The Height Production Function from Birth to Early Adulthood*

Elisabetta De Cao†

University of Rome "Tor Vergata"

CEIS Working Paper No. 165

Abstract

In this paper I specify a height production function in order to study the determinants of height from birth to early adulthood in the Philippines. I use a rich longitudinal data set on Filipino children born in 1983 and followed for more than 20 years. The structure of the production function allows height to be the result of the accumulation of inputs over time. The results show that inputs from conception to birth are relevant at each age of the children. Nutrition inputs have a positive but small effect on the child's height. The shorter the distance between the age when the nutrition input is applied and the age when height is measured, the higher the impact on height. The older the child, the bigger the impact with a peak during the growth spurt. The earlier disease inputs are experienced, the stronger their negative effect on height. The older the child, the stronger the effects of past diseases.

*I would like to thank my supervisor, Franco Peracchi for his patience and excellent advice. I thank Jere Behrman for his precious comments. I also thank Kenneth Wolpin, Petra Todd and Flavio Cunha for their helpful suggestions. I thank Linda Adair and the Cebu team for the data and constant help. Special thanks to Nicola Barban for his continous support and encouragement. I also thank the empirical micro club participants at the University of Pennsylvania (academic year 2008-2009), in particular Shalini Roy, Nirav Mehta and Seth Richards, for their valuable help. I really thank Lara Patricio Tavares, Melody Garcia and Tania De Renzis for correcting my English in the paper. All errors are my own.

†Department SEFEMEQ, University of Rome "Tor Vergata," Italy. Faculty of Economics; Address: Via Columbia 2, I - 00133 Roma, Mobile: +39 349 0071191, Tel: +39 06 7259 5624, Fax: +39 06 2040 219, Email: elisabetta.decao@gmail.com.

1 Introduction

Starting in the 1970s, anthropometric measures have increasingly been used in the social sciences as indicators of social well-being. Since then, adult height has been considered an indicator of the general health status in life, of the relative risk of survival, and labor productivity (Fogel, 1986). The early childhood investments, basically early nutrition and childhood living conditions, affect the growth of the person (Schultz, 2002). In particular, height-for-age reflects the accumulation of past outcomes, and thus it is a long-run measure of nutritional and health status (Falker and Tanner, 1986). Adult height may also predict late-life morbidity and mortality through health in childhood (Bozzoli et al., 2009), and also mental health at older ages (Case and Paxson, 2008a). Finally, adult height has been found to be positively correlated with earnings and labor productivity. In particular, Case and Paxson (2008b) explain this result by showing that height is positively associated with cognitive ability. Using data from the United States and the United Kingdom, they document that taller children have higher average cognitive test scores and that these test scores explain a large portion of the height premium in earnings. This is due to the fact that both cognition and height are driven by early childhood investments and therefore cognitive achievements are correlated with height, and wages are affected by cognitive skills (Strauss and Thomas, 2008).[1] Poor health can explain both low height and low labor productivity. This is more evident in developing countries where living conditions are poor.[2] According to the World Health Organization, in 2009, at a global level, about 178 million children are stunted, resulting from diseases, and a lack of food, vitamins and minerals. As growth slows down, brain development lags and stunted children learn less. Stunting rates among children are highest in Africa and Asia. In south-central Asia 41% of children are affected.

It is therefore necessary to investigate the factors driving height, since understanding the determinants of height is important in order to understand health (Deaton, 2007). The determinants of height can be divided into non-genetic fac-

[1] Maluccio, Hoddinott, Behrman, Quisumbing, Martorell, and Stein (2009) demonstrate that early childhood nutrition is causally linked to intellectual human capital in adulthood in the form of cognitive skills. Hoddinott et al. (2008) show that improvements in nutrition also have significant effects on adult male wage rates. Cunha and Heckman (2007) demonstrate that early-life environment not only plays a prominent role in the development of cognitive skills, but it also affects the non-cognitive skills formation.

[2] The fact that height increases wages or productivity is reported in: Immink and Viteri (1981); Behrman and Deolalikar (1989); Haddad and Bouis (1991); Steckel (1995); Thomas and Strauss (1997, 1998); Croppenstedt and Muller (2000); Schultz (2002, 2003); Dinda and Gangopadhyay (2006). Behrman et al. (2009b) report that once the endogenous choice of height is controlled, its significance seems to disappear. They claim that this could be due to the occupation composition in the sample, where only a small portion of the jobs appear to be physically demanding.

tors, genetic factors and the age when height is measured. The principal non-genetic factor is net nutrition. Net nutrition is the difference between food intake and the losses to activities and to diseases, most frequently diarrheal diseases, but also fevers or respiratory infections (Bozzoli et al., 2009). In developed countries there is evidence that genetic factors explain 80% of the variation in adult height and the rest is due to non-genetic factors. The proportion of the variation due to genetics seems to be less important when environmental stress is strong, for example in developing countries (Silventoinen, 2003). The age when height is measured must also be considered. In fact, most of the common historical and modern evidence looks at adult height (Steckel, 2009). However, human growth spans approximately 20 years, from conception to maturity, with the critical periods of infancy and adolescence in between.[3] Adult height is just the final result of a not-well-known process of growth that involves many different mechanisms and variables. The economic literature on height lacks studies in which individuals are followed from conception to maturity.

In this paper I study the determinants of height from birth to maturity, covering the entire human growth path. To do that I build and estimate a height production function. I use data from a developing country: the Cebu Longitudinal Health and Nutrition Survey (CLHNS), which is a rich longitudinal survey of a cohort of Filipino children followed from conception, in 1983-84, to 2005. This paper contributes to three areas. The first is the area of studies on the determinants of height in developing countries, the second regards health production functions, and the third focuses on models of human growth.

The determinants of height have been studied in several different fields, such as medicine, biology, demography and economics. In the economic literature on height in poor countries, many papers look at the effects on height of socio-economic factors such as maternal education, income, poverty, child labor, political oppression or schooling.[4] My paper, instead, looks at the direct biological determinants of height carefully considering different estimation problems that arise when trying to estimate the causal effects of these determinants or inputs on health. Many studies have focused on the effects of childhood health and nutrition on adult height at the population level (Deaton, 2007; Akachi and Canning, 2007), confirming that birth and adolescence are critical periods. However, the evolution of height should be analyzed in an individual longitudinal framework, as this paper does, to explain more clearly all of the mechanisms behind a person's growth.

The second area of studies focuses on health production functions, and two

[3]Comparing the same population over time, the age at which adult height is attained varies (A'Hearn et al., 2009). This is also true between developed and developing countries. Deaton (2008) notes that in India, people reach their adult height only in their twenties, several years later than in contemporary rich countries.

[4]See Steckel (2009) and Silventoinen (2003) for the complete references.

papers, in particular, use the same longitudinal data from the Philippines. The first is by Cebu-Study-Team (1992) who estimated a child health production function, analyzing four different outcomes: gestational age, weight, diarrhea and respiratory infection. They find that individual, household and community factors affect the outputs considered. Their paper focuses on the first year of an infant's life. The second paper by Liu, Mroz, and Adair (2009) considers infants from birth to age 24 months. They specify a dynamic optimization model of parents' investment in their children's health and they estimate a set of parents' demand functions for health inputs in conjunction with a set of health production functions for how a child's physiological development responds to these inputs. In this paper, I extend the production function up to when the child is a young adult, but focus on the height outcome only.

To motivate the specification of a height production function, I follow Steckel (2009) reasoning:

> *It is useful to think of the body as a biological machine, which consumes food as fuel - a blend of calories, protein, micronutrients and other ingredients. This machine expends fuel to breathe, keep warm, circulate the blood and so forth, and in physical effort, fighting infection and physical growth. The body's first priority is to survive, and growth stagnates or takes a back seat under conditions of inadequate net nutrition...*

Similar to the production process of a firm, the body can be considered a machine that combines different inputs through a particular technology to produce an output that in this case is height. The reason for estimating a production function is to find the *ceteris paribus* effects of each of the inputs. If I consider caloric intake and diarrhea as two of the inputs, the questions to answer are: "How does an exogenous change in caloric intake, holding all other inputs constant, affect height?" And "How does an exogenous change in diarrhea episodes, holding all other inputs constant, affect height?" And "At which age are those changes more relevant?" My intention is to find the technological parameters that answer the previous questions. Todd and Wolpin (2003) consider different methods for modelling the production function for cognitive skills to account for the fact that child development is a cumulative process depending on the history of family, on school inputs and on innate ability.[5] Todd and Wolpin (2007) apply this general modelling framework to study the determinants of children's scores on tests of

[5]Using a simple model of achievement, Todd and Wolpin (2003) underline the difference in the parameters estimated in non-experimental and experimental studies. In particular, by using non-experimental data it is possible to decompose the net impact that comes through a *ceteris paribus* biological effect of an input *versus* the indirect effects of parents' changing other inputs in response. Nevertheless non-experimental studies cannot by themselves answer questions such as "What is the total effect of an exogenous change in caloric intake on height, not holding all other inputs constant?". This is a direct policy question that looks at the *non-ceteris-paribus*

cognitive achievement in math and reading. Therefore, given that height results from the accumulation of several factors over time, and that the focus of the paper is to find the direct effects of the determinants of height, I apply the same approach as in Todd and Wolpin (2003, 2007) to the case of height formation.

Finally, this paper contributes to the literature on models of human height growth. A model of human height growth is a representation of the height growth curve. Height is the accumulation of height growth increments from birth.[6] The aim of these models is to fit optimally the human height path, usually using only the person's age as a regressor. Many different growth models have been developed since the 1940s. The large number of parameters to estimate, the poor fit in particular parts of the curve and the lack of interpretability with the biological process are known problems associated with these models. Darrell Bock et al. (1973) were the first to estimate a whole growth curve from infancy to maturity. After that, Preece and Baines (1978) developed a family of mathematical functions that better describes the whole growth curve with few parameters. Ten years later, Shohoji and Sasaki (1987) and Jolicoeur et al. (1988) used alternative curves to better fit human growth from birth or origin (egg fertilization) to adulthood. In general, these models do not focus on the determinants of height. Their main purpose is descriptive or predictive. This paper, instead, aims only to identify the separate effects of the determinants of height. The similarity with this class of studies is the temporal window, since I am also looking at the process of height formation from birth to early adulthood.

The paper is organized in six parts. Section 2 presents the data and a detailed description of the variables used. The data allow the derivation of a height production function,[7] including all of the past observed inputs prior to any height measurements. In section 3 I develop a theoretical model for studying the process of height formation. I consider both the child's age when height is measured and the distance between the application of the input and the outcome measure as relevant elements of the model. I present two specifications and their assumptions for empirically estimating the technology parameters of the height production function. The technology used is very simple: a parametric linear model with a quadratic trend for age. The model is introduced after the data, because it strongly depends on the characteristics of the data set used. Section 4 describes the em-

effect caused by changes in inputs. To answer it by using non-experimental data, it is necessary to estimate both the production parameters and the family input decision rules. On the other hand, experiments do not generally estimate production function parameters, but rather policy effects. See Todd and Wolpin (2003) for further details and examples.

[6]In the growth literature there are studies on cross-sectional or individual measurements. The short review here considers the individual growth curves only.

[7]The survey used is an observational study, and it allows the determination of the technological parameters of the height production function.

pirical results. I keep boys and girls separate due to their different growth paths. I also consider two types of height production function: (i) the infants' production function and (ii) the adolescents' production function. In section 5, I present some robustness checks and model specification tests. Finally, some conclusions are drawn in section 6.

2 Data

The data come from a joint project between the University of North Carolina and the University of San Carlos in Cebu, Philippines.[8] The Cebu Longitudinal Health and Nutrition Survey (CLHNS) is a longitudinal survey of a cohort of Filipino women who gave birth between May 1, 1983 and April 30, 1984. A stratified and single stage sampling procedure was used to randomly select 33 communities or barangays from the metropolitan Cebu area. Of them, 17 are urban communities and 16 are rural communities. The baseline survey includes 3327 women who were interviewed during the 6th to 7th month of pregnancy. All pregnant women of the barangay and the births were identified, and 3,080 non twin live births were consequently followed in the survey. Around 2,600 households were analyzed for the first 2 years. The children who were born during that period, their mothers, other caretakers, and selected siblings were followed through subsequent surveys conducted in 1991-2, 1994-5, 1998-9, 2002 and 2005. Apart from those last surveys, bimonthly surveys were conducted after birth for 24 months.

The initial focus of the survey was to collect information about the infants' feeding patterns. Later on, when the children were followed through adolescence and into young adulthood, the objective changed to a longitudinal intergenerational study of health.

It is important to notice that individuals are not surveyed at the same age. The waves of the panel are not evenly spaced. Table 2 reports the children's age at the time of the different follow-ups. For simplicity, later on I will use the range of ages, labeling them with the round mean age. The age ranges underlined with a bullet ($\bullet$) are the waves considered in the adolescent height production function, while for infants, I use the waves underlined with a star ($\star$).

Table 3 reports descriptive statistics for the variables considered. The values reported refer to the children at birth and when they are mean age 1, 2, 8.5 and 18.7 years. The following paragraphs show the advantages of the data, with a description of the variables I am going to use, and the disadvantages. Readers interested in the details of the variables are invited to read Appendix A.

[8] For further information, go to www.cpc.unc.edu.

2.1 Advantages

The two big advantages of this survey are: its length and its contents. There are 18 waves in total, 13 collected during infancy and early childhood and 5 during middle childhood, adolescence and early adulthood. The data spans over 20 years and covers issues such as health, nutrition, water quality and sanitation. It contains detailed information about the mother's health and behavior during pregnancy, such as health care practices or smoking behavior, children's education, household and individual economic situation and job, demographic information, family planning, intra-household relationships, reproductive health and sexual behaviors.

Of special interest for my study is the rich collection of anthropometric measurements from birth to age 22, as well as complete disease and nutrition information. Since the data have information at the individual, household and community levels, it is possible to study the longterm effects of prenatal and early childhood nutrition and health on later adult outcomes, matching physical and socio-economic information.

The *outcome* of this paper is raw height reported in centimeters. Height and weight were measured every two months for the first two years of life, and later during childhood and adolescence by the field staff in Cebu. The measurements were taken by specialists and this is a great advantage compared to the self-reported heights common of many datasets. Reliability checks were made to avoid heaping and other errors in the measurements. The distribution of height by age and sex is shown in Graph 1.

The inputs of the height production function can be divided in inputs from birth to early adulthood and inputs from conception to birth.

Inputs from birth to early adulthood Many medical papers suggest that approximately 60% of height variation in a population depends on genetic factors, but it is not clear which is the underlying process (see, e.g., Ginsburg et al., 1998) nor is the relationship between genetics and environmental factors clear. I use the *mother's height* as a proxy for genetics.[9] I assume that the rest of the genetic impact is captured by an individual's biological endowment included in the model. It represents the genetic inheritance and gene-environment interactions that are unobserved factors (Case and Paxson, 2008b).

Typically, medical studies use nutrients as nutrition factors, but I include

[9]Subramanian et al. (2009) found a negative association between maternal height and poor health in children (in particular, child mortality, anemia and anthropometric failure) in India. This suggests an intergenerational transfer of health from the mother to her offspring. Behrman et al. (2009a) claim that the intergenerational associations of anthropometrics may reflect genetics, but also may reflect non-genetic factors such as maternal early-life nutrition.

caloric intake,[10] which is a good aggregate indicator of nutrition, even if it does not capture the role of micronutrients (Branca and Ferrari, 2002). In many papers the energy intake is an approximation derived from the family food expenditure. In this case it is necessary to consider the household's size to get the individual's energy intake. The CLHNS data, instead, provide precise information about the individual's diet based on 24-hour dietary recalls or quantitative food frequency questionnaire. I consider the calories from nutrients different from the calories in human milk for infants. This is not sufficient to capture the infant's nutrition, therefore, I also consider whether *human milk* was given to the baby and an interaction term between breast milk and caloric intake. For older children and adolescents I use the total daily caloric intake.

In developing countries, there is clear evidence about the effect of diseases on human growth. In fact, some diseases reduce the absorption of nutrients, prevent food intake, produce nutrient losses or increase metabolic requirements (Stephensen, 1999). I distinguish between *infants* and adolescents because *diseases* have a different impact on a person, depending on age. For infants, I consider diarrhea, measles and other important diseases such as dengue fever, chicken pox, TB, primary complex, and worms or other parasites.[11] I use a dummy variable that indicates whether the infant had episodes of diarrhea or other important diseases in the seven days preceding the survey or had measles in the last two months. For the baseline I instead use a dummy variable that indicates if the infant had *feeding problems* in the few hours after birth. For adolescents I need to consider other diseases that affect height. I consider only *chronic diseases*, that is, diseases of long duration and generally slow progression, because they capture the intensity of the sickness. It could be that other reported illnesses are temporary or do not strongly affect height. I create a dummy variable that indicates whether a person suffered at least one of these critical diseases[12]. It is important to note, for example, that the disease dummy when the child is 8 years old indicates the presence of chronic illnesses that may have started years before.

I consider the *location* of the household, and whether the child lives in a rural or urban location as an exogenous input. Rural and urban places may differ in access to stores, population density, access to health care institutions, schools, among many other things. The Cebu-Study-Team (1992) found that the Cebu urban areas are prone to water contamination or to feces exposure, due to population density. However in rural places during heavy downpours, rainwater finds its way into the sewer networks and the springs. Contamination of water contributes to

[10] I am thankful to Linda Adair who provided me with the caloric intake computed by using the Food Composition Table owned by the Food and Nutrition Research Institute in the Philippines.

[11] Bozzoli, Deaton, and Quintana-Domeque (2009) and Akachi and Canning (2007) use infant mortality as a measure of disease in childhood. I prefer to look at the different diseases.

[12] See Appendix A for a complete list of the diseases considered.

the spread of human diseases.

Inputs from conception to birth Many researchers suggest that growth in utero may play an important role in determining health in adult life (Barker, 1998). Living in a poor country cannot be underestimated because of the possible malnutrition of the mother. Undernourishment is due to a low supply of nutrients or to a high demand of a fetus that is growing very fast. This can lead to permanent changes in the metabolism, morphology and physiology of the embryo, and consequently of the infant and of the adult (Barker, 1998). It would be ideal to have all of this information and to try to understand what is really happening between mother and fetus from conception to birth by computing the nutrient intake of the fetus. Because of the complexity of the problem and a lack of data, I use the following variables.

I include in the production function the infant's *birth weight* measured in grams one hour after birth. The quality of the pregnancy relates to the baby's length at birth and, his birth weight (a strong predictor of adult height). Low-birth-weight infants have a higher risk of dying from infectious diseases, or higher risk of developing serious illnesses in later life. The importance of the birth weight is well known and there is a huge literature about it in medicine and economics (e.g., Rosenzweig and Schultz (1983); Behrman and Rosenzweig (2004)).

The problems of prematurity are very similar to those of low birth weight. A baby who is both pre-term and small for his length of gestation[13] may later have deficits in size and ability. I computed the *duration of the gestation* as the difference between the last menstrual period and the date of birth. I include a categorical variable indicating whether the child had normal weight and normal term, low birth weight for his gestational age or simply pre-term but with normal weight for his gestational age.

Birth order has also been found to be a significant and independent predictor of adult height (Steckel, 1995). First-borns children are, during childhood, taller than children born later, since have had a period in which they were alone.

2.2 Disadvantages

The Cebu Longitudinal Health and Nutrition Survey is not a representative sample of the Philippines population. Therefore, the results of this paper are specific of the Cebu context.

The second disadvantage is that the survey spans over 20 years, but there is a lack of follow-ups between age 2 and 8.5 years of the children. More than 6 years

[13]The normal length of gestation is from 37 to 42 weeks.

of early childhood are missing, and this can be critical from the point of view of the empirical analysis of this paper.

Furthermore, the data, even if extremely rich, lacks important information such as the father's height. The energy intake for infants is exclusive of breast milk, since the Cebu Team was not able to compute calories from breast milk, given the complexity and almost impossibility of doing that.[14] The calories from nutrients different from the calories in human milk for infants and the breast milk can be bad proxies for the infant's nutrition. Suppose, for example, that the infant is exclusively breast fed, then the caloric intake from other nutrients is null, and the breast milk dummy is equal to one. But since, especially in the first year of life, almost all of the babies are exclusively breast feed, there is not much variability in the nutrition across infants, given the proxies used. Besides that, malnourished mothers may be unable to produce sufficient quantities of good quality milk to satisfy the baby's needs (Scott and Duncan, 2002). In this case, being breast feed may result negative for the growth of the kid.

The variables indicated as inputs from conception to birth are not exactly inputs, but the results of pre-birth inputs that are not available (e.g., birth weight). As explained before it is not possible to know the fetus' nutrition and diseases inputs. In any case, to exclude this phase of life would be an error given its well-known importance for a person's growth.

3 Theoretical model and empirical specifications

In this section I present a model for the height production function. This model is strictly related to the papers of Todd and Wolpin (2003, 2007). They estimate a model by considering different specifications of a skill production function for children. I modify their specifications by adapting them to the height formation study and to my data limitations. The similarity of the two studies is that the processes of both height formation and achievement are cumulative processes that depend on the history of inputs chosen by the families, are due to the environment or are simply inherited genetically.

I develop a height production function, since I am interested in technological parameters such as the effect of an exogenous change in one input, keeping all others constant. The technology that links inputs and output is fixed. It is created by nature and cannot be controlled. Economic agents play a negligible role in

[14]The composition of human milk is not uniform and changes during the course of lactation. The greatest change in composition occurs during the first 10 days post-partum when colostrum changes to mature milk. In addition, the calories contained in breast milk depend on the mother, on her's nutrition, on the quantity of breast milk the infant sucks, and on the time of day and it also changes during a single feed (Scott and Duncan, 2002).

choosing some of the inputs. The inputs they can choose are nutrition and diseases in the sense of prevention of diseases. They cannot choose either the age or the timing of children's growth.

It is widely known that height depends on the current age and on past inputs, such as health care practices, nutrient intake, disease incidence and genetic factors. A person's height is therefore a cumulative indicator because growth is a cumulative process by which past inputs and genetic endowment are combined in order to obtain height.[15]

Let me define the height production function that relates the height measured at age t to all previous investments in the child. Suppose that for $t = 0, \ldots, T$ and $i = 1, \ldots, N$ I have:

- H_{it} the observed height for child i at age t,
- $f(t)$ an age trend,
- $\mathcal{X}_{i,t-1}$=$(X_{it-1}, X_{it-2}, \ldots, X_{i1}, X_{i0})$ the vector of the observed inputs for child i from birth to age $t-1$,
- $\mathcal{V}_{i,t-1}$=$(V_{it-1}, V_{it-2}, \ldots, V_{i1}, V_{i0})$ the vector of the unobserved inputs for child i from birth to age $t-1$,
- μ_i the child's biological endowment,
- ϵ_{it} the measurement error for child i at age t.

Then the **height production function** is given by:

$$H_{it} = h_t[f(t), \mathcal{X}_{i,t-1}, \mathcal{V}_{i,t-1}, \mu_i, \epsilon_{it}]$$

The age trend is necessary because the child's age at each wave is different, and if I consider two height measurements for the same individual, their difference also depends on the ages. I use a linear and a quadratic term. [16]

The inputs $\mathcal{X}_{i,t-1}$ are nutrition and diseases. They enter the production function with a lag because I assume that the body needs time to absorb them. Thus I also assume that contemporaneous inputs are not appropriate for this study. I am assuming, for example, that the caloric intake of today does not affect height today, but height depends on past consumption of calories. This temporal lag depends on the structure of the data, in particular, on the distance between waves.

[15] Cuff (2004) writes that "adult stature is a cumulative indicator of net nutritional status over the growth years, and thus reflects command over food and access to healthful surroundings."

[16] I also tried a cubic splines term but there were not huge differences.

The child's biological endowment is determined at conception and it is constant over time.[17] Being unable to identify the inputs from conception to birth I use birth weight and other variables.[18] Those variables are, themselves, the results of different inputs and an endowment determined at conception.[19]

Finally, the measurement error ϵ_{it} is included because there can be errors in the height measurement, even when made by specialists, or measurement errors in the inputs.

To study empirically the height production function, I assume that it is linear in the inputs and in the unobserved endowment, and that the effects of the inputs depend both on the child's age and on the distance between the application of inputs and the height measurement. The true technology that links inputs and output is unknown, therefore I have chosen the simplest functional form.[20] This functional form implies perfect substitutability between inputs, and this is a strong assumption. There is not complementarity and this means that all investments should be concentrated in one period, during the high-return period, and no investments should be made when the returns are low. This is in line with most biomedical and epidemiological studies in the "early influences" literature. They show that investments in early childhood produce effects on adult outcomes. But the effects may be bigger as individuals age, because the child's development is divided in different stages that have different influences on the adult outcomes. It also seems plausible that there should be interactions among inputs, but their inclusion in the model is empirically intractable due to the limited number of observations. Hence, I obtain the following model:

$$\begin{aligned} H_{it} &= f(t) + X_{it-1}\beta_{t,t-1} + X_{it-2}\beta_{t,t-2} + \cdots + X_{i1}\beta_{t,1} + X_{i0}\beta_{t,0} + \\ &\quad + V_{it-1}\rho_{t,t-1} + V_{it-2}\rho_{t,t-2} + \cdots + V_{i1}\rho_{t,1} + V_{i0}\rho_{t,0} + \\ &\quad + \mu_i + \epsilon_{it}, \end{aligned} \tag{1}$$

[17] Case and Paxson (2008b) hypothesize an endowment determined at birth that changes according to the child's age. But their time-invariant individual effect also includes the environmental factors that in my study are observed and considered as further regressors in the model. Furthermore, I suppose that the gene-environment interactions are the same for each age of the child.

[18] See section 3.3.

[19] Suppose that height is given by: $h_{it} = \beta X + \gamma Y + \mu_i + \epsilon_{it}$. where X are the post-birth inputs, Y the pre-birth inputs and μ_i the endowment at conception. Now consider the birth weight B given by: $B_i = \gamma^* Y + \mu_i + \xi_i$. Substituting Y into the height equation I get:

$$h_{it} = \beta X + \gamma \left(\frac{B_i}{\gamma^*} - \frac{\mu_i}{\gamma^*} - \frac{\xi_i}{\gamma^*} \right) + \mu_i + \epsilon_{it}$$

Therefore, in my estimations, I am implicitly assuming that $\gamma = \gamma^*$, and the birthweight is perfectly capturing the pre-birth inputs that are unobserved.

[20] The results are not sensitive when I instead use a *log-linear* model.

and for simplicity I define:

$$e_{it} = V_{it-1}\rho_{t,t-1} + V_{it-2}\rho_{t,t-2} + \cdots + V_{i1}\rho_{t,1} + V_{i0}\rho_{t,0} + \epsilon_{it}.$$

The time-varying coefficient $\beta_{t,x}$ depends on the age t of the child and the distance x between the time the inputs were applied and the time of the height measurement. Consider for example the child's height at age $t = 8$. The effect of the diseases when the child is 5 years old may be different from the effect when the child is 2 years old. This is due to the different distances x ($\beta_{8,5} \neq \beta_{8,2}$). The effect of the diseases when the child is 5 years old on height at age 8 may be different than the effects on height at age 6. This is instead due to the different time of the height measures ($\beta_{8,5} \neq \beta_{6,5}$). Since I am interested in the effects of the inputs at different ages of the child and at different distances between the application of the inputs and the height measurement, I create an age dummy for each range of ages. The fourth column of Table 2 reports the mean ages that correspond to the range of ages I considered.[21]

Due to the structure of the data and the different growth patterns for boys and girls,[22] and for infants and older children, I estimate the production function for different subsamples. I analyze boys and girls separately and I also split-up "infants" and "adolescents." "Infants" are children between birth and age 2 years, while "adolescents" are people observed between ages 8 and 20 years approximately.[23]

The curse of dimensionality is an important issue in this model. Suppose that I had information at each point in time. In that case the number of parameters to estimate would be equal to $T * (T-1)/2$ (pooling all the observations). If I use all the waves available (18) the number of β's become $(18 * 17)/2 = 153$ plus the parameters of additional variables (ex: age, age squared, etc.). Table 1 shows the general structure of the β's parameters. In the empirical application I consider each child's age and I estimate as many production functions as the child's ages considered.

3.1 Different specifications

Behrman and Deolalikar (1988) summarize the most common estimation problems

[21] I do observe people at mean age 22 years, but I do not include them in the working sample, because most of them have already completed their height growth.

[22] At birth the typical boy grows faster than the typical girl, but the velocities become equal around 7 months and then girls grow faster until age 4. There are no differences until they reach adolescence. The typical girl is slightly shorter than the typical boy at all ages until adolescence. She is taller during her adolescence spurt because it takes place two years before the male spurt (Tanner, 1990).

[23] The labels are not technically correct since an infant is between 1 month and 12 months old, while an adolescent is usually between 13 and 19 years old.

in health and nutrition studies in developing countries, and I report the most relevant for this study. The first problem is endogeneity; e.g., the inputs and the outcome are determined simultaneously. The second problem regards the omitted variables. The exclusion of important variables can give biased estimates. A third problem is sample selectivity. This arises when the people in the sample are selected. For example, the weakest children die and thus only the strongest individuals stay in the sample. A fourth problem is the errors-in-variables problem. A fifth problem is the appropriate lag structure for the variables included in the model. Using only current data may not be sufficient when the effect of past variables are relevant. The CLHNS data contain detailed information on past inputs, and including them makes a considerable contribution to the existing health production functions.

In order to estimate (1), I consider two different specifications based on two estimation methods. They rely on different assumptions aimed at solving the open issues listed above.

Cebu-Study-Team (1992) and Liu, Mroz, and Adair (2009) use another specification that includes lagged values of the outcome and contemporaneous inputs in the model instead of the historical inputs. In the cognitive skills literature this specification is called "value added" specification. When data on past inputs are missing, the use of lagged outcome is quite common. However, the lagged outcome is correlated with the measurement error by construction, and additional lagged outcome measures can be used as instruments. Since past inputs are available, and I am specifically interested in their impact on height, I do not consider this specification. A lagged measure of height would capture almost all of the variability and it would not allow me to distinguish between the effects of nutrition and non-nutrition inputs.

3.1.1 Naive specification

My first specification imposes the following set of assumptions:

(a) Included inputs are not correlated with endowment

$$E[\mu_i|X_{it-1}, X_{it-2}, \ldots, X_{i0}] = 0.$$

(b) The $\mathcal{X}_{i,t-1}$ are strictly exogenous conditional on the unobserved effect μ_i

$$E[\epsilon_{it}|X_{i0}, \ldots, X_{it-1}, X_{it}, X_{it+1}, \ldots, X_{iT}, \mu_i] = 0. \qquad t = 0, 1, \ldots, T$$

(c) Omitted inputs are orthogonal to the included inputs

$$E[V_{it-1}, V_{it-2}, \ldots, V_{i0}|X_{it-1}, X_{it-2}, \ldots, X_{i0}] = 0.$$

These assumptions allow me to estimate the β's by OLS. This specification is called naive because it imposes very strong restrictions on the production function (1).

Let me analyze these assumptions one at a time:

Assumption (a) fails if the choice of inputs is correlated with the endowment. It is plausible that investments in children are correlated with the unobserved endowment.

Assumption (b) fails in the case of endogenous inputs. If you think that parents (or the child when he is old enough) choose the inputs to improve health outcomes according to a dynamic or sequential process, then the exogeneity of the measurement error fails. The reason is that these choices are likely to be based on past outcomes or past shocks. Therefore, assumption (b) imposes that later input choices are invariant to prior own height.[24]

Assumption (c) fails if there are omitted variables that are correlated with the observed inputs. Suppose, for example, that the protein intake is unobserved. Protein intake is likely to be correlated with caloric intake, which is an observed input.

3.1.2 Within-child fixed effects specification

This specification is feasible because the children are observed more than once, and it is possible to rely on several outcome and input measurements.[25] Consider the technology (1) at two different ages, t and s, with $t > s$:

$$\begin{aligned} H_{it} &= f(t) + X_{it-1}\beta_{t,t-1} + X_{it-2}\beta_{t,t-2} + \cdots + X_{is+1}\beta_{t,s+1} \\ &\quad + X_{is}\beta_{t,s} + X_{is-1}\beta_{t,s-1} + \cdots + X_{i1}\beta_{t,1} + X_{i0}\beta_{t,0} + \mu_i + e_{it} \\ H_{is} &= f(s) + X_{is-1}\beta_{s,s-1} + X_{is-2}\beta_{s,s-2} + \cdots + X_{i1}\beta_{s,1} + X_{i0}\beta_{s,0} + \mu_i + e_{is} \end{aligned}$$

Grouping the inputs applied at the same age and differentiating gives:

$$\begin{aligned} H_{it} - H_{is} &= f(t) - f(s) + X_{it-1}\beta_{t,t-1} + X_{it-2}\beta_{t,t-2} + \cdots + X_{is+1}\beta_{t,s+1} \\ &\quad + X_{is}\beta_{t,s} + X_{is-1}(\beta_{t,s-1} - \beta_{s,s-1}) + \cdots + X_{i1}(\beta_{t,1} - \beta_{s,1}) \\ &\quad + X_{i0}(\beta_{t,0} - \beta_{s,0}) + e_{it} - e_{is} \end{aligned} \tag{2}$$

[24] The strict exogeneity assumption can also be stated in terms of H_{it} as:

$$E[H_{it}|X_{i0}, \ldots, X_{it-1}, X_{it}, X_{it+1}, \ldots, X_{iT}, \mu_i] = X_{it-1}\beta_{t,t-1} + \cdots + X_{i0}\beta_{t,0} + E[\mu_i|X_{it-1}, X_{it-2}, \ldots, X_{i0}],$$

and this implies that both contemporaneous and future inputs are not correlated with the current height.

[25] A within family specification would be interesting, but the data contain anthropometric measurements of some siblings but not all the information about siblings' net nutrition.

The parameters resulting from equation (2) are the specific age input effects for the inputs applied between the two ages and the difference in the parameters for the inputs applied before the younger height H_{is}.

Consider the following example. Suppose that I consider a child's height at 2 and 3 years and I group inputs applied at the same age. Suppose the information is collected once a year. The difference in height between the two ages becomes:

$$H_{i3} - H_{i2} = f(3) - f(2) + X_{i2}\beta_{3,2} + X_{i1}(\beta_{3,1} - \beta_{2,1}) + X_{i0}(\beta_{3,0} - \beta_{2,0}) + e_{i3} - e_{i2}$$

The parameters in this equation are the specific age input effect $\beta_{3,2}$ for the inputs applied at age 2 and the difference in the parameters $\beta_{3,1} - \beta_{2,1}$ and $\beta_{3,0} - \beta_{2,0}$ for the inputs applied at ages 1 and 0, respectively.

The coefficients associated with the inputs applied before the younger height cannot be identified.[26] In the empirical application I always consider the first age $s = 0$ and therefore $H_{is} = H_{i0}$ is the height at birth. Doing so, I can identify all the parameters of the inputs applied between the two ages and I can compare them with the parameters estimated using the naive specification.

This second specification imposes the following set of assumptions:

(a) Included inputs are not correlated with the endowment: *not need be assumed.*

(b) The $\mathcal{X}_{i,t-1}$ are strictly exogenous conditional on the unobserved effect μ_i

(c) Differenced omitted inputs are orthogonal to the differenced included inputs.

These assumptions allow me to estimate the β's by fixed effects (FE). I explain them as follow:

Assumption (a) needs not be assumed because the within-child fixed effect estimator eliminates the endowment from equation (2). This estimator deals with the endowment heterogeneity.

Assumption (b) deals with the potential endogeneity of the inputs. The endogeneity can be due to feedback effects. The fixed effects allow a permanent change in the inputs, if the kid is born very small, then it is plausible that the parents will always give him more food, and this goes into μ_i. While, for example, if a child results very small certain point in time, and the parents will give him more food to help his growth, then this is not captured by the fixed effect and produces endogeneity. Using an instrumental variables approach, it could be possible to estimate the input demand equations for the endogenous inputs. I tried to use different sets of instruments as robustness checks.

Assumption (c) means that omitted inputs are constant over time and by the fixed effect estimators they are eliminated.

[26] This includes also the inputs applied from conception to birth for which I use proxies, for example, birth weight. Birth weight is a time-invariant variable, and it is not identified from the within-child estimator.

4 Empirical results

I estimate infants and adolescents' height production functions. The effects of the different inputs change according to the age and sex of the child. Growth in utero, infancy and the pre-puberty years turn out to be critical stages in the process of height formation. In general, the results show that inputs from conception to birth are relevant and quite constant at each age of the children. For example, an increase of the birth weight by 100 g increases height by 0.2-0.3 cm in a kid 2 years old or 18 years old; or having one older sibling decreases height by 0.2-0.3 cm in boys, and 0.1-0.3 cm in girls. The highest negative variation in boys'height is due to diseases experienced in early life. The earlier disease inputs are experienced, the stronger their negative effect on height. The older the child, the higher the reduction in height because of past sickness: for example, -0.7 cm if a boy is one year old, but -5 cm if the child is 11.5 years old. While, diseases experienced during infancy especially matter in a girl's height when she is younger than 1 year, but not later on. Nutrition inputs have a positive but small effect on the child's height. In most of the cases, the shorter the distance between the age when the nutrition input is applied and the age when height is measured, the higher the impact on height. The older the child, the bigger the impact, even if, nutrition effects start to decline after the adolescent's growth spurt. Consider a 11-year-old child. If the caloric intake, applied at age 8 years, increases by 100 kcal, then height increases by 0.2 cm. Looking at a 15-year-old child, an increase of 100 kcal, applied at age 11 years, produces an increase in height of about 0.15 cm.[27]

4.1 Estimates of infant height production function

To estimate the infant production function I use only four monthly spaced waves, starting from birth using information at the mean ages 0, 4, 8, 12, 16, 20 and 24 months. This reduces the number of parameters to include in the specification and enlarges the temporal window between two measurements from 2 months to 4 months.

The empirical model for the infants includes inputs from conception to birth and from birth to the age before the age when height is measured. Net nutrition for the infants consists of diseases, caloric intake, breast milk and breast milk times caloric intake. They all have the same lagged structure. I include the interaction term because the caloric intake of the infants excludes the caloric intake of breast milk. Table 4 presents the parameters of the production function for the caloric intake with the relative age of the child and the age at which inputs are applied.

[27]Unfortunately, there are not other estimates available for other countries to allow the comparison of my results.

The number of parameters depends both on the distance between the outcome and the input, and on the child's age. The calorie intake, breast feeding and disease inputs enter the model multiplied by the child's age dummies.

The other covariates are mother's height, location, birth weight, length of the pregnancy and birth order. They are time-invariant variables, constant for the entire life of the child, except location, which is a contemporaneous time-varying input. None of them are interacted with the child's age.

In the naive specification I estimate the infant height production function separately foreach age of the child.[28] Doing so, is like working with a cross-section data because I observe the infant in a specific range of ages only once. This allows me to get specific parameters that account both for the infant's age and for the distance between height measurement and application of the inputs.

I then use the within-child fixed effects specification. To get the 18 parameters reported in Table 4 using the within-child fixed effects specification, I estimate the fixed effect estimator on a couple of ages, where one of the two is always the height at birth. Hence I am able to identify the effects of inputs at specific ages for each age of the infant and I can compare them with the naive estimates.[29]

Tables 6 and 7 show the results of the naive and fixed effects specifications separately for boys and girls. I report the production functions only for infants who are 12 and 24 months old.[30] Many coefficients are not significant, but nutrition coefficients result jointly significant, and also diseases inputs coefficients.

The *inputs from conception to birth* are time-constant variables and therefore are dropped in the fixed effects specification. The naive specification shows that birth weight and birth order are always significant for each age and sex of the infant. An increase of 100g in birth weight produces an increase in the infant's height of about 0.2 cm at all ages, with slight variations between boys and girls. The older the infant, the bigger the negative impact of the parity. If a boy is 12

[28] I do not pool all of the observations due to the large number of parameters that need to be estimated.

[29] I run the following regressions where I report only the inputs time-varying (in $\alpha*$ there are trends and time-varying variables that enter only at the current time e.g.location):

$$\begin{aligned}
H_{24} - H_0 &= \alpha^{24} + X_{i20}\beta_{24,20} + X_{i16}\beta_{24,16} + \cdots + X_{i4}\beta_{24,4} + X_{i0}\beta_{24,0} \\
H_{20} - H_0 &= \alpha^{20} + X_{i16}\beta_{20,16} + \cdots + X_{i4}\beta_{20,4} + X_{i0}\beta_{20,0} \\
H_{16} - H_0 &= \alpha^{16} + X_{i12}\beta_{16,12} + \cdots + X_{i4}\beta_{16,4} + X_{i0}\beta_{16,0} \\
H_{12} - H_0 &= \alpha^{12} + X_{i8}\beta_{12,8} + \cdots + X_{i4}\beta_{12,4} + X_{i0}\beta_{12,0} \\
H_8 - H_0 &= \alpha^{8} + X_{i4}\beta_{8,4} + X_{i0}\beta_{8,0} \\
H_4 - H_0 &= \alpha^{4} + X_{i0}\beta_{4,0}
\end{aligned}$$

[30] The height production function estimates for infants 4, 8, 16 and 20 months old are available upon request.

months old, having one older sibling reduces his height by 0.26 and by 0.34 if he is 24 months old (in the naive specification). In girls the effect is smaller. The fact of being premature or premature and small-for-dates with respect to a normal baby does not appear to affect infant's height, probably because the birth weigh captures almost the entire effect.

The *mother's height* is always positive and significant. An increase of 1 cm in mother's height produces an increase in the infant's height of about 0.18 cm if the infant is a 24-month-old boy and of about 0.16 cm if the infant is a girl. The younger the infant, the lower the effect of mother's height.

Living in an *urban area* instead of in a rural one does not seem to have a significant impact on the infant's height. *Nutrition* has a positive and significant effect on height if applied at a short distance from the time of height measurement. The older the infant, the bigger the effect of the nutrition inputs. Nutrition inputs do not have a relevant effect on girls' height when they are 12 months old (or younger). This could be to the fact that there are not good measures of the quality of the breast milk,in most of the cases is exclusively given to the infants, and an indicator if the infant is breast fed or not might not be sufficient. If a 12-month-old boy was breast-fed in the previous months, that increases their height by 0.8 cm (Table 6). Between 12 and 24 months, caloric intake and breast milk have positive effects on both boys' and girls' growth. Consider a 24-month-old boy and that the nutrition inputs are applied at age 20 months (Table 7). If the caloric intake increases by 100 kcal and the boy is not breast-fed, then height increases by 0.09 cm (0.1 cm in the fixed effects specification); if he is also breastfed height increases by 0.096cm+1.576cm-0.170cm=1.502 cm (0.901 cm in the fixed effects specification). When the individual's endowment is eliminated by using the fixed effect, the total nutrition effect falls from 1.502 cm to 0.901 cm. If the boy is breastfed height increases by 1.576 cm in the naive specification and only 0.951 cm (not significant) in the fixed effect specification. This can be explained by the fact that stronger boys, as captured by a higher endowment μ_i, are more likely to be breast fed. In the ols estimation, a large part of the breast feeding premium reflects the fact that boys who are breastfed would increase more in height even if they were not breastfed.

Diseases always have a significant and negative impact on infants' height when the diseases occur at early ages: at birth or when the infant is 4 months old. The only exception is when the girls is 24 months old, and diseases experienced recently, at age 20 months, have an higher impact on height. Experiencing at least one disease[31] decreases boys' height by 0.6-1.6 cm, and girls' height by 0.5-2.3 cm depending on the specification and infant's age[32]. The older the boy, the more

[31] At least one of the diseases listed in the previous section.

[32] These values come from all of the infant's production functions, even if I report the produc-

he is affected by sicknesses. The younger the girl, the more she is affected by sicknesses.

4.2 Estimates of adolescent height production function

In the data used to estimate the adolescent height production function, the waves are no longer approximately equally spaced. There are intervals of about 3-4 years and a jump of 6 years between infancy and middle childhood that starts approximately at the age of 8. Table 2 reports underlined with a bullet the ranges of ages considered. The oldest people are adolescents, mean age 18.7 years. Most of them have completed their growth and therefore their height production function gives the complete picture of what determines their height from birth to early adulthood.

The variables entering the production function are inputs from conception to birth, net nutrition, location and mother's height. Table 5 reports the parameters estimated for caloric intake. According to the time when the input was applied, the variables included with the lagged structure are caloric intake for nutrients different from breast milk, breast milk, their interaction term, total caloric intake, infants' diseases and adolescents' diseases. Those variables are multiplied by the age dummies to account for the different child's age and distance between height measurement and application of the input.

As for the infants height production function, in order to apply the within-child fixed effects estimator I pair ages, with one always the height at birth.

Tables 8, 9, 10 and 11 report the results separately for gender and specifications. They represent the height production functions for people mean age $t = (8.5, 11.5, 15.5, 18.7)$. The inputs considered are all of the past inputs from birth to the previous height measurement age.

The *inputs from conception to birth* are still significant. An increase in birth weight of 100g produces an increase of about 0.25-0.3cm in height, with small variations between boys and girls. This is true for all adolescents analyzed. A unit increase of the birth order reduces the adolescent's height by about 0.2-0.3cm, according to age and sex of the person. Being pure pre-term has a positive effect on height in boys 8.5, 11.5 and 15.5 years old, which is unexpected, but it is probably due to a strong catch-up.[33] *Mother's height* is always positive and significant with bigger coefficients than in the infant production function. The older the child, the stronger the relation between his height and his mother's height.

Nutrition appears to be positive and significant in boys and girls when applied at short distance to the height measurement. Sometimes breast milk has a negative

tion functions only for infants who are 12 and 24 months old.

[33]Eckhardt et al. (2005) and Adair (1999) found that Filipino children exhibit catch-up growth after age 2.

effect on height, and that is not an expected result, but it depends on the age of the infant and thus it can be only a temporary effect. Besides that, the variable used for the breast feeding is not an indicator of the quality of the milk. A negative sign could indicate a breast milk that does not contain all of the nutrients needed by the infant, possibly due to the poor diet of the mother. Nutrition assumed in infancy has an effect on boys only until age 8.5, but not after they become teenagers. Whereas, if a 8.5 or 11.5-year-old girl was breast-fed when 4-month-old, that decreases height by about 2 cm. It is not clear why, but again, it could be that the breast milk was not nutrient enough. Nutrition has a higher impact when adolescents are 11.5 or 15.5 years old. Increasing caloric intake by 100 kcal increases height by 0.1-0.2cm according to sex and specification. In the fixed effect specification the parameters are slightly bigger than in the naive specification. This period coincides with the adolescence spurt and that could explain the greater importance of nutrients.

Diseases experienced in early life negatively affect boys' height (-3 or -5cm) when they are 8.5 or 11.5 years old. For boys, the reduction in height due to diseases is bigger in the naive specification then in the fixed effect specification. Girls do not seem to be affected by the diseases. This can be explained by the relatively small sample analyzed (keeping boys and girls together might produce different results) or by the different endowment in boys and girls. Whereas, boys with higher μ_i are less likely to experience the diseases, and in the naive specification, a much higher and negative effect reflects the fact that boys who were sick would have been shorter even if they had not experienced the disease. During later adolescence (ages 15.5 and 18.7) the disease inputs applied at any age are still not relevant for girls. The boys are, instead, negatively affected by chronic illnesses occurred at age 8.5. Those diseases of long duration may have started years before (mostly during early childhood, between ages 2 and 8). The reduction in height caused by those diseases is about -2.5cm in the naive specification and -3.3cm in the fixed effects specification.

5 Robustness

5.1 Potential endogeneity of observed inputs

Two papers analyze the endogeneity issues when estimating infants' health production function using the Cebu data. The first paper is by the Cebu-Study-Team (1992) and they analyze four different outcomes: gestational age, weight, diarrhea and respiratory infection. They focus only on the first year of life. They use a "value-added" specification, where each outcome is regressed on the previous outcome measure and on contemporaneous inputs. Thus, all the past inputs collapse

in the past outcome measure. An instrumental variables approach is used to find the effects of the contemporaneous endogenous inputs. Since I, instead, consider all the past inputs I should find instruments for all of them. Among their endogenous inputs there are also birth weight, gestational age, diarrhea, weight, febrile respiratory infection, different breast-feeding patterns, preventive health care for infants and personal hygiene.

The second paper is by Liu, Mroz, and Adair (2009), who look at children in the first two years of life. They estimate an empirical model that includes parents' demand equations and the child's health production functions. Their analysis is based on a dynamic model of parental investments. They jointly estimate four health outcomes: weight, height, incidence of diarrhea and incidence of severe respiratory infections. They use, as in the Cebu-Study-Team (1992), a "value-added" specification of the health production function. The endogenous inputs are breast-feeding choices, caloric intake for supplement food, prenatal care, mother's working and preventive health care. The authors find evidence of compensatory parental behavior, but they argue that it was impossible to observe all aspects of parents' behavior. Hence, it was impossible to disentangle the effects of the observable behavioral responses from the effects of the unobservable behavioral responses due to the choices of missing health inputs. Furthermore, they consider diarrhea and respiratory infections as health outcomes, but they do not consider them as inputs in the height production function. Since I consider only height as an outcome and diseases as relevant inputs, my model is not directly comparable with Liu, Mroz, and Adair's (2009) model.

To address the endogeneity issue, I also apply an instrumental variables approach. The inputs that are potentially endogenous are caloric intakes, breast feeding and diseases.[34] In this paper the endogeneity may be caused by feed back effects.[35] The parents may change their behavior depending on their child's health status. If, for example, the child is malnourished, his parents may increase his nutrient intake, or if the child is sick, his parents may improve the preventive health care. To estimate the input demand equations, I consider two sets of instrumental variables: one for the behavioral inputs of the infants' height production function and one for the behavioral inputs of the adolescents' height production function. The instrumental variables are reported in the Appendix in Tables 13 and 14.[36] Some of them are the same as in the empirical work by Cebu-Study-Team (1992).

[34]One can also argue that birth weight, birth order and gestational age are endogenous. If the mother has older sons, she can change her behavior during the pregnancy according to the past experience (Rosenzweig and Wolpin, 1995).

[35]Endogeneity due to simultaneity is not an issue here, because the model does not include contemporaneous inputs.

[36]I thank the National Statistics Office of the Philippines, which provided me with the CPI and inflation rates used to deflate prices.

These instrumental variables can affect the parents' choices in terms of food consumption and health care, but at the same time they are not correlated with the child's (adolescent's) height. I need at least one instrument per each lagged input. I use lagged values of the instrumental variables to have instruments and inputs measured at the same point in time.

Unfortunately, the instruments are weak, especially for the diseases inputs. Table 15 and the tests following the table show, as an example, the weakness of the instruments used in the height production function of an infant 24 months old estimated through the naive specification.[37] There can be several reasons for that: the lack of inclusion of some important prices, the common food home production that may be less influenced by the market's prices, the high correlation between instruments close in time, or the fact that individuals of the same family may be affected differently by changes in prices or village characteristics, just to mention a few. Consider, for example, the price of rice in January 1983. It is likely to be very similar to the price in May 1983. Or suppose, for example, that the prices of powdered milk increases, then parents may reduce their consumption of other goods to provide the same amount of milk for their infant.

Furthermore, I am aware of the potential endogeneity of net-nutritional inputs, but it can be that parents' compensatory behaviors after infancy are not relevant. Moreover, if missing inputs depend on investments in children, they are likely to be correlated with both the instruments and the observed inputs (Todd and Wolpin, 2003). It is also important to note that if the omitted inputs are correlated with the included inputs, then the IV procedure is not valid.

5.2 Omitted variables bias

To account for the omitted variables bias I estimate a *hybrid production function*[38] that includes in the naive specification annual household income and mother's education.[39] In general, the hybrid health production functions are production functions that contain some of the health inputs and the determinants of the other non-available inputs. In this case the health outcome is height, therefore, I estimate the height production functions.

[37] Almost all the production functions estimated (according to the different ages of the child and the specifications,) with the IV approach present bad instruments, even when I consider different sets of instruments for each lagged input.

[38] To solve the omitted variables bias it would also be possible to express the missing inputs as functions of current and past family income, prices and preference shocks. Once those omitted input demand equations are specified they can substitute the missing inputs in the production function (Todd and Wolpin, 2007).

[39] Household income is a categorical variable with 6 categories that correspond to the percentiles of the distribution of annual household income. The mother's education is a categorical variable with 5 categories according to the number of years of schooling.

I test the possibility of omitted variables bias by looking at the coefficients of household income and mother's education in the hybrid production functions. If they are significantly different from zero, then they need to be included in the functions. Family income is positive and significant in girls of mean age 8.5 years, while a better socio-economic situation of the family, with both higher income and higher mother's education, is positive and significant for boys when they are on average 11.5 years old. That means that omitted inputs seem to affect height when children begin their growth spurt.

Once controlled for income and mother's education the hybrid effect of the inputs on height is generally a biased estimate of the true technical relationship (other inputs held constant) embodied in the health production function (Rosenzweig and Schultz, 1983). Despite that it is likely that the inclusion of income and mother's education does not address satisfactorily the omitted variables problem. The hybrid production functions are available upon request.

5.3 Sample selection bias due to attrition

The last wave considered in the empirical exercise contains 2023 people (1071 males, 952 females). Between the baseline and the last wave around 35% of the children are lost, and 18% before the age of 2. This high attrition is common in long-term longitudinal studies and in data that come from developing countries.[40] The highest attrition rate is between wave 12 and 13, that is, between infancy and mean age 8.5. From wave 12 conducted in 1985-86 and wave 13 conducted in 1991-92 there are approximately 6-7 years of missing information, the children's childhood.

The two main reasons for attrition are death and migration. Two hundred and twenty-five (7.3% of the sample) children die: 167 (5.4%) in the first 2 years of life, 44 between the ages of 2 and 8, 14 children die during adolescence. In total 129 boys and 96 girls die.

The remaining 28% are mainly lost because of migration. The Cebu-Study-Team (1992) tested for selectivity of infants and the results show that the omitted variables that influence migration decisions do not greatly overlap with those that determine child health. I tested for selection on unobservables of adolescents who migrate and the conclusions are the same. Therefore I consider that attrition due to migration can be disregarded.

Eckhardt et al. (2005) used the same longitudinal survey to study the compensatory growth that occurred after 2 years of age among the children who were

[40] Alderman et al. (2001) study attrition in three different longitudinal data sets from developing countries. Even if attrition is very high, it is not a general problem for obtaining consistent estimates.

stunted in infancy. They report that people who died tended to be of very low socio-economic status and grew up poorly until their deaths, while those who moved tended to be of higher socio-economic status and from urban areas. In their study, attrition does not represent a problem.[41] Unfortunately, there is no information about the interviewers or other exclusion restrictions that could be used to test and correct the selection on unobservables that determine death. Hence, given the rather low percentage of children who died, I keep them in the sample.

5.4 Model specification

The naive and fixed effect specifications give similar results. That is true for the coefficients and for the predictions, as can be seen in Figure 2. The big difference about the naive and within-child fixed effect specifications is assumption (a) about the endowment. The fixed effects allow the inputs to be endogenous with respect to the endowment. It is plausible that investments in children are correlated with the unobserved endowment. To select between the two specifications I look at the out-of-sample goodness of fit.

I use the cross validation procedure to test the goodness of fit out of sample. This procedure can be used to compare non-nested models and, under some statistical assumptions, the root mean square error (RMSE) is used to summarize the errors. I do the cross validation using random holdout samples. I randomly divide the sample in 8 subsamples of equal size. I estimate the model on 7 of the 8 subsamples and I compute the RMSE for the left-out subsample. I repeat that, considering each time a different holdout subsample. I then sum the RMSE obtaining the total RMSE for that model. I repeat the subsetting 5 times and I compute the average total RMSE.

Table 12 shows the RMSE of the different specifications per each age of the person. The model with the smallest RMSE performs best with a data set that is independent of the data used to train the model. The naive specification always reports the lowest RMSE, and therefore, it predicts better out-of-sample than the fixed effects specification. However, the naive specification is based on a strong assumption that requires the endowment to not be correlated with the observed inputs. And one may argue that the cross-validation is based on level predictions. It would be interesting in the future to study growth in height and see which is the specification that performs better.

[41] The reason being that only a small proportion of children in the poorest health conditions were lost.

6 Conclusions

In this paper I study the determinants of height in a different way than in the existing literature. I clearly specify a model of height formation, building a height production function. I consider the cumulative nature of physical development, taking into account both the age of the child and the distance between application of the inputs and height measurement, and carefully considering the biological inputs from conception to early adulthood. In the model I include not only non-genetic inputs, such as nutrition and diseases' inputs, but also a genetic endowment, because both are important for development. I define the identification strategies, and I estimate the height production function using two different specifications.

The results show a systematic effect of net nutrition on the child's height, looking at different ages of the child. The older the child, the bigger the nutrition effects and the stronger the disease effects. I found that in most of the cases nutrition has an impact at a short distance, but nutrition effects start to decline after the adolescent's growth spurt. Diseases, instead, affect height if applied at early ages or at a long distance from the height measurement. Boys' height is negatively influenced by diseases, and that is true at any age, with a peak at age 11.5 years. While, diseases experienced during infancy especially matter in a girl's height when she is younger than 1 year, but not later on.

The model, indeed, shows the importance of including past inputs and of studying their effects according to different ages and sex of the children. This analysis is not sufficient to derive policy implications that look at the *non-ceteris-paribus* effect caused by changes in inputs. That would require a behavioral model of how parents make decisions about the inputs for their child and how the child makes decisions for himself when older. Besides that, the results can be important to know what factors determine a person's height at different stages of her life in a poor country, which periods are critical, and which inputs have the strongest impact on growth. In fact, growth *in utero*, infancy and the pre-puberty years turn out to be critical stages in the process of height formation. In particular, the magnitude of the inputs' effects during *in utero* period and early childhood are higher than later on. Looking at the literature on cognitive achievement, Cunha and Heckman (2008) find that there are sensitive periods in the child's skills development. Parents' investments are differently effective according to the stage of the child's life cycle (Cunha et al., 2006). Therefore, the results of this paper are in line with the increasing literature on the long term effects of early childhood conditions. Some things that happen before age 5 may have persistent effects on health and human capital accumulation (Almond and Currie, 2010).

References

Adair, L., 1999. Filipino children exhibit catch-up growth from age 2 to age 12. *Journal of Nutrition*, 129:1140–1148.

A'Hearn, B., F. Peracchi, and G. Vecchi, 2009. Height and the normal distribution: Evidence from Italian military data. *Demography*, 46:1–25.

Akachi, Y. and D. Canning, 2007. The height of women in Sub-Saharan Africa: The role of health, nutrition, and income in childhood. *Annals of Human Biology*, 34:349 – 410.

Alderman, H., J. Behrman, H. Kohler, J. Maluccio, and S. Watkins, 2001. Attrition in Longitudinal Household Survey Data. *Demographic Research*, 5(4).

Almond, G. and J. Currie, 2010. Human capita development before age five. *NBER Working Paper No. 15827*.

Barker, D. J. P., 1998. *Mothers, Babies and Health in Later Life*. Churchill Livingstone, Edinburg.

Behrman, J. R., M. C. Calderon, S. Preston, J. Hoddinott, R. Martorell, and A. D. Stein, 2009a. Nutritional supplementation of girls influences the growth of their children: Prospective study in Guatemala. *American Journal of Clinical Nutrition*, 90:1372–1379.

Behrman, J. R. and A. B. Deolalikar, 1988. Health and nutrition. In H. Cheney and T. N. Srinivasan, editors, *Handbook of Development Economics*, volume I, chapter 14. Elsevier Science Publishers B.V.

Behrman, J. R. and A. B. Deolalikar, 1989. Wages and labor supply in rural India: The role of health, nutrition and seasonality. In D. E. Sahn, editor, *Causes and implications of seasonal variability in household food security*. The Johns Hopkins University Press, Baltimore, MD.

Behrman, J. R., J. Hoddinott, J. A. Maluccio, and R. Martorell, 2009b. Brains versus brawn: Labor market returns to intellectual and health human capital in a poor developing country. *Middlebury College Economics Discussion Paper N. 0907*.

Behrman, J. R. and M. R. Rosenzweig, 2004. Returns to birthweight. *Review of Economics and Statistics*, 86(2):586–601.

Bozzoli, C., A. Deaton, and C. Quintana-Domeque, 2009. Adult height and childhood disease. *Demography*, 46(4).

Branca, F. and M. Ferrari, 2002. Impact of micronutrient deficiencies on growth: The stunting syndrome. *Annals of Nutrition and Metabolism*, 46(1).

Carter, C. O. and W. A. Marshall, 1978. The genetics of adult stature. In F. Falkner and J. M. Tanner, editors, *Human Growth*. New York and London: Plenum Press.

Case, A. and C. Paxson, 2008a. Height, health and cognitive function at older ages. *American Economic Review Papers and Proceedings*, 98(2).

Case, A. and C. Paxson, 2008b. Stature and status: height, ability, and labor market outcomes. *Journal of Political Economy*, 116(3).

Cebu-Study-Team, 1992. A child health production function estimated from longitudinal data. *Journal of Development Economics*, 32:323–351.

Croppenstedt, A. and C. Muller, 2000. The impact of farmers' health and nutritional status on their productivity and efficiency: Evidence from Ethiopia. *Economic Development and Cultural Change*, 48(3):475–502.

Cuff, T., 2004. Historical Anthropometrics. *EH.Net Encyclopedia*.

Cunha, F. and J. Heckman, 2007. The technology of skill formation. *American Economic Review*, 97:31–47.

Cunha, F. and J. Heckman, 2008. Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation. *The Journal of Human Resources*, XLIII(4):738–782.

Cunha, F., J. J. Heckman, L. Lochner, and D. V. Masterov, 2006. Interpreting the evidence on life cycle skill formation. In E. A. Hanushek and F. Welch, editors, *Handbook of the Economics of Education*, volume 1, chapter 12. Elsevier Science Publishers B.V.

Darrell Bock, R., A. Petersen, D. Thissen, J. Murray, and A. Roche, 1973. A parameterization for individual human growth curves. *Human Biology*, 45(1):63.

Deaton, A., 2007. Height, health, and development. *Proceedings of the National Academy of Sciences*, 104(33):13232–13237.

Deaton, A., 2008. Height, health, and inequality: The distribution of adult heights in India. *American Economic Review: Papers and Proceedings*, 98(2):468–474.

Dinda, S. and P. Gangopadhyay, 2006. Height, weight and earnings among coalminers in India. *Economics and Human Biology*, 4(3):342–350.

Eckhardt, C. I., P. Gordon-Larsen, and L. S. Adair, 2005. Growth patterns of Filipino children indicate potential compensatory growth. *Annals of Human Biology*, 32(1):3–14.

Falker, F. and J. M. Tanner, 1986. *Human Growth: A Comprehensive Treatise*. New York, United States: Plenum Press, second edition edition.

Fogel, R., 1986. *Human growth*, volume 3, chapter Growth and economic well being: 18th and 19th centuries. eds F. Falkner, J.M. Tanner, New York: Plenum.

Ginsburg, E., G. Livishits, K. Yakovenko, and E. Kobyliansky, 1998. Major gene control of human body height, weight and BMI in five ethnically different populations. *Annals of Human Genetics*, 62(4).

Haddad, L. J. and H. E. Bouis, 1991. The impact of nutritional status on agricultural productivity: Wage evidence from the Philippines. *Oxford Bulletin Economics and Statistics*, 53:45–68.

Hoddinott, J., J. A. Maluccio, J. R. Behrman, R. Flores, and R. Martorell, 2008. The impact of nutrition during early childhood on income, hours worked, and wages of Guatemalan adults. *Lancet*, 371:411–416.

Immink, M. and F. E. Viteri, 1981. Energy intake and productivity of Guatemalan sugarcane cutters: An empirical test of the efficiency wages hypothesis, Parts I and II. *Journal of Development Economics*, 92:251–287.

Jolicoeur, P., J. Pontier, M. Pernin, and M. Sempè, 1988. A lifetime asymptotic growth curve for human height. *Biometrics*, 44:995–1003.

Liu, H., T. Mroz, and L. Adair, 2009. Parental compensatory behaviors and early child health outcomes in Cebu, Philippines. *Journal of Development Economics*, 90:209–230.

Maluccio, J. A., J. Hoddinott, J. R. Behrman, A. Quisumbing, R. Martorell, and A. D. Stein, 2009. The impact of nutrition during early childhood on education among Guatemala adults. *Economic Journal*, 119:734 – 763.

Morgan, J., 1999. *Milk-Feeding and Weaning*, pages 1097–1108. (eds Sadler, M.J. and Strain, J.J. and Caballero, B.) Academic Press, San Diego, California.

Preece, M. A. and M. J. Baines, 1978. A new family of mathematical models describing the human growth curve. *Annals of Human Biology*, 5(1):1–24.

Rosenzweig, M. R. and T. P. Schultz, 1983. Estimating a household production function: Heterogeneity, the demand for health inputs, and their effects on birth weight. *Journal of Political Economy*, 91(5):723–746.

Rosenzweig, M. R. and K. I. Wolpin, 1995. Sisters, siblings, and mothers: The effect of teen-age childbearing on birth outcomes in a dynamic family context. *Econometrica*, 63(2):303–326.

Schultz, T. P., 2002. Wage gains associated with height as a form of health human capital. *American Economic Review Papers and Proceedings*, 92(2).

Schultz, T. P., 2003. Wage rentals for reproducible human capital: Evidence from Ghana and the Ivory Coast. *Economics and Human Biology*, 1:331–366.

Scott, S. and J. Duncan, 2002. *Demography and Nutrition: Evidence from Historical and Contemporary Populations*. Wiley-Blackwell.

Shohoji, T. and H. Sasaki, 1987. Individual growth of stature of Japanese. *Growth*, 51:432–450.

Silventoinen, K., 2003. Determinants of variation in adult body height. *Journal of Biosocial Science*, 35:263–285.

Steckel, R. H., 1995. Stature and the standard of living. *Journal of Economic Literature*, 33(4):1903–40.

Steckel, R. H., 2009. Heights and human welfare: Recent developments and new directions. *Explorations in Economic History*, 46(1).

Stephensen, A., 1999. Burden of infection on growth failure. *Journal of Nutrition*, 129:534–538.

Strauss, J. and D. Thomas, 2008. Health over the Life Course. In T. P. Schultz and J. A. Strauss, editors, *Handbook of Development Economics*, chapter 54, pages 3375–3474. Elsevier.

Subramanian, S. V., L. K. Ackerson, G. D. Smith, and N. A. John, 2009. Association of maternal height with child mortality, anthropometric failure, and anemia in India. *Journal of American Medical Association*, 301(16).

Tanner, J. M., 1990. *Fetus Into Man: Physical Growth from Conception to Maturity*. Harvard University Press.

Thomas, D. and J. Strauss, 1997. Health and wages: Evidence on men and women in urban Brazil. *Journal of Econometrics*, 77:159–185.

Thomas, D. and J. Strauss, 1998. Health, nutrition and economic development. *Economic Literature*, 36:766–817.

Todd, P. and K. I. Wolpin, 2003. On the specification and estimation of the production function for cognitive achievement. *Economic Journal*, 113:3–33.

Todd, P. and K. I. Wolpin, 2007. The production of cognitive achievement in children: Home, school and racial test score gaps. *Journal of Human Capital*, 1(1):91–136.

Victora, C. G., F. C. Barros, B. R. Kirkwood, and J. P. Vaughan, 1990. Pneumonia, diarrhea and growth in the first 4 y of life: A longitudinal study of 5914 urban Brazilian children. *American Journal of Clinical Nutrition*, 52:391–396.

Walter, T., M. Olivares, F. Pizarro, and C. Muñoz, 1997. Iron, anemia and infection. *Nutrition Review*, 55:111–124.

7 Tables and Graphs

Table 1: Lagged Inputs Parameters - β's

Age at height measure (years)	Age at 1 Period Lagged input	Age at 2 Period Lagged input	...	Age at (t-1) Period Lagged input	...	Age at (T-2) Period Lagged input	Age at (T-1) Period Lagged input
0	...	...	...	...	...	...	...
1	0 ($\beta_{1,0}$)	...	...	...	...	...	...
2	1 ($\beta_{2,1}$)	0 ($\beta_{2,0}$)	...	...	...	...	...
3	2 ($\beta_{3,2}$)	1 ($\beta_{3,1}$)	...	...	...	...	...
...	...	...	...	...	...	...	...
t	t-1 ($\beta_{t,t-1}$)	t-2 ($\beta_{t,t-2}$)	...	0 ($\beta_{t,0}$)	...	...	...
...	...	...	...	...	...	...	...
T-1	T-2 ($\beta_{T-1,T-2}$)	T-3 ($\beta_{T-1,T-3}$)	...	T-t ($\beta_{T-1,T-t}$)	...	0 ($\beta_{T-1,0}$)	...
T	T-1 ($\beta_{T,T-1}$)	T-2 ($\beta_{T,T-2}$)	...	T-t+1 ($\beta_{T,T-t+1}$)	...	1 ($\beta_{T,1}$)	0 ($\beta_{T,0}$)

Number of β's parameters to estimate according to the person's age at height measurement.

Table 2: Panel structure and ranges of ages

Surveys	**Min age**	**Max age**	**Mean age**	**St. dev.**	**N**
Months					
•⋆Delivery Survey 1983-4	0	0	0	0	3080
Follow-up n.1	1.4	3.4	2.05	.15	2878
•⋆Follow-up n.2	3.4	5.7	4.05	.15	2802
Follow-up n.3	5.4	7.3	6.05	.15	2717
⋆Follow-up n.4	7.5	9.1	8.04	.13	2665
Follow-up n.5	9.4	11.2	10.07	.15	2625
•⋆Follow-up n.6	11.6	13.1	12.07	.17	2594
Follow-up n.7	13.0	15.0	14.07	.17	2549
⋆Follow-up n.8	15.5	17.2	16.07	.17	2507
Follow-up n.9	17.7	19.1	18.06	.17	2507
⋆Follow-up n.10	19.6	21.1	20.06	.18	2498
Follow-up n.11	21.7	23.0	22.04	.16	2460
•⋆Follow-up n.12	23.6	25.2	24.05	.15	2448
Years					
• 1991-2	7.5	8.7	8.5	.05	2264
• 1994-5	10.6	13.0	11.5	.40	2182
• 1998-9	14.1	16.8	15.5	.67	2089
• 2002	17.9	19.8	18.7	.34	2023
2005	20.7	22.3	21.5	.30	1888

The ⋆ indicates the waves used in the infant production function.

The • indicates the waves used in the adolescent production function.

Table 3: Descriptive statistics of the principal variables.

	At birth		Mean Age 1 year		Mean age 2 years		Mean age 8.5 years		Mean age 18.7 years	
Boys' Variables*	Means	SD	Means	SD	Means	SD	Means	SD	Means	SD
Height (cm)	49.457	2.152	71.449	2.883	79.868	3.581	117.738	5.537	162.593	5.901
Mother's height (cm)	151.474	5.039								
Calorie intake (kcal)	0.000	0.000	461.833	382.131	754.377	433.239	1504.863	601.003	2105.662	978.661
Breast milk	0.991	0.095	0.603	0.489	0.127	0.334	0.000	0.000	0.000	0.000
Diseases	0.017	0.130	0.211	0.408	0.159	0.365	0.026	0.158	0.165	0.372
Location (1 if Urban)	0.770	0.421	0.757	0.429	0.757	0.429	0.746	0.436	0.744	0.437
Birth weight (kg)	3.023	0.456								
Pregnancy's length	0.146	0.442								
Birth order	2.586	2.457								
Household Income (pesos)	1989.773	3521.077	1487.985	2629.005	1693.638	2942.903	2331.194	3128.405	3433.004	4723.217
Mother's education (years)	7.949	3.824								
Girls' Variables*										
Height (cm)	49.009	2.108	69.923	2.841	78.315	3.629	117.624	5.573	150.964	5.454
Mother's height (cm)	151.434	5.042								
Calorie intake (kcal)	0.000	0.000	421.677	338.826	695.381	390.131	1354.057	505.668	1510.551	716.109
Breast milk	0.991	0.094	0.630	0.483	0.142	0.349	0.000	0.000	0.000	0.000
Diseases	0.015	0.120	0.180	0.384	0.145	0.352	0.028	0.165	0.165	0.372
Location (1 if Urban)	0.758	0.428	0.742	0.437	0.745	0.436	0.741	0.438	0.737	0.440
Birth weight (kg)	2.961	0.432								
Pregnancy's length	0.159	0.462								
Birth order	2.503	2.388								
Household Income (pesos)	2014.533	4595.366	1387.258	2120.290	1512.491	2065.838	2280.632	2464.449	3432.117	5609.358
Mother's education (years)	7.869	3.774								

*The detailed description of the variables is reported in section (4) of the paper.

Mother's height, birth weight, pregnancy's length, birth order and mother's years of education are time-invariant.

Figure 1: Height distribution by age and sex of the child

Table 4: *Infants' Height Production Function* - Calorie intake and ages

Wave	Mean Monthly Age at height measure (t)	Age at 1 Period Lagged kcal	Age at 2 Period Lagged kcal	Age at 3 Period Lagged kcal	Age at 4 Period Lagged kcal	Age at 5 Period Lagged kcal	Age at 6 Period Lagged kcal
0	0	...	...	...	...	...	...
2	4	0 ($\beta_{4,0}$)	...	...	...	...	...
4	8	4 ($\beta_{8,4}$)	0 ($\beta_{8,0}$)	...	...	...	...
6	12	8 ($\beta_{12,8}$)	4 ($\beta_{12,4}$)	0 ($\beta_{12,0}$)	...	...	...
8	16	12 ($\beta_{16,12}$)	8 ($\beta_{16,8}$)	4 ($\beta_{16,4}$)	0 ($\beta_{16,0}$)	...	...
10	20	16 ($\beta_{20,16}$)	12 ($\beta_{20,12}$)	8 ($\beta_{20,8}$)	4 ($\beta_{20,4}$)	0 ($\beta_{20,0}$)	...
12	24	20 ($\beta_{24,20}$)	16 ($\beta_{24,16}$)	12 ($\beta_{24,12}$)	8 ($\beta_{24,8}$)	4 ($\beta_{24,4}$)	0 ($\beta_{24,0}$)

Table 5: *Adolescent's Height Production Function* - Calorie intake and ages

Wave	Mean Yearly Age at height measure (t)	Age at 1 Period Lagged kcal	Age at 2 Period Lagged kcal	Age at 3 Period Lagged kcal	Age at 4 Period Lagged kcal	Age at 5 Period Lagged kcal	Age at 6 Period Lagged kcal	Age at 7 Period Lagged kcal
13	8.5	24m ($\beta_{8.5,24m}$)	1 ($\beta_{8.5,12m}$)	4m ($\beta_{8.5,4m}$)	0 ($\beta_{8.5,0}$)	...	...	...
14	11.5	8.5 ($\beta_{11.5,8.5}$)	2 ($\beta_{11.5,24m}$)	12m ($\beta_{11.5,12m}$)	4m ($\beta_{11.5,4m}$)	0 ($\beta_{11.5,0}$)	...	...
15	15.5	11.5 ($\beta_{15.5,11.5}$)	8.5 ($\beta_{15.5,8.5}$)	24m ($\beta_{15.5,24m}$)	12m ($\beta_{15.5,12m}$)	4m ($\beta_{15.5,4m}$)	0 ($\beta_{15.5,0}$)	...
16	18.7	15.5 ($\beta_{18.7,15.5}$)	11.5 ($\beta_{18.7,11.5}$)	8.5 ($\beta_{18.7,8.5}$)	24m ($\beta_{18.7,24m}$)	12m ($\beta_{18.7,12m}$)	4m ($\beta_{18.7,4m}$)	0 ($\beta_{18.7,0}$)

Table 6: Height production function of an infant 12 months old.

	Boys		Girls	
	Naive	Fixed Effects	Naive	Fixed Effects
Inputs				
Birth Order	−0.267***		−0.225***	
I-Urban	−0.099	−0.757	−0.110	1.204
Birth weight	2.154***		2.275***	
I-Premature&Small-for-dates	0.060		0.098	
I-Premature	0.004		0.300	
Mother's height	0.140***		0.127***	
Calorie Intake at 8m	0.065*	0.051	−0.023	−0.028
Calorie Intake at 4m	0.079	0.059	−0.001	−0.002
Breast milk at 8m	0.854**	0.649*	0.417	0.262
Breast milk at 4m	−0.038	0.068	−0.680	−0.455
Calorie Intake*Breast milk at 8m	−0.048	−0.075	0.074	0.084
Calorie Intake*Breast milk at 4m	0.037	−0.015	0.111	0.052
Disease at 8m	0.079	0.208	0.052	−0.035
Disease at 4m	−0.661**	−0.730**	0.004	−0.341
Disease at birth	−0.652	−0.222	−1.644	−2.256*
N	1364	4414	1224	4274
R Sq Adj 0.267	0.985	0.244	0.984	

Cluster standard error not reported. Signif. codes: (*) if p < .05, (**) if p < .01, (***) if p < .001
Each model (each column) includes the age in months and the age in months squared.
The letter m in the inputs indicates months (for example, 4m means 4 months).
A change of one unity in caloric intake corresponds to 100 kcal.
† means that the parameters are 0 or not well estimated.

Table 7: Height production function of an infant 24 months old.

	Boys		Girls	
	Naive	Fixed Effects	Naive	Fixed Effects
Inputs				
Birth Order	−0.344***		−0.302***	
I-Urban	−0.265	0.323	−0.267	−0.278
Birth weight	1.865***		1.879***	
I-Premature&Small-for-dates	0.010		−0.659	
I-Premature	0.034		−0.421	
Mother's height	0.186***		0.168***	
Calorie Intake at 20m	0.096***	0.107***	0.081**	0.069*
Calorie Intake at 16m	0.089**	0.081*	0.091*	0.112**
Calorie Intake at 12m	0.074	0.080	0.133*	0.128*
Calorie Intake at 8m	0.085	0.083	−0.054	−0.036
Calorie Intake at 4m	−0.008	−0.046	0.011	−0.012
Breast milk at 20m	1.576***	0.951	1.186*	1.079*
Breast milk at 16m	0.388	0.678	0.563	0.493
Breast milk at 12m	0.768	0.645	0.953*	0.836
Breast milk at 8m	0.245	0.244	−0.097	−0.127
Breast milk at 4m	−0.763	−0.642	−0.753	−0.481
Calorie Intake*Breast milk at 20m	−0.170***	−0.157*	−0.041	−0.044

Continued...

	Boys		Girls	
Calorie Intake*Breast milk at 16m	−0.009	−0.085	−0.066	−0.071
Calorie Intake*Breast milk at 12m	−0.026	−0.015	−0.040	−0.024
Calorie Intake*Breast milk at 8m	−0.088	−0.126	0.075	0.072
Calorie Intake*Breast milk at 4m	0.203**	0.160*	0.283**	0.199
Disease at 20m	−0.005	−0.199	−0.538**	−0.470*
Disease at 16m	−0.478*	−0.676*	−0.387	−0.297
Disease at 12m	−0.275	−0.342	−0.143	−0.211
Disease at 8m	0.085	0.263	−0.118	−0.137
Disease at 4m	−0.883**	−0.976***	0.015	−0.338
Disease at birth	−1.628*	−0.880	−1.187	−1.888
N	1286	4336	1160	4210
R Sq Adj	0.300	0.988	0.291	0.988

Cluster standard error not reported. Signif. codes: (*) if p < .05, (**) if p < .01, (***) if p < .001
Each model (each column) includes the age in months and the age in months squared.
The letter m in the inputs indicates months (for example, 4m means 4 months).
A change of one unity in caloric intake corresponds to 100 kcal.

Table 8: Height production function of a child 8.5 years old.

	Boys		Girls	
	Naive	Fixed Effects	Naive	Fixed Effects
Inputs				
I-Urban	0.182	0.385	−0.022	−0.221
Birth Order	−0.320***		−0.281***	
Birth weight	2.651***		2.428***	
I-Premature&Small-for-dates	0.676		−0.298	
I-Premature	2.200*		−0.408	
Mother's height	0.377***		0.332***	
Calorie Intake at 24m	0.103*	0.158***	0.096*	0.097*
Calorie Intake at 12m	0.068	0.048	−0.009	−0.007
Calorie Intake at 4m	0.128	0.066	−0.049	−0.041
Breast milk at 24m	1.627	1.320	0.981	0.407
Breast milk at 12m	0.204	−0.301	−0.247	−0.465
Breast milk at 4m	−0.379	−0.556	−1.790*	−1.827*
Calorie Intake*Breast milk at 24m	−0.258	−0.374*	−0.248	−0.41
Calorie Intake*Breast milk at 12m	−0.073	−0.029	0.139	0.185
Calorie Intake*Breast milk at 4m	−0.011	−0.004	0.445*	0.424*
Disease at 24m	−0.286	−0.613*	0.267	0.264
Disease at 12m	0.180	0.179	−0.386	−0.781
Disease at 4m	−1.225**	−1.283**	0.730	0.257
Disease at birth	−4.157**	−3.478*	−2.511	−3.182
N	1191	2808	1065	2498
R Sq Adj	0.250	0.980	0.207	0.994

Cluster standard error not reported. Signif. codes: (*) if p < .05, (**) if p < .01, (***) if p < .001
Each model (each column) includes the age in years and the age in years squared.
The letter m in the inputs indicates months (for example, 4m means 4 months).
A change of one unity in caloric intake corresponds to 100 kcal.

Table 9: Height production function of a child 11.5 years old.

	Boys		Girls	
	Naive	Fixed Effects	Naive	Fixed Effects
Inputs				
I-Urban	0.449	−0.013	−0.140	−0.294
Birth Order	−0.367***		−0.232*	
Birth weight	2.446***		3.324***	
I-Premature&Small-for-dates	0.601		−0.273	
I-Premature	3.324***		−1.318	
Mother's height	0.406***		0.372***	
Calorie Intake at 8.5	0.182***	0.208***	0.201***	0.215***
Calorie Intake at 24m	0.081*	0.135***	0.142*	0.139*
Calorie Intake at 12m	0.101	0.075	−0.011	−0.021
Calorie Intake at 4m	0.147	0.087	−0.227	−0.214
Breast milk at 24m	0.820	0.543	0.877	0.191
Breast milk at 12m	−0.016	−0.618	−0.256	−0.464
Breast milk at 4m	−0.329	−0.401	−2.690**	−2.732***
Calorie Intake*Breast milk at 24m	−0.125	−0.247	−0.277	−0.236
Calorie Intake*Breast milk at 12m	−0.154	−0.109	0.184	0.227
Calorie Intake*Breast milk at 4m	−0.141	−0.152	0.506**	0.496**
Disease at 8.5	−1.098	−1.453	−0.767	−0.520
Disease at 24m	−0.154	−0.426	−0.103	−0.027
Disease at 12m	0.092	0.043	−0.407	−0.849
Disease at 4m	−1.006	−1.095	0.903	0.444
Disease at birth	−5.474***	−4.698**	−2.095	−3.086
N	1137	2754	1037	2470
R Sq Adj	0.323	0.995	0.252	0.994

Cluster standard error not reported. Signif. codes: (*) if p < .05, (**) if p < .01, (***) if p < .001
Each model (each column) includes the age in years and the age in years squared.
The letter m in the inputs indicates months (for example, 4m means 4 months).
A change of one unity in caloric intake corresponds to 100 kcal.

Table 10: Height production function of a child 15.5 years old.

	Boys		Girls	
	Naive	Fixed Effects	Naive	Fixed Effects
Inputs				
I-Urban	0.582	0.884	−0.701	−0.116
Birth Order	−0.221**		−0.129	
Birth weight	2.919***		2.528***	
I-Premature&Small-for-dates	1.041		−0.467	
I-Premature	2.371**		0.753	
Mother's height	0.502***		0.429***	
Calorie Intake at 11.5	0.129***	0.174***	0.158***	0.162***
Calorie Intake at 8.5	0.103**	0.113**	0.088**	0.091**
Calorie Intake at 24m	0.054	0.117*	0.007	−0.001
Calorie Intake at 12m	0.068	0.029	−0.045	−0.061
Calorie Intake at 4m	0.056	−0.004	−0.077	−0.067
Breast milk at 24m	1.096	0.960	0.741	0.390

Continued...

	Boys		Girls	
Breast milk at 12m	−0.239	−0.693	0.980	0.782
Breast milk at 4m	0.515	0.170	−1.172	−1.228∗
Calorie Intake*Breast milk at 24m	−0.073	−0.217	−0.222	−0.229
Calorie Intake*Breast milk at 12m	−0.077	−0.022	0.032	0.103
Calorie Intake*Breast milk at 4m	−0.128	−0.110	0.325∗	0.343∗
Disease at 11.5	0.913	1.050	−0.353	−0.303
Disease at 8.5	−2.277∗	−2.911∗∗	−1.581	−1.391
Disease at 24m	−0.441	−0.720	0.139	0.207
Disease at 12m	−0.033	−0.187	0.289	−0.199
Disease at 4m	−0.865	−0.959	0.652	0.048
Disease at birth	−2.694	−1.817	−1.016	−1.891
N	1051	2668	989	2422
R Sq Adj	0.300	0.997	0.292	0.998

Cluster standard error not reported. Signif. codes: (*) if p < .05, (**) if p < .01, (***) if p < .001
Each model (each column) includes the age in years and the age in years squared.
The letter m in the inputs indicates months (for example, 4m means 4 months).
A change of one unity in caloric intake corresponds to 100 kcal.

Table 11: Height production function of a child 18.7 years old.

	Boys		Girls	
	Naive	Fixed Effects	Naive	Fixed Effects
Inputs				
I-Urban	−0.521	−0.956	−0.192	0.980
Birth Order	−0.192∗		−0.276	
Birth weight	2.759∗∗∗		2.487∗∗∗	
I-Premature&Small-for-dates	0.544		−0.276	
I-Premature	1.471		0.854	
Mother's height	0.474∗∗∗		0.467∗∗∗	
Calorie Intake at 15.5	0.007	0.016	0.059∗	0.065∗
Calorie Intake at 11.5	0.054∗	0.089∗∗	0.079∗	0.093∗
Calorie Intake at 8.5	0.049	0.050	0.058	0.067
Calorie Intake at 24m	0.026	0.088	−0.016	−0.018
Calorie Intake at 12m	0.002	−0.051	−0.057	−0.095
Calorie Intake at 4m	0.147	0.117	0.016	0.007
Breast milk at 24m	0.479	0.582	1.176	1.335
Breast milk at 12m	−0.217	−0.501	1.077∗	0.684
Breast milk at 4m	0.532	0.631	−0.398	−0.616
Calorie Intake*Breast milk at 24m	0.030	−0.123	−0.268	−0.380∗
Calorie Intake*Breast milk at 12m	0.001	0.015	−0.027	0.107
Calorie Intake*Breast milk at 4m	−0.189	−0.197	0.307	0.340∗
Disease at 15.5	−0.056	0.313	−0.054	0.357
Disease at 11.5	1.067	0.889	−0.195	−0.602
Disease at 8.5	−2.387∗∗	−3.283∗∗∗	−1.141	−0.827
Disease at 24m	−0.379	−0.660	0.343	0.419
Disease at 12m	0.348	0.268	0.346	−0.014
Disease at 4m	−0.568	−0.521	0.640	0.042
Disease at birth	−2.573	−1.805	−0.494	−0.873
N	1067	2684	948	2381
R Sq Adj	0.273	0.998	0.296	0.998

Continued. . .

	Boys	Girls
Cluster standard error not reported. Signif. codes: (*) if p < .05, (**) if p < .01, (***) if p < .001		
Each model (each column) includes the age in years and the age in years squared.		
The letter m in the inputs indicates months (for example, 4m means 4 months).		
A change of one unity in caloric intake corresponds to 100 kcal.		

Table 12: Cross Validation Results

	Boys		Girls	
	Naive	Fixed Effects	Naive	Fixed Effects
Mean Age				
18.7 years	73.291†	75.753	74.724†	77.409
15.5 years	50.323†	56.748	67.330†	70.323
11.5 years	53.892†	57.134	51.199†	53.998
8.5 years	40.523†	43.155	40.041†	43.145
24 months	25.749†	28.342	25.895†	28.328
20 months	24.811†	27.372	24.966†	27.361
16 months	23.084†	25.855	23.215†	25.803
12 months	34.783†	36.959	34.953†	36.952
8 months	47.349†	48.624	47.451†	48.520
4 months	60.508†	61.100	60.537†	61.184
Random holdout sample. A number of 8 subsamples of equal size are created.				
The model is estimated on 7 of the 8 subsamples and the RMSE is computed.				
This is repeated, considering each time a different holdout subsample.				
The total RMSE for that model is obtained summing all the RMSE.				
The average total RMSE is obtained replicating this procedure 5 times.				
† corresponds to the lower RMSE per age of the person.				

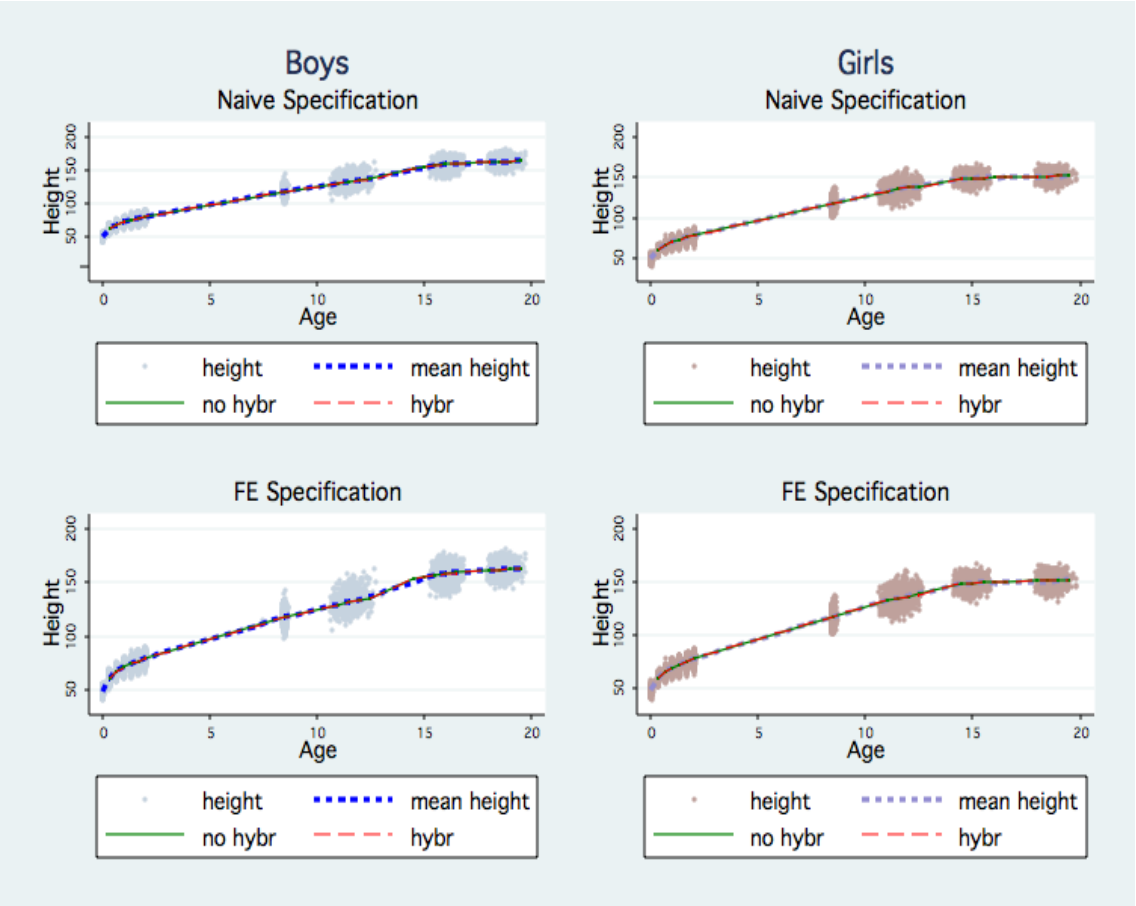

Figure 2: Height predictions per each age of the child

A Details of the variables used

A.1 Outcome

Height was measured using infantometers. The infantometer is a length measurement used for infants and children who are unable to stand. For children and teenagers, ordinary meter sticks were used to measure their height.

In 1984 and 1985, reliability checks were made on the anthropometric measurements of infants and their mothers taken by the field staff in Cebu. The procedure used was the Habicht procedure, according to which, some of the subjects are measured twice by a number of observers and one supervisor. The test results, which are supposed to provide indications of (1) measurement accuracy and (2) measurement precision, are based on comparisons of (a) the two measures of each observer and (b) the measurements of the observers with those of the supervisor. The latter is supposed to produce the most reliable measurements on account of his/her experience. The tests did not report systematic bias in the anthropometric measurements.

A.2 Inputs from birth to age 22

Genetics Individuals inherit both chromosomes and genes, but it is not known, for example, if the X chromosome has a bigger effect than the Y chromosome or vice versa (Carter and Marshall, 1978). DNA is inhered but height can be developed.

Caloric intake Daily energy intake is calculated from 24-hour dietary recalls during the surveys from birth to 2 years and in 1994, 1998, 2002 and 2005. During the 1991-92 survey, the child's intake is based on quantitative food frequency questionnaire, with items derived from a list based on 24-hour food recalls from women in the sample.

Breast feeding Feeding patterns in the first year of life have probably the strongest impact on infant development (Morgan, 1999). Wrong timing in the introduction of a food or inadequate diet can cause development and health problems for the child.

There are some standard rules for the age of weaning: the WHO recommends exclusive breastfeeding for six months, introducing age-appropriate and safe complementary foods at six months, and continuing breastfeeding for up to two years or beyond. Weaning is one of the most crucial events in a child's diet.

Early lactation protects the infant from the environment both through the nutrients and the protective factors that breast milk provides. Besides the nutritional importance of human milk, breastfeeding can lengthen the duration of postpartum amenorrhea and thereby delay subsequent births. In developing countries, this is a common contraceptive method.

Breast milk contains several growth factors, although the physiological importance of these elements is not clear (Morgan, 1999).

Diseases The association between undernutrition and diseases is likely to be bidirectional. Thus, diseases not only affect nutrition, but undernutrition also predisposes to diseases (Victora et al., 1990; Walter et al., 1997).

Silventoinen (2003) lists the diseases that, in the literature, have been found to affect human growth. In particular, inflammatory diseases hinder the growth of long bones, diarrhea can lead to malnutrition, and pneumonia has been found to have similar effects on growth. Some evidence exists that diabetes and asthma could be associated with a lower growth rate, while infections may interact with nutrition and chronic diseases, including severe conditions such as congenital heart disease. All of these diseases are likely to have an influence on growth as well as on final body height.[42]

Infants' diseases Diarrhea is commonly the result of an infection, bacterial or viral. Weaning diarrhea occurs in a background of poor sanitation, when supplement foods are introduced in addition to breast milk and the infant is exposed to new organisms. Children who constantly suffer from diarrhea are unlikely to be well-nourished (Bozzoli et al., 2009). The consequences are dehydration, malabsorption of nutrients, poor appetite and loss of nutrients. If it is severe and protracted, it can lead to death. In fact it is still among the leading causes of death in early childhood in poor countries.[43]

Measles can cause serious complications, including blindness, encephalitis, severe diarrhea, ear infections and pneumonia, particularly in malnourished children.

Adolescents' diseases To identify the adolescents' diseases I do the following. I match the historical information in the data about a child's diseases with the diseases listed by Silventoinen (2003) that have been found to negatively affect height. To do this, I refer to the International Statistical Classification of Diseases and Related Health Problems of the World Health Organization.[44]

In particular, I distinguish the following groups of diseases:

1. Certain infectious and parasitic diseases (e.g., TB, polio, dengue, measles).
2. Endocrine, nutritional and metabolic diseases (e.g., diabetes).
3. Diseases of the respiratory system (e.g., pneumonia, asthma, weak lungs, tonsilitis).
4. Diseases of the digestive system (e.g., ulcers).
5. Congenital malformations, deformations and chromosomal abnormalities (e.g., heart disease).

[42] Skerry, 1994; Martorell et al., 1975; Rowland, Cole and Whitehead, 1977; Victora et al., 1990; Brush, Harrison and Waterlow, 1997; Vercauteren and Susanne, 1976; Tattersall and Pyke, 1973; Herber and Dunsmore, 1988; Thon et al., 1992; Wise, Kolbe and Sauder et al., 1992; Cole, 2000; Poskitt, 1993. For the complete list and explanation, see Silventoinen (2003).

[43] The first one is pneumonia, the second is diarrhea, the third is malaria, the fourth is measles and the fifth is AIDS/HIV (WHO, 2006).

[44] The version used is the 2007. For details, go to www.who.int/classifications/en/.

A.3 Inputs from conception to birth

The supply of nutrients to the fetus depends on the nutrient store of the mother, on what she eats during pregnancy, on her body size and composition, and on the transmission of nutrients through the placenta (Scott and Duncan, 2002). Furthermore, it is well known that particular vitamins and minerals, such as vitamin A, E, and iron, e.g., play an important role in fetal development.

Birth weight Birth weight is a trait that seems to be inherited probably through the characteristics of the maternal utero. This is the case of normal birth weight. A baby has low birth weight or is small for gestational age when his mass at birth is below a defined limit at any gestational age. The are many reasons for giving birth to a low-birth-weight infant: mother's malnutrition before and during pregnancy, smoking behavior of the mother during pregnancy, the alcohol consumption during gestation, or some maternal diseases (e.g., German measles).

Moreover, women who are small at birth have double the risk of having, in turn, babies of low birth weight, which implies intergenerational effects.

Pregnancy's length A premature birth has a significant impact on the baby's ability to metabolize nutrients. Most pre-term infants with regular weight for their age catch up perfectly, given the right environmental conditions (Tanner, 1990).

Birth order More mouths to feed and more ease in contracting infections due to the presence of siblings reduce the growth velocity of the child (Tanner, 1990).

B Instrumental variables

Table 13: Instrumental variables for infants' endogenous inputs.

Variable	Description	Obs Age 0	Mean Age 0	Std. Dev. Age 0	Obs Age 12	Mean Age 12	Std. Dev. Age 12	Obs Age 24	Mean Age 24	Std. Dev. Age 24
Community level										
Dirty roads	does barangay have dirt roads?	2918	.27	.44	2430	.28	.45	2284	.33	.47
Elevaton	brgy's aver. elevation above sea level	2918	21.81	48.16	2430	23.37	50.92	2284	23.40	51.34
Gravel roads	does barangay have gravel roads?	2918	.51	.50	2430	.51	.50	2284	.52	.50
Population density	population density of barangay	2918	13,513.82	16,958.64	2430	12,976.1	16,703.57	2284	13,033.29	16,907.86
% of electricity	approx % hhs in brgy with electricity	2693	76.21	17.87	2223	74.33	17.23	2079	74.01	17.49
Rating of water supply	informant rating of h2o supply in brgy	2918	1.83	1.07	2430	1.68	.84	2284	1.65	.81
Kerosene price	price of 1 lt of kerosene	1318	3371.23	563.61	1141	3,509.33	671.54	1071	3,313.94	702.15
Cooking oil price	price of 75cl cooking oil	1386	726.51	608.29	1141	869.47	693.79	1084	556.70	290.94
Evaporated price	price of 100 g evaporated milk	1,383	749.24	121.08	1126	901.54	96.83	1047	937.53	116.34
Formula price	price of 400 g formula milk	1103	11,996.6	1413.09	769	11,559.33	1,057.68	796	11659.28	1466.84
Banana price	price of 1 banana	1198	149.48	38.79	1030	121.89	27.01	869	112.55	24.04
Corn price	price of 1 kg corn	1370	1521.85	140.14	1083	1,686.05	171.15	1010	1,800.28	104.37
Powdered price	price of 350 g powdered milk	1195	,8349.68	765.85	946	8514.50	539.79	957	8457.74	605.64
Cereal price	price of 350 g cereal	1061	669.37	409.95	669	7180.32	550.19	791	7874.07	550.38
Rice price	price of 1 kg rice	1345	1,991.60	93.43	1058	2103.11	149.73	1005	2197.84	122.41
Pork price	price of 1 kg pork	590	11,986.14	1,210.82	391	12,864.84	1,324.99	463	12014.89	391.52
Fish dry price	price of a piece of dried fish	1243	212.29	60.68	1131	251.74	111.56	1066	253.89	138.39
Egg price	price of medium size egg	1325	444.79	37.81	1058	478.36	33.67	1008	468.90	31.70
Household level										
Piped water	1 if the household has piped water	2797	.86	.35	2420	.85	.36	2193	.85	.36
Season	1 if the season is wet	2918	.58	.49	2539	.58	.49	2468	.58	.49
Only baseline info										
Distance to road	distance (m) to nearest vehicular road	2918	448.42	2,082.46	2539	465.57	2,109.45	2468	483.65	2,150.33
Refrigerator	do you/hh own a refrigerator?	2918	.07	.26	2539	.066	.25	2468	.07	.25
Walk to canned store	mins walk to canned/powdered milk store	2918	26.24	64.91	2539	26.63	65.08	2468	26.70	65.13
Walk to formula store	mins walk to nearest bb formula store	2918	65.44	100.25	2539	67.11	100.95	2468	67.19	100.94
Walk to cereal store	mins walk to nearest bb cereal store	2918	83.17	109.44	2539	85.47	110.16	2468	85.36	110.08
Travel time to priv hf	mins walk to nearest private health facility	2918	208.90	383.39	2539	205.025	380.27	2468	202.83	378.37
Travel time to pub hf	mins walk to nearest public health facility	2918	37.29	130.29	2539	37.77	129.83	2468	38.02	130.15
Travel time to trad hf	mins walk to nearest traditional health facility	2918	150.74	331.21	2539	138.53	317.11	2468	133.59	311.72

All the prices are monthly and deflated. I report the mean prices that correspond to a specific age of the infant.
The temporal window for the infants is 2 years and I suppose that there have been not relevant changes in the household's characteristics measured only in the baseline.
There are two community surveys for infants, collected in 1983 and 1986. For the central years I use the mean of the two measurements per each barangay's variable.

Table 14: Instrumental variables for teenagers' endogenous inputs.

Variable	Description	Obs Age 8.5	Mean Age 8.5	Std. Dev. Age 8.5	Obs Age 21.5	Mean Age 21.5	Std. Dev. Age 21.5
Community level							
Population density	population density of barangay	1857	16,072.65	20,750.42	386	12,835.03	20,219.48
Gravel roads	does barangay have gravel roads?	1867	.29	.45	386	.23	.42
Dirty roads	does barangay have dirt roads?	1867	.52	.50	386	.27	.45
Health priv	n° of private health facilities in the brgy	1867	.62	.56	386	2.14	3.23
Health pub	n° of public health facilities in the brgy	1867	3.18	1.68	386	6.72	5.53
Elevaton	brgy's aver. elevation above sea level	1946	21.17	43.90	1444	20.58	41.73
Cooking oil price	price of 75cl cooking oil	1867	3.07	1.44	1886	2.75	.47
Banana price	price of 1 banana	1498	1.80	.32	1731	1.37	.30
Corn price	price of 1 kg corn	1788	18.47	1.77	1872	14.54	.99
Rice price	price of 1 kg rice	1492	21.04	1.48	1717	20.89	.79
Pork price	price of 1 kg pork	615	131.23	6.63	1343	105.51	4.83
Chick price	price of 1 kg chicken	638	111.94	4.18	1065	79.01	6.67
Beef price	price of 1 kg beef	576	159.28	8.10	865	114.06	7.10
Cabbage price	price of 1 cabbage	1278	24.33	13.38	1250	9.97	3.46
Egg price	price of medium size egg	1758	4.15	.40	1888	3.00	.32
Salt price	price of 1 kg of salt	1846	13.77	24.54	1841	11.20	6.10
Household level							
Season	1 if the season is wet	2271	.49	.50	2271	.33	.47
Location water	location of drinking water source	2251	.21	.40	1545	.46	.50
Toilet	toilet facilities in the house or garden	2251	.57	.49	1871	.82	.38
Garbage	garbage close to the house	2251	.29	.45	1871	.62	.48
Walk store	mins walk to food store	2249	44.98	65.12	1870	36.38	46.95
Houses	n° of houses close to the house	2251	16.82	5.59	1871	16.91	5.59
Distance of house to road	distance (m) to nearest vehicular road	2242	3.47	8.67	1871	2.56	5.87
Electric	electricity in the house	2251	.74	.44	1871	.94	.23
Refrigerator	refrigerator in the house	2251	.22	.41	1871	.42	.49
Piped water	1 if the household has piped water	2251	.83	.37	1871	.72	.45

All prices are monthly and deflated. I report the mean prices that correspond to a specific age of the teenager.

Table 15: Summary results for first-stage regressions. Height production function of an infant 24 months old. Naive specification with instrumental variables.

Endogenous variables	Shea Partial R2	Partial R2	F(31,1163)	P.Value
Calorie Intake at 12m	0.013	0.144	5.890	0.000
Calorie Intake at 8m	0.005	0.126	5.810	0.000
Calorie Intake at 4m	0.005	0.089	3.310	0.000
Breast milk at 20m	0.019	0.041	1.710	0.010
Breast milk at 16m	0.014	0.060	3.000	0.000
Breast milk at 12m	0.011	0.066	3.070	0.000
Breast milk at 8m	0.020	0.084	3.860	0.000
Breast milk at 4m	0.038	0.072	2.710	0.000
Calorie Intake*Breast milk at 20m	0.006	0.024	1.260	0.159
Calorie Intake*Breast milk at 16m	0.016	0.036	1.980	0.001
Calorie Intake*Breast milk at 12m	0.006	0.029	1.250	0.165
Calorie Intake*Breast milk at 8m	0.008	0.037	1.490	0.042
Calorie Intake*Breast milk at 4m	0.020	0.076	2.820	0.000
Disease at 20m	0.011	0.036	1.920	0.002
Disease at 16m	0.010	0.033	1.270	0.146
Disease at 12m	0.013	0.032	1.320	0.113
Disease at 8m	0.013	0.040	1.850	0.003
Disease at 4m	0.012	0.029	1.460	0.049
Disease at birth	0.013	0.023	0.530	0.984

Included instruments: Age; Age Sq; Birth Order; I-Urban; Birth weight; I-Premature&Small-for-dates; I-Premature; Mother's height

Excluded instruments: Kerosene price at 20m, Cooking oil price at 16m, Evaporated milk price at 12m, Formula price at 8m, Banana price at 4m, Corn price at birth, Powdered milk price at 20m, Cereal price at 16m, Rice price at 12m, Pork price at 8m, Dried fish price at 4m, Egg price at birth, Dirt road at 20m, Elevaton at 16m, Pop density at 8m, % of houses with electricity at 4m, Water sup at birth, Water sup at 20m, Distance road at birth, Refrigerator at birth, Walk cann at birth, Walk formula at birth, Walk inft at birth, Time priv hf at birth, Time pub hf at birth, Time trad hf at birth, Father's age at birth, Med insurance at birth, Religion at birth, Season at 16m, Gravel road at 12m

Underidentification tests

Ho: matrix of reduced form coefficients has rank=K1-1 (underidentified)
Ha: matrix has rank=K1 (identified)
Kleibergen-Paap rk LM statistic Chi-sq(11)=2.70 P-val=0.9941
Kleibergen-Paap rk Wald statistic Chi-sq(11)=2.68 P-val=0.9943

Weak identification test

Ho: equation is weakly identified
Kleibergen-Paap Wald rk F statistic 0.084

Weak-instrument-robust inference

Tests of joint significance of endogenous regressors B1 in main equation
Ho: B1=0 and overidentifying restrictions are valid
Anderson-Rubin Wald test F(31,1163)=3.22 P-val=0.0000
Anderson-Rubin Wald test Chi-sq(31)=103.17 P-val=0.0000
Stock-Wright LM S statistic Chi-sq(31)=86.96 P-val=0.0000

Overidentification test of all instruments

Hansen J statistic: 7.445
Chi-sq(10) P-val = 0.6829