Student Paper

Young Lives

An International Study of Childhood Poverty

Determinants of Malnutrition among Children in Andhra Pradesh, India

Adnan M.S. Fakir

Paper submitted in part fulfilment of the requirements for the degree of MSc in Economics, School of Business, Management and Economics, University of Sussex.

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

Young Lives is funded by UK aid from the Department for International Development (DFID) and co-funded by the Netherlands Ministry of Foreign Affairs from 2010 to 2014 and by Irish Aid from 2014 to 2015.

UNIVERSITY OF SUSSEX

SCHOOL OF BUSINESS, MANAGEMENT AND ECONOMICS

DISSERTATION COVER SHEET

Candidate Number: 103847
Course Code: 838L1
Title of Dissertation: Determinants of Malnutrition among Children in Andhra Pradesh, India.
Masters Course: MSc Economics
Number of Words: 14,958
Date: 02nd September, 2013

Abstract:

This essay looks into two main determinants of child health: income and education, using a panel dataset from India. Impact of per capita consumption expenditure on child nutritional status is investigated using a number of estimation methods including two stage least squares and panel methods. Income effect is found to explain only between 0 to 34 percent of the improvement in child health. Maternal education effect on child nutritional status is found to be stronger in urban areas and among the wealthier. Paternal and community level education are also found to have significant impact on child health. Finally, paternal education is found to have significant positive impact for smaller communities, while maternal education for larger communities.

Candidate Number: 103847

Contents

List of Tables: **Page**

Additional Figures and Tables in Appendix

Preface

I would like to express my sincere gratitude to Prof. Andy Mckay for his guidance and understanding. I would also like to thank Amrita Saha and Vanika Grover for fruitful discussions, Dr. Torfinn Harding for econometric suggestions, and Maisha M Khan and Anika Ali for their help in proof-reading the essay. Finally, this essay was written during some troubling times and I am indebted to my father, Md. Golam Samdani Fakir, for his constant inspirational words that eventually lead to its completion.

Disclaimer:
The data used in this publication come from Young Lives, a 15-year study of the changing nature of childhood poverty in Ethiopia, India (Andhra Pradesh), Peru and Vietnam (www.younglives.org.uk). Young Lives is funded by UK aid from the Department for International Development (DFID) and co-funded from 2010 to 2014 by the Netherlands Ministry of Foreign Affairs. The views expressed here are those of the author(s). They are not necessarily those of Young Lives, the University of Oxford, DFID or other funders.

1) Introduction

The economic improvements in South Asian countries over the past decade have not been adequately reflected in improvements in child nutrition (Claeson, Bos et al., 2000; Glewwe, Agrawal and Dollar, 2004; Thang and Popkin, 2003). While mild to moderate malnutrition has been recognized to account for more than half of the 10 million children dying each year from preventable diseases (Black, Morris and Bryce, 2003), it is even more disquieting that half of the world's malnourished children are concentrated in three countries - Bangladesh, Pakistan and India. Not only in numbers but the proportion of children affected by malnutrition is much higher in South Asia, with an estimated 57 out of every 100 children being malnourished (El-Ghannam, 2003).

In fact, even with similar levels of purchasing power and per capita food production as sub-Saharan Africa, malnutrition persistently remains higher in South Asia. This has been dubbed as the "South Asian Enigma" by Ramalungaswami, Jonsson and Rohde (2003) and the authors attribute this differential in child growth primarily to lack of care and nutrition. Malnutrition causes a child to be more vulnerable to infectious diseases with several identified long term impacts. Poor educational attainment (Hall, Khanh et al., 2001), delayed cognitive development (Mendez and Adair, 1999) and lower intellectual and physical abilities in adult life leading to lower lifetime earnings (Strauss and Thomas, 1998) have been all linked to deprived nutritional status.

As a broad function, malnutrition is determined by health inputs, the local health environment and the child's genetic endowment (Glewwe, Koch and Nguyen, 2002). This essay scrutinizes the role of two major health input determinants, income and education, on a child's growth as measured by height-for-age and weight-for-age. As is generally understood and according to the UNICEF framework of the determinants of child under-nutrition, both income and education are part of the basic causes that lead to the underlying and immediate sources of short term and long term consequences of malnutrition.

A wide array of factors play as health inputs such as nutrient intake, level of care, quality of medical services, sanitation and toilet facilities and drinking water purity, many of which are correlated with both income and education. Thus the 'income effect' and 'education effect' transmit at least partially the effects of the health inputs on a child's health status. Given the multitude of pathways through which the two variables can establish their relationship with child status, it is important to understand their mechanisms in order to establish effective policies.

The Young Lives three round panel dataset between 2002 and 2009 from Andhra Pradesh, India is used to perform the analyses. World Bank reports that during this time India had maintained an annual real growth rate of above 8 percent on average. Even with the impressive growth, malnutrition rates remain high making it particularly relevant to identify the role played by income on child's health.

The contributions of this essay are as follows: (i) The use of a panel dataset to estimate the impact of income on child health while controlling for unobserved time-invariant individual and household characteristics. Concerns of endogeneity, omitted variables and measurement bias are also systematically addressed. (ii) The role of wealth index vis-a-vis consumption expenditure as proxies of income is also explored. (iii) Investigation of the differential impact of maternal education by age, urbanity and wealth on child nutritional status, while accounting for confounders, some of unobserved heterogeneity and using community fixed effects. (iv) Finally, the role of paternal education and how community level education augments with parental education in regard to the education effect on child health is studied.

There are conflicts regarding the role of each of these attributes on child health in the available literature. Where household income was overemphasized in the past, the significance of other household attributes, such as female literacy (Behrman and Wolfe, 1984), autonomy (Caldwell, 1993), and maternal health knowledge (Glewwe, 1999) have been gradually recognized. The role of these attributes is discussed in the literature review and when we introduce the variables for this paper. Majority of existing studies either use nutrient calorie intake or anthropometric measures for malnutrition analysis. Our analysis involves the latter. Height-for-age and weight-for-age measure child growth relative to its potential reflecting chronic and acute nutritional deprivation (Kynch and Maguire, 1998). The specific anthropometric measures, their advantages and usage in this paper are discussed in the literature review and when we introduce the Young Lives dataset.

The structure of this paper is as follows: Section II presents a literature review summarizing the understandings of anthropometric measures to date, contradictory findings of income and education in their relationship to child health as well as the various problems faced in their estimation. Section III outlines a brief theoretical model of reduced child health demand function and discusses the advantages and drawbacks of different econometric specifications and how they tackle the problems of estimation mentioned in the literature review. Section IV introduces the Young Lives dataset including their sampling methodology, a descriptive statistics of the dataset and introduces the variables to be used for analyses in this study. Finally sections V and VI present and discuss the results of the role of income and education on child health respectively. Section VII concludes the essay with a succinct review of findings, limitations and policy suggestions.

2) Literature Review

2.1) Anthropometric measures of malnutrition

Three main concepts have been introduced by Waterlow, Buzina et al. (1977) in regard to measuring malnutrition: stunting (low height-for-age), underweight (low weight-for-age) and wasting (low weight-for-height). Stunting is generally considered an irreversible, longer term cumulative indicator of poor physical growth due to slow growth in height compared to the reference healthy population. Glewwe, Koch and Nguyen (2002) point out that in developing countries such chronic malnutrition is a consequence of recurring episodes of diarrhea, childhood infectious diseases and inappropriate diet.

Longitudinal studies in India and Vietnam have shown that children who were stunted were at a disadvantage in terms of later cognitive, well-being and psychological outcomes. In Peru, children who were stunted at 2 years showed lower levels of cognitive ability at age 5 (Sanchez, 2009) and in Ethiopia, at age 12, stunted children were nearly one whole grade below compared to others (Dercon, 2008). There is also a clear link between socio-economic status and stunting. In Peru, over 50 percent of children from the poorest quintile were found to be stunted compared to fewer than 10 percent in the richest quintile (Pells, 2011).

Wasting on the other hand is attributed to shorter-term acute malnutrition and is considered to be reversible with improved conditions, permitting faster responses than stunting. Wasting however does have the disadvantage of classifying children with lower height as 'normal', something that ideally should be controlled with the stunting measure.

Being underweight can reflect both stunting and wasting due to lack of protein energy intake and micronutrient deficiencies. Due to its nature of not distinguishing between acute and chronic malnutrition, interpreting underweight is not as straightforward as the other two. Underweight children in India are amongst the highest in the world, double to that in Sub-Saharan Africa. According to World Bank's Health, Nutrition and Population Study (Gragnolatti, Shekar and Das Gupta, 2005: 1), "47 percent of children under three were underweight or severely underweight, and a further 26 percent were mildly underweight", totaling to about three-quarter of children in India suffering from malnutrition.

UNICEF (2013) stresses the importance of obtaining optimal growth before 24-months of age. Over the past decade, it has become clear that improvements in nutrition post 24-months usually do not recover the lost potential due to nutritional deprivation in the first 24 months. Cognitive and neural developments in children take place from pre-natal till 24-months child age period. Nutritional deficiencies during this time can thus lead to long term consequences. This redirects attention towards undernourished mothers who are more likely to give birth to underweight babies, who in turn are prone to the long term effects of stunting.

Another consequence of undernourishment that is becoming clearer is coined the "dual burden" of malnutrition or "nutrition transition" (Shetty, 2002). Stunted children who have rapid weight gain later in life are more prone to coronary heart disease, stroke, hypertension and diabetes (UNICEF, 2013). This obesity burden constitutes the other spectrum of the dual burden of malnutrition. This thesis will however, focus on the issue of the former.

The anthropometric measures are expressed as z-scores calculated from comparison with a reference healthy population selected by the National Center for Health Statistics, in accordance with the WHO recommendations (WHO, 1986). For example, the height-for-age z-scores for stunting are calculated as follows:

$$z\ score = \frac{H_i - H_r}{SD_r}$$

where, H_i is the height of the child in question, H_r is the median height of the reference healthy population of children of same age and gender and SD_r is the standard deviation of the height of the group of children in question from the reference population. Similarly the z-scores for underweight are obtained by comparing the weight of the child in question with the median weight of a reference population. In case of wasting the weight of the child in question is also compared with the median weight of a reference population but those with the same height as the child in question.

Out of the three indexes, stunting and wasting are normally the preferred measures of child nutritional status because of their distinguishable property. Because stunting takes into account the long run social conditions and cumulative nutritional status, WHO (1986) recommends stunting as the most reliable measure of child health. It is also worth noting here, that accurate data collection on age is especially problematic in rural areas of developing countries and can lead to biased results (Bairagi et al., 1982). Reliability of the data is crucial which makes anthropomorphic data collection a slow and expensive initiative.

2.2) Role of income on child health

As expected, income plays an important role in child health determination. While the causal effect of income on child health has been questioned by several authors (Glewwe, Koch and Nguyen, 2002, Kuehnle, 2013), unarguably a large portion (if not all) of income's role on child health operate through various other factors, such as the health inputs. Many inputs, such as food intake, household sanitation, quality of medical care received that are correlated with child health, are also correlated with income (Behrman and Wolfe, 1984). In other words, the estimated income coefficient will reflect at least a portion of the effect of the health inputs on child health. Income on the other hand also depends on level of education. Thus the estimated income coefficient also potentially captures an "income effect" portion of the "total education effect" (Thomas, Strauss and Henriques, 1991). One should ideally control for these (unobserved) heterogenic factors to better understand the determinants of child health.

Sarmistha (1999) points out that the instrument of choice used to measure income is crucial in analysis as estimates often differ based on the variable of choice. Per capita income, consumption expenditure and wealth indexes all provide, to various extents, measures of the socioeconomic status and resource availability of the household, each with their advantages and problems. While permanent income or household expenditure tends to be better estimates of resource availability, they are difficult and expensive to measure. Hence even though current income tends to have transitory components it is often used as a proxy.

Thomas, Strauss and Henriques (1991) point out some measurement problems related to current income, namely (i) respondents may be unwilling to disclose their income, (ii) income from self-employment is hard to measure and (iii) if only the mother of the household is surveyed, she may not know the total income of the household. It is likely that the measure of current income in our dataset also faces these common problems. Some of the measurement problems pointed out above can also translate to consumption expenditure measurement, which can bias regression estimates. However expenditure data is likely to be more accurate and better reflect permanent income, and thus is more suitable for such studies.

Furthermore, income is endogenous in nature which gives rise to the possibility of simultaneity bias. For example parents with children who are ill may decide to work more hours to pay for better medicine, thus earning more income. Since such negative shocks could increase household income, the impact of income on child health would be underestimated. The opposite might also be true. Parents might work less hours due to child's illness thereby earning lower income causing ordinary least squares to overestimate its impact on child health (Glewwe, Koch and Nguyen, 2002). To avoid this one can use non-labor income or value of household assets in place of current income that are more robust to shocks (Thomas, Strauss and Henriques, 1991). However, the most popular method to avoid the simultaneity bias is to use instrument variables for measures of income or consumption.

Using a two round panel Vietnam Living Standards Survey (VLSS) from 1993 and 1998, Glewwe, Koch and Nguyen (2002) estimate the role of growth in household income on child health for Vietnam to find small and questionable impact. The instruments they used to tackle the endogeneity problem in cross sectional analysis consist of types of agricultural land allocated to the household and non-labour income. For the panel estimates, because first difference regression was used and an instrument that predicts changes in household expenditure over time is more appropriate (Deaton, 1997), the authors used "changes in household income" as an instrument. However, this only tackles part of the measurement bias problem and assumes that expenditure can be considered exogenous.

Alderman, Hoogeveen and Rossi (2005) in a longitudinal study in Tanzania also instrument per capita consumption using per capita household income and quality of household roof under the same assumption of expenditure exogeneity. Under the circumstances that the measurement bias is comparatively a larger issue than income endogeneity, the regression could lead to near accurate estimates. However such very well might not be the case. It is very difficult to find relevant and dependable instruments that tackle the endogeneity problem especially in panel estimates. The authors also used random effects regression instead of fixed effects which assumes that the unobserved individual heterogeneity, such as parental preferences, is uncorrelated with the covariates in question (Clark and Linzer, 2012). This particular issue will be discussed further when the results of the panel regression are presented.

The Young Lives dataset used in this study contain wealth index information for all three rounds and consumption expenditure information for the latter two rounds, which are used for the analysis. We introduce the variables as well as the instrument variables used in more detail in section IV.

2.3) Role of education on child health

The role of education in relation to child health has been closely scrutinized in the available literature. Smith and Haddad (1999) report that almost half the reduction in underweight children between 1970 and 1995 is explained by increases in female adult literacy between 1970 and 1995. In societies where mothers are the main caregivers of the child, maternal education has been shown to have a stronger and significant effect on child health than paternal education (Behrman and Wolfe, 1987; Murthi, Guio and Dreze, 1995; Bishai, 1996). As majority of the studies focus on maternal education, this literature review begins with the role of maternal education on child health in focus.

While many studies report strong correlation between maternal education and child health, cross-section estimates do not establish a causal link between the two. Even though randomized experiments are the most powerful methodology to establish causal links, ethical issues arise when assigning determinants of malnutrition, such as education, as a random component in an experiment. Hence the studies are mainly from observational data and more than often have been from cross-sectional data rather than longitudinal.

Furthermore, Behrman and Rosenzweig (1994) mention that many of the large-scale studies showing a link between maternal education and child nutrition are not as informative because they do not set up controls for inter-generationally correlated genetic endowments. Since taller parents are more likely to have children with better health endowments, one way of controlling for this is by using the heights of parents, as demonstrated by Glewwe (1999).

The causal link between mother's education and child health has been strongly questioned. Studies have found a strong correlation between maternal education and indicators of care, such as better sanitation practices, better child feeding, timely immunization etc. (Bishai 1996; Bloom, Wypij and das Gupta, 2001; Desai and Alva, 1998). Cebu Study Team (1991) used a Philippines longitudinal data to find that educated mothers are better at recognizing threats to the health of their children. The causal link is that mother's education induces behavioral changes which lead to lowered prevalence of childhood diarrhea.

In an interesting cross-country study, Desai and Alva (1998) find that including regional and community fixed effects lowers the impact of education. They conclude that mother's education level thus acts as a proxy for geological area of residence and socioeconomic status, which must be controlled for in order to investigate the role of parental education. Thomas, Strauss and Henriques (1991) on the other hand showed that almost all the impact of maternal education on child survival could be explained by how easily the mother's could access information, such as read newspapers, watch TV and/or listen to radio.

Glewwe (1999) proposed that maternal literacy operate on child health primarily through three mechanisms: (i) through formally taught health knowledge, (ii) through numeracy and literacy skills that bring about cognitive development, and (iii) by making women more receptive to modern medicine. Using a Moroccan dataset, he concluded that health knowledge, acting as a mediator, does improve child health while numeracy and literacy skills promote uptake of health knowledge outside the classroom bringing about behavioral change. The finding is consistent with that of Desai and Alva (1998) and Frost, Forste and Haas (2005) who also concluded that maternal education influences health seeking behavior which in turn influences child health.

In a study in Bangladesh, Bhuiya, Streatfield and Meyer (1990) identify specific types of health knowledge associated with higher levels of education, including (i) washing hands after latrine use, (ii) use of oral rehydration therapy to treat diarrhea, (iii) awareness of boiling water, and (iv) contagions as a means of spread of disease. Such health knowledge leads to the behavioral changes that result in better household health-seeking practices and subsequently better child health.

It is important to see how endogeneity plays a role here. Knowledge about health and nutrition can also be passed down over generations, from the grandmothers to the mothers, as another form of "genetic endowment." Parents who have less nutrition knowledge are more likely to have sickly children and thus are more likely to seek out help. Similarly, a mother whose child has severe diarrhea is more likely to know about oral rehydration salts (ORS) than a mother whose child does not have diarrhea (Moestue, 2005). Thus endogeneity occurs when such unobserved 'genetic endowments' to be sick or short are causally related to both knowledge and child nutrition. This can also be seen as a case of reverse causality. Glewwe (1999) has shown that ignoring these 'effects' can lead to an underestimation of the impact of knowledge and is once again controlled for, to some extent, by including parental heights.

Another factor that can heavily influence maternal education role is their labour-force participation, leading to a differential impact of education, in this case based on socioeconomic wealth status. Educated mothers tend to be more employed in the labour force without simultaneously being able to maintain adequate child care leading to a negative effect on child's nutrition. This is shown by a study in Benin where a negative association was found (Reed, Habicht and Niameogo, 1996). Thus it would be naive to assume that improved maternal education always impacts child health positively. However there can be positive consequences as well for workaholic mothers. Barrera (1990) shows this in Philippines that even though educated mothers weaned their children sooner, they provided better health care leading to better overall child nourishment.

Summarizing the above discussion, one can safely conclude that the causal link between maternal education and child health can take place through various mechanisms, including but not restrictive to health knowledge. Geography of residence, socioeconomic status, health knowledge, general awareness, attitude towards medical services, degree of autonomy, practiced reproductive behavior all form a part of the maternal effect on child health (Frost, Froste and Haas, 2004).

Finally it is worth noting that while education increases awareness regarding the value of health including the ability to better communicate with health care providers, education and empowerment, and thereby better child health through empowerment, does not necessarily go hand in hand (Caldwell, 1979, LeVine, LeVine and Schnell, 2001). In other words, education does not automatically translate to empowerment. For example, in conservative areas of India, especially where the concept of arranged marriage is still very prevalent, level of education attainment has become an issue of social status in order to marry men of suitable socioeconomic standing. After marriage, the educated women may not have enough autonomy to have as significant an impact on child health as expected. (Moestu, 1996). This can also be seen as an example of heterogeneity in parental attitude towards children, which if additive in nature, can be controlled through fixed effects estimation.

One other issue needs to be elaborated before moving on to introducing the theoretical model. It is regarding the differential effect of maternal education, especially in regard to child gender and geographical residence. Sarmistha (1999) uses a probit model of nutritional status on a 1987-89 dataset from West Bengal to find that female literacy improves the nutritional status of boys at the cost of girls, strongly implying that male and female children follow different health functions. On the other hand, Bourne and Walker (1991) using 1981 Indian census data find that maternal education reduces mortality rates of girls more than boys.

The rural-urban differential of maternal education is also often stressed, with a greater impact of education in urban areas. While Bicego and Boerma (1993) attribute this due to availability of broader social and economic support, Glewwe, Koch and Nguyen (2002) and Caldwell (1994) suggest that a synergistic interaction between health services and education play a more important role. An example would be from Caldwell (1979), who shows that in Nigeria the benefit of maternal education is greater in villages with access to a hospital than those without it. On this understanding and to explore such an effect, "distance to nearest medical facility" is often added as part of the regression (Glewwe, Koch and Nguyen, 2002).

Finally studies have also shown that maternal education can have a differential impact based on the overall community level of education. This is based on the concept of a spillover effect, a positive externality, from educated individuals to uneducated individuals living within close proximity. If such a community level correlation exists, it may interfere with the causal pathway between education and child nutrition and should be controlled for (Desai and Alva, 1998).

Moestue (2005) uses the first round of Young Lives dataset to explore this and has found parental education to have a stronger impact on child nutrition where community level maternal literacy is higher. Cleland and Jejeebhoy (1996) report lower fertility rates amongst women with little or no education but living in educated communities, compared to similarly educated women living in communities with lower overall education level. These findings emphasize two important factors. The first is to include community fixed effects when analyzing maternal education causal pathways. The second is the need to investigate the role of community level education on child health, along with parental education.

This essay will be focused on using the Young Lives dataset which although contains rich individual, household and community level data on child nutrition and education, does not contain rich maternal health knowledge information. Consequently, differential impact of parental education is emphasized instead in this study while controlling for community-level, child-level and genetic endowment factors[1].

3) Theoretical model and econometric concerns

A theoretical model of a child's health status is first presented in this section followed by a discussion of the different econometric problems faced in its estimation and ways on how to handle them. The best place to begin modeling to understand the determinants of a child's health is from the child's health production function. The model presented here is adapted from Glewwe, Koch and Nguyen (2002).

$$H_{ict} = f(HI_{ict}, E_{ct}, \alpha_i)$$

The health of an individual child *i*, living in community *c* at time *t*, H_{ict}, can be determined by primarily three types of variables: a vector of time variant and time invariant observable individual health inputs HI_{ict}, such as pre-natal care, food nutrient intake, quality of medical care, medicine, household sanitation and toilet facilities, drinking water quality etc.; a vector of the local health environment E_{ct}, at community level *c* and time *t,* which includes the characteristics of the local community that directly affect the child's health status such as local disease prevalence, air and water pollution levels etc.; and finally time invariant child's genetic health endowment, α_i, inherited form his/her parents that also directly affects his/her health. This function essentially represents child health at a macro level grouping the elements of individual health inputs into vectors of economic, social and environmental factors.

The time invariant genetic health endowment is normally considered exogenous to child's health. The local health environment is also exogenous although, as Glewwe, Koch and Nguyen (2002) point out, can be argued to be endogenous given that households "migrate to healthier environments or take measures to improve the local health environment." This health environment variation at the community level can be controlled for using community fixed effects which is done so in this study. This is discussed in more detail later in this section.

[1] As a further study into the maternal health knowledge causal pathways, a separate "Knowledge and Networks" cross-sectional dataset is available that uses a sub section of 302 women from the Young Lives study in Andhra Pradesh.

Complete information to estimate a health production function, H_{ict}, using all variables $HI_{ict}, E_{ct}, \alpha_i$, is rarely available, and that can often lead to omitted variable bias. Furthermore, a health function is dynamic, meaning past health can impact current health, and should take into account the cumulative process of a child's growth, thus requiring information on past time periods (Strauss and Thomas, 2008). An alternative route is to consider the determinants of health inputs and substitute that into the production function.

$$HI_{ict} = g(Y_{ict}, MS_{ict}, FS_{ict}, \eta_i, E_{ct}, \alpha_i)$$

As discussed in the literature review, the health inputs depend on household income, Y_{ict}. In addition, parental education levels, mother's schooling MS_{ict}, and father's schooling , FS_{ict}, as well as time invariant unobservable parental preferences η_i, determine both quantity and quality of health inputs that a child receive. Definitions of E_{ct} and α_i remain as before.

One variable clearly left out of the equation is the presence of other siblings. The number of children a family is willing to have is an endogenous decision process. Not only does the decision to take a second child depend on the health of the firstborn but educated parents normally prefer to have fewer children (Qian, 2009). However David, Moncada and Ordonez (2004) point out that if family planning is not adequately practiced, this endogeneity is likely to be small. Considering the focus on family planning being carried out in India by various organizations over the past years, this variable is taken as endogenous and thus left out.

As discussed earlier in the literature review, income is also an endogenous variable. Parents adjust their work hours, depending on the health of their children, which affects their income. This problem is handled using instrument variables as will be soon discussed. Substituting the Health Inputs equation into the health production function thus yields:

$$H_{ict} = f(Y_{ict}, MS_{ict}, FS_{ict}, \eta_i, E_{ct}, \alpha_i)$$

If we use child's height-for-age z-score as an indicator of child's health status, H_{ict}, the reduced form health demand function becomes:

$$zscore_{ict} = \beta_0 + \theta_0 time_t + \beta_1 Y_{ict} + \beta_2 MS_{ict} + \beta_3 FS_{ict} + \beta_x X_{ict} + \nu_{ict}$$

$$where, \nu_{ict} = \eta_i + E_{ct} + \alpha_i + \mu_{ict}$$

All definitions remain as before; μ_{ict} is an idiosyncratic error term and X_{ict} is a vector of exogenous child and household characteristics, such as child age and child gender. This is the primary estimated equation of interest for section V where we investigate the role of income on child health, with β_1, β_2 and β_3 being the main parameters of interest. In section VI, where the role of education on child health is investigated, the same estimation problem holds except now the variables are gradually added, including a community level education variable, while looking into differential impacts. Instead of per capita income, the Young Lives dataset contain per capita consumption expenditure data as well as a constructed wealth index as proxies for household socioeconomic status. Per capita consumption expenditure is likely to be more accurate and better reflect a household's permanent income and hence is used as the main variable of interest.

Several points need to be addressed here. Ordinary least squares (OLS) estimates will be consistent and unbiased only if $E(\nu_{ict} | Y_{ict}, MS_{ict}, FS_{ict}, X_{ict}) = 0$. However, if any one of the η_i, E_{ct} or α_i variables is correlated with the covariates such that $E(\nu_{ict} \mid Y_{ict}, MS_{ict}, FS_{ict}, X_{ict}) \neq E(\nu_{ict})$, OLS estimates will be biased and inconsistent. A good example of this is with parental preferences for child's health, η_i. Some parents may be more responsible than others thus putting a higher weight on their child's health in their utility function. Being more responsible also often implies having higher income. Due to this correlation between η_i (which is contained in the residual) and Y_{ict}, OLS would lead to an overestimation of the role of income on child health, also picking up the parental care for their children and not *only* the impact of income (Kirchberger, 2008).

This problem is addressed in several ways. Parental preferences often depend on ethnicity and religion which is controlled for using dummy variables in each of the regressions for analyses, unless otherwise mentioned. This partially approximates parental preferences across individual children. In model specifications where instrument variables are used for income, the bias due to correlation between income and the unobserved preferences for child health will also be reduced, if not eliminated. However, care should be taken in interpreting the accuracy of the coefficient as it will depend on the explanatory power of the instruments used. This is discussed in more detail a little later.

First difference and fixed effects panel regressions are also estimated. Random effects estimations is not used as it assumes the unobserved heterogeneity to be random and uncorrelated with the predictor variables which does not match our economic theory. It is also rejected by the robust Hausman test (results reported in section V under panel estimation). Assuming that parental preferences can be treated as an additive fixed effect in that they are time invariant with no effect on growth, the differencing will remove the variable out of the estimation.

However, fixed effects estimation has a number of shortfalls. Firstly, Deaton (1997) points out that measurement bias is greatly aggravated when differences in variables are regressed on each other. Secondly, all time invariant variables drop out of the estimation, making the procedure not adequate to estimate the effect of such a variable, for example mother's education, on child health[2].

Thirdly, both first difference and fixed effects estimates are biased unless the assumption of strict exogeneity is upheld. Theoretically, we know that we have endogeneity in our model which will cause problems. Finally, in the presence of autocorrelation, that is when the error term tends to be correlated over time, first difference gives us more appropriate estimates. For these various possibilities, both the first difference and fixed effects estimates are reported.

In order to control for differences across communities, including differences in the health environment E_{ct}, community fixed effects are used in all regressions, unless otherwise mentioned. The other concern is to control for child's genetic endowment α_i. As discussed in the literature review, to control for this mother's height is included in the regression models. Since girls are also typically healthier than boys, inclusion of child gender, in addition to mother's height, also partially acts as a control. Furthermore, when using panel estimation this is treated as a fixed effect and differenced out. It is also worth noting that heteroskedasticity tests have rejected the null of homoskedasticity for the Young Lives dataset, hence robust standard errors are reported in all cases. Results of Breusch-Pagan and White tests are provided in section V, under cross-sectional estimation.

The last problems to address in this section are biases due to income endogeneity and measurement errors. In theory, an ideal instrument should be able to remove both the endogeneity and measurement biases. The instruments should be able to sufficiently explain the variation in household income but not be correlated with the unobserved determinants of child health or the measurement bias in consumption expenditure.

Following Glewwe, Koch and Nguyen (2002) and Alderman, Hoogeveen and Rossi (2005), in this study, the following instruments are used for cross-section analyses: (i) amount of irrigated land

[2] While this can be adjusted for by including time dummy variables, the time invariant variables and their interaction terms, the methodology is not pursued in this study.

owned, (ii) amount of non-irrigated land owned, (iii) number of relatives living abroad, (iv) whether the household received any non-labour income from social security or social subsidy in the last 12 months, and (v) whether the household received any rent in the last 12 months. The reason why we can use variable (iii) as an instrument is because the higher the number of relatives living abroad, the more remittance the household is likely to receive, which in turn influences the household expenditure pattern.

As will be seen in section V, although these instruments have statistically significant predictive power according to the F-test on excluded instruments, they are relatively weak in explaining per capita consumption expenditure in terms of its first stage R-square statistic leading to large standard errors. Finally, it should be noted that in difference based panel estimates, it is preferable to identify instruments that explain changes in consumption expenditure over time (Glewwe, Koch and Nguyen 2002). Unfortunately, no instrument that can handle such endogeneity at panel level was identified. Thus, although panel estimates remain free from unobserved heterogeneity and omitted variable bias, they are reported without handling for simultaneity or measurement biases.

4) Data and Descriptive Statistics

4.1) Young Lives Dataset

A three round longitudinal Young Lives dataset for Andhra Pradesh, India spanning between 2002 and 2009 was used for this study. Young Lives is an international project being carried out by University of Oxford in partnership with London School of Hygiene and Tropical Medicine, Save the Children UK and local organizations in four countries: India, Vietnam, Peru and Ethiopia. The project is being funded by UK's Department for International Department (DFID).

Andhra Pradesh is the fifth largest state in India with only 27 percent of its population living in urban areas. Poverty estimates for rural Andhra Pradesh is relatively lower at 11.2 percent compared to the national average of 28 percent. This is interesting because per capita expenditure in rural areas is only about 5 percent higher questioning the underlying poverty measurement process (Young Lives, 2008). Even though Andhra Pradesh has been gradually shifting away from agriculture to service sector, thanks to its IT revolution, significant disparities remain in terms of urbanity, religion, caste and ethnicity, as is also reflected by the Young Lives data.

In Andhra Pradesh, approximately 2,000 children belonging to a 'younger cohort' and another 1,000 children belonging to an 'older cohort' were sampled from 20 sites. The younger cohort of 1-year-old children were surveyed from the age of 6 to 18 months and every 3 years following, while children from the older cohort of 8-year-olds were surveyed from the age of 7.5 to 8.5 years and every 3 years following. The first round of data was collected in 2002, round two in 2006 and round three in 2009.

The 20 sites or geographic clusters were selected semi-purposely with an intention of over-sampling the poor. The 20 sites include 1 site from an urban slum in Hyderabad and 19 sites from 6 districts of Andhra Pradesh: Anantapur and Cuddapah in Rayalaseema; Karimnaga and Mahboobnager in Telangana; and West Godavari and Skrikakulam in the Coastal region. While these sites do cover a wide array of population characteristics by urbanity, ethnicity, religion and levels of development, the Young Lives dataset cannot be said to be nationally representative. Details about the fieldwork can be found in Attawell (2003).

Random sampling methodology was applied to select study households within a sample of administrative units or communities located in each site. In Andhra Pradesh, this amounted to 102 communities. Total number of children surveyed in round 01 is 3019 from both cohorts, 2944 from round 02 and 2937 from round 03. Although this makes the panel dataset unbalanced to a degree, there is relatively low attrition rate over the three rounds; only 2.7 percent from round 01 to round 03 including child deaths, and 2.2 percent excluding child deaths (Galab et al., 2011).

What makes the Young Lives dataset especially attractive is the depth of data collected and its focus on data accuracy, especially in that of recording anthropometric measures. Three main types of questionnaires were used, which varied from round to round according to the age of the child. The first is child level questionnaires containing data on child health and nutritional status, child care, delivery environment, breastfeeding practices, child activities, child education and child work. The second is household level questionnaires containing information on education for each of the household members, caretaker background, socio-economic status, livelihoods and access to services and social safety nets. Finally community level questionnaires were also used with information on demographics and social and environmental characteristics, including education and health.

Certain aspects regarding the dataset need to be noted here. The Young Lives dataset builds a wealth index of household prosperity based on an average of three other indices, namely, housing quality (wall, roof and floor materials of the house, including the number of rooms) , ownership of consumer durables (radio, bicycle, TV, motorbike, motorized vehicle, landline telephone, bed or table) and access to basic services (access to safe drinking water, improved sanitation, electricity and cooking fuel). Thus the wealth index is a long term indicator which is relatively more static than consumption expenditure (Kumra, 2008). Please note that ownership of land is not included as part of the wealth index. Also the wealth index gives equal weight to the three sub-indices in their impact on child health, which might not be the case.

While the Young Lives dataset calculates and reports the wealth index for all three rounds, per capita consumption expenditure is reported only for rounds 02 and 03. Therefore, for the analyses carried out in section V that involves per capita consumption expenditure, data from only rounds 02 and 03 are used. Data from all three rounds are however used when looking into panel regressions using wealth index as well as in section VI.

Height-for age (stunting) measures are reported for both cohorts for all three rounds. Weight-for-height (wasting) indicator is reported for children only up to 60 months old. Thus, although desirable, this severely restricts the sample size. Wasting data is also available only for the first round corresponding to the child age when per capita consumption expenditure information is not available. For these reasons, weight-for-height is not used as a child health indicator in this study. Finally, weight-for-age (underweight) measure is not reported for the older cohort in rounds 02 and 03. As discussed in the literature review, because height-for-age is considered a better longer term measure of child health, most of the analyses done in this essay focuses on stunting, while in some cases in section VI underweight is also investigated.

From the combined dataset, a total of 189 observations were dropped either because the height-for-age data were missing (N=80), the height-for-age data were flagged as extreme (N=46) or mother's education data were missing (N=77). With some overlap, the total amounted to 189 dropped observations. The before and after child count by cohort is provided in the Appendix under A-Table 01. All analyses conducted in this study is using this dataset containing a total of 8711 observations, out of which 5806 are from rounds 02 and 03 combined. It should be noted that a further 237 observations have missing values for consumption expenditure from round 02 and 88 observations from round 03. Stata's automated 'listwise deletion' protocol is allowed to handle this[3].

There has also been some commuting in terms of urbanity, region and community that is accounted for between rounds 02 and 03, although the numbers are relatively small. A total of 51 observations were found to have changed urbanity (rural to urban or urban to rural), 12 observations changed their region (movement between Coastal Andhra Pradesh, Rayalaseema and Telangana) and 62 observations changed their community (as per their community id in the dataset). The changes were taken into account in the dataset. Finally, 321 observations have missing values for their mother's height. To avoid dropping these observations which could lead to selectivity bias, a dummy variable with mean mother's height for the missing cases was added to the regressions.

4.2) Descriptive Statistics

We first check for the distribution of z-scores in all three rounds of the dataset. The density plots of height-for-age and weight-for-age by round and cohort are provided in the Appendix under A-Figure 01. From the shape of the plots we can see that the distribution is indeed normal and for both height-for-age and weight-for-age, a majority of the children's scores are well below the stunting and underweight lines, as indicated by the median value from the reference population.

Tables 01a and 01b report the descriptive statistics of the dataset. The mean height-for-age z-score for round 01 is -1.39, which increases in round 02 to -1.61 followed by a decrease in round 03 to -1.51. Respective stunting percentages are 31.4, 35.0 and 31.3 (mean of 32.6 over the three periods). The mean weight-for-age follows a similar trend although we do not see the recovery in round 03 that we see for the stunting measure. The mean weight-for-age z-score for round 01 is -1.67, which increases in round 02 to -1.86 and to -1.87 in round 03. Respective underweight percentages are 37.4, 44.3 and 46.0 (mean of 42.6 over the three periods).

[3] Under listwise deletion, for the specific regression, Stata will remove any observation which is missing on the outcome variable or any of the covariates.

Table 01a: Height-for-age descriptive statistics

Variables		Round 01, 2002 [N=2905]			Round 02, 2006 [N=2914]			Round 03, 2009 [N=2892]		
		Mean	p-value	Stunt	Mean	p-value	Stunt	Mean	p-value	Stunt
Sex of Child	Female	-1.34		29.2	-1.60		34.1	-1.50		29.7
	Male	-1.43	0.08	33.4	-1.62	0.55	35.8	-1.53	0.42	32.9
Maternal Education	None	-1.58		37.1	-1.79		41.5	-1.73		38.9
	Primary	-1.34		30.0	-1.50		30.9	-1.39		26.8
	Secondary	-1.18		22.5	-1.42		27.4	-1.24		21.4
	Higher	-0.89	<0.001	22.3	-1.24	<0.001	22.4	-1.08	<0.001	18.5
Wealth Index	Quartile 01	-1.66		40.5	-1.89		46.6	-1.84		42.6
	Quartile 02	-1.41		34.4	-1.77		40.0	-1.68		36.1
	Quartile 03	-1.39		30.4	-1.57		31.3	-1.51		28.9
	Quartile 04	-1.09	<0.001	20.5	-1.22	<0.001	22.2	-1.02	<0.001	18.0
Own Land	Yes	-1.44		33.7	-1.62		35.0	-1.65		36.1
	No	-1.33	<0.001	28.4	-1.49	0.01	33.3	-1.33	<0.001	24.8
Own Livestock	Yes	-1.48		34.7	-1.71		38.2	-1.68		36.8
	No	-1.32	<0.001	28.8	-1.54	<0.001	32.9	-1.40	<0.001	27.4
No. of Adults in Household	1-2	-1.41		30.9	-1.61		34.5	-1.53		30.6
	3-4	-1.44		32.8	-1.64		37.4	-1.52		31.8
	> 4	-1.25	0.02	30.3	-1.58	0.49	32.3	-1.49	0.26	31.4
Type Site	Urban	-1.14		21.3	-1.29		24.2	-1.11		19.2
	Rural	-1.46	<0.001	34.7	-1.72	<0.001	38.7	-1.66	<0.001	35.6
Child Religion	Hindu	-1.40		31.8	-1.63		35.7	-1.54		32.1
	Muslim	-1.21		24.8	-1.35		25.6	-1.21		20.8
	Other Religion	-1.68	<0.001	37.5	-1.59	<0.001	35.5	-1.41	<0.001	32.3
Child Ethnicity	Scheduled Castes	-1.47		34.1	-1.71		37.6	-1.70		36.3
	Scheduled Tribes	-1.76		44.5	-1.79		41.0	-1.88		43.5
	Backward Class	-1.39		32.0	-1.66		37.3	-1.50		31.2
	Other Castes	-1.10	<0.001	20.4	-1.31	<0.001	24.0	-1.16	<0.001	20.3
Region	Coastal	-1.42		30.9	-1.50		31.1	-1.36		27.2
	Rayalaseema	-1.08		27.0	-1.52		35.0	-1.57		32.2
	Telangana	-1.63	<0.001	35.6	-1.73	0.29	39.0	-1.62	0.73	34.9
Total		-1.39		31.4	-1.61		35.0	-1.51		31.3
(sd)		(1.35)		(46.4)	(1.01)		(47.7)	(0.03)		(46.4)

Table 01b: Weight-for-age descriptive statistics

Variables		Round 01, 2002 [N=2905]			Round 02, 2006 [N=1937]			Round 03, 2009 [N=1922]		
		Mean	p-value	Under-weight	Mean	p-value	Under-weight	Mean	p-value	Under-weight
Sex of Child	**Female**	-1.59		33.5	-1.83		42.0	-1.77		41.3
	Male	-1.74	<0.001	40.9	-1.89	0.16	46.4	-1.97	<0.001	50.2
Maternal Education	**None**	-1.83		42.7	-2.01		50.3	-2.12		55.3
	Primary	-1.67		37.5	-1.89		43.7	-1.92		44.2
	Secondary	-1.42		27.8	-1.68		37.4	-1.59		36.1
	Higher	-1.32	<0.001	29.6	-1.48	<0.001	31.7	-1.21	<0.001	24.6
Wealth Index	**Quartile 01**	-1.89		45.0	-2.11		55.7	-2.21		59.1
	Quartile 02	-1.72		38.1	-1.98		49.6	-2.05		50.0
	Quartile 03	-1.71		38.9	-1.84		42.4	-1.93		47.9
	Quartile 04	-1.36	<0.001	27.8	-1.52	<0.001	29.3	-1.31	<0.001	26.9
Own Land	**Yes**	-1.73		39.0	-1.86		44.3	-2.05		53.1
	No	-1.60	0.002	35.3	-2.47	0.26	66.7	-1.64	<0.001	36.6
Own Livestock	**Yes**	-1.74		38.6	-1.93		47.4	-2.06		53.5
	No	-1.62	0.003	36.4	-1.82	0.02	42.4	-1.75	<0.001	41.0
No. of Adults in Household	**1-2**	-1.71		37.8	-1.85		44.3	-1.87		42.2
	3-4	-1.71		38.8	-1.94		45.9	-1.93		48.7
	> 4	-1.51	0.35	34.1	-1.79	0.08	42.6	-1.82	0.15	44.0
Type Site	**Urban**	-1.37		28.5	-1.60		33.6	-1.39		28.8
	Rural	-1.77	<0.001	40.3	-1.95	<0.001	48.1	-2.04	<0.001	52.2
Child Religion	**Hindu**	-1.68		37.4	-1.88		44.7	-1.90		46.6
	Muslim	-1.53		35.9	-1.68		39.9	-1.60		39.6
	Other Religion	-1.62	0.09	46.7	-1.52	0.003	42.1	-1.64	0.002	44.4
Child Ethnicity	**Scheduled Castes**	-1.80		41.4	-1.94		45.7	-2.03		50.6
	Scheduled Tribes	-1.88		43.9	-2.03		50.6	-2.12		53.8
	Backward Class	-1.70		38.0	-1.92		46.3	-1.94		48.9
	Other Castes	-1.39	<0.001	28.9	-1.58	<0.001	35.0	-1.44	<0.001	30.8
Region	**Coastal**	-1.56		32.5	-1.77		39.8	-1.73		39.0
	Rayalaseema	-1.65		38.5	-1.85		41.3	-1.85		43.0
	Telangana	-1.80	<0.001	41.2	-1.98	0.67	51.6	-2.05	0.87	56.0
Total		-1.67		37.4	-1.86		44.3	-1.87		46.0
(sd)		(1.09)		(48.4)	(0.93)		(49.7)	(1.06)		(49.9)

De Onis, Monteiro et al. (1993) report that in Asia, for children under 5 years of age, 47 percent are stunted and 42 percent are underweight. Comparing the estimates, we see that our Young Lives sample has much a lower stunting percentage but is similar to the Asian average in terms of underweight. The issue that the Young Lives dataset is not nationally representative and contains data from children in older age groups is also at play here.

Note that for rounds 02 and 03, the weight-for-age measure is not provided for the older cohorts which could cause the jump between rounds 01 and 02. However when we check the statistics by cohort, the mean weight-for-age z-score for the younger cohort is -1.52 and the corresponding underweight percentage is 32.2. In other words, there is indeed a drop in the underweight measure between rounds 01 and 02.

A mean height-for-age and weight-for-age z-score line plots by cohort is provided in the Appendix under A-Figure 02 to investigate this. In the literature review the importance of growth in the first 24 months of a child was emphasized. This is reflected in the sharp drop in height-for-age and weight-for-age z-scores between the ages of 6-17 months. Note that the drop in height-for-age is much sharper than that for weight-for-age. Post this period, z-scores are relatively steady with slight fluctuations for both the cohorts. Simple t-tests indicate that on average older cohorts are more malnourished than the younger cohort both in terms of height-for-age ($P<0.001$) and weight-for-age ($P<0.001$). The recovery in period 03 height-for-age can also be seen in the figure which seemed to have occurred mainly for the younger cohort.

Tables 01a and 01b provide the patterns and relationships between different variables and child z-scores for each of the rounds. Except child gender, number of adults in household and region where the household is located, all other variables are strongly associated with child nutritional status ($P<0.01$). In general, patterns are similar for height-for-age and weight-for-age z-scores. Female children are better nourished than male children (mean height-for-age z-score over three rounds for male is -1.56 compared to -1.48 for female and that of weight-for-age z-score is -1.86 for male and -1.76 for female) and there is a stronger association between weight-for-age and child health (P-values of <0.001, 0.16, <0.001 for the three rounds respectively) than between height-for-age and child health (P-values of 0.08, 0.55, 0.42 for the three rounds respectively).

There is decrease in stunting and number of underweight children with increasing maternal education and wealth, while owning land and livestock indicate lower child health status. This is most likely because of the correlation of the two variables with living in rural areas. Urban children are on average less malnourished than rural children (mean height-for-age z-score over three rounds for urban is -1.18 compared to -1.61 for rural and that of weight-for-age z-score is -1.45 for urban and -1.92 for rural). A greater number of adults in the household, specifically greater than four, also indicates better nourishment, although it is not statistically significant in several of the cases. Finally, Muslim children and children born into other castes (which includes upper castes), have better nourishment and child health.

Focusing on maternal education, roughly 47 percent of the mothers have at least primary education in the Young Lives dataset, 22 percent at least secondary and around 9 percent higher than secondary. For clarification, it is important to note the education classification followed in India. Classes 1-5 is referred to as primary education, classes 6-10 as secondary, 11-12 as senior secondary and above the 12th grade is university level education.

As one would expect, per capita consumption expenditure is strongly positively associated with both height-for-age ($P<0.001$) and weight-for-age ($P<0.001$). That noted, we look into the relationship between wealth index and per capita consumption expenditure. This is given provided Table 02 where the mean values of wealth index and per capita consumption expenditure (at 2006 base prices) are presented by urbanity, cohort, ethnicity religion and region.

There is a growth of roughly 13.7 percent in per capita consumption expenditure (from Rs. 814 in 2006, Round 02 to Rs. 943 in 2009, Round 03) over the three years. Comparatively, wealth index also increased by roughly 10.8 percent (from 0.463 in 2006, Round 02 to 0.519 in 2009, Round 03) during the same time period. Urban consumption and wealth index is notably higher than its rural counterpart. However, while we see similar growth in per capita consumption expenditure for urban and rural households (14.5 percent for urban from Rs. 961 to Rs. 1,124 and 13.5 percent for rural from Rs. 762 to Rs. 881) between 2006 and 2009, there is a greater growth in rural wealth index (3.0 percent for urban from 0.66 to 0.68 and 13.0 percent for rural from 0.40 to 0.46).

We also see higher consumption and wealth for the older cohort possibly reflecting a later stage in their life-cycle, and thus more resource availability for growth. Rate of growth is also greater for the older cohort in terms of consumption (16 percent for the older cohort from Rs. 906 to Rs. 1,078 and 12 percent for the younger cohort from Rs. 766 to Rs. 872) but is roughly similar in terms of wealth index (9.5 percent for the older cohort from 0.47 to 0.52 and 10.0 percent for the younger cohort from 0.46 to 0.51).

Similar to findings from Table 01, per capita consumption expenditure and wealth index is the highest for other castes (which includes upper castes) and Muslim households. While it is difficult to say which region is better off based on wealth index, per capita consumption expenditure is highest in Telangana. However all three regions have similar growth rates of approximately 17 percent in terms of per capita consumption expenditure. If we look at the growth in wealth index however, over the past seven years between round 01 and round 03, Rayalaseema had the least growth of roughly 15 percent while both Coastal Andhra Pradesh and Telangana enjoyed growth of about 23 percent.

Table 02: Wealth index and real per capita consumption

	Wealth Index			**Real per capita consumption (2006 Base Prices)**	
	Round 01	Round02	Round03	Round02	Round03
By Typesite:					
Urban	0.644	0.661	0.681	960.82	1,123.91
	(0.131)	(0.146)	(0.119)	(605.46)	(663.62)
Rural	0.331	0.395	0.460	762.05	881.12
	(0.159)	(0.165)	(0.156)	(590.59)	(569.58)
By Cohort:					
Younger Cohort	0.408	0.459	0.513	765.67	871.86
	(0.202)	(0.197)	(0.178)	(465.76)	(465.27)
Older Cohort	0.408	0.469	0.524	905.94	1,078.38
	(0.205)	(0.198)	(0.173)	(789.62)	(788.63)
By Ethnicity:					
Scheduled	0.347	0.393	0.448	757.65	839.89
Castes	(0.158)	(0.153)	(0.149)	(545.29)	(596.78)
Scheduled Tribes	0.261	0.318	0.377	590.00	664.64
	(0.176)	(0.198)	(0.187)	(400.78)	(410.72)
Backward	0.406	0.471	0.529	794.11	962.01
Classes	(0.197)	(0.186)	(0.160)	(506.93)	(562.28)
Other Castes	0.549	0.589	0.630	1,021.74	1,141.85
	(0.182)	(0.175)	(0.147)	(829.05)	(708.90)
By Religion:					
Hindu	0.394	0.451	0.508	808.40	933.78
	(0.200)	(0.196)	(0.177)	(615.09)	(0.594)
Muslim	0.586	0.602	0.633	903.63	1,067.19
	(0.164)	(0.154)	(0.127)	(395.20)	(722.96)
By Region:					
Coastal	0.406	0.480	0.532	662.50	801.72
	(0.235)	(0.227)	(0.203)	(530.64)	(598.77)
Rayalaseema	0.427	0.461	0.504	766.30	918.11
	(0.164)	(0.163)	(0.153)	(519.88)	(737.70)
Telangana	0.394	0.446	0.512	876.13	1060.94
	(0.200)	(0.191)	(0.165)	(484.12)	(639.23)
Total	**0.408**	**0.463**	**0.519**	**813.88**	**942.57**
	(0.203)	**(0.196)**	**(0.177)**	**(600.76)**	**(603.93)**

Finally, because panel estimates will be reported in this study, panel descriptive statistics showing within and between variation for selected outcome and exposure variables and other time variant variables are reported in A-Table 02, in the appendix. Education variables are not reported as they do not vary over time, and hence contain zero within variation.

Some points can be noted from the table. Within variation is higher for height-for-age z-scores (0.600) than weight-for-age z-scores (0.463), which is interesting given that stunting is a measure of longer term child health. This indicates that perhaps for our dataset, there has been less variation in child wasting over this time period than compared to stunting. The other thing to note is that although within changes for the wealth index is low (0.090), it is relatively higher for the individual sub-indices. As noted earlier, the wealth index as a linear composite index might not be a fair representative of the variation in the sub-indices.

4.3) Variables

The variables used in this essay are divided into three main groups as per their role in the analysis. The first is the 'main exposure and outcome' group, which includes measures for health outcome, the child height-for-age and child weight-for-age, and the main variables of interest, which includes per capita consumption expenditure, maternal education, paternal education and community level education.

The second group is a set of baseline variables that have potential confounding abilities with the relationships of interest. This group includes the child age, child gender, ownership of land, ownership of livestock, number of adults in the household, location of the household in a rural or urban setting and finally the individual sub-indices of the wealth index: housing quality index, consumer durables index and access to services index. Please note that when per capita consumption expenditure is included as the main variable of interest, the sub-indices are not included in the regression, as they all measure socioeconomic status of a household and have high correlation.

The third group is a group of time invariant control variables which include mother's height, mother's height missing and a set of dummies for child ethnicity (scheduled tribes, scheduled castes, backward class and other castes), religion (Hindu, Muslim and other religions) and region (coastal Andhra Pradesh, Rayalaseema and Telangana) of the household. In case of community

Table 03: List of variables

Name	Description	Type
zhfa	child height-for-age z-score	continuous
zwfa	child weight-for-age z-score	continuous
logexpend	log of per capita consumption expenditure	continuous
logmumed	log of mother's education; in section VI, mother's education enters the regression as an ordered variable where, none=0 (reference category), primary=1, secondary=2 and higher=3.	continuous; categorical
logdaded	log of father's education; in section VI, father's education enters the regression as an ordered variable where, none=0 (reference category), primary=1, secondary=2 and higher=3.	continuous; categorical
comedu	community average of the percentage of adults in an household who have completed high school or above	continuous
agemon	age of child in months	continuous
agemon2	age of child in months squared	continuous
sex	= 1 if child is male, 2 if child is female	categorical
hhadults	number of adults in household	continuous
ownland	= 1 if household owns land, 0 if not	binary
animals	= 1 if household owns livestock, 0 if not	binary
typesite	= 1 if household is in an urban setting, 2 if in a rural setting	categorical
wealth	wealth index	continuous
housing	housing quality index	continuous
durables	consumer durables index	continuous
services	access to services index	continuous
services* typesite	interaction term between access to services index and urban/rural setting of the household	continuous
mtheight	mother's height in cm	continuous
mtheightmiss	= average mother's height if data on mother's height is missing	continuous
stribe, bclass, ocaste	child ethnicity dummy variables; stribe = scheduled tribes, bclass = backward classes, ocaste = other castes (reference category is scheduled castes)	binary
muslim, otherrel	child religion dummy variables; muslin = child is Muslim, otherrel = child follows other religion such as Christianity or Buddhism (reference category is Hindu)	binary
rayal, coastal	region dummy variables; rayal = household is located in Rayalaseema, coastal = household is located in Coastal Andhra Pradesh (reference category is Telangana)	binary
comage	community average of child's age in months	continuous
comadults	community average of the number of adults in households	continuous
comhq	community average of housing quality index	continuous
comcd	community average of consumer durables index	continuous
comsv	community average of access to services index	continuous
comland	community average of the no. of households that own land	continuous
comanimals	community average of the no. of households that own livestock	continuous

level regression, a separate list of dummy variables representing community level fixed effects is also added.

Table 03 provides a more detailed description of the data. Please note that in order to save space, when regression results are reported in sections V and VI, the variable name from table 03 will be reported and not its full descriptive name. That is, tables will label 'log of per capita consumption expenditure' simply as 'logexpend' and so forth.

5) Analysis: Role of income on child health

In this section the reduced form equation from section III is estimated and the impact of income on child nutritional status is attempted to be approximated as accurately as possible. As mentioned earlier, per capita consumption expenditure is used as the main variable of interest with height-for-age as the measure for chronic health outcome. Separate regressions are run for urban and rural households. Data from round 02 and 03 are used for the analysis and the cross-sectional estimates are first presented followed by panel estimates.

5.1) Cross-sectional estimation

Tables 04 and 05 provide the different model estimates for round 02 and round 03 respectively, separated by rural and urban areas. The first column for each area, provides the OLS estimates. Although the estimates are likely to be biased from endogeneity and unobserved heterogeneity, it provides a good starting point for the analysis. The second column controls the same regression for community fixed effects, handling unobserved heterogeneity at the community level. The estimation however still suffers from unobserved heterogeneity at the child level and also from endogeneity problems.

The third column provides two stage least squares fixed effects (2SLSFE) estimation that in addition to the previous, controls for endogeneity and measurement bias using instrument variables. As mentioned in section III, five instruments are used. This includes (i) amount of irrigated land owned, (ii) amount of non-irrigated land owned, (iii) number of relatives living abroad, (iv) whether the household received any non-labour income from social security or social subsidy in the last 12 months, and (v) whether the household received any rent in the last 12 months.

Before looking into the estimates the dataset was checked for heteroskedasticity in both the rounds. For round 03, both the Breusch-Pagan test (P = 0.0094) for linear associations and White

Table 04: Round 02 (2006), Height-for-age by typesite

	Urban			Rural			
	OLS	Comm. Fixed Effects	2SLSFE	OLS	Comm. Fixed Effects	2SLSFE	2SLSFE[1]
agemon	-0.008	-0.0001	-0.005	-0.013	-0.010	-0.008	-0.121
	(0.023)	(0.026)	(0.026)	(0.014)	(0.015)	(0.015)	(0.053)
agemon2	0.001	0.001	0.001	0.001	0.002	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
hhadults	0.012	0.010	0.016	0.018	0.028*	0.026***	0.028**
	(0.020)	(0.021)	(0.023)	(0.002)	(0.010)	(0.010)	(0.010)
mtheight	0.035***	0.033***	0.032***	0.023***	0.022***	0.022***	0.021***
	(0.011)	(0.011)	(0.012)	(0.004)	(0.004)	(0.004)	(0.004)
mtheightmiss	0.059***	0.057***	0.055***	0.038***	0.036***	0.036***	0.036***
	(0.020)	(0.020)	(0.020)	(0.007)	(0.007)	(0.007)	(0.007)
sex	0.035	0.070	0.078	-0.012	-0.028	-0.030	-0.032
	(0.075)	(0.077)	(0.078)	(0.042)	(0.042)	(0.043)	(0.043)
logmumed	0.200***	0.172***	0.117*	0.025	0.040*	0.049*	0.032
	(0.051)	(0.056)	(0.065)	(0.023)	(0.025)	(0.027)	(0.027)
logdaded	-0.030	-0.028	-0.067	0.042*	0.036*	0.058*	0.025
	(0.051)	(0.055)	(0.059)	(0.021)	(0.022)	(0.030)	(0.032)
logexpend	0.340***	0.404***	1.027***	0.224***	0.179***	-0.060	0.327
	(0.090)	(0.089)	(0.356)	(0.045)	(0.048)	(0.243)	(0.282)
stribe	0.192	0.136	0.129	-0.169**	-0.175*	-0.166*	-0.184*
	(0.230)	(0.259)	(0.281)	(0.076)	(0.097)	(0.098)	(0.099)
bclass	0.243*	0.201	0.163	-0.195***	-0.115*	-0.086	-0.138*
	(0.126)	(0.155)	(0.165)	(0.056)	(0.063)	(0.070)	(0.071)
ocaste	0.173	0.148	0.018	0.107	0.146*	0.188*	0.110*
	(0.148)	(0.174)	(0.195)	(0.077)	(0.083)	(0.096)	(0.081)
muslim	-0.004	0.058	0.208	0.044	0.067	0.050	0.081
	(0.122)	(0.140)	(0.162)	(0.143)	(0.146)	(0.146)	(0.148)
otherrel	0.396**	0.328*	0.510**	-0.414**	-0.514**	-0.544**	-0.498**
	(0.171)	(0.197)	(0.212)	(0.191)	(0.216)	(0.220)	(0.219)
rayal	0.100	-0.065	-0.222	0.033	-1.262***	-1.509***	-1.123***
	(0.104)	(0.305)	(0.345)	(0.056)	(0.153)	(0.279)	(0.307)
coastal	0.112	-0.330	-0.041	0.336***	-0.595**	-0.770*	0.480
	(0.099)	(0.588)	(0.607)	(0.059)	(0.285)	(0.418)	(0.314)
constant	-6.409	-7.060		-4.075*	-2.149		
	(3.318)	(3.621)		(1.760)	(1.810)		
N	698	698	679	1979	1979	1967	1967
R-sq	0.159	0.209	0.070	0.100	0.164	0.069	0.078
Sargan-Hansen Overidentification Test (p-value)			0.648			0.084	0.145
F-test on excluded instruments			24.36			13.43	16.85
Endogeneity Test (p-value)			0.044			0.467	0.591

[1] Only amount of (i) irrigated land owned, and (ii) non-irrigated land owned are used as instruments.

Robust standard errors in parentheses.

*p<0.10, **p<0.05, ***p<0.01

Table 05: Round 03 (2009), Height-for-age by typesite

	Urban			Rural		
	OLS	Comm. Fixed Effects	2SLSFE	OLS	Comm. Fixed Effects	2SLSFE
agemon	-0.055*	-0.048	-0.047	-0.009	-0.011	-0.014
	(0.029)	(0.031)	(0.030)	(0.017)	(0.018)	(0.018)
agemon2	0.0002*	0.0002	0.0002	0.00002	0.00003	0.00004
	(0.0001)	(0.0001)	(0.0001)	(0.00006)	(0.00006)	(0.00006)
hhadults	0.026	0.028	0.033	0.005	0.013	0.014
	(0.018)	(0.019)	(0.019)	(0.008)	(0.009)	(0.009)
mtheight	0.030***	0.031***	0.031***	0.021***	0.022***	0.021***
	(0.009)	(0.020)	(0.010)	(0.004)	(0.004)	(0.004)
mtheightmiss	0.051***	0.055***	0.055***	0.035***	0.036***	0.035***
	(0.017)	(0.018)	(0.017)	(0.007)	(0.007)	(0.007)
sex	-0.088	-0.101	-0.077	0.054	0.052	0.041
	(0.074)	(0.077)	(0.082)	(0.040)	(0.041)	(0.042)
logmumed	0.214***	0.183***	0.153**	0.031	0.032	0.020
	(0.051)	(0.054)	(0.060)	(0.022)	(0.024)	(0.025)
logdaded	-0.062	-0.071	-0.101*	0.040*	0.038*	0.022
	(0.052)	(0.056)	(0.058)	(0.021)	(0.022)	(0.024)
logexpend	0.309***	0.288***	0.625**	0.175***	0.168***	0.436**
	(0.081)	(0.088)	(0.292)	(0.042)	(0.045)	(0.187)
stribe	-0.068	0.092	0.042	-0.344***	-0.229**	-0.241**
	(0.203)	(0.258)	(0.270)	(0.073)	(0.095)	(0.094)
bclass	0.407***	0.466***	0.448**	-0.080	-0.026	-0.068
	(0.136)	(0.173)	(0.180)	(0.057)	(0.065)	(0.069)
ocaste	0.332**	0.384**	0.325*	0.249***	0.312***	0.242***
	(0.148)	(0.180)	(0.198)	(0.073)	(0.077)	(0.087)
muslim	0.023	-0.043	0.023	-0.052	-0.068	-0.035
	(0.118)	(0.133)	(0.139)	(0.118)	(0.125)	(0.128)
otherrel	0.295	0.243	0.263	-0.334	-0.277	-0.274
	(0.186)	(0.204)	(0.218)	(0.171)	(0.190)	(0.194)
rayal	-0.182*	0.532*	0.503*	0.063	0.118	0.091
	(0.104)	(0.284)	(0.294)	(0.053)	(0.245)	(0.227)
coastal	0.050	0.723***	0.807***	0.406***	0.326	0.339*
	(0.097)	(0.278)	(0.298)	(0.0534)	(0.203)	(0.194)
constant	0.816	-0.780		-3.661	-5.312	
	(6.400)	(6.761)		(3.484)	(3.784)	
N	709	709	695	2095	2095	2092
R-sq	0.180	0.243	0.134	0.117	0.170	0.076
Sargan-Hansen Overidentification Test (p-value)			0.930			0.137
F-test on excluded instruments			17.70			23.78
Endogeneity Test (p-value)			0.213			0.111

Robust standard errors in parentheses

*p<0.10, **p<0.05, ***p<0.01

test (P = 0.0001) for non-linear associations reject the null of homoskedasticity. For round 02, while the Breusch-Pagan test do not reject the null (P = 0.3001), the White test does (P = 0.0039). Finding indications towards the presence of heteroskedasticity in round 01 as well, all regressions estimated in this essay report robust standard errors.

In general, estimates from round 02 and round 03 follow similar trends with some differences. We first discuss the estimates from round 02. Firstly, mother's height is significantly correlated with child health status which partially controls for the child's genetic endowment. In order to allow non-linear flexibility in the relationship between child age and child health, a quadratic term was added to the regressions. However the relationship is not significant for either of the models due their relatively high standard errors. Since we are dealing with data from only rounds 02 and 03, meaning where child age is in the post 24 months period, these results are not as surprising as it would seem. In fact, given that we see an improvement in the height-for-age z-score in the descriptive statistics over the two rounds, one can expect child health to improve with child age. We will see this exact relationship in the panel regressions when we control for unobserved heterogeneity at the child and household levels.

Girls in urban areas are better nourished than boys while we see the opposite scenario in rural areas. However, none of these estimates is statistically significant. Mother's education level, however, is statistically significant in all urban regressions and two of the rural level regressions. The estimated coefficients are also larger for urban areas, meaning maternal education not only plays a differential impact, but possibly interacts with other variables found in urban areas (such as better medical facilities or better communication) that boosts its effect on child health. This issue has been discussed in the literature review and the differential roles of maternal education will be investigated further in section VI. Paternal education, interestingly, is significant for rural areas only.

We also note differences in the control variables. In rural areas, compared to the reference category of schedules castes, samples from scheduled tribes and backward classes has a significant negative association with child health while samples from other castes (which includes upper castes) has a significant positive association. Although samples from scheduled tribes and backward classes show opposite association in urban areas, this is not statistically significant when controlling for endogeneity and community level heterogeneity. Compared to Hindu children, children from other religions have significantly improved health in urban areas, but degraded health in rural areas.

We now discuss the impact of per capita consumption expenditure on child health. Ordinary least squares provide statistically significant estimates of 0.340 (with standard error 0.090 and t-stat 3.81) for urban areas and 0.224 (with standard error 0.045 and t-stat 5.02) for rural areas, holding other predictor variables constant.. This is however, without controlling for endogeneity or community heterogeneity. An example of community differences can be that wealthier communities have better health facilities. Community fixed effects will remove any bias that arise because of such differences (Glewwe, Koch, Nguyen, 2002).

It is worthwhile to note here that because the per capita consumption expenditure variable is log transformed, to better estimate a non-linear relationship, a 100 percent change in per capita consumption expenditure will bring about the estimated change. Hence for the OLS estimates, a 100 percent change in per capita consumption expenditure will, on average, improve the height-for-age z-scores by 0.340 in urban areas and 0.224 in rural areas. Similar interpretation extends to other the log variables, log of mother's education and log of father's education.

As expected. adding community fixed effects to the regression lowers the estimate for rural areas to 0.179 (with standard error 0.048 and t-stat 3.76) but increases that for urban areas to 0.404 (with standard error 0.089 and t-stat 4.53). Note that both the estimates are still statistically significant. Finally we add the instrument variables to account for endogeneity problems.

Based on the results of the F-test on excluded instruments which gives values that are higher than 10 in each case, it can be said that the instruments used have sufficient explanatory power. However, it should be noted that they do not explain a large percent of the variation in per capita consumption expenditure. A regression of the excluded instruments on per capita consumption expenditure yields R-square statistics of 0.094 and 0.045 for urban and rural areas respectively in round 02 and 0.148 and 0.099 for urban and rural areas respectively in round 03.

The other important statistic is Sargan-Hansen test statistic of overidentifying restrictions which determines the validity of the instruments used. The joint null hypothesis is that the instruments are correlated with the error term. Thus if the null hypothesis is rejected, it means that the instruments are correctly excluded from the estimated equation. In round 02, while the null hypothesis is comfortably rejected for the urban areas, we cannot reject the null hypothesis for rural areas at the 10 percent level.

Suspecting that some of the instruments have low explanatory power compared to the others, we re-run the rural regression with only the instruments with highest explanatory power on per capita consumption expenditure (that is only using amount of (i) irrigated land owned, and (ii) non-irrigated land owned as instruments). This result is provided in the fourth column of the round 02 rural estimates. With fewer instruments having higher explanatory power, this time the null hypothesis is rejected satisfying the validity of the instruments. Based on the F test and overidentification test, it can be seen that while the selected instruments are proper for urban areas, they are relatively weaker instruments for rural areas.

The 2SLSFE estimate for urban areas increases the coefficient to 1.027 (with a higher standard error of 0.356, as would be expected from instrument estimation, and t-stat of 2.90), which is still statistically significant. In case of rural areas, the re-run 2SLSFE regression with instruments passes the validity test and is reported in column four. This provides an estimate of 0.327 (with standard error 0.282 and t-stat 1.18) which is not statistically significant.

This lack of significance can be due to the weak instruments, or because after controlling for endogeneity, per capita consumption expenditure no longer plays a significant role in improving child health in rural areas. If we consider the Wu-Hausman endogeneity test on per capita consumption expenditure, the null hypothesis that per capita consumption expenditure can actually be treated as exogenous is not rejected for rural areas (P-value of 0.591) but is rejected for the urban areas (P-value of 0.044). In that case controlling for endogeneity might not be required. However, this test result might not be valid given the weak instruments for rural areas. Hence, based on economic reasoning, the presence of endogeneity is best not neglected.

Round 03 estimates follow the general trend of round 02 estimates. Some of the differences are noted here. Maternal education is only significant for urban areas while paternal education is significant for two out of the three models in rural areas. Paternal education is also found to be significantly negatively associated with child health in urban areas when in the 2SLSFE estimation model. In urban areas, compared to schedules castes samples , samples from backward classes or other castes have significantly increased chances of better child health. The same holds for other castes in rural areas.

Finally the most noteworthy difference is that the 2SLSFE estimate for rural areas is statistically significant; indicating that after controlling for endogeneity and measurement bias, per capita consumption expenditure does have an impact on child health. The estimated coefficient is 0.436 (with standard error 0.187 and t-stat 2.34). In fact, all estimates of per capita consumption expenditure in all three models of both urban and rural areas are statistically significant.

Similar to maternal education, per capita consumption expenditure has a larger impact on child health in urban areas. Furthermore, the estimated coefficients for round 02 are larger than those for round 03. In urban areas the estimates are 0.340 vs. 0.309 for OLS, 0.404 vs. 0.288 for Community Fixed Effects and 1.027 vs. 0.625 for 2SLSFE. Similarly, in rural areas the estimates are 0.224 vs. 0.175 for OLS, 0.179 vs. 0.168 for Community Fixed Effects and 0.327 vs. 0.436 for 2SLSFE. Round 03 estimates however, have lower standard errors and are likely to be more accurate than round 02 estimates.

5.2) Panel estimation

The robust Hausman test (Wooldridge, 2002: 288) for fixed effects estimation versus random effects estimation clearly reject random effects for both rounds (P<0.001). Given that economic theory also support fixed effects procedure, the results for first difference (FD) estimation and within fixed effects (FE) estimation are reported in Table 06. It is worth noting that the Breusch-Pagan LM Test also rejected the use of pooled regression (P<0.0001) in both cases.

Table 06: Panel, Height-for-age by typesite

	Urban		Rural	
	First Difference	Within Fixed Effects	First Difference	Within Fixed Effects
agemon	0.023***	0.024***	0.014***	0.014***
	(0.003)	(0.003)	(0.004)	(0.001)
agemon2	-0.00009***	-0.00009***	-0.00005***	-0.00005***
	(0.00001)	(0.00001)	(0.000007)	(0.000007)
hhadults	0.109**	0.125**	-0.014	-0.015
	(0.051)	(0.053)	(0.024)	(0.024)
ltconsrpc	0.079	0.094	0.006	0.007
	(0.064)	(0.065)	(0.028)	(0.028)
N	672	1407	1895	4074
rho		0.756		0.798
R-square	0.11	0.12	0.051	0.052

Robust standard errors in parentheses
*p<0.10, **p<0.05, ***p<0.01

Before discussing the estimates, it is worthwhile to recall from the literature review that, panel estimation controls for individual child-level and household-level heterogeneity such as parental preferences η_i and child's genetic health endowment, α_i. This assumes that the unobservables are additive in nature and are time invariant. Inclusion of community fixed effects in the estimation controls for unobserved heterogeneity at the community as well. As noted earlier, problems of endogeneity and measurement bias however remain.

The only variables that significantly change over time are child age in months, the number of adults living in the household and per capita consumption expenditure. The FD and FE estimates are very similar for both urban and rural areas. After controlling for the individual heterogeneity, child age is now significantly associated with height-for-age health status as would be expected. The number of adults in the household represents the earning capacity of a household. From the panel estimates, we can see that the variable is significantly and positively associated with child health for urban areas. No significant association is however found for rural areas.

The rho value, also known as the interclass correlation, of the fixed effects regression provides us with the percentage of variation that is explained by individual specific effects in the model. The high rho values suggest that majority of the 'within' variation in the panel data is explained by individual specific effects and not idiosyncratic effects. Thus controlling for the individual specific effects leads to more accurate estimation.

None of the estimates for per capita consumption expenditure is statistically significant for the panel estimates. While the fixed effects estimates of 0.094 (with standard error 0.065 and t-stat 1.43) for urban areas come close to statistical significance, the rural area estimates are far from being statistically significant. In fact, the standard errors for the rural area estimates are at least 4 times larger than the coefficients themselves (0.006 with standard error 0.028 for FD and 0.007 with standard error 0.028 for FE) making it difficult to make any inference. While the standard errors for urban area estimates are still large, they are at least smaller than the coefficients.

This lack of significance can be a result of endogeneity and aggravated measurement bias that is not controlled for, or because per capita consumption expenditure no longer plays a significant role in improving child health after controlling for unobserved heterogeneity. Assuming the case of the former and given the cross-sectional estimates, it may be suggested that per capita consumption expenditure has a greater impact on child health in urban areas than rural areas. Finally even though all the estimates are positive, unfortunately we cannot conclude with certainty that per capita consumption expenditure positively impacts child health based on the panel estimates.

5.3) Impact estimation

In this section we translate the estimates coefficients in Tables 04-06 to see how much of the change in height-for-age and the change in stunting is explained by per capita consumption expenditure. The analysis in this section follows that of Glewwe, Koch and Nguyen (2002) where the authors do a similar calculation using 1993 and 1998 Vietnam Living Standard Surveys.

Table 07: Impact of per capita consumption on child health

	Mean height-for-age z-score		Stunting	
	Urban	Rural	Urban	Rural
2002	-1.141	-1.291	21.3	34.7
2006	-1.291	-1.723	24.2	38.7
2009	-1.107	-1.656	19.2	35.5
Change [R2 - R1]	-0.150	-0.432	2.9	4.0
Change [R3 - R2]	0.184	0.067	-5.0	-3.2
	Per Capita Consumption Expenditure		Change in Stunting	
OLS	0.325 [0.022][1]	0.120 [0.008]	-0.9	0
Community Fixed Effects	0.346 [0.024]	0.174 [0.011]	-0.9	-0.5
2SLSFE	0.826 [0.056]	0.382 [0.024]	-1.7	-0.8
2SLSFE (Upper Bound of 95 percent CI)	1.15 [0.078]	0.616 [0.039]	-2.1	-1.1
First Difference (Panel)	0.079 [0.005]	0.006 [0.0004]	-0.05	0
Within Fixed Effects (Panel)	0.094 [0.006]	0.007 [0.0004]	-0.05	0

[1]Change in mean height-for-age z-score is provided within brackets.

Table 07 reports the impact calculation. The first five rows show the changes in mean height-for-age z-score and stunting between the three rounds. The change between round 02 and round 03 is of interest in this calculation. During this period, mean height-for-age z-scores dropped by 0.184 and 0.067 for urban and rural areas respectively while stunting reduced by 5 percent and 3.2 percent in urban and rural areas respectively.

During this time, the log change in urban consumption expenditure is calculated to be log(1123.91) - log(960.82) = 0.0681 and that in rural consumption expenditure is calculated to be log(881.20) - log(762.05) = 0.0631. The change in mean height-for-age z-score is thus calculated by simply multiplying the specific estimated coefficient with the change in consumption expenditure. This change is then added to individual height-for-age z-scores of 2006 and the predicted change in stunting is calculated.

The calculated coefficients for per capita consumption expenditure from different models are reproduced in Table 07 with the calculated change in mean height-for-age z-score provided in brackets. For cross-sectional estimates, the average of the two coefficients is provided. Corresponding changes in stunting percentage is also presented in the same row along the coefficients.

One of the first things to consider is which of the estimates can be deemed reliable. Each of the models has its strengths and weaknesses and given the lack of a significant panel 2SLSFE estimate, the answer will depend on which of the assumptions in model weaknesses we are willing to accept. Do we assume that individual heterogeneity does not bias the estimates excessively, such that the panel non-significance is mainly due to measurement errors, and accept the 2SLSFE estimates? Or do we assume that endogeneity and measurement error do not bias the panel estimates and accept the FD or the FE coefficients (even though they are non-significant)?

In presence of autocorrelation, FD estimates will be more accurate than FE estimates (Wooldridge, 2002). The user written *xtserial* command in Stata, which reports the Wooldridge test for autocorrelation in panel data, is used to investigate this (Drukker, 2003). The null hypothesis of no first order serial correlation is easily rejected (P<0.001) indicating that for our dataset the FD estimates are more reliable than the FE estimates.

From Table 07 we can see that the highest values in predicted changes are provided by the 2SLSFE estimates. In urban areas, the mean height-for-age z-score is increased by 0.056 (compared to 0.184 in the dataset) while stunting is reduced by 1.7 percent (compared to 5.0 percent in the dataset). Similarly, in rural areas, the mean height-for-age z-score is increased by 0.024 (compared to 0.067 in the dataset) while stunting is reduced by 0.8 percent (compared to 3.2 percent in the dataset).

Based on the estimates and calculation presented in Table 07. the maximum change in height-for-age and change in stunting that is explained by income is as follows: (i) according to height-for-age z-score it is 30 percent and 36 percent for urban and rural areas respectively, and (ii) according to stunting it is 34 percent and 25 percent for urban and rural areas respectively.

If we stretch and take the upper bound of the 95 percent confidence interval of the 2SLSFE estimates, then the percentages increase to (i) according to height-for-age z-score: 42 percent and 58 percent for urban and rural areas respectively, and (ii) according to stunting: 42 percent and 34 percent measure for urban and rural areas respectively. Because, the significance of the relationship is still under scrutiny, as was seen from the panel results, it is safer to stay with the average and not consider the upper bound of the confidence interval.

A suitable conclusion to this analysis would thus be that, during this period, increases in household income only contributed to a portion to the reduction in child stunting. While we cannot state with certainty the estimated impact, under set assumptions we can say that the income effect explains between zero to 34 percent of the change in stunting only.

5.4) Prologue to section VI

Education is likely to be another variable that contributed to the change in stunting. We have already seen positive and differential effects of education on child health from the cross-sectional estimations made in this section. Before moving onto section VI where the role of education is further investigated, we carry out a few panel estimations, similar to those in Table 06 but using the wealth index and the sub-indices. Because this information is available in all three rounds, the full dataset is used in the estimation. The results are provided in A-Table 03 in the Appendix.

The panel estimations indicate that wealth index significantly and positively impacts child health in rural areas but not urban areas, where the impact is only positive but not significant. Breaking this down further, into the sub-indices that form the wealth index, we find that the consumer durables index is both significant and highly positive for both urban and rural areas. Access to services index on the other hand is positive and significant for rural areas only. The housing quality index seems to have the least effect on child health. Due to this differential impact of the sub-indices which is not properly captured by wealth index as a composite measure, in the regressions carried out in section VI, the different sub-indices are included as covariates.

6) Analysis: Role of education on child health

The analysis of the determinants of child health is continued. Ee look into the role of maternal, paternal and community level education in this section. In order to identify the level of education that has the most impact on child health, we break down parental education into categorical variables as mentioned in Table 03.

In the young lives dataset, there is a high correlation between maternal and paternal education. Among mothers with a secondary degree, over 80 percent of their husbands have at least a secondary degree. Similarly, among mothers with a degree higher than secondary, over 57 percent of their husbands have at least a higher than secondary degree. The Pearson correlation coefficient between maternal and paternal education is found to be 0.545 ($P<0.0001$) exhibiting this relationship.

Variations in maternal education with respect to selected indicators are presented in Table 08. Given that education levels are assumed to not change over the three rounds, the calculation is reported based on a single round.

Table 08: Maternal education patterns

	Mother's education (percentage)[1]				
	None	Primary	Secondary	Higher	P-value
Urban	12.8	23.8	48.5	45.0	<0.001
Female child	47.1	49.6	48.5	50.0	0.385
Top wealth Quartile	8.7	23.1	54.6	51.7	<0.001
Bottom wealth quartile	35.7	18.5	8.6	12.2	<0.001
Paternal Education > Primary	44.0	76.4	91.6	92.7	<0.001
Community-level higher education > 50%	35.6	57.7	78.5	78.7	<0.001
N	1574	411	643	286	

[1] Percentages are reported based on Round 02 data. They however almost identical for all three rounds with the minor differences coming from missing observations in different rounds.

Generally expected trends related to maternal education can be seen from Table 08. All of the reported variables are strongly associated with maternal education ($P<0.001$) except for child gender ($P = 0.385$), as would be anticipated. Levels of maternal education are higher in urban areas and in households that fall in the top wealth quartile according to the Young Lives reported wealth index. Conversely, we see very low percentages of high maternal education in households that fall in the bottom wealth quartile.

Husband's education level also increases with higher levels of maternal education. Finally, the calculated community level higher education is divided into two percentiles, and the final row of Table 08 reports the associated maternal education level percentages of households living in the higher percentile communities. Here we see higher percentages of educated mothers living in more educated communities.

We begin looking into maternal education with three different specifications for both height-for-age and weight-for-age. Table 09 reports pooled regression estimates of the different specifications as a starting point. Note that all three specifications are run with community fixed effects to control for unobserved heterogeneity at the community level. The first specification is a crude association where the only covariate is maternal education. Although the relationship has been established previously, this is reported to note the fall in maternal education coefficients in the subsequent specifications.

Please note that an interaction term has been added to this estimation, based on the results from the panel regression conducted on wealth index and its sub-indices that is reported in A-Table 02 in the Appendix. The estimates of access to services index were positive and significant for rural areas but negative and non-significant for urban areas. This suggests differential growth or impact of the services index based on rural-urban differences. The interaction term is added to account for this.

The results indicate that the height-for-age z-scores for children whose mothers have completed at least primary education would, on average, be 0.218 greater than children whose mothers have not. Similarly, the weight-for-age z-scores for children whose mothers have completed at least primary education would, on average, be 0.070 greater than children whose mothers have not. The reference category in for all parental education estimates that follow is the 'no education or less than primary' category. Note that the estimated coefficient for primary and secondary maternal education, is larger for both height-for-age than weight-for-age. The coefficients are however similar in case of higher maternal education.

The second specification includes the baseline variables that adjust for confounding relationships in the maternal education estimates. Please note however, that the child age and child sex covariates are independent risk factors to child health. Finally, the third specification includes the time invariant control variables that partially controls for unobserved parental preferences as well as unobserved differences by region. The maternal education coefficients for all three levels fall in the second and third specification as we include the controls.

Table 09: Maternal education and child health, Pooled estimation

		height-for-age z-score			weight-for-age z-score		
Maternal Literacy	primary	0.218***	0.151***	0.111**	0.070*	0.006	-0.023
		(0.037)	(0.037)	(0.0365)	(0.037)	(0.036)	(0.036)
	secondary	0.262***	0.102***	0.067*	0.247***	0.096**	0.062*
		(0.033)	(0.035)	(0.035)	(0.035)	(0.036)	(0.051)
	higher	0.445***	0.258***	0.202***	0.445***	0.250***	0.208***
		(0.049)	(0.050)	(0.050)	(0.051)	(0.051)	(0.051)
Baseline Variables	agemon		-0.006***	-0.006***		-0.013***	-0.013***
			(0.001)	(0.001)		(0.002)	(0.0018)
	agemon2		0.000***	0.000***		0.000***	0.000***
			(0.000)	(0.000)		(0.000)	(0.000)
	sex		0.050**	0.048**		0.139***	0.142***
			(0.024)	(0.023)		(0.024)	(0.024)
	hhadults		0.006*	0.005		0.002	0.001
			(0.004)	(0.003)		(0.002)	(0.002)
	housing		0.071	0.038		0.023	0.001
			(0.050)	(0.050)		(0.050)	(0.050)
	durables		1.764***	1.647***		1.803***	1.643***
			(0.272)	(0.267)		(0.313)	(0.311)
	services		0.199**	0.168**		0.171**	0.145*
			(0.078)	(0.077)		(0.080)	(0.080)
	services* typesite		-0.559***	-0.541***		-0.582***	-0.547***
			(0.153)	(0.150)		(0.171)	(0.169)
	ownland		-0.045	-0.047		0.005	0.001
			(0.030)	(0.029)		(0.035)	(0.034)
	animals		-0.002	-0.009		0.042	0.034
			(0.029)	(0.028)		(0.028)	(0.028)
	typesite		0.110	0.141*		0.056	0.058
			(0.083)	(0.081)		(0.088)	(0.087)
Time invariant control variables	mtheight			0.024***			0.019***
				(0.002)			(0.002)
	mtheightmiss			0.040***			0.031***
				(0.004)			(0.004)
	stribe			-0.193***			-0.071
				(0.060)			(0.061)
	bclass			-0.020			-0.028
				(0.038)			(0.038)
	ocaste			0.124***			0.189***
				(0.046)			(0.049)
	muslim			-0.037			-0.153**
				(0.062)			(0.065)
	otherrel			-0.208**			0.041
				(0.093)			(0.150)
	rayal			0.199			0.290
				(0.124)			(0.152)
	coastal			0.182			0.098
				(0.114)			(0.112)
	constant	-1.637***	-1.898***	-5.543***	-1.895***	-2.002***	-4.866***
		(0.0173)	(0.166)	(0.368)	(0.018)	(0.173)	(0.403)
	N	8711	8710	8710	6763	6762	6762
	R-sq	0.090	0.116	0.146	0.111	0.162	0.183

Robust standard errors in parenthesis

*p<0.10, **p<0.05, ***p<0.01

Baseline variables that remain significant include child age, child gender, access to services index and the interaction term between services index and typesite. The coefficients on child sex indicate that overall females are better off than males, and the relationship is more profound in case of the weight-for-age indicator than for the height-for-age indicator. The significance of interaction term captures the differential impact of the services index based on rural-urban difference. This is explored further when the differential impacts of maternal education is explored.

While all education levels remain significant for height-for-age z-scores, only secondary and higher education remain significant for weight-for-age z-scores. Furthermore, in the third 'complete' specification we can see that the coefficients fall by more than half the crude association (higher maternal education falls from 0.445 to 0.202 for height-for-age z-scores and from 0.445 to 0.208 for weight-for age z-scores). The statistical significance however, reasserts the importance of the relationship.

To be sure that maternal education retains its significance in cross-sectional estimates and not only in pooled estimates, we run the third specification from Table 09 for all of the rounds for both height-for-age z-scores and weight-for-age z-scores. The results are reported in A-Table 04 in the appendix. Overall, much of the results remain similar. Primary level maternal education remains significant for two of the rounds in height-for-age z-scores but for none in weight-for-age z-scores. The crucial difference is that none of the estimates for secondary level maternal education is significant. Finally, the coefficients for higher maternal education remain similar and significant but are the largest for round 03 (compared to that of round 01 and 02) in both the health measures. This hints towards an increasing impact of higher maternal education amongst the Young Lives households.

It should be noted that individual-level heterogeneity is not controlled for, which could lead to biased estimates. Also, given the differential impact of education we have seen in section V as well as in the literature review, it is natural to question the estimates presented here and how they will differ based on urban or rural locations, wealth distribution and child cohorts.

6.1) Differential impact of maternal education on child health

Following Moestu (2005), the LR test is used to identify significant differences in the role of maternal education based on typesite, wealth, cohort and sex differences. The LR test essentially reports the difference in likelihood (that follows a chi-square distribution) between two models, one with and one without a dummy interaction term. The interaction term is between maternal education and the variable in question (which is either typesite, wealth, child cohort or child sex). The continuous maternal education variable is used for the LR analyses. It is worth noting that, for example, when differences in maternal education by cohort is investigated, cohort is added as a variable in both the models if it is already not included as a covariate. The same is done regarding sex, wealth and typesite.

The results from the LR test for all three rounds is reported in A-Table 05 in the appendix. Since, typesite and the wealth index have high correlation (Pearson correlation coefficient of -0.587 with P<0.0001) both interaction terms of wealth with maternal education and typesite with maternal education is included when performing the LR test for wealth differences. This reports effects independent of the correlation with typesite.

The LR test suggests that maternal education do not have a different effect by sex for both height-for-age or weight-for-age for each of the rounds. Consequently, no separate regressions are run by child gender. All other p-values are however found to be significant, except that for round 01 height-for-age by wealth (P = 0.2039). This confirms the differential impact of education on nutrition. Because of the similarity in the results, estimations are run separately by typesite and by wealth quartile for only round 03 of the dataset, and by cohort for only round 01 of the dataset[4]. The results are provided in Tables 10, 11 and 12 respectively. As usual, community fixed effects are applied in each of the regression estimation and robust standard errors are reported.

Table 10 reports that maternal education has higher and more significant impact for urban areas compared to rural areas both in terms of height-for-age and weight-for-age health indicators. This result is consistent with the one we had found from section V analysis, reported in Tables 04 and 05. This difference in impact is more profound for height-for-age z-scores than for weight-for-age z-scores. For example, while change in height-for-age associated with higher maternal education is

---

[4] We use round 01 data for cohort analysis because the older cohort does not contain weight-for-age z-score data for rounds 02 and 03.

Table 10: Maternal education by typesite, Round 03

	height-for-age z-scores				weight-for-age z-scores			
	Urban		Rural		Urban		Rural	
	coef.	robust se	coef.	robust se	coef.	robust se	coef.	robust se
primary	0.131	0.148	0.111*	0.061	-0.110	0.174	-0.039	0.072
secondary	0.186*	0.108	0.001	0.066	0.078	0.142	-0.050	0.078
higher	0.490***	0.142	0.045	0.090	0.288	0.191	0.233*	0.114
agemon	-0.044	0.031	-0.005	0.017	1.003	0.628	0.158	0.315
agemon2	0.000	0.000	0.000	0.000	-0.005	0.003	-0.001	0.002
sex	-0.036	0.077	0.065	0.041	0.094	0.109	0.227***	0.048
hhadults	0.019	0.019	0.006	0.009	0.048*	0.027	0.010	0.011
housing	0.397*	0.248	-0.026	0.084	0.983***	0.335	-0.058	0.104
durables	0.694**	0.279	0.544***	0.149	0.941**	0.401	0.710***	0.174
services	0.198	0.297	0.398***	0.152	0.185	0.399	0.458**	0.187
ownland	0.144	0.119	-0.037	0.054	0.026	0.183	-0.011	0.064
animals	0.198	0.200	-0.021	0.048	0.405	0.263	-0.026	0.057
N	749		2143		504		1418	
R-square	0.240		0.176		0.250		0.219	

*p<0.10, **p<0.05, ***p<0.01

Unreported variables are the time invariant control factors: schedules tribe, backward class, other castes (omitted group is schedules caste), muslim, other religion (omitted group is hindu), rayalaseema, coastal (omitted group is telangana) and mother's height.

Table 11: Maternal education by wealth quartile, Round 03

	height-for-age z-scores				weight-for-age z-scores			
	Lowest Quartile		Top Quartile		Lowest Quartile		Top Quartile	
	coef.	robust se	coef.	robust se	coef.	robust se	coef.	robust se
primary	0.217*	0.116	0.146	0.161	0.099	0.123	-0.040	0.204
secondary	0.140	0.139	0.307**	0.125	0.139	0.173	0.184	0.178
higher	0.136	0.152	0.608***	0.148	0.410**	0.201	0.415*	0.219
agemon	0.028	0.032	-0.042	0.033	1.13**	0.500	0.780	0.715
agemon2	-0.000	0.000	0.000	0.000	-0.006**	0.003	-0.004	0.004
sex	-0.030	0.079	-0.059	0.080	0.135	0.085	0.094	0.121
hhadults	-0.011	0.020	0.029	0.018	-0.030	0.022	0.012	0.028
housing	-0.363	0.224	0.348	0.333	-0.019	0.250	1.102**	0.537
durables	0.848***	0.313	0.038	0.355	1.311***	0.376	1.211**	0.504
services	-0.778**	0.365	-0.129	0.426	-0.386	0.425	0.465	0.628
ownland	-034	0.104	0.098	0.117	-0.039	0.115	-0.178	0.188
animals	0.066	0.089	0.046	0.144	0.099	0.091	0.144	0.208
N	709		723		492		480	
R-square	0.234		0.290		0.337		0.292	

*p<0.10, **p<0.05, ***p<0.01

Unreported variables are the time invariant control factors: schedules tribe, backward class, other castes (omitted group is schedules caste), muslim, other religion (omitted group is hindu), rayalaseema, coastal (omitted group is telangana) and mother's height.

Table 12: Maternal education by cohort, Round 01

	height-for-age z-scores				weight-for-age z-scores			
	Younger Cohort		Older Cohort		Younger Cohort		Older Cohort	
	coef.	robust se	coef.	robust se	coef.	robust se	coef.	robust se
primary	0.075	0.095	0.143	0.108	0.012	0.075	-0.006	0.110
secondary	0.080	0.086	0.127	0.100	0.090	0.070	0.125	0.106
higher	0.304**	0.126	0.241*	0.140	0.256***	0.098	0.059	0.144
agemon	0.090	0.066	-0.161	0.533	0.003	0.051	-0.134	0.515
agemon2	-0.007***	0.003	0.001	0.003	-0.002	0.002	0.001	0.003
sex	0.114*	0.062	-0.014	0.066	0.166***	0.047	0.192***	0.067
hhadults	0.001	0.004	-0.017	0.026	-0.002	0.002	0.016	0.025
housing	0.091	0.130	-0.056	0.141	0.073	0.102	-0.076	0.136
durables	0.912***	0.219	0.632**	0.294	0.523**	0.176	0.653**	0.295
services	0.160	0.200	-0.105	0.236	0.175	0.165	0.087	0.240
ownland	0.125	0.085	-0.083	0.092	0.100	0.067	-0.009	0.096
animals	-0.212***	0.009	0.180**	0.047	-0.005	0.938	0.134	0.138
N	1912		993		1911		993	
R-square	0.268		0.197		0.187		0.219	

*p<0.10, **p<0.05, ***p<0.01

Unreported variables are the time invariant control factors: schedules tribe, backward class, other castes (omitted group is schedules caste), muslim, other religion (omitted group is hindu), rayalaseema, coastal (omitted group is telangana) and mother's height.

0.490 in urban areas and 0.045 in rural areas, change in weight-for-age associated with higher maternal education is 0.288 for urban and 0.233 for rural.

The wealth index was broken down into quartiles and estimates are reported for the lowest and top quartile in Table 11. A clear difference can be seen for height-for-age z-scores, where the coefficient for higher maternal education is significant and much larger (0.608) for the top quartile than that of the lowest quartile (0.136). While we see a similar trend for secondary education, interestingly primary education seems to have a larger impact for the lowest wealth quartile households. This can be because primary education is not sufficient enough to have a strong impact in top wealth quartile households, while it is so for lower wealth quartiles.

Stratification by younger and older cohorts is essentially running estimations for separate age groups. As already mentioned, only results for round 01 are reported. In round 01, the younger cohort consisted of 1-year-old children and the older cohort consisted of 8-year-old children. Results from Table 12 indicate that higher maternal education level has a larger impact for the younger cohort (estimates are 0.304 for height-for-age and 0.256 for weight-for-age) compared to that for the older cohort (estimates are 0.241 for height-for-age and 0.059 for weight-for-age). It is difficult to infer anything from primary and secondary level education because of their high standard errors.

Whether the differential education impact found above is due to age difference or simply due to cohort difference is a matter of concern. The Young Lives dataset allows us to compare the younger cohort at age 8 (in round 03) with the older cohort also at age 8 (in round 01). The results for this are presented in A-Table 06 in the appendix. Results are similar to the ones reported in Table 12. Higher maternal education level still has a larger impact for the younger cohort (estimates are 0.283 for height-for-age and 0.306 for weight-for-age). This indicates that maternal education effects for the two cohorts differently, possibly due to differences in parental preferences.

6.2) Role of paternal and community level education on child health

In this section we go over three analyses: (i) whether paternal education significantly plays a role in child health estimation, (i) the role of community level education on child health, and (iii) how community size complements with parental education in impacting child health.

Table 13 reports the role of paternal education on child health using two specifications. The first is the independent effect of paternal education controlled for confounders using base variables as well using a comprehensive set of controls. The second specification adds maternal education to the first specification. As usual, both specifications are run with community fixed effects.

The results suggest that secondary and higher paternal education has significant positive impact on child health for both height-for-age and weight-forage. The relationship is however no longer significant for height-for-age when maternal education enters the specification. This indicates a confounding relationship between the role of maternal and paternal education on child health. This also means that the coefficients presented in previous tables that do not include parental education overestimated the role of maternal education.

Comparing with the pooled regression presented in Table 09, adding parental education into the model, reduces the maternal education coefficients. While secondary education impact becomes non-significant, higher education coefficient drops from 0.202 to 0.175 for height-for-age and from 0.208 to 0.170 for weight-for-age.

Table 13: Paternal education and child health, Pooled estimation

		height-for-age z-scores				weight-for-age z-scores			
		(1)		(2)		(3)		(4)	
		coef.	rob. se	coef.	rob. se	coef.	rob. se	coef.	rob. se
Paternal Literacy	primary	-0.009	0.035	-0.024	0.036	0.026	0.035	0.027	0.035
	secondary	0.075**	0.033	0.054	0.035	0.089***	0.033	0.085**	0.034
	higher	0.088**	0.040	0.046	0.042	0.155***	0.041	0.119***	0.042
Maternal Literacy	primary			0.104***	0.038			-0.036	0.037
	secondary			0.049	0.037			0.031	0.038
	higher			0.185***	0.052			0.170***	0.053
Baseline Variables	agemon	-0.006***	0.001	-0.006***	0.001	-0.013***	0.002	-0.013***	0.002
	agemon2	0.000***	0.000	0.000***	0.000	0.000***	0.000	0.000***	0.000
	sex	0.049*	0.024	0.047*	0.023	0.141***	0.024	0.141***	0.024
	hhadults	0.005	0.004	0.005	0.003	0.001	0.002	0.001	0.002
	housing	0.033	0.051	0.034	0.050	0.010	0.050	-0.010	0.050
	durables	1.730***	0.271	1.626***	0.268	1.618***	0.312	1.618***	0.311
	services	0.170***	0.078	0.155*	0.077	0.124*	0.080	0.124	0.080
	services* typesite	-0.583***	0.153	-0.542***	0.150	-0.562***	0.171	-0.562***	0.170
	ownland	-0.048	0.030	-0.049*	0.029	-0.003	0.035	-0.003	0.034
	animals	-0.008	0.029	-0.006	0.028	0.039	0.028	0.039	0.028
	typesite	0.151	0.084	0.145	0.082	0.062	0.088	0.062	0.087
	N	8710		8710		6762		6762	
	R-sq	0.114		0.147		0.161		0.184	

*p<0.10, **p<0.05, ***p<0.01

Unreported variables are the time invariant control factors: schedules tribe, backward class, other castes (omitted group is schedules caste), muslim, other religion (omitted group is hindu), rayalaseema, coastal (omitted group is telangana) and mother's height.

Role of community level education is investigated in a similar manner. A variable indicating the community average of the percentage of adults in a household who have completed high school or above is used as an indicator for community level education. The mean of the constructed community level education variable is 0.195 with standard deviation 0.130, and ranges between 0.013 and 0.627. Results from pooled regression are reported in A-Table 06 in the appendix. Community level education is found to be significant for both height-for-age (estimated coefficient is 0.812 with standard error 0.202) and weight-for-age (estimated coefficient is 0.760 with standard error 0.215).

Interpreting the coefficients is straightforward. A 10 percent increase in community level education would, on average, improve height-for-age z-scores by 0.0812 and weight-for-age z-scores by

0.0202, holding everything else constant. Worthy to note is that the variable remains consistently significant and positive in cross-sectional regressions for all the rounds (results not reported).

This consistency is however broken once we take the community size into account. The final analysis in this section involves stratifying the sample by community size. Any community with household samples less than 20 (in a single round) is identified as a small community, while the rest are identified as large communities. This of course assumes that the number of samples themselves reflects the size of the communities and is appropriately captured by the sampling procedure. Out of 102 communities, 32 were found to be small. Table 14 presents the results.

Table 14: Community size effect on education, Pooled estimation

		Small Communities				Large Communities			
		height-for-age z-scores		weight-for-age z-scores		height-for-age z-scores		weight-for-age z-scores	
		coef.	rob. se	coef.	rob. se	coef.	rob. se	coef.	rob. se
Community Level Literacy	comedu	0.259	0.707	0.587	0.711	0.842***	0.220	0.802***	0.235
Paternal Literacy	primary	0.047	0.090	0.214**	0.090	-0.023	0.039	0.014	0.038
	secondary	0.308***	0.081	0.304***	0.085	0.014	0.038	0.047	0.037
	higher	0.255**	0.103	0.206*	0.107	0.004	0.045	0.092*	0.046
Maternal Literacy	primary	0.059	0.084	-0.039	0.084	0.100**	0.041	-0.032	0.039
	secondary	0.014	0.091	-0.051	0.093	0.079**	0.040	0.071*	0.040
	higher	-0.034	0.146	0.377**	0.163	0.238***	0.053	0.190***	0.054
Baseline Variables	agemon	-0.002	0.002	-0.006	0.004	-0.006***	0.001	-0.013***	0.002
	agemon2	0.000	0.000	0.000	0.000	0.000***	0.000	0.000***	0.000
	sex	0.085	0.059	0.154**	0.061	0.049**	0.025	0.154***	0.026
	hhadults	-0.006	0.007	0.001	0.005	0.007*	0.004	0.001	0.002
	housing	0.019	0.117	-0.027	0.122	0.044	0.056	-0.002	0.055
	durables	2.748**	1.391	3.608**	1.467	1.618***	0.272	1.679***	0.319
	services	-0.085	0.174	0.005	0.178	0.200**	0.086	0.126	0.090
	services* typesite	-1.198*	0.707	-1.640**	0.748	-0.532***	0.156	-0.601***	0.176
	ownland	-0.085	0.080	-0.013	0.086	-0.039	0.032	-0.008	0.038
	animals	0.021	0.066	0.100	0.067	-0.010	0.032	0.028	0.032
	typesite	0.414	0.337	0.180	0.356	0.155*	0.081	0.179**	0.082
	N	1250		974		7460		5788	
	R-sq	0.127		0.185		0.130		0.159	

*p<0.10, **p<0.05, ***p<0.01

Unreported variables are the time invariant control factors: schedules tribe, backward class, other castes (omitted group is schedules caste), muslim, other religion (omitted group is hindu), rayalaseema, coastal (omitted group is telangana) and mother's height, and the community level fixed effects.

Compared to the pooled estimates in A-Table 06, community level education coefficients for the small communities are smaller and non-significant (0.259 for height-for-age and 0.587 for weight-for-age). On the other hand, compared to the pooled estimates, the estimates for large communities further increases from 0.812 to 0.842 for height-for-age and from 0.760 to 0.802 for weight-for-age. This suggests that small communities tend to dilute the effect of community level education on child health. The result is consistent with Moestu (2005).

However, what is intriguing is that, even after applying all controls, the effect of paternal education on child health is mostly significant and positive for smaller communities, while the same is found for maternal education in larger communities. In the absence of causal pathway analysis, the reason for this can only be speculated. Perhaps larger communities have better facilities that complement with maternal education allowing it to have a stronger and more significant effect. We have seen similar indications while looking into the differential impact of maternal education. It can be further speculated that in absence of such facilities in smaller communities, paternal education picks up the gap and plays a more dominant and significant role.

7.0) Conclusion

This essay is essentially divided into two parts looking into two main determinants of child health: income and education. The Young Lives panel dataset between 2002 and 2009 from Andhra Pradesh is used in the study. The principal findings, implications as well as limitations are summarized in this concluding section.

Due to data restrictions, the income analysis is done for the 2006 to 2009 time period when there is a reduction in stunting. Per capita consumption expenditure was used as a measure of income growth, and its impact on child nutritional status was investigated using a number of estimation methods. A comprehensive set of controls was employed to (partially) control for individual observed and unobserved heterogeneity. Community level heterogeneity, such as the health environment, was controlled for using community fixed effects. Presence of endogeneity and measurement errors were controlled for using instrument variables. Finally time invariant unobserved heterogeneity was removed using panel fixed effects procedure.

Significant and positive association between per capita consumption expenditure and height-for-age z-scores was found for all estimation procedures except for fixed effects panel estimation. Although the panel estimates do not control for endogeneity and measurement biases, which could be the reason for the lack of significance, this finding precludes us from confirming with certainty that income plays a significant role on child health.

Finally, based on the significant estimations, it is concluded that the income effect only explains between 0 to 34 percent of the change in stunting, however only under set assumptions. Therefore other actors also play significant roles in the reduction of stunting. It should be noted using this to argue against the benefits of income growth will be misleading. Glewwe, Koch and Nguyen (2002) point out that income growth also typically increases government budget as well as overall community wealth which can affect child health through improved health services.

Next, we turn the analysis towards education and look into a number of issues in the role of education on child health. Substantial differential impacts of maternal education on child health were noted and an overall estimate was deemed inappropriate as it would mask the heterogeneity. The effect of maternal education was found to be stronger in urban areas than rural areas, among the more wealthy than among the poor and on the younger cohort compared to the older cohort. It is speculated that the first two differential impacts is mainly due to maternal education complementing with modern services available in urban areas and more resource availability. No gender based difference was found.

Bicego and Boerma (1993) also report stronger education impact in urban areas and Bairagi (1980) reports stronger education impact for wealthier households. The idea that modern services available in urban areas complement maternal education has been also suggested by several authors including Caldwell (1979), Caldwell (1994) and Glewwe, Koch and Nguyen (2002), as discussed in the literature review. This complementary effect implies two things: (i) emphasis should be put on research that identifies the causal pathway of the complementary effects, and (ii) policies should focus on improving the complementary variables which can have dual impact on the society and child health through its complementary nature.

Role of education beyond the mother-child pair was also investigated. While paternal education was found to have significant positive impact for only weight-for-age health measures, community level education was found to be significant for both child health measures. By segregating the sample according to community size, paternal education was found to have significant positive impact for smaller communities, while the same was found for maternal education but for larger communities. The idea of educated communities complementing maternal education is speculated to cause this difference.

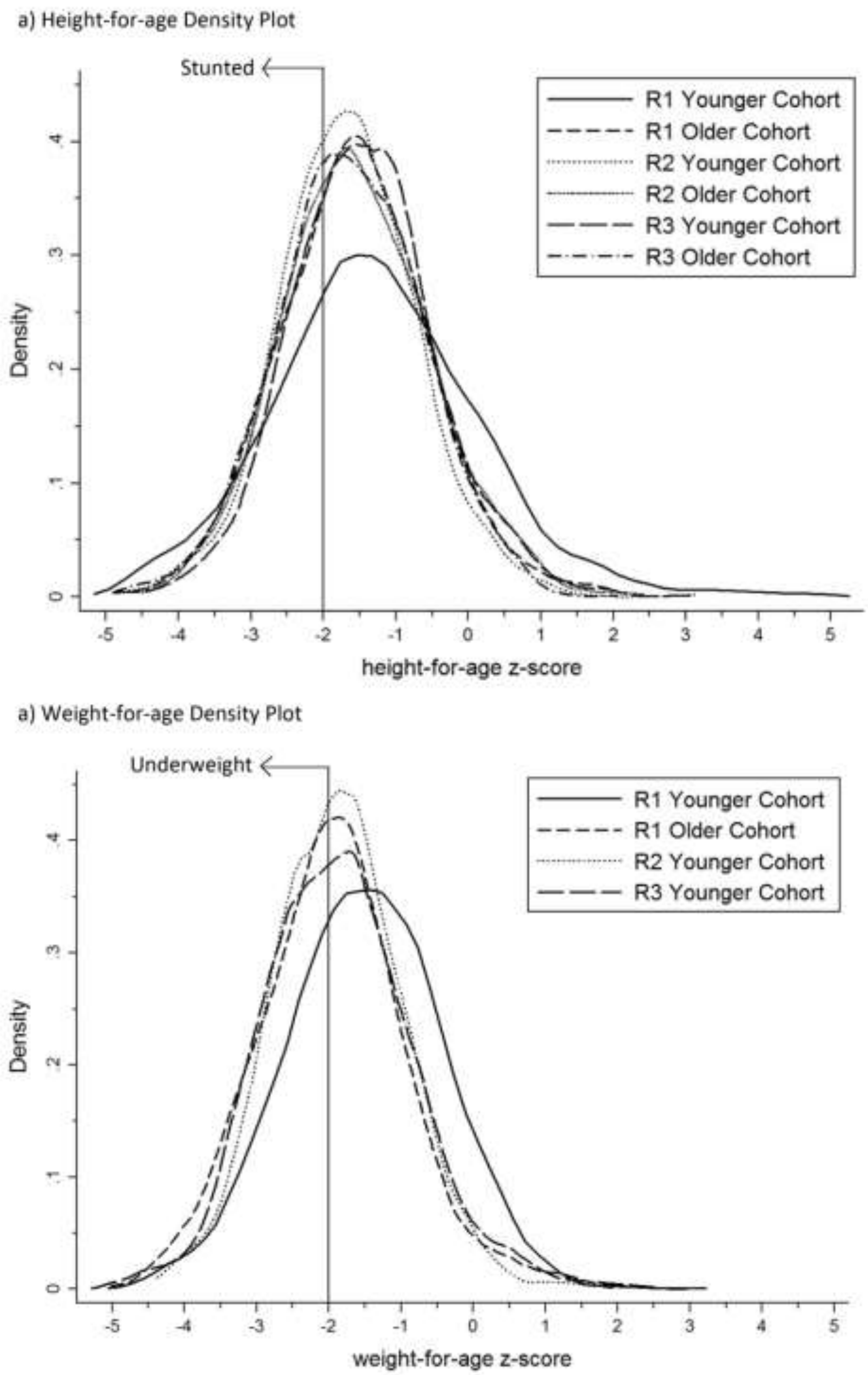
Some limitations in the education results should be noted. Majority of the analyses was done with categorical education variables which lumps the first few years of education, below primary level attainment, with the no education category. This ignores the importance of few years of educational attainment, which have been shown to be crucial in skills acquisition by other authors (Basu and Stephenson, 2005).

In absence of a panel estimation, individual heterogeneity was not controlled for which can influence maternal education impacts, as indicated in the discussion related cohort and same age analysis (Table 12 and A-Table 06). Causal pathways, especially those of health knowledge, are important elements of child health determination and were left for future investigation.

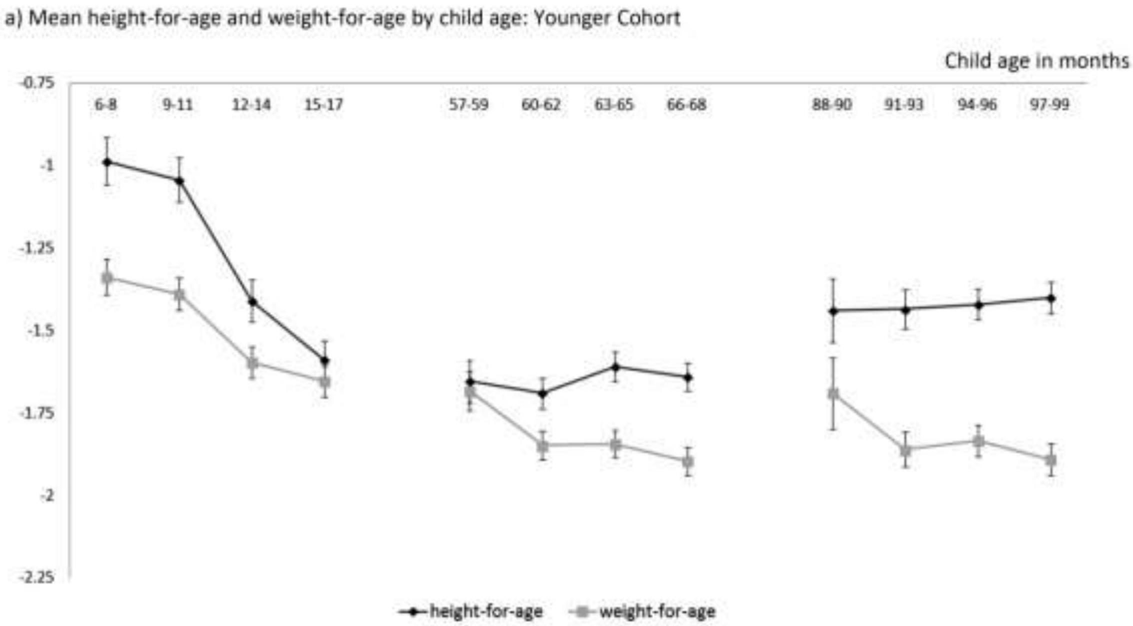
The Young Lives dataset note several shocks that took place within the seven year period, including shocks, family deaths, livestock death etc. which was not controlled for in the regressions. Controlling for this would result in more accurate estimates. Finally, as was stressed earlier, given that the Young Lives dataset is not nationally representative, any extrapolation of any of the findings will have to be adjusted with appropriate weights and should be treated with caution.

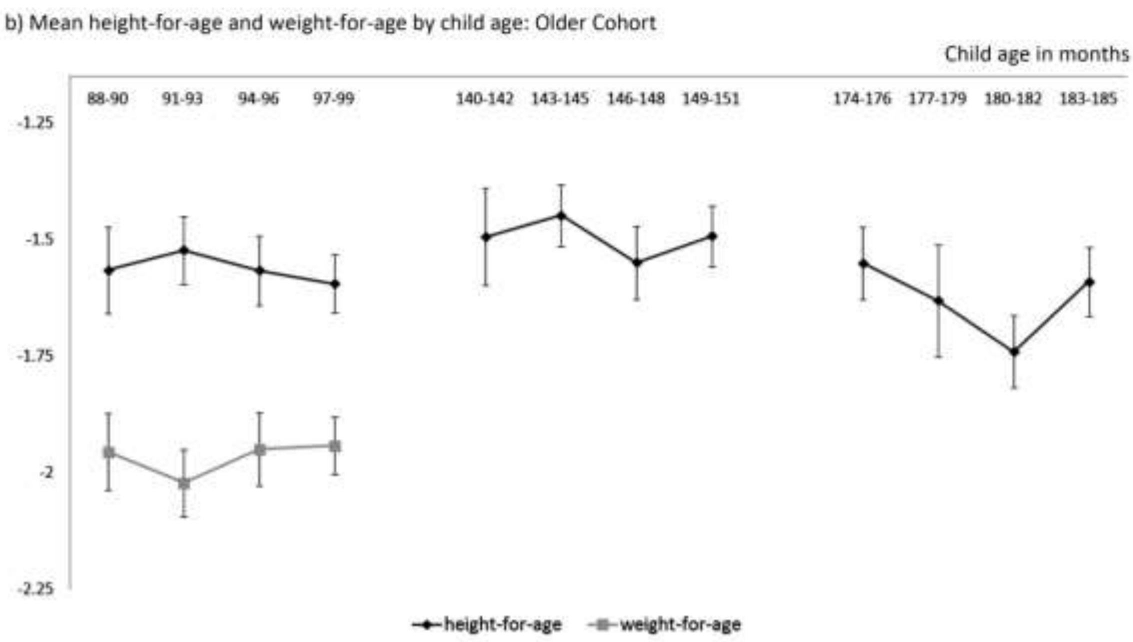
Appendix: Additional Figures and Tables

A-Figure 01 KDensity Plots:



A-Figure 02 Z-score line plot:





A-Table 01: Observations count of before and after dropping of missing values

	Younger	Older	Total
Before Dropping			
2002	2011	1008	3019
2006	1950	994	2944
2009	1951	986	2937
Total	5912	2988	8900
After Dropping			
	Younger	**Older**	**Total**
2002	1912	993	2905
2006	1937	977	2914
2009	1922	970	2892
Total	5771	2940	8711

A-Table 02: Panel descriptive statistics

Variable		Std. Dev.	Mean	Observations	
zhfa	overall	1.143		N	8711
	between	0.977	-1.505	n	2943
	within	0.600		T-bar	2.960
zwfa	overall	1.042		N	6763
	between	0.961	-1.784	n	2942
	within	0.463		T-bar	2.300
agemon	overall	52.524		N	8711
	between	39.714	85.925	n	2943
	within	34.362		T-bar	2.960
hhadults	overall	4.310		N	8711
	between	2.948	3.752	n	2943
	within	3.133		T-bar	2.960
tconsrpc	overall	605.755		N	5481
	between	506.853	879.711	n	2914
	within	348.303		T-bar	1.881
wealth	overall	0.198		N	8710
	between	0.176	0.463	n	2943
	within	0.090		T-bar	2.960
housing	overall	0.281		N	8711
	between	0.228	0.541	n	2943
	within	0.163		T-bar	2.960
durables	overall	0.194		N	8710
	between	0.163	0.243	n	2943
	within	0.106		T-bar	2.960
services	overall	0.256		N	8711
	between	0.226	0.604	n	2943
	within	0.121		T-bar	2.960

A-Table 03: Panel estimation with wealth index and sub-indices by typesite

	WI: Urban		WI: Rural		Indices: Urban		Indices: Rural	
	First Difference	Fixed Effects	First Difference	Fixed Effects	First Difference	Fixed Effects	First Difference	Fixed Effects
agemon	-0.005***	-0.001	-0.008***	-0.006***	-0.005***	-0.001	-0.008***	-0.006***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
agemon2	0.000***	0.000	0.000***	0.000***	0.000***	0.000	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
hhadults	0.059	0.071*	0.007**	0.005*	0.055	0.070*	0.007**	0.005*
	(0.037)	(0.038)	(0.003)	(0.003)	(0.037)	(0.039)	(0.003)	(0.003)
wealth	0.189	0.299	0.295***	0.453***				
	(0.245)	(0.242)	(0.115)	(0.124)				
housing					-0.058	0.069	0.037	0.091
					(0.110)	(0.111)	(0.056)	(0.059)
durables					0.505**	0.479***	0.205*	0.205*
					(0.178)	(0.185)	(0.104)	(0.115)
services					-0.098	-0.167	0.146*	0.232***
					(0.152)	(0.156)	(0.077)	(0.084)
N	1480	2205	4258	6505	1480	2205	4258	6505
R-square	0.013	0.004	0.036	0.031	0.019	0.009	0.036	0.032

Robust standard errors in parentheses

*p<0.10, **p<0.05, ***p<0.01

A-Table 04: Maternal education and child health, Cross section estimation by round

	height-for-age z-score			weight-for-age z-score		
	Round 01	Round 02	Round 03	Round 01	Round 02	Round 03
continuous[1]	0.003	0.004*	0.007**	0.003	0.007**	0.011***
	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)
primary	0.089	0.138***	0.117**	0.0003	-0.027	-0.052
	(0.231)	(0.010)	(0.041)	(0.060)	(0.058)	(0.067)
secondary	0.085	0.067	0.072	0.083	0.020	0.033
	(0.069)	(0.054)	(0.054)	(0.059)	(0.062)	(0.067)
higher	0.203**	0.211***	0.265***	0.184**	0.148*	0.306***
	(0.099)	(0.077)	(0.072)	(0.080)	(0.089)	(0.095)
agemon	-0.074***	0.012	-0.013	-0.042***	0.229	0.435
	(0.009)	(0.012)	(0.015)	(0.007)	(0.171)	(0.285)
agemon2	0.001***	-0.000	0.000	0.000***	-0.002	-0.002
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)
sex	0.082*	0.012	0.044	0.176***	0.041	0.191***
	(0.048)	(0.035)	(0.036)	(0.038)	(0.040)	(0.044)
hhadults	0.004	0.070	0.007	-0.001	0.016	0.019*
	(0.004)	(0.009)	(0.008)	(0.002)	(0.010)	(0.010)
housing	0.053	-0.000	0.055	0.038	-0.041	0.088
	(0.102)	(0.077)	(0.079)	(0.081)	(0.088)	(0.099)
durables	0.835***	0.917***	0.628***	0.636***	0.776***	0.766***
	(0.177)	(0.131)	(0.130)	(0.150)	(0.154)	(0.166)
services	0.076	-0.075	0.319**	0.124	-0.029	0.289*
	(0.157)	(0.120)	(0.132)	(0.133)	(0.138)	(0.166)
ownland	0.071	-0.405	0.002	0.069	0.508*	-0.006
	(0.066)	(0.327)	(0.049)	(0.055)	(0.306)	(0.061)
animals	-0.088	0.014	-0.021	0.033	0.035	-0.007
	(0.064)	(0.042)	(0.045)	(0.050)	(0.047)	(0.055)
N	2905	2913	2892	2904	1936	1922
R-square	0.168	0.189	0.214	0.175	0.202	0.256

[1] This is a continuous maternal education variable from a separate regression, with the same baseline and control variables as reported here.

Robust standard errors in parentheses

*p<0.10, **p<0.05, ***p<0.01

Unreported variables are the time invariant control factors: schedules tribe, backward class, other castes (omitted group is schedules caste), muslim, other religion (omitted group is hindu), rayalaseema, coastal (omitted group is telangana) and mother's height.

A-Table 05: LR test p-values for differential impact of maternal education

		Round 01	Round 02	Round 03
sex	height-for-age z-score	0.2768	0.8110	0.4943
	weight-for-age z-score	0.1401	0.2026	0.1720
cohort	height-for-age z-score	<0.0001	0.0174	0.0086
	weight-for-age z-score	0.0006	n/a	n/a
typesite	height-for-age z-score	0.0992	<0.0001	<0.0001
	weight-for-age z-score	0.0024	0.0089	<0.0001
wealth index	height-for-age z-score	0.2039	<0.0001	<0.0001
	weight-for-age z-score	0.0049	0.0168	0.0010

A-Table 06: Same age analysis

	height-for-age z-scores				weight-for-age z-scores			
	Round 1 Older Cohort (8 years old)		Round 3 Younger Cohort (8 years old)		Round 1 Older Cohort (8 years old)		Round 3 Younger Cohort (8 years old)	
	coeff	robust se	coeff	robust se	coeff	robust se	coeff	robust se
primary	0.143	0.108	0.0627	0.070	-0.006	0.110	-0.052	0.067
secondary	0.127	0.100	0.0525	0.063	0.125	0.106	0.033	0.067
higher	0.241*	0.140	0.283***	0.087	0.059	0.144	0.306***	0.095
agemon	-0.161	0.533	0.708***	0.258	-0.134	0.515	0.435	0.258
agemon2	0.001	0.003	-0.004***	0.001	0.001	0.003	-0.002	0.001
sex	-0.014	0.066	0.071*	0.043	0.192***	0.067	0.191***	0.044
hhadults	-0.017	0.026	0.015	0.009	0.016	0.025	0.019*	0.010
housing	-0.056	0.141	0.172*	0.096	-0.076	0.136	0.089	0.099
durables	0.632**	0.294	0.452***	0.157	0.653**	0.295	0.766***	0.166
services	-0.105	0.236	0.436***	0.163	0.087	0.240	0.289*	0.166
ownland	-0.083	0.092	0.034	0.062	-0.009	0.096	-0.006	0.061
animals	0.180**	0.047	0.021	0.055	0.134	0.138	0.007	0.055
N	993		1922		993		1922	
R-square	0.197		0.264		0.219		0.256	

*p<0.10, **p<0.05, ***p<0.01

Unreported variables are the time invariant control factors: schedules tribe, backward class, other castes (omitted group is schedules caste), muslim, other religion (omitted group is hindu), rayalaseema, coastal (omitted group is telangana) and mother's height.

A-Table 07: Community level education and child health, Pooled estimation

		height-for-age z-scores		weight-for-age z-scores	
		coef.	robust se	coef.	robust se
Community Level Literacy	comedu	0.812***	0.202	0.760***	0.215
Paternal Literacy	primary	-0.018	0.035	0.042	0.035
	secondary	0.048	0.034	0.079**	0.033
	higher	0.040	0.042	0.113***	0.042
Maternal Literacy	primary	0.093**	0.037	-0.034	0.035
	secondary	0.064*	0.036	0.050	0.037
	higher	0.209***	0.050	0.197***	0.052
Baseline Variables	agemon	-0.005***	0.001	-0.012***	0.002
	agemon2	0.000***	0.000	0.000***	0.000
	sex	0.049**	0.023	0.148***	0.024
	hhadults	0.005	0.003	0.001	0.002
	housing	0.034	0.050	-0.011	0.050
	durables	1.683***	0.264	1.733***	0.309
	services	0.016**	0.077	0.125	0.080
	services* typesite	-0.577***	0.148	-0.633***	0.168
	ownland	-0.049*	0.030	-0.011	0.035
	animals	-0.006	0.029	0.042	0.029
	typesite	0.162**	0.077	0.148*	0.078
Community Level Fixed Effects	comage	-0.001	0.002	-0.001	0.002
	comadults	-0.065***	0.018	-0.083***	0.017
	comhq	0.207	0.158	0.479***	0.155
	comcd	-0.470	0.394	-0.505	0.391
	comsv	-0.199	0.241	-0.275	0.254
	comland	-0.276	0.183	-0.360*	0.193
	comanimals	0.244**	0.124	0.208*	0.126
	N	8710		6762	
	R-sq	0.1225		0.1551	

*p<0.10, **p<0.05, ***p<0.01

Unreported variables are the time invariant control factors: scheduled tribes, backward classes, other castes (omitted group is schedules castes), muslim, other religion (omitted group is hindu), rayalaseema, coastal (omitted group is telangana) and mother's height.

References

Alderman, H., Hoogeveen, H. & Rossi, M. (2005). "Reducing Child Malnutrition in Tanzinia." World Bank Policy Research Working Paper 3567. April 2005.

Attawell, K. (2003). "International longitudinal research on childhood poverty: practical guidelines and lessons learned from Young Lives". Oxford: Young Lives. [Online] Available from: http://www.younglives.org.uk. [Accessed: 19th June 2013]

Bairagi, R. (1980). "Is income the only constraint on child nutrition in rural Bangladesh?" Bulletin of the World Health Organization. 58(5). p. 767-72.

Bairagi, R., Aziz, K. M., Chowdhury, M. K. & Edmonston, B. (1982). "Age Misstatement for Young Children in Rural Bangladesh." Demography. 19(4). p. 447-458.

Barrera, A. (1990). "The role of maternal schooling and its interaction with public health programs in child health production." Journal of Development Economics. 32(1). p. 69-91.

Basu, A. M. & Stephenson R. (2005). "Low levels of maternal education and the proximate determinants of childhood mortality: a little learning is not a dangerous thing." Social Science & Medicine. 60(9). p. 2011-23.

Behrman, J. R. & Rosenzweig, M. R. (1994). "Caveat emptor: Cross-country data on education and the labour force." Journal of Development Economics. 44. p. 147-171.

Behrman, J. R. & Wolfe, B. L. (1984). "Evidence on nutrition demand." Journal of Development Economics. 14. p. 105-128.

Behrman, J. R. & Wolfe, B. L. (1987). "How does mother's schooling affect family health, nutrition, medical care usage, and household sanitation?" Journal of Econometrics. 36(12). p. 185-204.

Bhuiya, A., Streatfield, K. & Meyer, P. (1990). "Mother’s hygienic awareness, behaviour, and knowledge of major childhood diseases in Matlab, Bangladesh." In Caldwell J., Findley S., Caldwell P., Santow G., Cosford W., Braid, J. & Broers-Freeman D. (eds). What we know about health transition: The cultural, social, and behavioral determinants of health (p. 462–477). Canberra, Australia: The Australian National University Printing Service.

Bicego, G. T. & Boerma J. T. (1993). "Maternal education and child survival: a comparative study of survey data from 17 countries." Social Science & Medicine. 36(9). p. 1207-1227.

Bishai, D. (1996). "Quality time: how parents' schooling affects child health through its interaction with childcare time in Bangladesh." Health Economics. 5(5). p. 383-407.

Black, R. E., Morris, S. S. & Bryce J. (2003). "Where and why are 10 million children dying every year?" The Lancet. 361(9376). p. 2226-2234.

Bloom, S. S., Wypij, D. & das Gupta, M. (2001). "Dimensions of women's autonomy and the influence on maternal health care utilization in a North Indian city." Demography. 38(1). p. 67-78.

Caldwell, J. (1994). "How is greater maternal education translated into lower child mortality?" Health Transition Review. 4. p. 224-229.

Caldwell, J. C. (1979). "Education as a factor in mortality decline: an examination of Nigerian data." Population Studies. 33(3). p. 395-413.

Caldwell, J. C. (1993). "Health transition: The cultural, social, and behavioural determinants of health in the Third World." Social Science & Medicine. 36(2). p. 125–135.

Claeson, M., Bos, E. R., Mawji, T. & Pathmanathan, I. (2000). "Reducing child mortality in India in the new millennium." Bulletin of the World Health Organization. 78(10). p. 1192-1199.

Clark, T. S. & Linzer D. A. (2012). "Should I Use Fixed or Random Effects?" The Society for Political Methodology Working Paper. [Online] Available from: http://polmeth.wustl.edu/media/Paper/ClarkLinzerREFEMar2012.pdf. [Accessed: 15th August 2013]

Cleland, J. & Jejeebhoy S. J. (1996). "Maternal schooling and fertility: evidence from census and surveys." In Jeffery, R. & Basu, A. M. (eds). Girls' schooling, women's autonomy and fertility change in South Asia. New Dehli: Sage Publications.

David V., Moncada M. & Ordonez F. (2004). "Private and public determinants of child nutrition in Nicaragua and Western Honduras", Journal of Economics and Human Biology. 2. p. 457–488.

de Onis, M., Monteiro, C. Akre, J. & Glugston G. (1993). "The worldwide magnitude of protein-energy malnutrition: an overview from the WHO Global Database on Child Growth." Bulletin of the World Health Organization. 71(6). p. 703-712.

Deaton, A. (1997). "The Analysis of Household Surveys." Baltimore: Johns Hopkins University Press.

Dercon, S. (2008). "Children and the Food Price Crisis." Policy Brief 5, Oxford: Young Lives. [Online] Available from: http://www.younglives.org.uk. [Accessed: 19th June 2013]

Desai, S. & Alva, S. (1998). "Maternal education and child health: is there a strong causal relationship?" Demography. 35(1). p. 71-81.

Drukker, D. M. (2003). "Testing for serial correlation in linear panel-data models." Stata Journal. 3. p. 168–177.

El-Ghannam, A. R. (2003). "The Global Problems of Child Malnutrition and Mortality in Different World Regions." Journal of Health and Social Policy. 16(4). p. 1-26.

Frost, M. S., Forste, R. & Haas, D. W. (2005). "Maternal education and child nutritional status in Bolivia: finding the links." Social Science and Medicine. 60. p. 395-407.

Galab, S., Kumar, A. V., Reddy, P. P., Singh R. & Vennam U. (2011). "The Impact of Growth on Childhood Poverty in Andhra Pradesh: Initial findings from India." Round 3 Survey Report. Oxford: Young Lives. [Online] Available from: http://www.younglives.org.uk. [Accessed: 19th June 2013]

Glewwe P., Koch S. & Nguyen, B.L. (2002). "Child nutrition, economic growth and the provision of health care services in Vietnam in the 1990s." Policy Research Working Paper Series 2776, The World Bank.

Glewwe, P. (1999). "Why does mother's schooling raise child health in developing countries? Evidence from Morocco." Journal of Human Resources. 34(1). p. 124-159.

Glewwe, P., Agrawal, N. & Dollar, D. (2004). "Economic Growth, Poverty, and Household Welfare in Vietnam." World Bank Regional and Sectoral Studies 29086. Washington DC: The World Bank.

Gragnolatti, M., Shekar, M. & Das Gupta, M. (2005). "India's Undernourished Children: A Call for Reform and Action." World Bank Health, Nutrition and Population Discussion Paper. August 2005.

Hall, A., Khanh, L. N., Son, T. H., Dung, N. Q., Lansdown, R. G., Dar D. T., Hanh, N. T., Moestu, H., Khoi, H. H. & Bundy D. A. (2001). "An association between chronic undernutrition and educational test scores in Vietnamese children." European Journal of Clinical Nutrition. 55(9). p. 801-804.

Kirchberger, M. (2008) "Do Household's Water and Sanitation Choices Really Matter for Child Health?" Student Paper. Oxford: Young Lives. [Online] Available from: http://www.younglives.org.uk. [Accessed: 19th June 2013]

Kuehnle, D. (2013). "The Causal Effect of Family Income on Child Health: A Re-Examination Using an Instrumental Variables Approach" Bavarian Graduate Program in Economics (BGPE) Discussion Paper No. 133. [Online] Available from: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2232853. [Accessed: 16th July 2013]

Kumra N. (2008). "An Assessment of the Young Lives Sampling Approach in Andhra Pradesh, India." Young Lives Technical Note No. 2. Oxford: Young Lives. [Online] Available from: http://www.younglives.org.uk. [Accessed: 19th June 2013]

Kynch, J. & Maguire, M. M. (1998) "Food and human growth in Palanpur." In Lanjouw, P. & Stern, N. (eds). Economic Development in Planpur over Five Decades. New Delhi: Oxford University Press.

LeVine, R. A., LeVine, S. E. & Schnell B. (2001). ""Improve the women": Mass schooling, female literacy, and worldwide social change." Harvard Educational Review. 71(1). p. 1-50.

Mendez, M. A. & Adair, L. S. (1999). "Severity and timing of stunting in the first two years of life affect performance on cognitive tests in late childhood." Journal of Nutrition. 129(8). p. 1555-1562.

Moestue H. (2005). "Adult Education and Child Nutrition in India and Vietnam: The Role of Family, Neighbours and Friends." Student Paper. Oxford: Young Lives. [Online] Available from: http://www.younglives.org.uk. [Accessed: 19th June 2013]

Moestue, H. (1996). "The role of education among middle-class women in Sivakasi, South India - a qualitative study." Undergraduate thesis. Geography Department, University of Durham.

Murthi, M., Guio, A-C. & Dreze J. (1995). "Mortality, Fertility and Gender Bias in India: A District-Level Analysis." Population Council. 21(4). p. 745-782.

Pells, K. (2011). "Poverty, Risk and Families' Responses: Evidence from Young Lives." Policy Paper 4. Oxford: Young Lives. [Online] Available from: http://www.younglives.org.uk. [Accessed: 19th June 2013]

Qian N. (2009). "Quantity-Quality and the One Child Policy: The Only-Child Disadvantage in School Enrollment in Rural China," NBER Working Papers 14973, National Bureau of Economic Research. [Online] Available from: http://www.nber.org/papers/w14973. [Accessed: 07th June 2013]

Ramalungaswami, V., Jonsson U. & Rohde J. (2003). "Commentary: The Asian Enigma." Raising Capable Children, A Symposium on Early Childhood Development. UNICEF.

Reed, B. A., Habicht, J. P. & Niameogo C. (1996). "The effects of maternal education on child nutritional status depend on socio-environmental conditions." International Journal of Epidemiology. 25(3). p. 585-592.

Sanchez, A. (2009). "Early Nutrition and Cognitive Achievement in Pre-School Children in Peru." Working Paper 57. Oxford: Young Lives. [Online] Available from: http://www.younglives.org.uk. [Accessed: 19th June 2013]

Sarmistha P. (1999). "An Analysis of Childhood Malnutrition in Rural India: Role of Gender, Income and Other Household Characteristics." World Development 27(7). p. 1151-1171.

Shetty, P. S. (2002). "Nutrition transition in India." Public Health Nutrition. 5(1A). p. 175-182.

Smith, L. C. & Haddad, L. (1999). "Explaining child malnutrition in developing countries: a cross-country analysis." International Food Policy Research Institute (IFPRI), Food Consumption and Nutrition Division Discussion Paper 60. [Online] Available from: http://www.ifpri.org/sites/default/files/pubs/pubs/abstract/111/rr111.pdf. [Accessed: 22nd June 2013]

Strauss, J. & Thomas, D. (1998). "Health, Nutrition, and Economic Development." Journal of Economic Literature. 36(2). p. 766-817.

Thang, N. M. & Popkin, B. (2003). "Child malnutrition in Vietnam and its transition in an era of economic growth." Journal of Human Nutrition and Dietetics. 16(4). p. 233-244.

The World Bank. (2013). World DataBank, World Development Indicators. [Online] Available from: http://databank.worldbank.org/data/views/reports/tableview.aspx. [Accessed: 02nd August 2013]

Thomas, D., Strauss, J. & Henriques, M-H. (1991). "How does mother's education affect child height?" Journal of Human Resources. 26(2). p. 183-211.

UNICEF. (2013). "Improving Child Nutrition. The achievable imperatives for global progress." New York: United Nations Publications.

Waterlow, J. C., Buzina, R., Keller, W., Lane, J. M., Nichaman, M. Z. & Tanner J. M. (1977). "The presentation and use of height and weight data for comparing the nutritional status of groups of children under the age of 10 years." Bulletin of the World Health Organization. 55(4). p. 489-498.

WHO. (1986). "Use and interpretation of anthropometric indicators of nutritional status." Bulletin of the World Health Organization. 64. p. 929-941.

Wooldridge, J. (2002). "Econometric Analysis of Cross Section and Panel Data." Cambridge, Massachusetts: MIT Press.

Young Lives. (2008). "Young Lives: Andhra Pradesh Round 2 Survey." Young Lives Country Report September 2008. Oxford: Young Lives. [Online] Available from: http://www.younglives.org.uk/publications/CR. [Accessed: 19th June 2013]