8
Did not participate	719	43.2
Total	1665	100
Mother's Education		
Some education	836	50.2
No education	826	49.6
Total	1665	100

[17] The exact original sample sizes were n=2,011 in Round 1, n=1,950 in Round 2, n=1,930 in Round 3, and n=1,915 in Round 4 (Singh and Mukherjee, 2016).

***Source**: Young Lives data, Round 2*

5.2 Advantages and limits of secondary analysis

There were several considerations that led to the selection of a secondary data set for this paper. The first was size: the practical consideration that in order to perform robust quantitative analysis, a certain sample size was desirable which would not have been achievable on an individual scale and with a time constraint. Additionally, in order to have subgroups of a large enough size to compare effectively, the original sample would need to be significantly large. A smaller or less diverse sample could be more prone to outliers, bias and other issues less often present in a large sample. Moreover, the sample size helped to ensure that variables discussed in this paper were normally distributed, which is optimal for quantitative analysis.

The second consideration was quality: YL was a reputable study with a large research team. Its data has been used by an array of researchers, and the World Bank has recently commissioned a study of early childhood education based on its data. As education research in India has at times been hampered by unreliable data (Kingdon, 1996), YL's study offered rare information collected with rigorous sampling and response management procedures, resulting in good quality data with minimal attrition (Singh and Mukherjee, 2016 and Young Lives, 2017).

Thirdly, making use of YL data for this paper allowed for the study of effects over time, as it was a longitudinal study. As educational outcomes can take years to manifest, and as analysis of the relationship between education and human capital development (with educational attainment serving as an indicator) is usually looked at over the span of decades rather than years, the immediacy of a cross-sectional study would not have sufficed as a sample for this research question (Bryman, 2014)

However, whilst there are definitive advantages of utilising secondary data, there were also limitations to consider. The first was a lack of firsthand familiarity with the data which is ideal for social science research (Bryman, 2014). Particularly because I only made use of YL's quantitative data (the qualitative data is not made available publicly), I could not engage as primary researcher would nor nuance the empirical results with qualitative

analysis. However, this limit is perhaps mitigated by the process of quantitative analysis itself, which prioritises objectivity in order to see patterns more clearly. Additionally, because this paper's interest is in broad indicators of concepts and how they interact, a lack of familiarity is a bias, but the subjectivity of firsthand research has its own biases in terms of replicability and subjectivity.

The constraints of this paper also include limitations common to longitudinal research, such as sample attrition. Children whose families had migrated, or who were not present on the days during which various elements of the survey were conducted, would not have been included in subsequent rounds. In order to mitigate against this limitation, I used a restricted sample which included only children whose information on critical variables was recorded from both rounds, and checked that those with missing information had no significant difference on key statistics.

Issues with generalisability also present a disadvantage. YL's information is restricted to only one state in India of a total of 29. As each state has its own geographical, socio-political, economic and historical makeup, analysis from Andhra Pradesh alone will not be generalisable to the whole of the Indian population (Alcott *et al*., forthcoming).

Bryman (2014) discusses two other major concerns with utilising secondary data: complexity and quality. Whilst the YL study does have analysis organised across community, household and individual levels in multiple cohorts and waves, it is in fact the complexity of this data that makes it fit for purpose for many avenues of inquiry, including mine, particularly as such minimal data on one sample is available for all of these variables. Regarding quality, because YL has been such a well-organised and well-funded study, this is essentially a non-issue. These latter concerns can therefore be dismissed.

5.3 Variable selection and description

The variables I selected are intended to be indicators of the key concepts I wish to analyse. The intention is to examine how these variables interact in order to draw conclusions that are meaningful in a broader context. As the research question is whether early childhood education affects later cognitive achievement, and if there are other factors at play that could be moderating that relationship, variables from the YL data have been selected that are as

close as possible to these concepts. The primary dynamic to determine is that between early childhood education and academic outcomes; however, gender and maternal education have also been included due to widespread evidence documenting their influence in the Indian context.

5.3.1 Early Childhood Education (ECE)

The core concept of this paper is early childhood education. It is therefore the primary explanatory variable of this study[18]. ECE refers to regularly attended education in a setting or establishment outside of a child's home for children approximately aged 3-5, during the period immediately preceding primary school (Woldehanna and Gebremedhin, 2012).
In the YL surveys, experiences of early childhood education were measured through a group of questions answered by parents and/or caregivers. These included whether their children were currently participating in preschool, whether their child had attended a creche before age 3, if the child had attended preschool at any time since age 3, reasons for attending preschool, and, if applicable, reasons the child stopped attending preschool[19]. Of these questions, the first provided the information closest to what the variable 'preschool participation' is trying to estimate in this paper; that is, if the child had participated in preschool, though responses were not broken down by preschool type (which would have been valuable, and was made available to other YL researchers such as Singh and Mukherjee). Therefore it was this question's responses that were used for the explanatory variable. 'Preschool' was thus coded as a binary variable for this study, with responses being either yes or no[20]. In my sample, 56.8% of children were attending preschool during Round 2, when they were approximately aged 5.

However, it is important to note that local educational policies have different ages of entry for primary school. In the case of Andhra Pradesh, it is possible that children surveyed in Round 2 were eligible to attend both preschool and primary school, being of the age of 5. This dichotomy is indicative of the previously mentioned phenomena of early entry and nonlinear

[18] It should not be assumed that I consider ECE to be the only explanatory variable accounting for cognitive development. Variance not explained by these models or other similar models could be due to genetic endowment (see Todd and Wolpin, 2004), parental education, SES, or other inputs.
[19] In the YL questionnaire 'preschool' referred to an *anganwadi*, a private preschool, or another similar early childhood education centre, but was not readily disaggregated.
[20] See Meghir and Rivkin (2011) for a discussion of methodologies for binary education choice, as analysed by Heckman, LaLonde and Smith, 1999.

trajectories between preschool and primary.[21] Therefore, this duality of eligibility posed a challenge to the estimation of this variable, as some students may have entered primary school early, which made it more difficult to ascertain the degree to which preschool participation impacted their academic performance. Eligibility aside, it is also common practice for parents to send children to primary school early to 'get ahead'. Therefore though 'is child currently attending preschool?' should have been a straightforward question, bias could have been introduced by responses of 'no' from parents whose children had attended preschool but already moved on. As this bias proved difficult to calculate or minimise, however, responses to this question were taken at face value.

5.3.2 Cognitive outcomes

Cognitive outcomes, as measured by academic performance, were selected as the main outcome variable for this study because of their strong link within the literature to poverty alleviation and other desirable socioeconomic outcomes. Academic performance is therefore intended to be interpreted as a medium-term indicator of whether the return on investment in ECE is likely to be favourable. Therefore, by assessing whether preschool is a good predictor of cognitive outcomes, this paper offers analysis that can help substantiate the evidence base for policy decisions on ECE that could tie into wider national development goals.

Whilst there are a variety of ways of quantifying academic performance, test scores are by far the most common. Most often, test results used in such research are children's scores on comprehension, reading or maths assessments, because research shows young children in education should have the skills which these tests assess (Indian Ministry of Women and Child Development, 2013). As such, test scores are usually selected because they are proxies for numeracy and literacy skills deemed to indicate cognitive achievement[22] (Manji, 2018). I took direction from precedent in the literature by utilising students' test scores to estimate educational attainment. In the case of YL research, results on only a small selection of tests are recorded, of which I used two: one verbal reasoning test (the PPVT, or Peabody Picture

[21] Despite the RTE Act's stipulation that children should enter Grade 1 at age 5-6, only at age 8 does correct enrolment by age stabilise (Kaul *et al*., 2017).

[22] Using test scores on a limited number of assessments could constrict the meaning of 'cognitive ability' as other outcomes from a more diverse range of school subjects are not included; however, maths, comprehension, and reading are usually deemed the most critical skills to assess and also the most comparable amongst various educational studies.

Vocabulary Test), and one maths test (Attanasio, Meghir and Nix, 2015). Scores from two tests rather than one were included in this paper's model in order to offer a broader scope of analysis and comparison.

Internationally, there is a spectrum of tests deemed appropriate for cognitive assessment and comparison, of which the PPVT is one. The PPVT measures receptive vocabulary and therefore serves as a measure of verbal cognitive development, while maths assessment test scores were noted to measure quantitative capability. The YL data recorded these results both in the form of raw scores and in the form of percentage correct. I used score by percentage correct as this is more readily understandable to those unfamiliar with the intricacies of the tests. Accordingly, two different measures of educational attainment were used as outcome variables to represent cognitive achievement: PPVT scores as measured by percentage correct, and maths test scores as measured by percentage correct[23]. Amongst this sample, children's mean score on the PPVT was 75.8% (SD = 13.61%), and mean score on the maths test was 44.6% (SD = 22.7%).

5.3.3 Gender

Literature suggests that ECCE interventions have the potential to compensate for the gender biases that have historically affected Indian school systems (Morabito, Vandenbroeck and Roose, 2013 and Magnuson *et al*., 2016). Consequently, I examined whether gender was associated with preschool attendance for this sample, as well as whether it was associated with variation in score results. I also included gender as a control variable in the multivariate OLS regressions to determine whether it made up any part of preschool's ability to predict results on the PPVT and maths tests.

Descriptive statistics on this paper's sample show that 53.5% of children in this sample were male, and 46.5% were female. 57.8% of boys surveyed were attending preschool, while 55.7% of girls were attending. Of the children who were attending preschool, 54.3% were boys and 45.7% were girls.

[23] See Woldehanna and Gebremedhin (2012) for similar justification using a mathematics and a verbal test type from YL data to determine preschool's effects.

5.3.4 Maternal education

Human capital research has positioned women's education as an intergenerational asset, by which parents' education is a determinant of their children's cognitive development (see Asadullah and Yalonetzky, 2012; Becker, Murphy and Tamura, 1990; Andrabi, Das and Khwaja, 2012; Magnuson, 2007 and Black *et al*., 2005). In research from India specifically, mother's education has been shown to be positively associated with children's educational outcomes (Jeong, Kim and Subramanian, 2018). Accordingly, including maternal education in the estimation models was based on the established literature highlighting its importance as an input in intergenerational educational and economic outcomes.

Construction of a variable to represent 'maternal education' necessitated interpreting the data YL gathered during Round 2 on household education statuses. The survey asked numerous questions about the education status in each household, including caregiver's, father's and mother's education levels, with responses measured by highest grade achieved. However, descriptive statistics on mothers' education levels showed that just under 50% responded with 'none', indicating no formal education whatsoever[24]. Taking the lead of other research such as the IECEI, which grouped mother's education into larger brackets (Kaul *et al*., 2017), I recoded this variable as binary for the purposes of this study, with mothers falling either into having 'some' education or 'none'. This allowed for a clearer picture of the role of maternal education.

One fundamental issue in estimating the effect of maternal education is the challenge of dealing with unmeasured abilities, or 'endowments', of children (Rosenzweig and Wolpin, 1994). The problem is that it is very difficult to account for the innate abilities of a child, exclusive of his or her mother's educational influence. Because this is a highly complex matter to calculate, this paper did not account for bias resulting from such endowments. Another issue was not accounting for other education levels in the household (fathers, caregivers). But as the majority of literature on intergenerational educational transmissions in LAMI contexts cites mother's education as a paramount factor, I opted to follow this precedent.

[24] This statistic also corroborates the fact that a large quantum of children in India are still first generation learners (Indian Ministry of Women and Child Development, 2013).

Amongst this sample, 49.6% of mothers reported having no education, while 50.2% of mothers reported having completed some education. Of the children attending preschool during the survey, 43.4% had mothers without any education, and 56.6% had mothers with some level of education.

5.4 Risks and issues in estimation

It is important to note that this paper is concerned with complex social phenomena which are in reality not monolithic. This complexity presents a challenge in identifying variables that can be perfect indicators, and in building models that can comprehensively test the relationships between the selected concepts. The analysis within this paper is also constrained because it relies on the data as it was gathered by the YL team. Therefore proxies have been used to estimate and measure the engagement between these variables as best as possible given the data available. The risk of using proxies is that there exists a gap between the picture painted by empirical analysis and the realities colouring how these variables interact in situ, for individual children. However, as much as possible, I have chosen to select variables that estimate concepts as closely as possible and which are in line with precedent set by the relevant literature.

One particular issue in estimation, discussed previously, was that of non-linear trajectories. A surprising 44% of children were reported in Round 2 of the YL survey to have started primary school even though only 20% had reached the appropriate age (5), which is the legal minimum age of entry (Woodhead, 2006). This issue with data is problematic in and of itself but also points to the problems faced by other researchers working in the Indian education context dealing with challenges of nonlinear pathways and overage children in early childhood and primary education (Singh and Mukherjee, 2016). To illustrate this issue further, data shows that 9% of all children in Andhra Pradesh were overage in primary school (Education-inequalities.org, 2018). This issue could have introduced bias to the findings, as children who were reported as ‘not currently attending preschool’ could have already completed some preschool and moved on to primary, but would have been missed within this paper. Moreover, determining which of them had completed preschool early versus were truly ‘not attending’ would have been complex, and impractical to recode meaningfully. Therefore, for the purposes of this paper, children who responded ‘no’ to ‘Is child currently

attending preschool' were included in the 'not attending preschool' category, but it is important to be aware of this possible risk in estimation.

6. Statistical methods

My initial approach to selecting methods was to explore the methodological possibilities that would best correspond to the research questions. As my primary aim was to estimate the effect of preschool on academic achievement, quantitative methods that employ Null Hypothesis Significance Testing (NHST) were deemed the most suitable option (Field, 2009). Whilst a number of methodologies might have been applicable, I selected three based upon precedent[25], constraints from the data and the possibility for further research to build upon this paper.

Firstly, I compared the mean test scores of the major groups (preschool attendance versus non-attendance, boys versus girls and children with uneducated versus educated mothers). This offered an initial picture of any significant differences amongst the groups and helped to set expectations for results from the multivariate analysis. Secondly, I conducted linear regression on preschool participation and test scores as an uncontrolled estimation of preschool's ability to predict score outcomes. Thirdly, I utilised multivariate OLS regression to estimate preschool's ability to predict score outcomes, but with gender and mother's education as conditions.

6.1 Bivariate testing

6.1.1 Overview of bivariate methods

I formulated the research problematic according to NHST formatting to explore the likelihood of a change in the predictor variable affecting a change in either or both of the outcome variables, with 'no effect' being the null hypothesis in each situation and a difference in means or frequencies representing the alternative hypotheses. The bivariate analysis involved a number of tests, including Chi-square and t-tests, as well as looking at

[25] Meghir and Rivkin (2011) provided a useful review of some of the methodologies used in the economics of education, including a discussion of models such as the Mincer model for estimating returns to education, and other methods that have been used in assessing the impact of education quality. A number of these methods have been used to identify 'causal' relationships, but as this paper is not evaluating the results of experimental or quasi-experimental research, these models would not be appropriate to replicate in this context. However, they offer important conceptual background.

resulting values for r, r^2, and other measures of effect size. The level of statistical significance used for each test was a p value of < .05. All tests were two-tailed tests.

6.1.2 Chi-square tests

I used Chi-square tests in instances for which variables were categorical: gender and maternal education[26]. The premise was to determine whether these predictor variables covaried with preschool participation using contingency tables. If this was found to be the case, and if either of these also covaried with scores, this would support the need to build a multivariate model which controlled for these variables.

6.1.3 Use of t-tests

Independent samples t-tests were also used to determine whether the independent variables (preschool participation, and subsequently gender and maternal education) were associated with test score outcomes by comparing the mean scores for each group of children. Because all predictor variables were dichotomous nominal variables, I recoded each pair as dummy variables (Field, 2009). A test statistic showing a significant difference in means between groups, coupled with an effect size of a particular level, would indicate that the explanatory variable could help to predict test score outcomes.

6.1.4 Measuring effect size

One of the challenges cited with using NHST is that 'statistical significance' does not always correspond to actual importance (Field, 2009). In order to gain a more calibrated understanding of the magnitude of the observed effect, effect sizes using either Cohen's d or Pearson's r were also examined. The magnitude of each effect was determined according to the widely used parameters set out by Cohen (1988). Allowance was made for biserial or point biserial correlation (Field, 2009).

6.2 Multiple regression using Ordinary Least Squares (OLS)

[26] Yates' correction was applied to the tests run to mitigate against Type 1 errors, as all variables mentioned above were dichotomous. (Field, 2009).

I selected linear regression as the main methodology for this paper due to its versatile functionality in summarising relationships between variables, and to extensive precedent in the literature[27]. An added benefit of linear regression was that estimating the coefficients in this model elucidated the direction, strength, and significance of each relationship, with the resulting coefficients measuring how variance in the predictor variable could predict variance in the outcome variables[28]. Accordingly, this was a method which offered a relatively uncomplicated way to interpret results. Lastly, the nature of the regression models would allow further research to build easily upon the models within this paper; adding additional explanatory variables to the model from YL data to see how they change the ability of the model to predict the outcome of interest could offer a straightforward plan for taking this analysis further.

Specifically, I utilised Ordinary Least Squares (OLS) to estimate the coefficients of the various models[29]. Initially, this included estimating the relationship between preschool and test scores, and I followed this with the creation of multivariate linear models to re-summarise these relationships whilst controlling gender and maternal education. The method of least squares was used to estimate the parameters associated with each of the explanatory variables. Each model was then assessed for its goodness of fit to determine whether it improved how well test scores could be predicted.

6.3 Assumptions and bias

There are a number of important assumptions for linear regression, as the relationships between assumptions and the behaviour of variables is important for the ability to construct conclusions. Whilst the majority of these assumptions were met (linearity, homoscedasticity, normally distributed errors and multicollinearity), there remained the assumption of

[27] OLS regression has often been used by studies exploring relationships between similar variables. One of many examples was the use of OLS regression to estimate a statistically significant relationship between years of parental education and child's educational outcomes, in which an increase in one year of a parent's education matched an increase by .2-.25 in years of child's education (see Black, Devereux and Salvanes, 2005 and also Jeong, Kim and Subramanian, 2018)

[28] All regression models for this paper were based on linear equations, and the assumptions for linear models were checked for.

[29] Woldehanna and Gebremedhin (2012) used propensity score matching substantiated with OLS and IV regression to estimate the effects of pre-school attendance on cognitive development at ages 5 and 8 in Ethiopia with YL data, a similar research question. As it uses data from the same wider study on the same topic type, it provided a good reference point for my use of OLS methodology. See Loeb *et al*., 2005 from the US context for other examples of use of OLS on a similar topic.

endogeneity, which in turn affects the ability to interpret any results as causal. The assumption regarding multicollinearity could have presented an issue, given there was a correlation shown between mother's education and preschool, which were both predictors in the multivariate models. However, Field (2012) advises that this assumption is only problematically violated if any of the input variables correlate highly, that is, they have a correlation of above .8 or .9. As this was not the case for these variables, this assumption held. This left only omitted variable bias (a type of endogeneity) and correlation with external variables as issues.

Whilst multicollinearity can be tested for using statistical software and corresponding analysis, endogeneity necessitates application of social science intuition. In the case of omitted variable bias, an unobserved variable related to both the explanatory and the dependent variable can bias the result (Schonemann and Steiger, 1976). This paper included only two variables as controls, maternal education and gender. But literature on ECE highlights socioeconomic status (SES) as an important predictor variable of both preschool participation and academic attainment (Woodhead, 2009). Therefore, SES represents a variable which has been shown to correlate with both the independent and dependent variable. And as it was not included in the multivariate models, it could have confounded results. Other variables such as urban/rural location, nutritional status, birth order place and caste have also been shown to impact academic outcomes and could therefore also have been 'omitted variables' associated with both the independent and dependent variables (Singh and Mukherjee, 2018), but were not included to reduce overburdening the model and overly detracting from the original research question.

However, omitted variable bias is more easily remediated than other kinds of endogeneity, such as simultaneous causality or self-selection. One way to reduce the bias would be to build a multivariate model that includes a greater selection of predictor variables, which could improve the validity of the results by showing, through partial correlations, which coefficients contribute the most to the model's ability to predict test scores, as well as improving the model's overall ability to predict scores, signified by a higher value for r^2 (or a statistically significant Δr^2 value). However, though such a model could offer more information on the roles played by important variables, it still would not be able to deduce causal relationships, as it would be difficult to include all variables which could produce an effect. Indeed, such a model might be overly cumbersome and difficult to analyse, as well as

detracting from the initial questions of interest. I therefore opted to select only two variables which past research has deemed important specifically for the intergenerational transmission of human capital, and intergenerational improvement of quality of life.

Correlation with external variables presented a related issue. As SES is a known correlative of other variables within the models, this represents a violation of the assumption of no correlation with external variables, similar to bias introduced from endogeneity. Accordingly, hypothetical correlation from variables in the error term to both independent and dependent variables could also have led to a distorted estimation of the other parameters in the model. However, problems of correlation within the error term were checked for using the Durbin-Watson test statistic, the result of which was not significant, therefore dismissing issues with this assumption (Durbin and Watson, 1950).

Another way to limit these biases, or at least think critically about them, would be to conduct further research utilising differential effects. This methodology could be used to compare each input variable's explanatory potential, which could help to correct unobserved biases associated with both variables (Rosenbaum, 2006). However, as the original research question was not which explanatory variable is the best predictor of scores, rather how these variables relate, this method was not pursued.

Despite these assumptions, the chosen methods are justified because the normality of the data made a good case for linear regression, and because these methods offered the best possible vehicle for analysis in the context of this specific research problematic and the related literature. However, any conclusions drawn must refer to associations rather than confirmations of causality. In the case of this paper, this also involves recognising that it is likely other socioeconomic characteristics are involved in the bigger picture of who attends preschool and scores well.

7. Results

The main theme explored by the analysis was early childhood education's relationship to academic outcomes. Subsidiary themes included the roles of maternal education and gender in educational attainment. The goal of the regression analysis was to estimate the role of preschool in predicting test score outcomes on tests taken by children in Andhra Pradesh at age 12, first in an uncontrolled and then in a controlled model. The findings show that, for this sample, preschool participation and gender were ineffective predictors of test score outcomes, but maternal education emerged as a significant predictor of test scores on both PPVT and maths tests. Findings are grouped below by the original research questions.

7.1 Was early childhood education a good predictor of test scores?

Initial results indicated that preschool participation did not have a statistically significant impact on PPVT scores, but did on maths scores, yet with only a small effect size. Figures 1a and 1b below depict mean scores on each test by children who did and did not attend preschool.

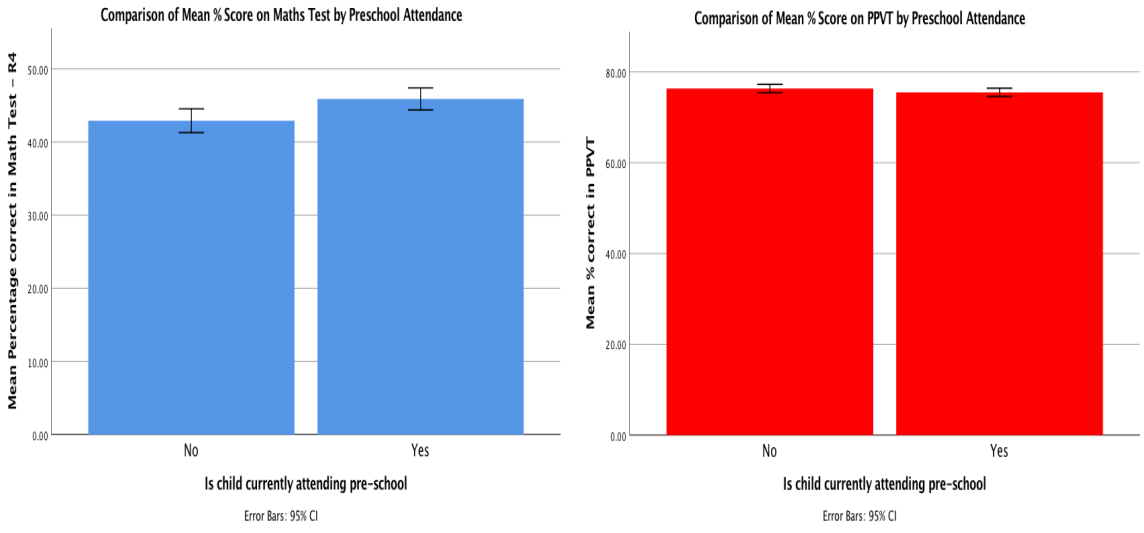

Figure 1a: Comparison of mean percent score on Maths test by preschool attendance (L)
Figure 1b: Comparison of mean percent score on PPVT by preschool attendance (R)
Source: Young Lives India, Rounds 2 and 4

These results suggest that a further level of analysis may be needed to understand why they contradict the existing body of evidence on ECE. Possible explanations could have to do with bias in the model due to endogeneity, or with the homogenisation of 'preschool' as a single 'treatment' rather than a further disaggregated variable. Table 7.1 below summarises the findings.

Table 7.1: Summary of test score means compared by preschool attendance

Is child currently in preschool?		
Test		
PPVT	Yes	No
n	942	718
Mean Score	75.49	76.31
SD	14.27	12.69
t = 1.23	r = -.030	Cohen's d = .06
p value = .220	r-squared = .0009	df = 1658
Maths	Yes	No
n	918	706
Mean Score	45.89	42.91
SD	23.27	22.03
t = -2.62	r = .065**	Cohen's d = .13
p value = .009**	r-squared = .004	df = 1622

** indicates statistical significance at 1% level

Source: Young Lives India, Rounds 2 and 4

As the (two-tailed) significance of the t-statistic for differences in means on the PPVT test was not less than .05, the null hypothesis that no difference would be observed could not be rejected. Moreover, the correlation coefficient also had an insignificant p-value, showing no observed association between preschool participation and PPVT test outcomes[30]. The values of Cohen's d and r-squared for this relationship confirmed that not even a weak effect could be observed, nor could preschool participation account for any variability in PPVT outcomes.

[30] In fact, there was a slightly negative relationship, as the mean PPVT test scores for children who attended preschool was, surprisingly, lower. This highlights the value of using two-tailed tests. However, as the correlation method used was point-biserial due to variable typologies, the sign is irrelevant, and at any rate the result was not statistically significant.

However, an association between preschool participation and maths test results was observed (t-statistic = -2.62; p-value = .009). This association was significant at the .001 level. But further interrogation revealed that despite this statistically significant result, the effect size was weak (Cohen's d = .13), and only 4% of the variance in maths test scores could be explained by preschool participation.

7.2 Did boys who attended preschool fare better than girls on these tests?

Subgroup analysis by gender revealed that boys were not more likely than girls to score higher on either the PPVT or maths test. Additionally, differences in preschool participation rates for boys and girls were not significant. This confirms analysis from YL India conducted by Streuli, Vennam and Woodhead (2011) and also mirrors World Bank data which indicates that pre-primary enrolment across India is now essentially on par (Data.worldbank.org, 2018). Table 7.2 below summarises the findings from the gender subgroup analysis, and Figures 2a and 2b below illustrate that there was no significant difference between mean boys' scores and mean girls' scores on either test.

Table 7.2: Summary of test score means compared by gender

	Gender	
Test		
PPVT	M	F
n	886	774
Mean Score	76.31	75.32
SD	13.18	14.09
t = 1.47	r = -.036	
p value = .142	r-squared = .001	df = 1658
Maths	M	F
n	871	753
Mean Score	44.33	44.89
SD	23.06	22.47
t = -.497	r = .012	
p value = .619	r-squared = .000	df = 1622

Source: Young Lives India, Rounds 2 and 4

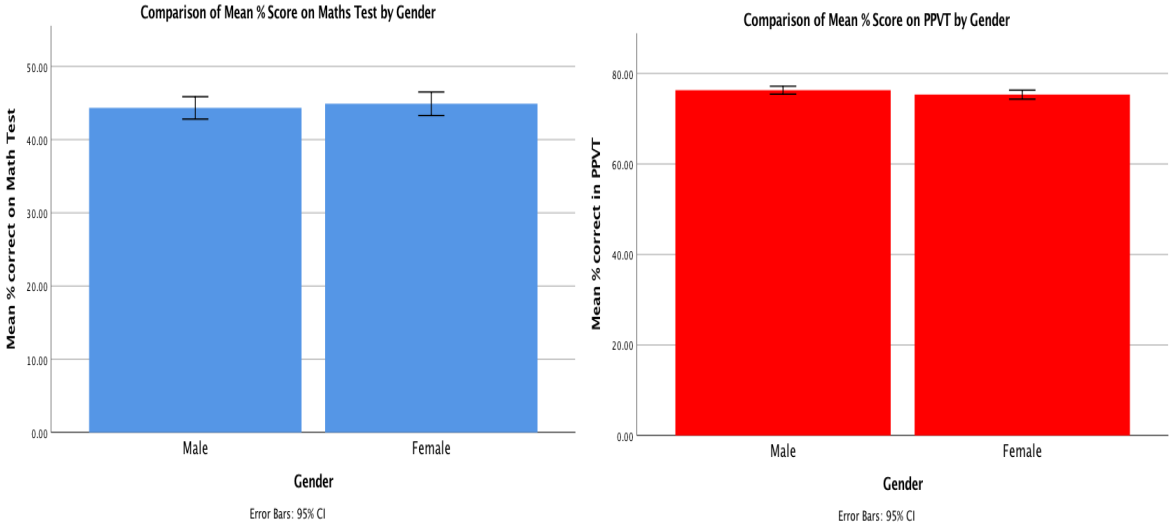

Figure 2a: Comparison of mean percent score on Maths test by gender (L)

Figure 2b: Comparison of mean percent score on PPVT by gender (R)

Source: Young Lives India, Rounds 2 and 4

The implications of these findings are that strides toward gender parity in education could well be taking effect in Andhra Pradesh, but it is also possible that the lack of further disaggregation of 'preschool' by type masked gender disparity in preschool attendance or results. Other research conducted by the YL team, for example, showed sons had higher attendance at private preschools than daughters (Galab, Reddy and Himaz, 2008).

7.3 Did children with educated mothers fare better on these tests?

Interestingly, maternal education emerged as the most significant predictor of test scores on both the PPVT and maths assessments. As illustrated by Figures 3a and 3b below, analysis found statistically significant differences between test score outcomes for children whose mothers were educated and those whose were not. The results summarised in Table 7.3 show that mother's education was significantly associated with test score outcomes on both tests; children whose mothers had some education outperformed those whose mothers did not. Effect sizes as measured by r also showed a medium effect in both instances, furthering the evidence substantiating maternal education's ability to predict outcomes amongst this sample.

Therefore, this justified the inclusion of this variable in the multivariate model as a control for the relationship between preschool and test score outcomes.

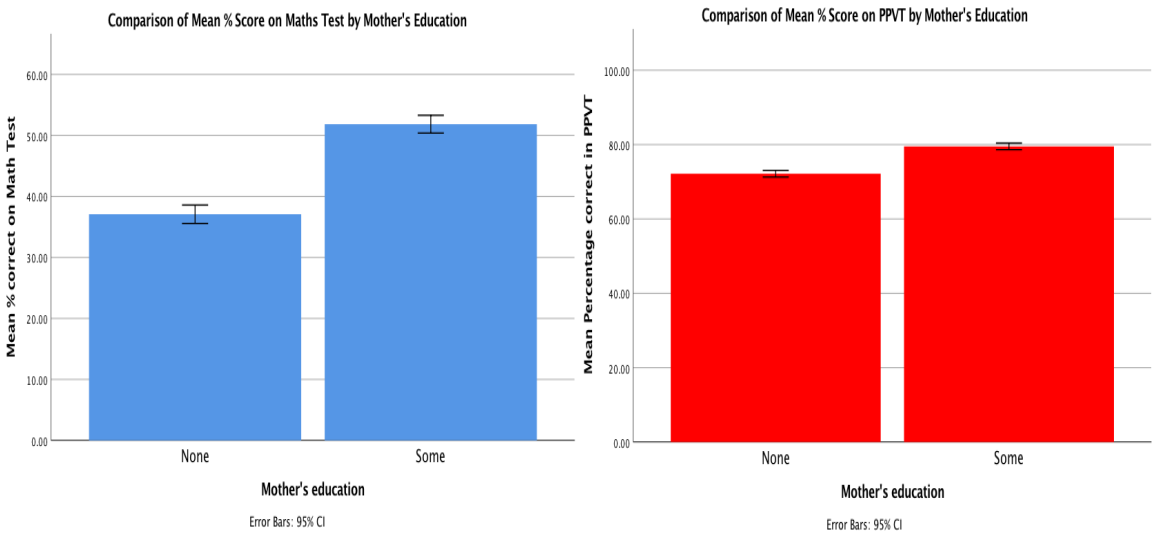

Figure 3a: Comparison of mean percent score on Maths test by mother's education (L)

Figure 3b: Comparison of mean percent score on PPVT test by mother's education (R)

Source: Young Lives India, Rounds 2 and 4

Table 7.3: Summary of test score means compared by maternal education

	Maternal Education Level	
Test PPVT	None	Some
n	823	834
Mean Score	72.16	79.52
SD	13.08	13.16
t = -11.42	r = .270**	
p value = .000**	r-squared = .073	df = 1655
Maths	None	Some
n	792	829
Mean Score	37.01	51.84
SD	21.66	21.39
t = -13.81	r = .325**	
p value = .000**	r-squared = .105	df = 1619

** indicates statistical significance at 1% level

Source: Young Lives India, Rounds 2 and 4

Mother's education was also found to be significantly associated with preschool participation. Results from Chi-square testing showed statistically significant differences in proportions of children who did and did not participate in preschool based on their mother's level of education. Of all the children attending preschool, 43.4% had mothers with no education, while 56.6% had mothers with some education. On the other hand, of children not attending preschool, 58% had mothers with no education, but only 42% had mothers with some education. These results reveal that the proportion of children that were not attending preschool was significantly higher amongst families with uneducated mothers. Mothers with some education were also more likely to have children attending preschool than not. These results indicate that mother's education level could be a useful predictor variable for preschool participation.

However, due to the issues of external variables discussed, it is possible that there was some bias in the coefficient for maternal education resulting from a correlation between SES and maternal education, and between SES and preschool participation. These results should also be interpreted with caution, as the sample was not representative given the original sampling procedure undertaken by the YL team. But as maternal education was the only predictor variable which showed a strong association with academic outcomes during the initial bivariate analysis and even after multivariate regression with partial correlations, it offers an opportunity for further analysis as an important input into human capital development outcomes.

7.4 Moving from bivariate to multivariate analysis

The methodological decisions made for the purpose of this analysis were taken from the perspective of a social scientist rather than that of a statistician. A statistician might frame the assumptions in such a way as to prioritise the empirical alone; however, my approach has been to interpret methodological considerations within the context of the overall research questions. It is for this reason that, despite analysis at the bivariate level demonstrating no statistically significant relationship between preschool and PPVT scores, and only a weak relationship between preschool and maths scores, I opted to continue on to multivariate analysis.

The primary reason for building multivariate models was the interest in understanding not just the unconditional but also conditional relationships between ECE and attainment, with gender and maternal education factoring in. As these have been shown to be statistically significant factors in evidence originating from the Indian context, understanding the role they do or do not play in this sample could point to important relationships from a policy point of view.

From the bivariate analysis, maternal education was deemed a significant factor, so building this variable into the multivariate model had empirical basis. Whilst gender was not deemed a significant factor at the bivariate level, it has been too important historically in the context of the Indian education system to discount on the sole basis of my bivariate analysis. If it were included in the multivariate model but not shown to have strong associations with attainment, this would only serve to ratify the earlier analysis from this paper. Therefore multivariate testing was conducted to determine the conditional relationship between ECE and academic outcomes, controlling for both gender and maternal education. However, it is important to recognise what the data from the bivariate analysis in this paper indicated, which was surprisingly little association between preschool and attainment. Given the overwhelming evidence that supports a relationship between these variables, it would be important to consider whether other analysis is statistically significant but actually has weak effect sizes which are not being reported along with results for the sake of being able to publish statistically significant findings.

7.5 Once gender and maternal education were controlled, was preschool a useful predictor of test scores?

The results of the multivariate regression indicated that gender, in this sample, should be discounted as a predictor, and preschool also did not really offer much value as a predictor, but maternal education offered a lot of value as a predictor for attainment. The models for predicting PPVT and maths scores are summarised in Table 7.4 below:

Table 7.4: Multivariate model summaries for predicting Maths and PPVT scores

Model Summary for Predicting PPVT Scores					
Variable	Coefficient	95% Confidence Interval	S.E.	t-ratio	sig.
Preschool attendance	-0.073	(-3.29, -0.72)	0.66	-3.05	.002**
Gender	-0.043	(-2.44, 0.08)	0.64	-1.83	0.067
Mother's education	0.282	(6.39, 8.94)	0.65	11.81	0.000
*F = 47.87***	*F sig = .000***	*r = .28*	*r-squared = .08*		
Model Summary for Predicting Maths Scores					
Variable	Coefficient	95% Confidence Interval	S.E.	t-ratio	sig.
Preschool attendance	0.013	(-1.54, 2.74)	1.09	0.55	0.581
Gender	0.006	(-1.81, 2.40)	1.07	0.28	0.783
Mother's education	0.322	(12.55, 16.79)	1.08	13.55	.000**
*F = 63.62***	*F sig = 000***	*r = .33*	*r-squared =.11*		

** indicates statistical significance at 1% level

Source: Young Lives India, Rounds 2 and 4

7.5.1 Summary of PPVT model findings

The overall result of the PPVT model found it to be a 'good fit' for estimating test score outcomes. However, much is revealed by breaking the model down into each individual coefficient. Surprisingly, the model showed that preschool attendance had a negative, statistically significant coefficient in the conditional relationship. This contradicts a large part of the existing body of evidence on the efficacy of ECE for academic outcomes. However whilst mathematically the coefficient was negative, it would be illogical to presume that such a result could be generalised, as suggesting that preschool participation has a 'real' negative effect on children's cognitive abilities or attainment would be a misinterpretation of these results. Accordingly, despite the surprising coefficient, this result is best interpreted within the context of the research question as a 'negligible' finding, as if it had been a one-tailed test, though it was not. Meanwhile, gender was shown to have a slightly negative coefficient in the PPVT model, but the significance value of the t-test indicated that it was not a significant contributor. This result confirmed the results of the bivariate analysis showing that there was not a strong relationship between gender and PPVT outcomes. Lastly, the model's results showed that mother's education had a highly significant correlation (p = .000) with test outcomes for the PPVT. In fact, it was the only variable in this model with a positive part correlation (.279), indicating that this was the only variable that still had a positive

association with PPVT scores when the other variables in the model were controlled for. Therefore, though the model as a whole was found to be a good method of estimation, the goodness of fit came primarily from the coefficient for maternal education.

7.5.2 Summary of Maths test model findings

In the initial bivariate analysis, linear regression showed a statistically significant association between preschool participation and maths test results. The findings of that analysis were consistent with those of Singh and Mukherjee, who also found that preschool was associated with improved maths scores (2017). However, results from the multivariate model showed that the conditional association (with gender and maternal education controlled) was no longer statistically significant. As with the PPVT model, this result contradicts a large part of the existing evidence body. Gender was also not shown to be a significant contributor to the model, confirming the original results of the bivariate analysis. However, as was the case with the PPVT model, maternal education emerged as a highly significant contributor ($p = .000$) to the model, and provided the bulk of the model's ability to predict maths score outcomes. Though the F-test results found this model to be a 'good predictor' of maths scores, it is important to note that the r-squared value demonstrates that the model could only account for 11% of the variance in score results, and also that the original research question was about whether ECE (not the other conditional variables) was a useful way to estimate score outcomes, which this model indicates it was not.

7.6 Conclusion of results

Overall, though both models were 'good fits', the findings did not support the original hypothesis that ECE would be a good predictor of test score outcomes. These results were multifaceted. Correlation between preschool and PPVT results was in fact slightly negative, a highly unexpected result. On the other hand, correlation between preschool and maths results was found significant at the bivariate level; however, once gender and maternal education were conditioned out, the coefficient was no longer significant. Furthering the conclusion that preschool was not a good predictor, r^2 values showed that no variation in PPVT outcomes and only 4% in maths outcomes could be explained by preschool participation. Additionally, as the effect size for the association between preschool and maths scores was only .13, what social scientists would deem 'small', this was not actually a very effective relationship.

These results contradict the profusion of evidence showing positive effects of preschool on academic attainment. Including additional covariates for the regression created better models, but the original bivariate relationships were weak even for maths outcomes to begin with. In fact, for this data, the best ability to predict test score outcomes would come not from the multivariate models, which were statistically significant but with limited contribution from two of the variables, but from bivariate models between maternal education and test results.

Therefore, though statistically the findings show both models to be 'good predictors' of test score outcomes, practically and intuitively, they are not. Firstly, effect sizes are negligible in some cases, and only small-to-medium in others. Secondly, the correlation coefficients themselves are small, indicating that interpreting a statistically 'strong' association should be done with caution, even when p-values are very small (<.001). Thirdly, the r^2 values are also small, meaning that variance in outcomes is likely to be better explained, or further explained, by other predictor variables not included in the model. Lastly, the lion's share of the 'predicting' comes from the wrong variable - mother's education. My initial hypothesis stated that, based on the theory, literature and evidence, the strongest predictor variable should be preschool participation. For these reasons, despite significant findings, these results must be interpreted with intuition and caution.

8. Discussion

Three major themes emerged from this paper's findings. The first was that preschool was not a valuable predictor of test scores amongst this sample, supporting the idea that the Indian evidence base for ECE needs further substantiation in order to help existing ECE policies and practice stay on target (see Alcott *et al*., 2018; Kaul and Sankar, 2009; Singh and Mukherjee, 2017 and Kaul *et al*., 2017). The second was that maternal education has significant value as a predictor of academic outcomes, confirming other evidence of its role in intergenerational educational attainment (see Black, Devereux and Salvanes, 2005; Rosenzweig and Wolpin, 1994; Becker, Murphy and Tamura, 1990; Andrabi, Das and Khwaja, 2012; Magnuson, 2007 and Black *et al*,. 2005, and also Jeong, Kim and Subramanian, 2018 for evidence from India and Singh and Mukherjee, 2015 for evidence using YL India data). The third was that gender disparity appears to be reducing in scope, as also indicated by current data from the World Bank (data.worldbank.org, 2018). As a group, these themes help to substantiate the evidence base on early education in India and clarify areas of progress and residual challenges.

8.1 Preschool was not a good predictor of academic outcomes

Surprisingly, the findings indicated that children amongst this sample who participated in ECE did not achieve the desired academic outcomes that would enable them to participate in socioeconomic betterment of self or society. In both bivariate and multivariate instances, variation in preschool had little or no association with variation in test outcomes, particularly once effect sizes, r^2 values, and part and partial correlations were considered[31]. The surprising conclusion is therefore that attending preschool did not 'make a difference' to how children performed academically later on[32].

These findings corroborate other evidence, but only to a degree. The significant association between preschool and maths scores in the bivariate analysis was consistent with other YL research in Peru, Ethiopia and Vietnam (Early Childhood Development: informing policy and

[31] Perhaps a problem with the hegemony of support for ECE is that researchers are reporting significant associations between ECE and academic outcomes without reporting effect sizes, because there is more fanfare around finding significant results than inconclusive ones. Reporting effect sizes could actually show, in some instances, that though findings are statistically significant they are not strong.

[32] This is anecdotal language; in reality, this finding shows lack of association, not lack of causation. But without lack of association, one at any rate cannot have causation.

making it a priority, 2018). Regarding the overall conclusion that preschool did not 'make a difference', other mixed results on ECE in India have also been reported by Chopra (2012), Pattnaik (1996) and Manhas and Qadiri (2010), who found that preschool was not associated with variance in academic outcomes in all cases. Reports from Save the Children (2009) and ASER (Beyond Basics, 2018) also reported low attainment on pre-literacy and pre-numeracy skills for 5-year-olds, which indicates inadequate preschooling.

However, overall, the results contradict a large majority of the evidence on ECE, which overwhelmingly finds positive associations between preschool and academic attainment. Evidence from experimental research such as the well-known Abecedarian Program and the HighScope Perry Preschool Project (see Currie and Thomas, 1995 and Campbell and Ramey, 1994) was historically very persuasive, as it showed causation. Additionally, as discussed in Section 3.2, a plethora of other research has shown preschool impacting or being associated with higher academic outcomes. Consequently, because this paper's findings appear to be out of character for the wider evidence body, it was important to explore why this may have been the case.

Two possibilities emerge as to why these findings do not match the patterns of the wider evidence base. Firstly, problems could have arisen with the transition from data collection to variable construction which obscured important indicators. Secondly, the wider evidence body itself, from which the 'pattern of positive association' stems, might not be heterogeneous enough to account for variation in patterns on ECE such as what this evidence shows.

Recapitulating the first reason, it is possible that these findings do not fit into wider patterns because the methodology obscured what might otherwise have yielded significant results. Initially, problems with construction of the primary explanatory variable came from the records of preschool participation: responses to 'is the child currently attending preschool' were not readily disaggregated by private versus public in the data I was able to access. Given evidence in the literature that preschool type has been associated with variations in attainment (Beyond Basics, 2018; Singh and Mukherjee, 2015 and Alcott and Rose, 2015),

this would have offered a valuable condition[33]. Consequently, had it been possible to further disaggregate 'preschool' by type, statistically significant findings might have ensued, which would have corroborated other evidence in the literature on India that indicates children who attend private preschools perform better later on (see Singh and Mukherjee, 2018 and Chopra, 2012).

Additionally, the original question asked by the YL survey, 'is the child currently attending preschool?', pertained to children roughly age 5. However, the minimum age of entry to primary school in Andhra Pradesh is also age 5; therefore, it is possible that families answered 'no' to this question because their children had attended preschool but had already moved on to primary school either early or 'on time', a documented phenomenon, by the time Round 2 took place (see Alcott *et al*., forthcoming)[34]. Had it been possible to generate a more accurate picture of which children actually completed a substantial portion of preschool of any kind, significant results for those children could have ensued, which would have corroborated the wider patterns on ECE improving academic attainment.

A useful methodological approach building on and further illuminating these findings could therefore entail conducting regression analysis that disaggregates preschool experience by type (as in the research conducted by Singh and Mukherjee, 2018), and also by other determinants of quality such as teacher qualification, structural assets, curriculum suitability, etc., (though this information is likely not available across India today) to determine how each of these categories varies with attainment outcomes. Such research might reveal under what circumstances, or in what conditions, ECE in India matches wider patterns of positive association with academic outcomes. Knowing which conditions make a difference would in turn offer important information for ECE policies.

Recapitulating the second reason which may explain why these findings do not match wider patterns, it is possible that this is the case because those wider patterns are not diverse enough

[33] However, documented issues with 'nonlinearity' of early educational trajectories might have complicated this anyway, as many children have been shown to complete some government preschool as well as some private preschool (Alcott *et al*., 2018). Therefore, preschool categories by type might also have been a challenge to estimate.

[34] The YL team should have mitigated against this response bias, given 'early entry' is a documented phenomenon. Perhaps a question worded along the lines of 'did this child complete at least two full years of ECE between the ages of 3-5?' would have allowed for a more accurate analysis of the effects of preschool on later attainment.

to account for variety of ECE experience and quality on a global level. Literature authenticating this possibility comes from Woodhead (2006), who has noted that there is a distinct lack of literature from LAMI countries on ECE's role in academic attainment, and also from YL researchers Woldehanna and Gebremedhin (2012), who have noted the same. Much of the literature on ECE has been comprised of samples from developed country contexts such as the US and UK (see aforementioned Currie and Thomas, 1995 and Campbell and Ramey, 1994, as well as Currie, 2001; Heckman, 2011; Sylva *et al*., 2011; Black *et al*., 2016; Rosenzweig and Wolpin, 1994; Harmon and Walker, 1995; DeCicca and Smith, 2013; Heckman and Masterov, 2007; Belfield, Nores and Schweinhart, 2006; Magnuson *et al*., 2004 and Goodman and Sianesi, 2005 for examples). These studies, though in some ways diverse, have had certain commonality in terms of the relative quality of ECE provision, ambient notions of education, patterns of parental cognition, minimal historical issues of gender and a myriad of other factors which would have made them distinct from the YL India sample. These other commonalities may have helped to contribute to the patterns which emerged. Accordingly, we can deduce that there is a possibility that patterns on ECE in LAMI countries could be different from those in developed contexts, but it is difficult to corroborate this without further evidence, the need for which has also been widely cited for India specifically (see Alcott *et al*., 2018; Kaul and Sankar, 2009; Chopra, 2012; Singh and Mukherjee, 2017 and Kaul *et al*., 2017).

If patterns in LAMI country contexts were to differ from the wider evidence body, as these findings indicate may be the case, it is likely that 'quality of ECE' divides them. This conclusion is based on evidence from Rao and Kaul, who found that *anganwadis* offered insufficient pre-primary education (2017) and Kaul *et al*. (2017) who noted that even private pre-primary is developmentally inadequate. Moreover, researchers such as Sylva *et al*. (2011), Chopra (2012) and Marope and Kaga (2015) have noted that poor-quality preschool provision can have adverse effects on educational attainment. Accordingly, overall poor quality ECE provision in India could be contributing to lack of statistically significant results that would mirror patterns found in contexts where the overall ECE quality is higher.

In summary, it is possible that the these findings do not match wider patterns on ECE because of unobserved differences in preschool experiences due to methodological challenges, or because wider patterns are not inclusive enough to make allowance for the low-quality preschool experience characteristic of some contexts. Ultimately, both circumstances point to

important considerations: firstly, that more nuanced studies on ECE's role in academic attainment in India which disaggregate preschool experiences along as many lines as possible would help to discern what elements of ECE are working well, and secondly, that more evidence is needed in order to determine whether the findings of this paper represent a new pattern, different to that of the existing evidence, which should be fed back into policymaking and practice.

8.2 The role of gender in predicting attainment appears to be diminishing

These findings demonstrate that gender is not a good estimator of academic attainment: the evidence on both participation and test scores shows that amongst this sample boys and girls were equally likely to participate in preschool and also to perform well or poorly on assessments at age 12. Additionally, including gender as a controlled condition in multivariate models looking at preschool's relationship to academic attainment added no value. These findings corroborate existing evidence on multiple fronts: evidence that gender does not play a role in estimating a range of education-related outcomes has also been found using YL India data by Vennam *et al*. (2009) and again by Streuli, Vennam and Woodhead (2011). Meanwhile, official World Bank data shows that pre-primary enrolment in India is effectively equal between boys and girls (Data.worldbank.org, 2018), which corroborates this paper's findings on preschool participation rates and also offers a backdrop to the evidence on equal attainment.

These results speak to the long-established international focus on reaching gender parity in education, which has been a particular priority since the inception of the MDGs (Un.org, 2018). They also suggest that efforts made by the Indian government to redress gender disparity in education, such as launching the National Programme of Education for Girls at Elementary Level, are making a difference (Streuli, Vennam and Woodhead, 2011). And whilst these results do not support the views of Kaul *et al.* (2017) that gender in education is still a point of disparity in India, or evidence from Magnuson *et al*. (2016) or Garcia, Heckman and Ziff (2018) who indicate that girls in some instances benefit less from pre-primary education, they represent an important finding that demonstrates progress is being made. Accordingly, international discourse, national policy and the array of interventions supporting girls' education should be encouraged by such results (see Unterhalter, 2007), but should still give importance to the well-documented history of gender disparity in the Indian

education system (see Asadullah and Yalonetzky, 2012 and Kaul *et al*., 2017), and should continue to disaggregate research on education by gender, as is the precedent within the literature.

A complexity in these findings could arise from the possibility that variation between preschool type and gender was masked by the methodology involved in documenting preschool attendance, which did not account for private versus public. For example, other research, including qualitative research from YL India, has indicated that parents are more likely to invest in private education for sons than for daughters (Woodhead *et al*., 2009).

8.3 Maternal education was a useful predictor of attainment

Results on maternal education found it to be a strong predictor of academic outcomes amongst this sample. These findings support wide-ranging evidence from the human capital tradition that investing in mothers' education pays intergenerational dividends (see Rosenzweig and Wolpin, 1994; Becker, Murphy and Tamura, 1990; Andrabi, Das and Khwaja, 2012; Magnuson, 2007 and Black *et al*,. 2005). They also corroborate evidence from India that maternal education has an important role to play in educational attainment (Jeong, Kim and Subramanian, 2018; Singh and Mukherjee, 2015 and Kaul *et al*., 2017). These findings provide clear evidence of the intergenerational nature of maternal education as an 'endowment'. They also reflect well on international discourse concerning the instrumental importance of educating girls (Schultz, 1993 and King and Hill, 1993).

Because 'maternal education' was recoded for this paper as a binary variable, it is interesting to note that having any education at all made a difference to variation in test score outcomes on both tests. This is an important point for policymaking in India, as the country still has an adult female literacy rate of only 59% and a large quantum of first generation learners (www.worldbank.org, 2018).

8.4 Complementarities and points of departure from recent findings

One of the most closely related, recent pieces of research to this paper is that of Singh and Mukherjee (2018), which uses data from YL India to estimate the role of private preschool in predicting cognitive achievement and subjective wellbeing at age 12. Both Singh and

Mukherjee's 2018 paper and mine examine ECE in India using the Young Lives data from the younger cohort. Both examine cognitive achievement as a main outcome variable, and both make use of OLS multivariate regression. Therefore, at least in part, there is a major overlap of both outcomes of interest and of methodologies. However, their paper focuses on the role of private preschool in particular, also includes subjective wellbeing as an outcome variable, and makes use of propensity score matching techniques to substantiate their findings. As such, they were looking more holistically at children's outcomes by including wellbeing than simply academic attainment, so it is possible that their research was less informed by human capital theory's notions of early education's instrumental value than mine. Policymakers who are concerned more with education's role as an early investment for future public returns might therefore be less inclined to make use of the totality of their evidence than mine, which has more straightforward conclusions for such a purpose. Additionally, their main point of inquiry was of the value of private preschool, whereas my paper does not differentiate preschool experience by type. Their paper therefore speaks to the debate about private versus public school quality in India, topical due to the rapidly rising private sector, whereas my paper, because it does not differentiate by type, is able to comment on the state of ECE in India (or Andhra Pradesh) more widely.

Singh and Mukherjee's findings demonstrate that children who attended private preschool had significantly higher maths scores and enhanced subjective wellbeing. They also show that early entry to preschool (before age 4) improves outcomes. These cover ground that my paper does not; however, we overlap in our demonstration of no impact of preschool on PPVT outcomes, both with and without controls. Together, these papers could offer insight regarding changes in ECE curriculum or pedagogy to better prepare children for success in verbal and comprehension skills.

Additional points of complementarity arise from other YL research on early childhood education and academic outcomes. YL, being an international study on childhood poverty, has also drawn samples from Ethiopia, Peru and Vietnam. Accordingly, it is valuable to understand whether there are patterns in the evidence on this topic emerging from these other contexts, because there is in general minimal evidence on ECE in LAMI countries. Working with YL data in Ethiopia using propensity score matching, Woldehanna and Gebremedhin (2012) found that preschool attendance had a significant impact on cognitive development at ages 5 and 8, as well as on enrolment and progression. Though Ethiopia faces challenges with

access to ECE as stratified by geography and wealth status, unlike India, these results indicate that Ethiopia's pattern on ECE matches wider international patterns, rather than resonating with evidence from this paper. Generating evidence across all four YL countries, Woodhead *et al*. found that all four younger cohorts showed evidence of inequalities in school readiness by the time they entered primary school at age 5-6 (2009). Moreover, they found that issues of quality persist across the board, indicating the ineffectiveness of existing services. Accordingly, this paper's findings on the Andhra Pradesh sample relate to theirs citing quality as an impediment to the effectiveness of ECE in varied LAMI contexts.

A number of other studies based on YL research have been cited in this paper which pertain to India specifically (Singh and Mukherjee, 2016, 2017 and 2018, and Galab, Reddy and Himaz, 2008 amongst others), but there is not an abundance of papers utilising four-country YL data to draw wider conclusions on ECE in LAMI contexts to date. This will change when the World Bank, who in 2016 commissioned YL to examine preschool's impact on numeracy skills at ages 5, 8 and 12 in all four countries, releases its findings this year. Accordingly, whilst the complementarities of results and points of departure are important amongst the YL-based papers, it is their ability to offer evidence of cross-country patterns that could be most promising for re-evaluation of useful frameworks on ECE practice relating to international development agendas.

9. Conclusions and further research

9.1 Conclusions of research

The aim of this study was to explore the impact of preschool participation on academic attainment, which is of interest because of educational attainment's link to outcomes important for development agendas such as improving SES and reducing intergenerational poverty. This research problem was addressed through layers of quantitative analysis, including the creation of multivariate models estimating the ability of preschool participation to predict test scores at age 12 whilst controlling the possible influences of gender and maternal education.

As pointed out by Meghir and Rivkin, the evidence arising from the multiplicity of methods applied to returns on education is often contradictory (2011). But despite conflicts in evidence, there remains a hegemony of perception that early childhood education is an unquestionably effective input for desirable academic and economic outcomes. The body of evidence posits that ECE should 'work': theory and literature both indicate that participating in early education should improve overall attainment by providing children with the skills to succeed academically. International and national policies relating to ECE and ECCE have therefore endorsed this view.

For these reasons, this paper's primary hypothesis was that participation in ECE should lead to improved academic outcomes. However, much of the evidence available is from developed contexts, which have very different parameters of experience for early years education. As a result, an additional line of inquiry alongside the original hypothesis was whether the patterns from existing literature would hold in this context, where respondents were from India and from a poorer background, and likely to have been accessing lower quality services. Woodhead (2006) also asks this question, noting that the generalisability of the existing evidence body is more limited than many actors choose to accept. The problem is that the specific conditions under which particular ECE interventions are conducted is unlikely to be replicated in the same way in other contexts, yet the assumption underpinning international policies is that ECE uniformly precipitates the same kinds of outcomes and benefits.

The findings indicate that this assumption did not hold. Despite being able to successfully create multivariate models that were good estimators of test results, little to no association was found between preschool participation and academic outcomes. These findings highlight firstly that ECE in India is facing issues of quality and secondly that evidence on ECE is being generalised because it is convincing on many fronts, and because there is not sufficient evidence from elsewhere to change this. Accordingly, this paper argues that ECE policies for LAMI countries should rely not just on precedent but must take into consideration evidence from the local sample populations they seek to serve. It also argues against viewing 'preschool' as a monolithic or uniform experience, because a plausible explanation for the inconclusive findings is that 'preschool' was not disaggregated into further categories which could have revealed specific elements or types of ECE contributing to variation in outcomes. However, considering all preschool experiences as 'uniform' for this sample also had value, because it demonstrated that as a group preschools in Andhra Pradesh are not of good enough quality to enhance cognitive outcomes. This evidence underscores the need to reform ECE services in India (Streuli, Vennam and Woodhead, 2011) in order to reap the ROI that the annual expenditure on ECCE should merit.

India's release of the National Curriculum Framework and Quality Standards in 2013 demonstrates that it is aware of these facts, yet it is still spending only 1.6% of its total educational expenditure on pre-primary, compared to 28.4% on primary and 41.4% on secondary (www.worldbank.org, 2018). Therefore, India's government needs to align its spending with its efforts to redress inadequate pre-primary services, particularly in light of the increasing competitiveness of the private market. But increased spending is only part of the solutions which have been proposed. Other researchers have cited the need for a regulatory body to help with governing quality standards (Chopra, 2012). Such standards could include a range of elements such as developmentally appropriate curricula, quantity of trained teachers and staff, appropriateness of setting as well as age-appropriate pedagogy. This recommendation is based upon evidence demonstrating that the higher the quality of the programme, the more likely children are to develop skills to help them succeed in later grades (Manji, 2018). Furthermore, related policies improving ECE could help to ensure that families for whom private provision is not an option do not miss out on quality early years education. A more competitive government provision could in turn help foster the development of better regulation of the private sector, improving quality across the board.

This notion of improving impact raises that of 'effectiveness'. If preschool participation has been shown here to be ineffective as a determinant of improved academic outcomes, then it follows that more 'effective' models of ECE are needed. Effective ECE programmes are those adapted to the local context, and that also respond to issues raised by research and analysis stemming from the local context (Alcott *et al*., forthcoming). India already does well to include ECE services alongside others for health and nutrition, and to conceive of ECCE as an important pathway to socioeconomically desirable outcomes (Streuli, Vennam and Woodhead, 2011), but a further focus on ensuring ECE offerings are of high enough quality to actually achieve these outcomes is vital. Increasing the effectiveness of India's ECE models can best be accomplished by garnering more reliable, relevant evidence to inform their improvement. Ultimately, education programs are interventions which are dependent upon a range of factors, and importantly upon their human participants, including teachers, staff, families and children themselves. The added complexity of local policies, politics, agendas and histories further nuances the possible impact and outcomes such interventions can achieve. It is therefore almost impossible to control for enough of these factors to examine outcomes without some bias. However, substantiating the evidence base with analysis that includes the considerations most relevant to local populations will be a start in the right direction, building on the first steps which have been taken already by YL and only a few others.

In this vein, this paper has been an endeavour to help substantiate the limited evidence body on ECE in India, provide additional nuance to existing debates on topics such as gender, and to increase awareness of the limits of generalisability of existing ECE literature. Though the findings are, in themselves, not generalisable, their importance lies in their departure from the results of other research showing ECE's effectiveness. This paper's conclusions support the notion that evidence from existing studies shows that ECE can have an impact, not that it will have an impact (Woodhead, 2006). There is no guarantee that in other settings or contexts with different conditions, this evidence is necessarily replicable without doing the groundwork to ascertain whether, and to what degree, that is true. As Woodhead (2006) reminds us, patterns of effect have specific historical and cultural contexts, and variables that have particular kinds of value in one circumstance may prove completely different in others.

Perhaps it is because there is ultimately a belief that education is good, and because there is overwhelming evidence that early education is great, that advocates of ECE have been

tempted to generalise; this has helped developing countries to craft important policies and programs that are on paper responsive to the evidence. And certainly, the dialectic between research and policy is complex and vulnerable to wider debates and priorities. But where emerging evidence on ECE does not conform to existing patterns, it should be taken as an opportunity to reflect and improve on policies and services, rather than to detract from the basic premise of early education's value nor the vast body of evidence that supports it. Such reflection reveals that the best way to interpret the evidence from this paper, within the wider context of related literature, is that ECE is a critical intervention point that is not being maximised.

9.2 Further inquiry

My research, along with a large portion of the existing evidence body, compares the effect of preschool participation against the control of no participation, but in order to determine how existing ECE services can be improved, further research into what elements or types of ECE covary with improved outcomes should be explored (Yoshikawa and Nieto, 2013; Engle *et al*., 2007 and Engle *et al*., 2011). This type of research could be considered 'second generation' research on the role of ECE (Chopra, 2012) which would respond to the calls for more nuanced evidence (Kaul *et al*., 2017) and concerns with poor quality (Indian Ministry of Women and Child Development, 2013).

Another opportunity for further research could be to widen the models I have created in this paper by adding other highly pertinent predictor variables, thereby increasing the validity of the analysis. Additional variables should correspond to key social or economic inputs in the Indian context, such as SES, caste or geographic setting, but there could be others. The analysis would be improved by being able to hold constant each of the variables in turn (using partial correlation) so as to better understand which variables are truly playing an explanatory role (Kline, 2004). This could also help to reduce some of the omitted variable bias or endogeneity (Woolridge, 2016) in my model, though it would be unlikely and impractical to include all explanatory variables within one model. Making use of propensity score matching, as used by Woldehanna and Gebremedhin (2012), could also offer a worthwhile methodology for further research on this problematic by which experimental-style research comparing the effects of various input variables could offer further insight.

Finally, a third way to build upon the research within this paper could be to build a statistical model that would take into account children's developing capacities at various stages, rather than take one set of scores as the only measure of attainment on which to base a value judgment about preschool participation. For example, if a model were to include preschool participation as an explanatory variable, and then measure test score outcomes at for example two points (age 8 and age 12) rather than just one, as this model has done, it might be possible to deduce more information about how the impact of ECE materialises over time. Such research could follow examples set by Woldehanna and Gebremedhin (2012), who measured preschool's effect on academic outcomes at ages 5 and 8, and by the forthcoming World Bank paper, which will examine preschool's effect on numeracy skills measured at ages 5, 8 and 12.

Though different in their approach and potentially the specificities of their findings, the goals of such research ideas would all resonate with the original foundation of this thesis, which is concerned with improving the quality of ECE for Indian children, particularly the more socioeconomically vulnerable, in order to help increase their chances at achieving successful life outcomes for themselves and their families.

References

Aakvik, A., Salvanes, K. and Vaage, K. (2003). Measuring heterogeneity in the returns to education in Norway using educational reforms. C.E.P.R. Discussion Papers: No. 4088.

Alcott, B. and Rose, P. (2015). Schools and learning in rural India and Pakistan: Who goes where, and how much are they learning?. *Prospects*, 45, pp.345-363.

Alcott, B., Banerji, M., Bhattacharjea, S., Ramanujan, P., and Nanda, M. (2018). One step forward, two steps back: transitions between home, pre-primary and primary education in rural India. *Compare: A Journal of Comparative and International Education*. DOI 10.1080/03057925.2018.1527214

Alcott, B., Bhattacharjea, S., Ramanujan, P., & Nanda, M. (Forthcoming). Participation trends in early childhood education programmes: Who goes where and why? In *Early childhood education in India; findings from a cohort study* (S. Bhattacharjea & V. Kaul, eds). Springer.

Andrabi, T., Das, J. and Khwaja, A. (2012). What did you do all day? Maternal education and child outcomes. *Journal of Human Resources,* 47(4).

Arora, S., Bharti, S., and Mahajan, A. (2006). Evaluation of non-formal pre-school educational services provided at *Anganwadi* Centres (Urban Slums of Jammu City). *Indian Journal of Social Science*, 12(2), pp.135-137.

Asadullah, M. and Yalonetzky, G. (2012). Inequality of educational opportunity in India: Changes over time and across states. *World Development*, 40(6), pp.1151-1163.

Attanasio, O., Meghir, C. and Nix, E. (2015). Human capital development and parental investment in India. Cowles Foundation Discussion Papers. New Haven, CT: Cowles Foundation for Research in Economics, Yale University.

Barro, R. (1989). Economic growth in a cross section of countries. Working paper no. 3120. Cambridge, Mass. NBER.

Becker, G. (1962). Investment in human capital: A theoretical analysis. *Journal of Political Economy*, 70(5), pp.9-49.

Becker, G. (1964). Human Capital: A Theoretical and empirical analysis, with special reference to education. National Bureau of Economic Research, New York. Distributed by Columbia University Press.

Becker, G. (1993). *Human Capital*. 3rd ed. Chicago: University of Chicago Press.

Becker, G., Murphy, K. and Tamura, R. (1990). Human capital, fertility, and economic growth. *Journal of Political Economy*, 98(5, Part 2), pp.S12-S37.

Belfield C., Nores, M., Barnett, W., and Schweinhart, L. (2006). The High/Scope Perry Preschool program: Cost-benefit analysis using data from the age-40 followup. *Journal of Human Resources,* 41(1), pp.162–90.

Ben-Porath, Y. (1967). The production of human capital and the life cycle of earnings. *Journal of Political Economy*, 75(1), pp.352-365.

Beyond Basics. (2018). Annual Status of Education Report (Rural) 2017. New Dehli: ASER Centre.

Black, M., Walker, S., Fernold, L. and Andersen, C. (2016). Early childhood development coming of age: science through the life course. *The Lancet*, 389(10064), pp.77-90.

Black, S., Devereux, P. and Salvanes, K. (2005). Why the apple doesn't fall far: Understanding intergenerational transmission of human capital. *The American Economic Review*, 95(1), pp.437-449.

Boserup, E. (1970). *Woman's role in economic development*. London: Earthscan.

Britto, P. (2018). Invest in children's futures; invest in Early Childhood Development. [online] UNICEF Connect. Available at: https://blogs.unicef.org/blog/invest-in-childrens-futures-invest-in-early-childhood-development/ [Accessed 2 Aug. 2018].

Britto, P., Engle, P. and Super, C. (2013). Early childhood development: Translating research to global policy. In: P. Britto, P. Engle and C. Super, ed., *Handbook of Early Childhood Development Research and its Impact on Global Policy*. Oxford: Oxford University Press, pp.3-23.

Britto, P., Engle, P. and Super, C. ed., (2013). *Handbook of Early Childhood Development Research and its Impact on Global Policy*. Oxford: Oxford University Press.

Bryman, A. (2014). *Social Research Methods*. 4th ed. Oxford: Oxford University Press, pp.312-315.

Campbell, F. and Ramey, C. (1994). Effects of early intervention on intellectual and academic achievement: A follow-up study of children from low-income families. *Child Development*, 65, pp.684-698.

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

Chopra, N. (2012). A study of early childhood education programmes in Dehli, India. *International Journal of Early Years Education*, 20(2), pp.159-174.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioural Sciences*. 2nd ed. New York: Lawrence Erlbaum Associates.

Cunha, F. and Heckman, J. (2007). The technology of skill formation. *American Economic Review*, 97(2), pp.31-47.

Cunha, F., Heckman, J., Lochner, L. and Masterov, M. (2006). Interpreting the evidence on life cycle skill formation. In: E. Hanushek and F. Welch, ed., *Handbook of the Economics of Education*. New York: Elsevier.

Currie, J. (2001). Early childhood education programs. *Journal of Economic Perspectives,* 15(2), pp.213–38.

Currie, J. and Thomas, D. (1995). Does a head start make a difference?. *American Economic Review*, 85(3), pp.341-64.

Das, J. and Zajonc, T. (2010). India shining and Bharat drowning: Comparing two Indian states to the worldwide distribution in mathematics achievement. *Journal of Development Economics*, 92(2), pp.175-187.

Data.worldbank.org. (2018). School enrollment, preprimary (% gross) | Data. [online] Available at: https://data.worldbank.org/indicator/SE.PRE.ENRR?locations=IN [Accessed 14 Jul. 2018].

Data.worldbank.org. (2018). School enrollment, primary (gross), gender parity index (GPI) | Data. [online] Available at: https://data.worldbank.org/indicator/SE.ENR.PRIM.FM.ZS?locations=IN [Accessed 3 Sep. 2018].

Datta, S., Boratne, A., Cherian, J., Joice, Y. and Vignesh, J. (2010). Performance of *anganwadi* centres in urban and rural area: a facility survey in coastal South India. *Indian Journal of Maternal and Child Health*, 12(4), pp.9.

Dearden, L., Stephen, M. and Howard, R. (1997). Intergenerational mobility in Britain. *Economic Journal*, 107(440), pp.47-66.

DeCicca, P. and Smith, J. (2013). The long-run impacts of early childhood education: Evidence from a failed policy experiment. *Economics of Education Review*, 36, pp.41-59.

Denison, P. (1985). Trends in American economic growth. Washington, DC: The Brookings Institution, pp.1929-1982.

Durbin, J. and Watson, G. (1950). Testing for serial correlation in least squares regression. *Biometrika*, 37(3-4), pp.409-428.

Early Childhood Development: informing policy and making it a priority. (2018). Young Lives Impact Case Study. Oxford: Young Lives.

Education-inequalities.org. (2018). *World Inequality Database on Education: India*. [online] Available at: https://www.education-inequalities.org/countries/india#?dimension=region&group=|Andhra%20Pradesh&year=latest [Accessed 10 Aug. 2018].

Engle, P., Black, M., Behrman, J., Mello, M., Gertler, P., Kapiriri, L., Martorell, R., Young, M. and the International Child Development Steering Group. (2007). Child development in developing countries 3: Strategies to avoid the loss of developmental potential in more than 200 million children in the developing world. *The Lancet,* 369, pp.229–42.

Engle, P., Fernald, L., Alderman, H., Behrman, J., O'Gara, C., Yousafzai, A., Cabral de Mello, M., Hidrobo, M., Ulkuer, N. and Ertem, I. (2011). Strategies for reducing inequalities and improving developmental outcomes for young children in low-income and middle income countries. *The Lancet,* 378 (9799), pp.1339-1353.

Field, A. (2009). *Discovering Statistics Using SPSS*. 3rd ed. London: SAGE.

Galab, S., Reddy, P. and Himaz, R. (2008). Young Lives Round 2 Survey Report Initial Findings: Andhra Pradesh, India. Oxford: Young Lives.

Garcia, J., Heckman, J. and Ziff, A. (2018). Gender differences in the benefits of an influential early childhood program. *European Economic Review*, 109, pp.9-22.

Garcia, J., Heckman, J., Ermini, D. and Prados, M. (2016). The life-cycle benefits of an influential early childhood program. NBER Working Papers. National Bureau of Economic Research.

Goodman, A. and Sianesi, B. (2005). Early education and children's outcomes: How long do the impacts last?. *Institute for Fiscal Studies*. 26(4), pp.513–48

Gormley Jr., W., Phillips, D., Newmark, K., Welti, K. and Adelstein, S. (2011). Social-emotional effects of early childhood education programs in Tulsa. *Child Development*, 82(6), pp.2095-2109.

Government of India, Ministry of Law and Justice (2009). The Right of Children to Free and Compulsory Education Act. New Dehli: Government of India Press.

Government of India. (2013). National Early Childhood Care and Education Policy. Ministry of Women and Child Development. New Delhi: GoI.

Grantham-McGregor, S., Cheung, Y., Cueto, S., Glewwe, P., Richter, L., Strupp, B., and the International Child Development Steering Group. (2007). Developmental potential in the first 5 years for children in developing countries. *The Lancet,* 36, pp. 60–70.

Handbook on Measuring Equity in Education. (2018). Montreal: UNESCO Institute for Statistics.

Hanushek, E. (2008). Schooling, gender equity, and economic outcomes. In: M. Tembon and L. Fort, ed., *Girls' Education in the 21st Century*. Washington, D.C.: The World Bank.

Harmon, C. and Walker, I. (1995). Estimates of the economic return to schooling for the United Kingdom. *The American Economic Review*, 85(5), pp.1278-1286.

Hauser, R. and Logan, J. (1992). How not to measure intergenerational occupational persistence. *American Journal of Sociology*, 97(6), pp.1689-711.

Heckman, J. (1995). Lessons from the bell curve. *Journal of Political Economy.* 103(5), pp.1091.

Heckman, J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science*, 312(5782), pp.1900-1902.

Heckman, J. (2011). The economics of inequality: The value of early childhood education. *American Educator*, 35(1), pp.31-35.

Heckman, J. and Masterov, D. (2007). The productivity argument for investing in young children. *Applied Economic Perspectives and Policy*, 29(3), pp.446–493.

Herz, B. and Sperling, G. (2016). *What Works in Girls' Education*. Washington, D.C.: The Brookings Institution.

Indian Ministry of Women and Child Development (2013). National early childhood care and education (ECCE) Curriculum Framework.

Jeong, J., Kim, R. and Subramanian, S. (2018). How consistent are associations between maternal and paternal education and child growth and development outcomes across 39 low-income and middle-income countries?. *Journal of Epidemiology and Community Health*, 72(5).

Karoly, L., Greenwood, P., Everingham, S., Houbé, J., Kilburn, M., Rydell, C., Sanders, M. and Chiesa, J. (1998). Investing in our children: What we know and don't know about the costs and benefits of early childhood interventions. Santa Monica, CA: RAND.

Kaul, V. and Sankar, D. (2009). Early Childhood Care and Education in India. Education for All - Mid-Decade Assessment. New Dehli: National University of Educational Planning and Administration (NUEPA).

Kaul, V., Bhattacharjea, S., Chaudhary, A., Ramanujan, P., Banerji, M. and Nanda, M. (2017). The India Early Childhood Education Impact Study. Dehli: CECED, Ambedkar University Dehli, and ASER Centre.

Kaul, V., Ramachandran, C. and Upadhyaya, G. C. (1993). Impact of Early Childhood Education on Retention in Primary Grades: A Longitudinal Study.

King, E. and Hill, A. (1993). *Women's Education in Developing Countries*. Baltimore: Published for the World Bank [by] the Johns Hopkins University Press.

Kingdon, G. (1996). Private schooling in India: Size, nature, and equity-effects. *Economic and Political Weekly*, 31(51), pp.3306-3314.

Kline, R. (2004). *Principles and Practice of Structural Equation Modelling*. 2nd ed. New York: The Guilford Press, p.29.

Kohlberg, L. 1968. Early education: A cognitive-developmental view. *Child Development*, pp.1013-1062.

Kumra, N. (2008). An assessment of the Young Lives sampling approach in Andhra Pradesh, India. Young Lives Technical Notes. Oxford, England: Young Lives.

Leibowitz, A. (1974). Home Investments in Children. *Journal of Political Economy*, 82(5), pp.1111-31.

Levin, H. and Schwartz, H. (2006). Costs of early childhood care and education programmes. Background paper prepared for the Education for All Global Monitoring Report 2007 Strong foundations: early childhood care and education. New York: National Center for the Study of Privatization in Education and UNESCO.

Loeb, S., Bridges, M., Fuller, B., Rumberger, R. and Bassok, D. (2005). How much is too much? The influence of pre-school centers on children's social and cognitive development. *Economics of Education Review,* 26(1), pp.52–66.

Magnuson, K. (2007). Maternal education and children's academic achievement during middle childhood. *Developmental Psychology*, 43, pp.1497–1512.

Magnuson, K., Kelchen, R., Duncan, G., Schindler, H., Shager, H. and Yoshikawa, H. (2016). Do the effects of early childhood education programs differ by gender? A meta-analysis. *Early Childhood Research Quarterly*, 36, pp.521-536.

Magnuson, K.A., Meyers M., Ruhm, C.J. and Waldfogel, J. (2004) Inequality in pre-school education and school readiness. *American Educational Research Journal,* 41(1), pp.115–57.

Manhas, S. and Qadiri, F. (2010). A comparative study of preschool education in early childhood education centres in India. *Contemporary Issues in Early Childhood*, 11(4).

Manji, S. (2018). Getting early grade reading right: A case for investing in quality Early Childhood Education programs. United States Agency for International Development (USAID).

Marope, P. and Kaga, Y. (2015). Investing against Evidence: The Global State of Early Childhood Care and Education. Paris: UNESCO.

Meghir, C. and Rivkin, S. (2011). Econometric Methods for Research in Education. In: E. Hanushek, S. Machin and L. Woessmann, ed., *Handbook of the Economics of Education*. New York: Elsevier.

Mincer, J. (1958). Investment in Human Capital and Personal Income Distribution. *Journal of Political Economy*, 66(4), pp.281-302.

Mohite, P. and Bhatt, N. (2008). From home to school: Mapping children's transition in the Indian context. *Research in Comparative and International Education*, 3(3), pp.308-316.

Morabito, C., Vandenbroeck, M. and Roose, R. (2013). 'The Greatest of Equalisers': A critical review of international organisations' views on early childhood care and education. *Journal of Social Policy*, 42(3), pp.451-467.

Mulligan, C. (1999). Galton versus the human capital Approach to inheritance. *Journal of Political Economy*, 107(6), pp.184- 224.

Nadeau, S., Elder, L., Kataoka, N., Neuman, M. and Valerio, A. (2011). Investing in Young Children : An Early Childhood Development Guide for Policy Dialogue and Project Preparation. Washington, D.C.: The World Bank.

Nagaraja, G. and Anil, M. (2014). A sociological study of children irregularity and dropout from *Anganwadi* Centre of Kolar district, Karnataka State. *International Journal of Health Sciences and Research*, 4(3), pp.23-28.

NCERT. (1996). Identifying Numeracy and Reading Readiness Levels of Entrants to Grade I: A Synthesis. Report under DPEP, Department of Elementary Education. New Delhi.

Patrinos, H. (2008). Returns to education: The gender perspective. In: M. Tembon and L. Fort, ed., *Girls' Education in the 21st Century*. Washington, D.C.: The World Bank.

Pattnaik, J. (1996). Early childhood education in India: History, trends, issues, and achievements. *Early Childhood Education Journal*, 24(1), pp.11-16.

Pratham (2010). Annual Status of Education Report, 2009. New Delhi, India: Pratham and ASER Centre.

Rao, N. and Kaul, V. (2017). India's integrated child development services scheme: challenges for scaling up. *Child: Care, Health and Development*, 44(1), pp.31-40.

Rao, N., Sun, J., Wong, J., Weekes, B., Ip, P., Shaeffer, S., Young, M., Bray, M., Chen, E. and Lee, D. (2013). Early Childhood Development and Cognitive Development in Developing Countries: A Rigorous Literature Review. London: DFID.

Richter, L., Daelmans, B., Lombardi, J., Heymann, J., Lopez Boo, F., Behrman, J., Lu, C., Lucas, J., Perez-Escarnilla, R., Dua, T., Bhutta, Z., Stenberg, K., Gertler, P. and Darmstadt, G. (2016). Investing in the foundation of sustainable development: pathways to scale up for early childhood development. [online] Available at: http://dx.doi.org/10.1016/ S0140-6736(16)31698-1 [Accessed 20 Sep. 2018].

Rosenbaum, P. (2006). Differential effects and generic biases in observational studies. *Biometrika*, 93(3), pp.573–586.

Rosenzweig, M. and Wolpin, K. (1994). Are there increasing returns to the intergenerational production of human capital? Maternal schooling and child intellectual achievement. *The Journal of Human Resources*, 29(2), pp.670-693.

Save the Children. (2009). Situational Analysis: Early Childhood Development in India. India, New Delhi: Save the Children.

Schonemann, P. H. and Steiger, J. H. (1976). Regression component analysis. *British Journal of Mathematical and Statistical Psychology*, 29, pp.175-189.

Schultz, T. (1993). Investments in the schooling and health of women and men: quantities and returns. *Journal of Human Resources*, 23(4), pp.694-734.

Shabana, K., Mohhamad, B., Sushil, D., Mahajan, H. and Muneer, M. (2013). A cross-sectional study on client satisfaction and utilization of ICDS services in Kashmir Division of Jammu and Kashmir State. *Indian Journal of Public Health Research and Development,* 4(4), pp.206.

Shonkoff, J. and Phillips, D. (Eds) (2000). *From Neurons to Neighborhoods: The Science of Early Childhood Development*. Washington D.C., USA: National Academy Press.

Shonkoff, J. and Richter, L. (2013). The powerful reach of early childhood development: A science-based foundation for sound investment. In: P. Britto, P. Engle and C. Super, ed., *Early Childhood Development Research and its Impact on Global Policy*. Oxford: Oxford University Press, pp.24-34.

Singh, R. and Bangay, C. (2014). Low fee private schooling in India – More questions than answers? Observations from the Young Lives longitudinal research in Andhra Pradesh. *International Journal of Educational Development,* 39, pp.142–50.

Singh, R. and Mukherjee, P. (2015). Determinants of Successful Completion of Secondary Education: Evidence from Andhra Pradesh. Young Lives Working Paper 143, Oxford: Young Lives.

Singh, R. and Mukherjee, P. (2016). Education Trajectories: From Early Childhood to Early Adulthood in India. Oxford, England: Young Lives.

Singh, R. and Mukherjee, P. (2017). Comparison of the Effects of Government and Private Preschool Education on the Developmental Outcomes of Children: Evidence from Young Lives India. Oxford, England: Young Lives.

Singh, R. and Mukherjee, P. (2018). Effect of preschool education on cognitive achievement and subjective wellbeing at age 12: evidence from India. *Compare: A Journal of Comparative and International Education*, pp.1-19.

Solon, G. (1999). Intergenerational mobility in the labor market. In: O. Ashenfelter and D. Card, ed., *Handbook of Labor Economics*. Amsterdam: Elsevier Science, pp.1761-1800.

Streuli, N., Vennam, U. and Woodhead, M. (2011). Increasing choice or inequality? Pathways through early education in Andhra Pradesh, India. The Hague, The Netherlands: Bernard van Leer Foundation.

Strong Foundations: Early Childhood Care and Education. (2006). EFA Global Monitoring Report. Paris: UNESCO.

Sylva, K., Melhuish, E., Sammons, P., Siraj-Blatchford, I. and Taggart, B. (2011) Pre-school quality and educational outcomes at Age 11: Low quality has little benefit. *Journal of Early Childhood Research,* 9(2), pp.109-124.

The Global Gender Gap Report 2017. (2017). [online] Geneva: The World Economic Forum. Available at: http://www3.weforum.org/docs/WEF_GGGR_2017.pdf [Accessed 2 Sep. 2018].

Todd, P. and Wolpin, K. (2004). On the specification and estimation of the production function for cognitive achievement. *Economic Journal,* 113(485), pp.F3–33.

Tooley, J. and Dixon, P. (2003). Private Schools for the Poor: A Case Study from India. Reading: CfBT.

Uis.unesco.org. (2018). India | UNESCO UIS. [online] Available at: http://uis.unesco.org/country/IN [Accessed 3 Sep. 2018].

Un.org. (2018). *United Nations Millennium Development Goals*. [online] Available at: http://www.un.org/millenniumgoals/ [Accessed 19 Sep. 2018].

UNESCO (2000). Dakar Framework for Action: Education for All. Meeting Our Collective Commitments. World Forum on Education, Dakar, Senegal, 26-28 April, 2000, UNESCO, Paris.

UNICEF (2007). Policies for Early Childhood Development, http://www.unicef.org/earlychildhood/index_40752.html (accessed 20 August 2018).

Unterhalter, E. (2007). *Gender, Schooling and Global Social Justice*. New York: Routledge.

Vennam, U., Komanduri, A., Cooper, E., Crivello, G. and Woodhead, M. (2009). Early Childhood Education Trajectories and Transitions: A Study of the Experiences and Perspectives of Parents and Children in Andhra Pradesh, India.

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

Woodhead, M. (2006). Changing perspectives on early childhood: theory, research and policy. Paper commissioned for the EFA Global Monitoring Report 2007, Strong foundations: early childhood care and education.

Woodhead, M. (2009) Pathways Through Early Childhood Education in Ethiopia, India and Peru: Rights, Diversity and Equity. Oxford: Young Lives.

Woodhead, M., Ames, P., Vennam, U., Abebe, W. and Streuli, N. (2009). Equity and quality? Challenges for early childhood and primary education in Ethiopia, India and Peru. Working Paper 55, The Hague: Bernard van Leer Foundation.

Woolridge, J. (2016). *Introductory Econometrics: A Modern Approach*. 6th ed. Boston: CENGAGE Learning.

www.worldbank.org (2018). *Education Statistics*. [online] Datatopics.worldbank.org. Available at: http://datatopics.worldbank.org/education/country/india [Accessed 12 Jul. 2018].

Yoshikawa, H., and Nieto, A., (2013). Paradigm shifts and new directions in research on early childhood development programs in low and middle income countries. In Britto, P.R., Engle, P.L. and Super, C. (2013), *Handbook of early childhood development research and its impact on global policy* (pp.487-497). Oxford: Oxford University Press.

Young Lives (2017). Cohort Maintenance – Tracking and Attrition. A Guide to Young Lives Research. [online] Oxford: Young Lives. Available at: http://www.younglives.org.uk/sites/www.younglives.org.uk/files/GuidetoYLResearch-S12-CohortMaintenance.pdf [Accessed 4 Oct. 2018].

Young Lives Survey Design and Sampling in India. (2014). [online] Young Lives. Available at: https://www.younglives.org.uk/content/young-lives-survey-design-and-sampling-andhra-pradesh-india [Accessed 20 Sep. 2018].

Young, M. ed., (2007). Early child development from measurement to action: a priority for growth and equity. Portland: The World Bank.

Younglives-india.org. (2018). Young Lives India Home | Young Lives India. [online] Available at: https://www.younglives-india.org/ [Accessed 29 Sep. 2018].