The height production function
from birth to age two

Elisabetta De Cao[*]

Centre for Health Service Economics & Organisation, University of Oxford
New Radcliffe House (2nd floor), Walton Street, Oxford, United Kingdom
(e-mail: elisabetta.decao@gmail.com; elisabetta.decao@phc.ox.ac.uk)

Abstract

Infancy is one of the most critical periods for the formation of adult height. This paper studies the determinants of height from birth to age two using rich longitudinal data on Filipino children. A height production function is specified where height is the result of the accumulation of inputs (i.e., nutrition and diseases) over time. The empirical specification allows the causal identification of the age specific effects of both nutrition and diseases on height. Considering gender differences in growth patterns, the results show that diseases play a major role in reducing height, and that girls are more strongly affected than boys.

JEL-Classification: I10; I12; O15; C13.

Keywords: height; early-life events; health production function; Philippines.

---

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, the relative risk of survival, and the labor productivity (Fogel, 1986). In particular, a seminal work by Case and Paxson (2008) explains the positive correlation between adult height and labor productivity by showing that height is positively associated with cognitive ability. The authors show that both cognition and height are driven by early childhood investments; therefore, cognitive achievements are correlated with height and wages are affected by cognitive skills. Poor health can explain both low height and low labor productivity. This is more evident in developing countries where living conditions are poor (see for example, Behrman and Deolalikar, 1989; Haddad and Bouis, 1991; Thomas and Strauss, 1997; Croppenstedt and Muller, 2000; Schultz, 2002; Dinda and Gangopadhyay, 2006; Price, 2013). Therefore, it is necessary to investigate the factors driving height, since understanding the determinants of height is important in understanding health (Deaton, 2007). The correlation between final height and height at birth is between .25 and .30, between .70 and .80 with height at age two, and then it increases slowly after age two (Schmidt et al., 1995). Therefore, a child's early nutrition and health conditions are critical in explaining adult height.

In this paper, I study the determinants of height from birth to age two in a developing country. To do that, I build and estimate a height production function.

To motivate the specification of a height production function, I follow Steckel's (2009, pg 7-8) 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?" My intention is to find the technological parameters that answer the previous questions.

This paper uses part of the Cebu Longitudinal Health and Nutrition Survey (CLHNS), which is a rich longitudinal survey of a cohort of Filipino children that are followed every two months for the first two years of life.[1] The data allow the derivation of a height production function from birth to age two.

In particular, I study height as the result of the accumulation of several factors

---

[1]Further waves have been collected from age 8 until age 25-26 of the child. These data have not been used in this paper.

over time, identifying the direct effects of its determinants.[2] The determinants of height can be divided into non-genetic factors, genetic factors, and the age at which height is measured. The principal non-genetic factor is net nutrition, which is the difference between food intake and the losses to activities and to diseases (Eveleth and Tanner, 1991). In developed countries, there is evidence that genetic factors explain 80 percent of the variation in adult height and the rest is due to non-genetic factors (Silventoinen, 2003). 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). Therefore, the interplay between nutrition and diseases, and the understanding of which are the critical growth periods, become crucial.

Different papers have used the CLHNS data to study the determinants of infants' health outcomes. An important paper by Adair and Guilkey (1997) studies the association between health inputs and child's stunting. The authors analyse the determinants of stunting in 2-year-old children, and show that stunting is positively associated to diseases, early supplemental feeding and low birth weight, while negatively associated with mother's height, breast-feeding, and preventive health care. However, the endogeneity of these health inputs is not considered and it likely leads to biased results. Two papers address this problem. The first paper by Cebu-Study-Team (1992) estimates four health production functions for the following outcomes: gestational age, weight, diarrhea, and respiratory infection. They find that indi-

---

[2]Todd and Wolpin (2003, 2007) consider different methods for modeling 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. They consider different specifications of the skill production function that rely on different assumptions and data limitation. I follow the same approach to study the process of height formation and I clearly explain the assumptions made to identify the technological parameters of the height production function.

vidual, household, and community factors affect the outputs considered. However, Cebu-Study-Team (1992)'s paper focuses on the first year of an infant's life and does not consider height as an outcome. A second paper by Liu et al. (2009) considers infants from birth to age two. 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. Both papers include in the empirical model lagged values of the outcome instead of the historical inputs.[3] This paper differs from the others because it clearly includes in the production function only the determinants of height, mainly past and present net nutrition factors.

There is extensive research that demonstrates the importance of early childhood investments for child health, growth, skills development, and labor outcomes later in life (see for example, Glewwe and King, 2001; Schultz, 2002; Cunha and Heckman, 2008; Maluccio et al., 2009; Almond and Currie, 2011). In particular, Glewwe and King (2001) uses the CLHNS data and shows that malnutrition in the second year of life is critical in determining later cognitive development. Glewwe and King (2001) does not focus on the determinants of growth, but considers changes in height[4] as well as birth weight as endogenous nutritional variables that explain the children's IQ score. This paper complements Glewwe and King (2001)'s work, showing that

---

[3]When data on past inputs are missing, the use of the lagged outcome is quite common, and since the lagged outcome is correlated with the shock by construction, an additional lagged outcome measures can be used as instruments to address endogeneity.

[4]The authors consider the change in height between 0 and 6 months, 6-12 months, 12-18 months, 18-24 months and 2-8 years.

infancy is a critical period because of the relative importance of diseases versus nutrition. The analysis shows that the magnitude of the input effects during the second year of life is the highest, especially for girls. This might be important to design policy interventions that target individuals early in life to improve their health and potentially their socio-economic outcomes later in life.

The paper is structured as follows. In Section 2, I develop a model for studying the process of height formation, and I present the empirical specification. Section 3 presents the data and a detailed description of the variables used. Section 4 describes the empirical results. In Section 5, I present some robustness checks. Finally, Section 6 concludes.

2   The height production function

In this section, I present a model for the height production function. I am interested in technological parameters, such as the effect of an exogenous change in one input, while 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 choosing some of the inputs, since 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. Therefore, 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.

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,

- $\mathcal{X}_{i,t} = (X_{it}, X_{it-1}, \ldots, X_{i1}, X_{i0})$ the vector of inputs for child i from birth to age t,

- μ_i the child's biological endowment,

- ϵ_{it} a shock to the height production for child i at age t.

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

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

where the inputs $\mathcal{X}_{i,t}$ are nutrition and diseases.

To study empirically the height production function, I make different assumptions.[5]

---

[5]See Todd and Wolpin (2003, 2007) for a detailed description of different specifications of a skill production function for children and the assumptions made for the empirical specifications. 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.

(A) I assume that the child's biological endowment is determined at conception and it is constant over time.[6]

(B) I assume that the height production function is linear in the inputs and in the unobserved endowment, and that the effects of the inputs depend on the child's age. The true technology that links inputs and output is unknown. This functional form implies that there is not complementarity, which 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 various influences on the adult outcomes.[7] Hence, I obtain the following model:

$$H_{it} = f(t) + X_{it}\beta_t + X_{it-1}\beta_{t-1} + \cdots + X_{i1}\beta_1 + X_{i0}\beta_0 + \mu_i + \epsilon_{it}. \tag{1}$$

(C) I also assume that the time-varying coefficient β_s $(s = 0, 1, \ldots, t)$ depends only on the child's age s. For example, the effect of the diseases experienced at birth on height at age one may be different than the effects on height at age two. [8]

---

[6]Case and Paxson (2008) 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.

[7]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.

[8]In a recent working paper, Griffen (2014) estimates a height production function using data from Guatemala on children up to age 7 years. He focuses on the impact of caloric intake on height, and controls for the measurement error in the caloric intake. His model relies on two extra

2.1 Empirical specification

In order to estimate (1), I consider a within-child fixed effects specification (FE).[9] This specification is feasible because the children are observed more than once, and several outcome and input measurements are available.[10] In particular, consider differencing (1) by age:

$$\Delta_{it} = H_{it} - H_{it-1} = f(t) - f(t-1) + X_{it}\beta_t + \epsilon_{it} - \epsilon_{it-1} \tag{2}$$

The βs parameters resulting from the above equation (2) are the specific input effects for the inputs applied between the two periods. The age trend is expressed as a linear and a quadratic term.

The within-child fixed effect estimator eliminates the endowment from equation (2), dealing with the endowment heterogeneity. However, there might be potential endogeneity of the nutrition and disease inputs. The fixed effects allow a permanent

different assumptions: both contemporaneous and lagged inputs have constant effects by age. The author does consider diseases in his production function, but he does not address their endogeneity. Puentes et al. (2014) use the CHLNS and data from Guatemala to estimate weight and height production function on children from 6 to 24 months. They specify a production function that allows the past inputs to change over time. However, they do not address the endogeneity of breast milk and diarrhea. Moreover, both Griffen (2014) and Puentes et al. (2014) consider boys and girls together and add a gender dummy in their models. However, since the growth patterns differ between boys and girls, in this paper I treat them separately.

[9]Cebu-Study-Team (1992) and Liu, Mroz, and Adair (2009) use the same data to estimate different health production functions. They adopt a specification that includes lagged values of the outcome in the model instead of the historical inputs. In the cognitive skills literature this specification is called the "value added" specification. 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.

[10]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.

change in the inputs. On the other hand, contemporaneous inputs could respond to previous shocks causing endogeneity because they are correlated to unobserved parental preferences regarding their children's nutrition and preventative care. If, for example, a child is very small at a certain point in time, and the parents give him/her more food to help his/her growth, it is not captured by the fixed effect and produces endogeneity.

I address endogeneity of both nutrition and diseases by using variation in village-level food prices, household characteristics, and climatic shocks as instrumental variables (IV) to estimate the production parameters via IVFE. A second motivation for the use of IVs is that both nutrition and diseases are measured with error.

The within-child fixed effect estimator assumes that differenced omitted inputs are orthogonal to the differenced included inputs or that omitted inputs are constant over time and the fixed effect estimators eliminate them. In the Appendix, I report the estimates of a hybrid production function where I include family income as a proxy for the time-varying omitted variables.

3   Data

The country of interest is the Philippines, and in particular, the Metropolitan Cebu or Metro Cebu. Cebu is a province in the Philippines and it consists of Cebu Island and 167 surrounding islands.

The CLHNS is a longitudinal survey of a cohort of Filipino women who gave birth between May 1, 1983 and April 30, 1984.[11] A stratified and single stage sam-

---

[11]For more information about the project and to download the data, visit

pling procedure was used to randomly select 33 communities or barangays from the Metropolitan Cebu. 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 two 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 in the first two years of life of the children.

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. The data spans over 20 years and covers issues such as health, nutrition, water quality and sanitation. It contains detailed information about the mothers' health and behavior during pregnancy, such as health care practices or smoking behavior, children's education, household and individual economic situation, demographic information, family planning, intra-household relationships, and reproductive health.

Given that the data lack information from age three to age seven, and that infancy is a critical period for the formation of height, I focus on the first two years.

Of special interest for my study is the rich collection of anthropometric measurements from birth to age two, as well as the complete disease and nutrition informa-

http://www.cpc.unc.edu.

11

tion. Since the data have information at the individual, household, and community levels, it is possible to study the long-term effects of prenatal and early childhood nutrition and health on later adult outcomes, matching physical and socio-economic information.

The data used is composed by 13 waves collected during infancy. It is important to notice that individuals are not surveyed at the same age, there are differences of several days.[12] Table 1 reports the children's age at the time of the different follow-ups.

Insert table 1 here.

The CLHNS is not a representative sample of the Philippines population, nor is it of all Cebu because of the criterion of selection based on fertility. However, Mendez and Adair (1999) find that the sample is representative of the women who were married with at least one child in the early '80s.

The *outcome* variable for this paper is raw height reported in centimeters. Height and weight were measured every two months for the first two years of life by the field staff in Cebu. Specialists took the measurements, which is a great advantage compared to the self-reported heights common to many datasets. Reliability checks were made to avoid heaping and other errors in the measurements.

The inputs of the height production function refer to the infancy period, and as previously specified, the most relevant non-genetic inputs are nutrition and diseases.

---

[12]This is controlled for in the analysis, by considering the difference in days between consecutive waves.

3.1   Inputs

The data are collected bimonthly from birth to age two. I aggregate the inputs between birth and age one year and between age one and two.[13] In particular, I consider *caloric intake*, which is a good aggregate indicator of nutrition, even if it does not capture the role of micronutrients. The CLHNS data provide precise information about the individual's diet based on 24-hour dietary recalls or a quantitative food frequency questionnaire. Daily energy intake is calculated from 24-hour dietary recalls during the surveys from birth to age two years.[14] The caloric intake for infants is exclusive of breast milk. I compute the average caloric intake in the first and second year of life. Since this does not entirely capture the infant's nutrition, I also consider breast-feeding.

Breast-feeding has been found to improve both cognitive ability and adolescent health and, therefore, positively affects long-term academic achievement (Rees and Sabia, 2009). Belfield and Kelly (2012) finds that breast-feeding for at least 6 months instead of formula feeding at birth is negatively associated with obesity and positively associated with cognitive performance. In the analysis, I consider if the child was breastfed in the first and second year.[15]

As for the diseases, I consider if the infant had feeding problems in the few

---

[13]Glewwe and King (2001) study the effects of malnutrition during infancy on children's cognitive development using the CLHNS data. They also aggregate the inputs over the first and second year of life, and in a second specification over 6-month periods. The shortest periods produce less precise estimates because they require an increase number of IVs, and that also apply to my analysis.

[14]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.

[15]I combine two questions: "Was breast milk given to infant yesterday?" and "Was breast milk fed to infant seven days ago?". The child is considered breastfed if the answer is yes to at least one of the questions.

hours after birth (baseline or wave 0) and diarrhea episodes later on (waves 1 to 12). In fact, some diseases reduce the absorption of nutrients, prevent food intake, produce nutrient losses, or increase metabolic requirements (Stephensen, 1999). In particular, I compute the total number of times the infant had feeding problems at baseline or experienced diarrhea episodes in his/her first and second year of life. For simplicity, I will refer to these infant diseases as diarrhea episodes.[16] This is likely an underreporting of the total number of diarrhea episodes experienced by the child.

Due to the different growth patterns of boys and girls, I estimate the production function by gender.[17]

Table 2 reports descriptive statistics for the main variables, by gender and age. The table also reports also some other inputs that affect the child's development: genetic and environmental inputs, as well as inputs from conception to birth. A proxy for the genetic inputs is *mother's height*.[18] I assume that the rest of the genetic impact is captured by the individual's biological endowment included in the model. It represents the genetic inheritance and gene-environment interactions that are unobserved factors (Case and Paxson, 2008). An extra variable that captures the environmental inputs is the *location* of the household, and in particular, the

---

[16]During the baseline there is a question: "What are the infant's health problems affecting feeding?", and I indicate as 1 if the infant has at least one the problems. In every wave from 1 to 12 there is a yes/no question: "Has the infant had diarrhea during the past seven days?".

[17]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).

[18]Many medical papers suggest that approximately about 60 to 80 percent of height variation in a population depends on genetic factors, but it is not clear what is the underlying process (see, e.g. Ginsburg et al., 1998; Silventoinen, 2003) nor is the relationship between genetics and environmental factors clear. The data do not contain father's height.

percentage of time the child has lived in an urban area from conception to age two. Moreover, table 2 shows descriptive statistics of inputs from conception to birth: the infant's *birth weight*, the *duration of the gestation* (a categorical variable indicating whether the child had normal weight and normal term, low birth weight for his gestational age, or pre-term but with normal weight for his gestational age), and the *birth order*.[19] Both the inputs from conception to birth and the genetic and environmental inputs are time invariant variables that are not identified by the fixed effect estimators. Table 2 shows that there are no relevant differences between boys and girls at birth: The birth weight is on average about 3 kg, 88 percent of the pregnancies have normal length, and the birth order is 2.5 for the girls and 2.6 for the boys. The percentage of time spent in an urban location is about 75 percent. During the first year of life, about 75 percent of the boys are breastfed compared to 77 percent of the girls, while there is a decrease in the second year to about 38 percent. The average caloric intake exclusive of breast milk is slightly higher for the boys in the first year (331 kcal versus 294 kcal for the girls), and it is the same for boys and girls in the second year (670 kcal). The number of times the infant experienced diarrhea are on average 0.65 in the first year for the boys and 0.61 for the girls. The diarrhea episodes increase on average to 0.92 in the second year of life for the boys and 0.78 for the girls.

---

[19]Many researchers suggest that growth *in utero* may play an important role in determining health in adult life (Barker, 1998). The importance of the birth weight is well known and there is a huge literature about it in medicine as well as in economics (e.g.Rosenzweig and Schultz (1983); Behrman and Rosenzweig (2004)). The problems of prematurity are very similar to those of low birth weight. Birth order has also been found to be a significant and independent predictor of adult height (Steckel, 1995). First-born children are, during childhood, taller than children born later, since they have had a period in which they were alone. These inputs from conception to birth are not exactly inputs, but the results of pre-birth inputs that are not available (e.g. birth weight).

Insert table 2 here.

3.2 Instrumental variables

Valid instruments must be uncorrelated with height and correlated with one or more of the endogenous variables (nutrition - caloric intake and breast milk - and diseases). Therefore, I need to find instruments that only operate through their impact on nutrition and diseases. The IVs that would seem to satisfy these requirements can be categorized in three groups: food local prices, household characteristics, and climatic shocks.[20] [21]

In particular, I consider local prices of the main food items (e.g. egg, banana, powder milk, evaporated milk, kerosene.)[22] [23] It is difficult to imagine how these prices could be correlated with height; hence, they should be uncorrelated with the error term in the second stage regression.[24] The food prices of the major food items are expected to be negatively correlated with food consumption. The caloric intake is exclusive of breast milk; therefore, most of the caloric intake reported for infants in their first year corresponds to breast milk substitutes. If the price of the formula goes up, then the mother might prefer to continue breast-feeding.

The household's characteristics are presence of an infant store in the neighbor-

---

[20]Some of the IVs used in this paper have been used in other studies (see for example, Cebu-Study-Team, 1992; Glewwe and King, 2001; Liu et al., 2009; Ugaz and Zanolini, 2011).

[21]The choice of the IVs has also been made avoiding the presence of instruments highly correlated or the problem of missing IVs for many observations.

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

[23]Note that kerosene is used for cooking.

[24]Using food prices as instruments is commonly done when estimating production functions (Todd and Wolpin, 2003).

hood, availability of piped water as the water source for the house, possession of a refrigerator, minutes walk to the nearest infant store, and distance to the nearest vehicular road. One can think that a family's location decision (where to live in the village) or what access to purchases depends on the child's height. To address this potential endogeneity of the household's characteristics variables, I average them over the same village of current residence and I use these averages as IVs. Sanitation variables (e.g. piped water) may affect the child's chances of coming into contact with pathogens, and if a household owns a refrigerator, it increases the possibility of preserving food. Accessibility to stores and roads are correlated with accessibility to food.

I consider two climatic instrumental variables. The first is season, and it is a dummy variable that indicates if the survey falls in the rainy season. This variable may be important for diseases: for example, in the case of diarrhea, extensive rainfall can contaminate the water supply with fecal pathogens. The second variable is the effect of one of the strongest typhoons, Nitang, which hit Cebu on September the 2nd, in 1984 and killed about 1,500 people.[25] Extreme weather conditions, such as typhoons, are associated with problems of malnutrition and waterborne diseases, and may also disrupt income-earning activities. Ugaz and Zanolini (2011) use the CLHNS data to investigate whether the exogenous weather shock caused by this typhoon had an impact on children's anthropometric outcomes later in life. They look at infants and at the effects of the typhoon during pregnancy and right after birth.[26] Variation in the exposure to the typhoon comes from the dates of birth of

---

[25]See http://www.pagasa.dost.gov.ph for details.

[26]Ugaz and Zanolini (2011) also consider the subsample of siblings available in the data. In

the children. This spread in the dates of birth offers an exogenous type of variation to exposure to the typhoon. Ugaz and Zanolini (2011) find that the likelihood of reporting diarrhea right after Nitang is greater. The typhoon hit the island when the children were between 4 and 16 months old, therefore, I consider as IV a dummy variable equal to one if the child was exposed to the typhoon in her/his first year of life, and zero otherwise.

The IVs are time-variant variables computed at the time of each interview. They are averaged over year one and year two.[27] I assume that all the IVs affect children at both ages. There is one variable that affects children only at age one: the typhoon. I then estimate as many IVFE specifications as many combinations of the IVs (combinations of 12 IVs for year one, and combinations of 11 IVs for year two). However, in the paper I present the results based on the best four sets of IVs that satisfy the quality tests (robust versions of the under and over identification tests (Kleibergen and Paap, 2006)) and have the highest Kleibergen-Paap Wald F-statistic.

Table A1 in the Appendix reports the descriptive statistics of the IVs considered. Section 5 reports detailed robustness analysis of the IVs.

4   Empirical results

The empirical results are shown in Tables 3-6. The tables report both the FE and IVFE estimates for each model. In the IVFE specifications, I consider both nutrition and diseases as endogenous inputs. The estimation of each change in height allows

---

particular, they study the younger siblings that were exposed to the typhoon during the pregnancy.

[27]The rainy season variable indicates the number of rainy season months experienced during year one or two. The variation in this variable is due to the different interviews dates.

me to derive the effect of all the inputs. In particular, I estimate the following equations to identify all the β's contained in model (1):

$$\Delta_{i1} = H_{i1} - H_{i0} = \Delta f(1) + X_{i1}\beta_1 + \epsilon_{i1} - \epsilon_{i0}$$
$$\Delta_{i2} = H_{i2} - H_{i1} = \Delta f(2) + X_{i2}\beta_2 + \epsilon_{i2} - \epsilon_{i1}$$

Insert table 3 here.

Insert table 4 here.

Insert table 5 here.

Insert table 6 here.

The FE provides evidence on the effect of exogenous input variables on height. The IVFE instead allows a causal interpretation of the effects of nutrition and diseases on height. One difference between FE and IVFE results is the magnitude of the diarrhea coefficients, where the effect is negative and larger in the IVFE specification than in the FE, which results in nutrition inputs that, in most of the cases, lose their statistical significance.

Tables 3-6 also report on the quality of the instruments used. In all models, the under identification test (Kleibergen-Paap test) rejects the null hypothesis, indicating that the models are identified and the excluded instruments are relevant and correlated with the endogenous variables. I also check the validity of the instruments. The estimates, in fact, always satisfy the over identification test (Hansen's J statistic), where the null hypothesis is never rejected, suggesting that the instruments are not correlated with the error term. In all models, the instruments have strong

19

predictive power for both nutrition and disease, as revealed by the F-test statistics (bottom of Tables 3-6), even if the F-test statistics are always higher for the nutrition inputs. Tables A2-A5 in Appendix report the first stage regression results for each endogenous variable in each model. The estimates present the expected signs as described in the Instrumental variables section.

To describe the second stage results, I mainly focus on the IVFE results that address the endogeneity problem. The effects of the different inputs vary according to the age of the child, and the magnitude and timing differ between boys and girls. The IVFE results are based on different sets of IVs that lead to very similar coefficients estimates.

In particular, once controlled for endogeneity, the inputs applied in the second year of life are significant and larger than the ones applied in the first year. This confirms the results found by Glewwe and King (2001) that using the same data underline the relevance of malnutrition on cognitive development in the second year of life. Once controlled for endogeneity, the importance of nutrition, both breastfeeding and caloric intake, vanishes, except at age two for boys where caloric intake is positively and statistically significant. Instead, the negative effect of diarrhea, instead, increases using the IVFE and remains statistically significant, with a higher magnitude for girls.

More specifically, diarrhea has a significant and negative impact on infants' height, except on boys of age one. If experienced in the first year of life, the results indicate that an increase by one of the diarrhea episodes decreases height of a girl by about 2.3/2.5 cm. In a two years old boy, an increase in the caloric intake

by 100 kcal results in an increase in height by about 0.3 cm. Experiencing an extra episode of diarrhea in the second year of life, reduces the height of a boy by 2/2.2 cm and the height of a girl by 3.2/3.5 cm.

Overall, these results show that growth in infancy is critical because diseases play a major role, and girls are more strongly affected.

5   Robustness

5.1   Selection of instrumental variables

In this section I present analysis that shows that my findings are robust to changes in the instrument set (Puentes et al., 2014).[28]

I consider all the combinations of the IVs, and estimate one IVFE specification for each combination. Since there are three endogenous variables, every specification estimated contains a minimum of three IVs to exactly or over identify the empirical model. See Table A1 for a detailed description of the IVs considered.

To estimate the height production function at age one, I consider 12 IVs, and therefore 4,017 available combinations.[29] At age two, I consider 11 IVs, producing 1,981 available combinations.[30] Out of these 4,017 (1,981) combinations for children aged one (two), I select the sets of IVs for which the IVFE specifications satisfy both the weak instrument test (Kleibergen-Paap) and the over-identification test (Hansen

---

[28]Further analyses are not reported, but are available upon request.

[29]In total there are $(2^{12} - 1) = 4,095$ combinations, but I exclude (12+66) combinations/sets, that contain one or two IVs and cannot identify my empirical model, to leave 4,017.

[30]In total there are $(2^{11} - 1) = 2,047$ combinations, but I exclude (11+55) combinations/sets, that contain one or two IVs and cannot identify my empirical model, to leave 1,981.

J).[31] This leaves 870 (302) IV sets for boys at age one (two), 111 (437) for girls at age one (two) and 25 (103) contemporaneously for both boys and girls at age one (two).

There is no single set of IVs that pass both the endogeneity and weak instrument tests for all four groups of analysis (boys at age one and two, girls at age one and two). However, there are different sets of IVs that satisfy the tests both for boys and girls at age one (25 sets), and at age two (103 sets). The sets of IVs slightly differ between years one and two because an extreme weather shock is considered as an IV only at age one (typhoon) and because there is a big difference in the nutritional inputs. In year one, most of the children are breastfed, but the proportion breastfed in year two is much lower. Therefore, IVs do not have the same predictive power across different ages and differently explain the endogenous inputs. Some IVs are more relevant for the weaning period, which is predominant in the second year of the child's life (for example, egg and kerosene prices).

Tables 3-6 present the results based on the four best sets of IVs for boys/girls of age one, selected from the 25 sets that satisfy the tests both for boys and girls at age one; and the four best sets of IVs for boys/girls of age two, selected from the 103 sets that satisfy the tests both for boys and girls at age two. These best IVFE specifications are those with the highest Kleibergen-Paap Wald F-statistic both for boys and girls. At age one, the best four models have a Kleibergen-Paap Wald F-statistic greater than 1.89, while at age two the best four models have a Kleibergen-Paap Wald F-statistic greater than 2.4.

---

[31]In specifications with three instruments, I only consider the under-identification test.

To show that the results (Tables 3-6) are robust to different sets of IVs, I report the distribution of the coefficients estimates. Figures A1-A4 show the coefficient point estimates (and 95% C.I.) for caloric intake, breast milk, and diarrhea when different sets of IVs are used that satisfy the tests for both boys and girls at age one (25 specifications), and at age two (103 specifications).

These results indicate that: caloric intake tends to be positively and significantly associated with height at age one for girls (in 20% of the specifications), and at age two for boys (in 53% of the specifications); Breast milk tends to be negatively and significantly associated with boys' height at age one (in 4% of the specifications), and positively, and in some specifications significantly associated with girls' height at age one (in 4% of the specifications), but breast milk is never statistically significant at age two neither for boys nor girls; Diarrhea is always negatively and significantly associated with height both at age one and age two, except for boys at age one where the diarrhea coefficient estimate is never statistically significant.

5.2 Omitted variables bias

To account for the omitted variables bias, I estimate a *hybrid production function* that includes in the empirical specification (both IV and IVFE) the annual household income in pesos.

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.

I test the possibility of omitted variables bias by looking at the coefficients of household income in the hybrid production functions. The hybrid production functions are reported in the Appendix, on Tables A6-A9. Overall, a comparison of Tables 3-6 and Tables A6-A9 shows that most of the estimated input effects are very similar across the non-hybrid and hybrid specifications.

Once controlled for endogeneity, the effect of family income is never statistically significant except in the one-year-old boys' models (Table A6). An increase of income in the first year of life of the child by 1000 pesos increases height by approximately 0.001 cm. This can be an indication of omitted variable bias in the one-year-old estimates for boys. However, the standard deviation of the income variable is quite high (Table 2), and it is likely that its inclusion does not satisfactorily address the omitted variables problem. Moreover, once income is included, 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).

5.3   Sample selection bias due to attrition

The initial sample consists of 3080 children (Table 1), but by the end of the second year, around 20 percent of the children are lost because of attrition.

The two main reasons for attrition are death and migration. 167 (5.4 percent of the sample) children die in the first two years of life. The remaining 14.6 percent are mainly lost because of migration.

It seems that the people who died tended to be shorter and in poor health.

Unfortunately, there are not plausible 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. Attrition due to mortality is claimed not to represent a big problem because only a small proportion of children in the poorest health conditions are lost (this is also claimed by Eckhardt et al., 2005, who uses the same data). If a selection mechanism is in place so that only the healthiest survive, then my estimates would be a lower bound of the true effect.

As for migration, Cebu-Study-Team (1992) tested for selectivity of infants and the results show that the omitted variables that influence migration decisions do not coincide with those that determine child health.

6 Conclusions

In this paper, I study the determinants of height by building a height production function from birth to age two. I consider the cumulative nature of physical development, taking into account the biological inputs that cover the entire process of height formation. I estimate an empirical specification for the height production function where the change in height between two consecutive measurements allows the reduction of the endogenous inputs. I use both FE and different IVFE specifications based on different sets of IVs. The IVFE allow the estimation of conditional demand equations for both nutrition and diseases, treated as endogenous inputs.

The results show that the effects of the different inputs vary according to the age

of the child, and that the magnitude and timing differ between boys and girls. In particular, girls' growth is strongly affected by diarrhea episodes at age one where experiencing an extra episode of diarrhea reduces height by 2.3/2.5 cm. However, diseases experienced in the second year of life, have the largest and most negative effects on height at age two. Experiencing an extra episode of diarrhea in the second year of life, reduces the height of a boy by about 2/2.2 cm and the height of a girl by 3.2/3.5 cm. Such high impact of diarrhea on growth is most likely due to episodes that are persistent and acute, and have been found to increase the risk of stunting (see for example, Checkley et al., 2008; Walker et al., 2011).

The model shows the importance of including past inputs and of studying their effects according to different ages of the children. Most importantly, the paper shows that infancy is a critical period for the height formation because diseases play a major role compared to nutritional intake. Bozzoli et al. (2009) claims that some form of scarring in infancy negatively affects lifetime health, as marked by adult height. Given that adult height is considered an indicator of the general health status in life, of the relative risk of survival, and labor productivity (Fogel, 1986), there is an economic rationale for investing in these early years and reducing the disease load. Some things that happen early in life may have persistent effects on health and human capital accumulation both in developed and developing countries (Case and Paxson, 2008, 2010; Cunha and Heckman, 2008; Chetty et al., 2011; Smith, 2009; Almond and Currie, 2011; Currie and Vogl, 2013; Gertler et al., 2014).

This paper confirms the importance of the first two years of life in determining height at age two, and the relevance of intervening in this time window to prevent

child undernutrition because this is a long-term investment (Victora et al., 2008, 2010). Future research should focus on later periods of life until the body matures. This would allow one to identify if there are (and which are) other critical periods that determine adult height, and also which inputs have the strongest impact.

References

Adair, L. S. and D. K. Guilkey, 1997. Age-specific determinants of stunting in Filipino children. *Journal of Nutrition*, 127(2):314–20.

Almond, G. and J. Currie, 2011. Human capital development before age five. In O. Ashenfelter, R. Layard, and D. Card, editors, *Handbook of Labor Economics*, volume 4B, chapter 15, pages 1315–1486. Amsterdam, The Netherlands: North-Holland.

Barker, D. J. P., 1998. *Mothers, Babies and Health in Later Life.* London: British Medical Journal Publishing Group.

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*, pages 107–118. Baltimore, MD: The Johns Hopkins University Press.

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

Belfield, C. R. and I. R. Kelly, 2012. The benefits of breast feeding across the early years of childhood. *Journal of Human Capital*, 6(3):251–277.

Bozzoli, C., A. Deaton, and C. Quintana-Domeque, 2009. Adult height and childhood disease. *Demography*, 46(4):647–669.

Case, A. and C. Paxson, 2008. Stature and status: Height, ability, and labor market outcomes. *Journal of Political Economy*, 116(3):499–532.

Case, A. and C. Paxson, 2010. Causes and consequences of early-life health. *Demography*, 47(1):S65–S85.

Cebu-Study-Team, 1992. A child health production function estimated from longitudinal data. *Journal of Development Economics*, 38(2):323–351.

Checkley, W., G. Buckley, R. H. Gilman, A. M. O. Assis, R. L. Guerrant, S. S. Morris, K. Mølbak, P. Valentiner-Branth, C. F. Lanata, R. E. Black, et al., 2008. Multicountry analysis of the effects of diarrhoea on childhood stunting. *International journal of epidemiology*, 37(4):816–830.

Chetty, R., J. N. Friedman, N. Hilger, E. Saez, D. W. Whitmore Schanzenbach, and D. Yagan, 2011. How does your kindergarten classroom affect your earnings? Evidence from project star. *The Quarterly Journal of Economics*, 126:1593–1660.

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.

Cunha, F. and J. Heckman, 2008. Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation. *The Journal of Human Resources*, 43(4):738–782.

Currie, J. and T. Vogl, 2013. Early-life health and adult circumstance in developing countries. *Annual Review of Economics*, 5:1–36.

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

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.

Eveleth, P. and J. M. Tanner, 1991. *World-Wide Variation in Human Growth*. Cambridge. Cambridge University Press, second edition.

Fogel, R., 1986. Physical growth as a measure of the economic well-being of populations: The eighteenth and nineteenth centuries. In F. Falkner and J. M. Tanner, editors, *Human Growth: A Comprehensive Treatise*, volume 3, pages 263–281. New York: Plenum Press.

Gertler, P., J. Heckman, R. Pinto, A. Zanolini, C. Vermeersch, S. Walker, S. M. Chang, and S. Grantham-McGregor, 2014. Labor market returns to an early childhood stimulation intervention in Jamaica. *Science*, 344(6187):998–1001.

Ginsburg, E., G. Livshits, 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):307–322.

Glewwe, P. and E. M. King, 2001. The impact of early childhood nutritional status on cognitive development: Does the timing of malnutrition matter? *The World Bank Economic Review*, 15(1):81–113.

Griffen, D., 2014. Height and calories in early childhood. *UPENN PSC Working Paper*.

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(1):45–68.

Kleibergen, F. and R. Paap, 2006. Generalised reduced rank tests using the singular value decomposition. *Journal of Econometrics*, 127(1):97–126.

Liu, H., T. Mroz, and L. Adair, 2009. Parental compensatory behaviors and early child health outcomes in Cebu, Philippines. *Journal of Development Economics*, 90(2):209–230.

Maluccio, J. A., J. Hoddinott, J. R. Behrman, R. Martorell, A. R. Quisumbing, and A. D. Stein, 2009. The impact of nutrition during early childhood on education among Guatemala adults. *Economic Journal*, 119(537):734 – 763.

Mendez, M. and L. S. Adair, 1999. Severity and timing of stunting in the first two

years of life affect performance on cognitive tests in late childhood. *The Journal of Nutrition*, 129(8):1555.

Price, G. N., 2013. The allometry of metabolism and stature: Worker fatigue and height in the Tanzanian labor market. *Economics & Human Biology*, 11(4):515–521.

Puentes, E., F. Wange, J. R. Behrman, F. Cuhna, J. Hoddinott, J. A. Maluccio, L. Adair, J. B. Borja, R. Martorell, and A. D. Stein, 2014. Early life heith and weight production functions with endogenous energy and protein inputs. *Mimeo.*

Rees, D. I. and J. J. Sabia, 2009. The effect of breast feeding on educational attainment: Evidence from sibling data. *Journal of Human Capital*, 3(1):43–69.

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.

Schmidt, I. M., M. Jørgensen, and K. F. Michaelsen, 1995. Height of conscripts in Europe: is postneonatal mortality a predictor? *Annals of human biology*, 22(1):57–67.

Schultz, T. P., 2002. Wage gains associated with height as a form of health human capital. *American Economic Review Papers and Proceedings*, 92(2):349–353.

Silventoinen, K., 2003. Determinants of variation in adult body height. *Journal of Biosocial Science*, 35(2):263–285.

Smith, J. P., 2009. The impact of childhood health on adult labor market outcomes. *The review of economics and statistics*, 91(3):478–489.

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):1–23.

Stephensen, A., 1999. Burden of infection on growth failure. *Journal of Nutrition*, 129(2):534–538.

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(1):159–185.

Todd, P. and K. I. Wolpin, 2003. On the specification and estimation of the production function for cognitive achievement. *Economic Journal*, 113(485):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.

Ugaz, J. and A. Zanolini, 2011. Effects of extreme weather shocks during pregnancy and early life on later health outcomes: The case of Philippines' typhoons. *Working Paper, University of Chicago, Harris School of Public Policy.*

Victora, C. G., L. Adair, C. Fall, P. C. Hallal, R. Martorell, L. Richter, H. S. Sachdev, et al., 2008. Maternal and child undernutrition: consequences for adult health and human capital. *The Lancet*, 371(9609):340–357.

Victora, C. G., M. de Onis, P. C. Hallal, M. Blössner, and R. Shrimpton, 2010. Worldwide timing of growth faltering: revisiting implications for interventions. *Pediatrics*, pages peds–2009.

Walker, S. P., T. D. Wachs, S. Grantham-McGregor, M. M. Black, C. A. Nelson, S. L. Huffman, H. Baker-Henningham, S. M. Chang, J. D. Hamadani, B. Lozoff, et al., 2011. Inequality in early childhood: risk and protective factors for early child development. *The Lancet*, 378(9799):1325–1338.

Tables

Table 1: Survey structure and range of ages by gender.

	Boys			Girls		
	N	Mean age	SD	N	Mean age	SD
Delivery 1983-4	1632	0	0	1448	0	0
Follow-up n.1	1525	2.051	.152	1353	2.051	.149
Follow-up n.2	1489	4.040	.139	1313	4.052	.170
Follow-up n.3	1439	6.051	.158	1278	6.045	.149
Follow-up n.4	1406	8.037	.126	1259	8.038	.141
Follow-up n.5	1386	10.068	.144	1239	10.068	.160
Follow-up n.6	1367	12.076	.169	1227	12.070	.164
Follow-up n.7	1342	14.072	.164	1207	14.073	.186
Follow-up n.8	1316	16.070	.178	1191	16.063	.171
Follow-up n.9	1310	18.068	.172	1197	18.054	.169
Follow-up n.10	1316	20.078	.190	1182	20.050	.161
Follow-up n.11	1302	22.047	.164	1158	22.041	.162
Follow-up n.12	1288	24.055	.153	1160	24.047	.153

[a] The age is in months.

Table 2: Descriptive statistics of the principal variables

	Birth		Age one		Age two	
	Mean	SD	Mean	SD	Mean	SD
Boys						
N	1618		1367		1.288	
Height (cm)	49.46	2.15	71.449	2.883	79.868	3.581
Age (years)	.0128	.013	1.006	.014	2.005	.013
Breast milk			.753	.383	.340	.369
Caloric intake (kcal)†			331.450	283.874	670.536	328.656
Diarrhea episodes			.649	.849	.916	1.055
Birth weight* (kg)	3.02	0.45				
Birth order*	2.59	2.46				
Normal pregnancy	88.94%					
Premature&Small*	7.53%					
Premature*	3.52%					
Mother's height*(cm)	151.58	5.02				
Prop. urban location*	.75	.42				
Family income (pesos)			93,658	153,084	84,373	130,173

	Birth		Age one		Age two	
	Mean	SD	Mean	SD	Mean	SD
Girls						
N	1433		1226		1160	
Height (cm)	49.01	2.11	69.923	2.841	78.315	3.630
Age (years)	.012	.013	1.005	.013	2.004	.013
Breast milk			.773	.375	.343	.370
Caloric intake (kcal)†			293.536	252.486	670.093	295.044
Diarrhea episodes			.613	.793	.778	.957
Birth weight* (kg)	2.96	.43				
Birth order*	2.50	2.39				
Normal pregnancy	88.08%					
Premature&Small*	7.95%					
Premature*	3.97%					
Mother's height*(cm)	151.47	5.03				
Prop. urban location*	.74	.42				
Family income (pesos)			93,874	21,571	75,976	100,580

[a] * corresponds to time-invariant variables.

[d] † The caloric intake is exclusive of breast milk.

Table 3: Boys' height production function at age one. Dependent variable: change in height ($\Delta_{Height}1$)

	FE	IVFE	IVFE	IVFE	IVFE
		IVSET1	IVSET2	IVSET3	IVSET4
Breast milk age one	0.726*	-4.875	-4.234	-3.635	-4.070
	[0.356]	[2.578]	[2.489]	[2.478]	[2.488]
Caloric intake age one	0.131**	-0.422	-0.285	-0.176	-0.260
	[0.0451]	[0.300]	[0.270]	[0.294]	[0.270]
Diarrhea age one	-0.142	-0.694	-0.957	-1.069	-1.076
	[0.0882]	[0.874]	[0.780]	[0.869]	[0.761]
Observations	1359	1201	1201	1184	1201
R^2	0.986	0.981	0.981	0.982	0.981
Kleibergen-Paap Underif. Test		14.52**	16.713**	14.994*	18.154**
p-value		0.006	0.005	0.020	0.006
Hansen J Overid. Test		2.684	3.961	3.324	4.488
p-value		0.443	0.411	0.650	0.481
Kleibergen-Paap Wald statistic		2.298	2.187	1.911	2.147
F-test statistic					
Breast milk		12.15***	13.12***	10.66***	11.79***
Caloric intake		24.24***	31.71***	14.65***	27.10***
Diarrhea		4.14***	3.58***	3.69***	3.51***

[a] Every model includes age and age squared between two consecutive waves.

[b] The kcal is exclusive of breast milk.

[c] A change of one unity in caloric intake corresponds to 100 kcal.

[d] Robust standard error in parenthesis.

[e] * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

[e] Signif. codes: (*) if p < .05, (**) if p < .01, (***) if p < .001.

[f] The Kleibergen-Paap Wald statistic is the robust version of the Cragg-Donald Wald F statistic. It is always below the Stock-Yogo critical values.

[g] IVset1 contains: rainy season+evaporated milk price+distance road+time to infant store+typhoon +powder milk price; IVset2 is IVset1+refrigerator; IVset3 is IVset1+piped water+banana price; IVset4 is IVset1+refrigerator+piped water.

Table 4: Girls' height production function at age one. Dependent variable: change in height ($\Delta_{Height}1$)

	FE	IVFE	IVFE	IVFE	IVFE
		IVSET1	IVSET2	IVSET3	IVSET4
Breast milk age one	1.144***	8.510	6.214	6.826	6.020
	[0.333]	[4.938]	[4.204]	[4.032]	[4.034]
Caloric intake age one	0.171***	0.949	0.586	0.892*	0.566
	[0.0470]	[0.539]	[0.417]	[0.400]	[0.401]
Diarrhea age one	-0.280**	-2.296*	-2.332*	-2.551*	-2.310*
	[0.0946]	[0.984]	[0.939]	[1.017]	[0.945]
Observations	1221	1080	1080	1065	1080
R^2	0.985	0.971	0.974	0.973	0.975
Kleibergen-Paap Underif. Test		10.975*	12.565*	13.385*	13.892*
p-value		0.027	0.028	0.037	0.031
Hansen J Overid. Test		2.884	4.532	4.926	4.630
p-value		0.410	0.339	0.425	0.463
Kleibergen-Paap Wald statistic		1.941	1.928	1.918	1.890
F-test statistic					
Breast milk		16.53***	14.85***	6.40***	12.97***
Caloric intake		22.08***	25.88***	13.44***	22.62***
Diarrhea		4.09***	3.53***	3.64***	3.10**

[a] Every model includes age and age squared between two consecutive waves.

[b] The kcal is exclusive of breast milk.

[c] A change of one unity in caloric intake corresponds to 100 kcal.

[d] Robust standard error in parenthesis.

[e] * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

[e] Signif. codes: (*) if p < .05, (**) if p < .01, (***) if p < .001.

[f] The Kleibergen-Paap Wald statistic is the robust version of the Cragg-Donald Wald F statistic. It is always below the Stock-Yogo critical values.

[g] IVset1 contains: rainy season+evaporated milk price+distance road+time to infant store+typhoon +powder milk price; IVset2 is IVset1+refrigerator; IVset3 is IVset1+piped water+banana price; IVset4 is IVset1+refrigerator+piped water.

Table 5: Boys' height production function at age two. Dependent variable: change in height ($\Delta_{Height}2$)

	FE	IVFE	IVFE	IVFE	IVFE
		IVSET1	IVSET2	IVSET3	IVSET4
Breast milk age two	0.938***	-1.174	-0.693	-1.033	-0.892
	[0.189]	[1.984]	[1.838]	[1.897]	[1.828]
Caloric intake age two	0.197***	0.298*	0.284*	0.318*	0.309*
	[0.0210]	[0.147]	[0.136]	[0.143]	[0.138]
Diarrhea age two	-0.278***	-2.214***	-2.041***	-2.120***	-2.132***
	[0.0540]	[0.593]	[0.510]	[0.584]	[0.572]
Observations	1250	1060	1060	1060	1060
R^2	0.944	0.876	0.890	0.881	0.882
Kleibergen-Paap Underif. Test		14.713*	14.489*	13.841*	15.286**
p-value		0.012	0.013	0.017	0.009
Hansen J Overid. Test		7.831	8.963	8.933	8.268
p-value		0.098	0.062	0.063	0.082
Kleibergen-Paap Wald statistic		2.450	2.454	2.420	2.493
F-test statistic					
Breast milk		5.00***	6.21***	5.23***	5.04***
Caloric intake		20.58***	20.97***	23.31***	18.73***
Diarrhea		4.91***	4.91***	5.78***	5.30***

[a] Every model includes age and age squared between two consecutive waves.

[b] The kcal is exclusive of breast milk.

[c] A change of one unity in caloric intake corresponds to 100 kcal.

[d] Robust standard error in parenthesis.

[e] * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

[e] Signif. codes: (*) if p < .05, (**) if p < .01, (***) if p < .001.

[f] The Kleibergen-Paap Wald statistic is the robust version of the Cragg-Donald Wald F statistic. It is always below the Stock-Yogo critical values.

[g] IVset1 contains: distance road+time to infant store+kerosene price+powder milk price+egg price +infant food store+piped water; IVset2 is IVset1-kerosene price+refrigerator; IVset3 is IVset1-time to infant store+banana price; IVset4 is IVset1-time to infant store+evaporated milk price.

Table 6: Girls' height production function at age two. Dependent variable: change in height ($\Delta_{Height}2$)

	FE	IVFE	IVFE	IVFE	IVFE
		IVSET1	IVSET2	IVSET3	IVSET4
Breast milk age two	1.066***	-0.561	-0.326	0.659	0.229
	[0.206]	[2.513]	[2.523]	[2.469]	[2.526]
Caloric intake age two	0.190***	0.204	0.187	0.200	0.270
	[0.0264]	[0.175]	[0.183]	[0.173]	[0.174]
Diarrhea age two	-0.269***	-3.171***	-3.178**	-3.294***	-3.493***
	[0.0676]	[0.949]	[1.062]	[0.919]	[0.976]
Observations	1119	942	942	942	942
R^2	0.937	0.833	0.834	0.828	0.812
Kleibergen-Paap Underif. Test		17.96**	15.664**	18.812**	17.905**
p-value		0.003	0.008	0.002	0.003
Hansen J Overid. Test		5.456	5.816	6.325	4.89
p-value		0.244	0.213	0.176	0.299
Kleibergen-Paap Wald statistic		2.658	2.431	2.700	2.567
F-test statistic					
Breast milk		4.17***	4.46***	4.06***	3.72***
Caloric intake		19.12***	20.11***	18.96***	18.63***
Diarrhea		3.86***	3.60***	3.55***	3.66***

[a] Every model includes age and age squared between two consecutive waves.

[b] The kcal is exclusive of breast milk.

[c] A change of one unity in caloric intake corresponds to 100 kcal.

[d] Robust standard error in parenthesis.

[e] * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

[e] Signif. codes: (*) if p < .05, (**) if p < .01, (***) if p < .001.

[f] The Kleibergen-Paap Wald statistic is the robust version of the Cragg-Donald Wald F statistic. It is always below the Stock-Yogo critical values.

[g] IVset1 contains: distance road+time to infant store+kerosene price+powder milk price+egg price +infant food store+piped water; IVset2 is IVset1-kerosene price+refrigerator; IVset3 is IVset1-time to infant store+banana price; IVset4 is IVset1-time to infant store+evaporated milk price.

Appendix

Table A1: Descriptive statistics of the instrumental variables used

| | | AGE ONE | | | |
| | | BOYS | | GIRLS | |
IV variables	Definitions	Mean	SD	Mean	SD
Rainy season	rainy season	0.514	0.022	0.515	0.025
Evaporated milk price	price of 100 g evaporated milk	830.214	81.098	833.029	84.663
Banana price	price of 1 banana	132.091	28.204	130.633	29.254
Distance road†	distance (m) to nearest vehicular road	263.084	493.205	273.193	514.314
Refrigerator†	own a refrigerator	0.066	0.042	0.062	0.042
Time to infant store†	minutes walk to nearet infant store	16.282	17.827	17.012	18.284
Kerosene price	price of 1 lt of kerosene	3571.445	514.417	3588.882	532.450
Powder milk price	price of 350 g powdered milk	8437.374	573.609	8474.985	605.092
Egg price	price of medium size egg	468.232	30.794	466.666	34.648
Infant food store†	presence of infant store close to home	0.709	0.212	0.692	0.226
Piped water†	piped water as water source	0.865	0.260	0.846	0.271
Typhoon Nitang	child exposed to the typhoon during the first year	0.639	0.481	0.659	0.474

| | | AGE TWO | | | |
| | | BOYS | | GIRLS | |
		Mean	SD	Mean	SD
Rainy season	rainy season	0.516	0.017	0.517	0.016
Evaporated milk price	price of 100 g evaporated milk	953.566	104.386	953.909	113.660
Banana price	price of 1 banana	116.789	28.429	117.435	29.065
Distance road†	distance (m) to nearest vehicular road	279.101	516.595	285.455	531.207
Refrigerator†	own a refrigerator	0.068	0.055	0.064	0.046
Time to infant store†	minutes walk to nearest infant store	16.746	18.421	17.236	18.445
Kerosene price	price of 1 lt of kerosene	3565.118	671.498	3556.290	704.747
Powder milk price	price of 350 g powdered milk	8626.369	543.801	8617.647	573.267
Egg price	price of medium size egg	473.714	25.502	472.550	27.587
Infant food store†	presence of infant store close to home	0.710	0.217	0.693	0.229
Piped water†	piped water as water source	0.859	0.268	0.842	0.279

[a] † indicates household characteristics averaged at barangay level.
[b] The instruments are averaged over year one or year two.

Table A2: Boys' age one first-stage estimates.

	IVset1			IVset2			IVset3			IVset4		
	Breast milk	Caloric intake	Diarrhea	Breast milk	Caloric intake	Diarrhea	Breast milk	Caloric intake	Diarrhea	Breast milk	Caloric intake	Diarrhea
$\Delta_{age}1$	2.047* [0.830]	-4.606 [5.493]	-1.171 [1.388]	2.200** [0.832]	-6.365 [5.465]	-1.13 [1.401]	1.984* [0.859]	-4.107 [5.696]	-1.088 [1.391]	2.169** [0.840]	-5.923 [5.469]	-1.325 [1.397]
$\Delta_{age2}1$	-1.429* [0.654]	7.015 [4.667]	-0.485 [1.134]	-1.223 [0.657]	4.649 [4.614]	-0.429 [1.131]	-1.188 [0.671]	4.503 [4.871]	-0.867 [1.161]	-1.229 [0.654]	4.733 [4.595]	-0.467 [1.144]
Rainy season	-1.561* [0.776]	13.66** [5.051]	5.531*** [1.540]	-1.408 [0.750]	11.90* [4.955]	5.572*** [1.541]	-1.361 [0.801]	11.84* [5.223]	4.972** [1.635]	-1.47 [0.770]	12.79* [5.101]	5.177** [1.616]
Evaporated milk price	0.0008** [0.00024]	-0.0043* [0.00182]	-0.000999 [0.000579]	0.0006* [0.000251]	-0.0023 [0.00188]	-0.0010 [0.000593]	0.00075** [0.000257]	-0.0039* [0.00194]	-0.00077 [0.00062]	0.0006* [0.00026]	-0.0026 [0.0019]	-0.0009 [0.0006]
Distance road†	0.00002 [0.00007]	-0.00005 [0.00049]	-0.00007 [0.00017]	0.000002 [0.00007]	0.0002 [0.00048]	-0.00007 [0.00018]	0.00004 [0.0008]	-0.00025 [0.000605]	-0.00014 [0.00019]	-0.000002 [0.00007]	0.0002 [0.00048]	-0.00009 [0.0002]
Time to infant store†	0.000996 [0.00278]	-0.0403* [0.0186]	0.00519 [0.00687]	-0.00011 [0.00278]	-0.0276 [0.0182]	0.00489 [0.00688]	0.00327 [0.00364]	-0.0757** [0.0278]	0.00564 [0.00888]	0.000423 [0.00309]	-0.0352 [0.0205]	0.00828 [0.0077]
Typhoon	0.0698** [0.0235]	-0.0658 [0.169]	0.126* [0.0493]	0.0686** [0.0233]	-0.0522 [0.167]	0.126* [0.0494]	0.0710** [0.0237]	-0.0565 [0.170]	0.114* [0.0498]	0.0683** [0.0232]	-0.0481 [0.167]	0.124* [0.0494]
Powder milk price	0.00003 [0.00002]	-0.00024 [0.00017]	0.00002 [0.000047]	0.000007 [0.00002]	0.000013 [0.00016]	0.00001 [0.00005]	0.000012 [0.000024]	-0.00004 [0.00018]	0.00004 [0.00005]	0.000008 [0.00002]	-0.00001 [0.00017]	0.00002 [0.00005]
Refrigerator†				-1.276*** [0.318]	14.66*** [2.303]	-0.344 [0.755]				-1.303*** [0.329]	15.05*** [2.373]	-0.517 [0.776]
Piped water†							-0.000788 [0.113]	-0.336 [0.968]	0.184 [0.270]	0.0363 [0.0853]	-0.521 [0.635]	0.231 [0.218]
Banana price							-0.000967* [0.00045]	0.0124*** [0.0031]	0.000365 [0.00103]			
Observations	1201	1201	1201	1201	1201	1201	1184	1184	1184	1201	1201	1201
R-squared	0.797	0.615	0.387	0.8	0.626	0.388	0.796	0.623	0.388	0.8	0.626	0.388

[a] † indicates household characteristics averaged at village level.
[b] Note: Every model includes age and age squared.

Table A3: Girls' age one first-stage estimates.

	IVset1			IVset2			IVset3			IVset4		
	Breast milk	Caloric intake	Diarrhea	Breast milk	Caloric intake	Diarrhea	Breast milk	Caloric intake	Diarrhea	Breast milk	Caloric intake	Diarrhea
$\Delta_{age}1$	1.491 [0.772]	-2.915 [5.533]	0.531 [1.480]	1.674* [0.781]	-5.328 [5.381]	0.482 [1.478]	1.706* [0.844]	-3.702 [5.915]	0.418 [1.569]	1.715* [0.823]	-4.313 [5.635]	0.207 [1.542]
$\Delta_{age^2}1$	-1.228 [0.694]	7.371 [4.752]	-0.338 [1.306]	-1.074 [0.698]	5.350 [4.582]	-0.380 [1.316]	-1.226 [0.717]	6.306 [4.827]	-0.447 [1.329]	-1.084 [0.704]	5.107 [4.634]	-0.314 [1.319]
Rainy season	-0.457 [0.417]	7.931 [4.594]	1.305 [0.907]	-0.434 [0.401]	7.623 [4.334]	1.299 [0.905]	-0.375 [0.419]	7.120 [4.596]	0.983 [0.846]	-0.420 [0.407]	7.956 [4.524]	1.209 [0.890]
Evaporated milk price	0.00062* [0.0002]	-0.0048** [0.0017]	-0.0007 [0.0005]	0.0005* [0.00025]	-0.0032 [0.0017]	-0.0007 [0.00048]	0.0005* [0.00026]	-0.0046* [0.0018]	-0.0006 [0.00049]	0.0005 [0.00026]	-0.0035 [0.0018]	-0.0006 [0.0005]
Distance road†	0.0002** [0.00007]	-0.0007 [0.00047]	-0.0002 [0.00016]	0.0002** [0.00007]	-0.0006 [0.00046]	-0.0002 [0.00016]	0.0002** [0.000074]	-0.0008 [0.0005]	-0.0003* [0.00016]	0.0002*** [0.00007]	-0.0006 [0.00046]	-0.0002 [0.00016]
Time to infant store†	-0.0051 [0.0028]	-0.0007 [0.0191]	0.0150* [0.00618]	-0.0064* [0.00279]	0.0162 [0.0189]	0.0154* [0.00629]	-0.00470 [0.00343]	-0.0279 [0.0224]	0.0104 [0.00702]	-0.00661* [0.00296]	0.0105 [0.0205]	0.0169* [0.00690]
Typhoon	0.0088 [0.0238]	0.118 [0.162]	0.175*** [0.0491]	0.0101 [0.0237]	0.101 [0.159]	0.175*** [0.0491]	0.0104 [0.0241]	0.0852 [0.162]	0.169*** [0.0489]	0.0101 [0.0237]	0.0995 [0.159]	0.175*** [0.0491]
Powder milk price	0.00003 [0.00002]	-0.0002 [0.00014]	0.00001 [0.00004]	0.000015 [0.00002]	0.00006 [0.00014]	0.000015 [0.00005]	0.000022 [0.00002]	-0.000042 [0.00015]	0.000036 [0.00005]	0.00001 [0.00002]	0.00003 [0.00015]	0.00002 [0.00005]
Refrigerator†				-1.074** [0.358]	14.11*** [2.287]	0.291 [0.759]				-1.059** [0.372]	14.49*** [2.378]	0.186 [0.792]
Piped water†							-0.0548 [0.0975]	-0.207 [0.651]	-0.00579 [0.208]	-0.0184 [0.0950]	-0.455 [0.641]	0.123 [0.206]
Banana price							-0.000488 [0.000463]	0.0103*** [0.00283]	0.00105 [0.000960]			
Observations	1080	1080	1080	1080	1080	1080	1065	1065	1065	1080	1080	1080
R-squared	0.812	0.609	0.393	0.814	0.621	0.393	0.809	0.615	0.395	0.814	0.621	0.393

[a] † indicates household characteristics averaged at village level.
[b] Note: Every model includes age and age squared.

43

Table A4: Boys' age two first-stage estimates.

	IVset1			IVset2			IVset3			IVset4		
	Breast milk	Caloric intake	Diarrhea	Breast milk	Caloric intake	Diarrhea	Breast milk	Caloric intake	Diarrhea	Breast milk	Caloric intake	Diarrhea
$\Delta_{age}1$	2.130 [1.617]	-37.46* [15.12]	-0.310 [4.911]	1.801 [1.622]	-33.67* [14.92]	-0.809 [4.917]	2.354 [1.581]	-36.85* [14.47]	0.993 [4.855]	2.204 [1.592]	-36.60* [14.91]	0.565 [4.876]
$\Delta_{age^2}1$	-0.717 [0.525]	16.47*** [4.853]	0.612 [1.608]	-0.624 [0.526]	15.41** [4.831]	0.745 [1.608]	-0.695 [0.520]	15.06** [4.697]	0.202 [1.611]	-0.762 [0.521]	16.05*** [4.788]	0.273 [1.620]
Distance road†	-0.0001 [0.0001]	-0.0033*** [0.00085]	-0.0005 [0.0003]	-0.0002 [0.0001]	-0.0022* [0.0009]	-0.0007* [0.00032]	-0.0001 [0.0001]	-0.0024* [0.00097]	-0.0002 [0.0002]	-0.0001 [0.00008]	-0.0032** [0.00114]	-0.0004* [0.0002]
Time to infant store†	0.00267 [0.00425]	0.0263 [0.0391]	0.0225 [0.0155]	0.00486 [0.00406]	0.000820 [0.0360]	0.0274 [0.0158]						
Kerosene price	0.00003 [0.00003]	-0.0003 [0.0002]	0.0001 [0.0001]				0.00004 [0.00003]	-0.0006* [0.0002]	0.00005 [0.0001]	0.00001 [0.00003]	-0.0004 [0.0003]	0.00002 [0.0001]
Powder milk price	0.00004 [0.00003]	-0.000001 [0.0002]	-0.0002*** [0.00006]	0.0001* [0.00003]	-0.0002 [0.0002]	-0.0002** [0.0001]	0.00002 [0.00003]	0.0003 [0.0002]	-0.0002** [0.00007]	0.00004 [0.00002]	0.000002 [0.00023]	-0.0002*** [0.0001]
Egg price	0.0002 [0.000517]	-0.0137*** [0.00389]	0.0014 [0.0016]	0.00012 [0.0005]	-0.0121** [0.00401]	0.001 [0.0016]	0.0003 [0.0005]	-0.0136*** [0.0038]	0.0015 [0.0016]	0.00012 [0.0005]	-0.0149*** [0.0043]	0.0006 [0.0017]
Infant food store†	-0.0385 [0.0977]	3.280*** [0.764]	1.021** [0.317]	0.0192 [0.0942]	2.624*** [0.779]	1.074*** [0.316]	-0.0312 [0.0909]	2.206** [0.705]	0.664** [0.257]	-0.0578 [0.0880]	3.077*** [0.672]	0.835*** [0.245]
Piped water†	-0.227* [0.105]	0.201 [0.792]	-0.424 [0.264]	-0.154 [0.0950]	-0.654 [0.732]	-0.230 [0.257]	-0.319** [0.105]	1.357 [0.788]	-0.445 [0.250]	-0.215* [0.108]	0.257 [0.785]	-0.428 [0.277]
Refrigerator†				-0.723** [0.234]	8.347 [4.331]	-1.354 [0.804]						
Banana price							-0.000749 [0.000491]	0.0164*** [0.00398]	0.00239 [0.00139]			
Evaporated milk price										0.0002 [0.00026]	0.0017 [0.0024]	0.0012 [0.0008]
Observations	1060	1060	1060	1060	1060	1060	1060	1060	1060	1060	1060	1060
R-squared	0.464	0.850	0.458	0.467	0.852	0.459	0.465	0.852	0.458	0.464	0.850	0.458

[a] † indicates household characteristics averaged at village level.
[b] Note: Every model includes age and age squared.

Table A5: Girls' age two first-stage estimates.

	IVset1			IVset2			IVset3			IVset4		
	Breast milk	Caloric intake	Diarrhea	Breast milk	Caloric intake	Diarrhea	Breast milk	Caloric intake	Diarrhea	Breast milk	Caloric intake	Diarrhea
$\Delta_{age}1$	3.262 [2.086]	-52.25** [16.21]	-1.738 [5.782]	2.967 [2.099]	-45.93** [16.41]	-2.477 [5.828]	2.659 [2.059]	-47.11** [15.89]	-0.162 [5.802]	2.865 [2.072]	-52.25** [16.18]	-0.437 [5.757]
$\Delta_{age^2}1$	-1.103 [0.691]	20.31*** [5.285]	0.961 [1.904]	-1.009 [0.695]	18.35*** [5.341]	1.172 [1.921]	-0.918 [0.693]	18.41*** [5.274]	0.829 [1.935]	-1.035 [0.694]	19.55*** [5.295]	0.711 [1.922]
Distance road†	0.0001 [0.0001]	-0.0036*** [0.0007]	-0.0009*** [0.0003]	0.00004 [0.0001]	-0.0025*** [0.0008]	-0.0009*** [0.0003]	-0.00001 [0.0001]	-0.0028*** [0.0007]	-0.0006** [0.0002]	0.00002 [0.00012]	-0.0037*** [0.0008]	-0.0006*** [0.0002]
Time to infant store†	-0.0068 [0.00426]	0.0519 [0.0304]	0.0240 [0.0127]	-0.0058 [0.00439]	0.0292 [0.0309]	0.0273* [0.0125]	0.00003 [0.00003]	-0.0007** [0.0002]	0.0002* [0.00008]			
Kerosene price	0.00002 [0.00003]	-0.0006** [0.0002]	0.00015 [0.00008]							0.00003 [0.00004]	-0.0011*** [0.00027]	0.0001 [0.00009]
Powder milk price	0.00002 [0.00003]	0.00002 [0.0002]	-0.0001 [0.00006]	0.00003 [0.00003]	-0.0002 [0.0002]	-0.00006 [0.00006]	0.000012 [0.00003]	0.0001 [0.0002]	-0.0001* [0.00006]	0.00003 [0.00003]	0.00008 [0.0002]	-0.00011 [0.0001]
Egg price	0.0009 [0.0005]	-0.0065 [0.0039]	-0.00102 [0.0015]	0.0009 [0.0005]	-0.00641 [0.0038]	-0.00047 [0.0014]	0.0009 [0.00054]	-0.0067 [0.0038]	-0.0012 [0.00148]	0.001 [0.0006]	-0.0110* [0.0044]	-0.0016 [0.0016]
Infant food store†	0.0965 [0.103]	1.874** [0.698]	0.827** [0.271]	0.123 [0.0941]	1.735* [0.688]	0.666** [0.253]	0.195* [0.0936]	1.032 [0.641]	0.577** [0.238]	0.169 [0.0931]	1.454* [0.616]	0.576* [0.234]
Piped water	-0.293** [0.106]	1.153 [0.702]	-0.412 [0.271]	-0.229* [0.0973]	-0.400 [0.676]	-0.149 [0.247]	-0.272** [0.101]	1.189 [0.697]	-0.693** [0.249]	-0.249** [0.115]	1.883* [0.787]	-0.527* [0.268]
Refrigerator†				-0.647 [0.422]	12.40*** [2.976]	-0.868 [0.911]						
Banana price							-0.0008 [0.000477]	0.0093** [0.00353]	-0.0006 [0.00121]			
Evaporated milk price										-0.0001 [0.0003]	0.0061** [0.0022]	0.0006 [0.0007]
Observations	942	942	942	942	942	942	942	942	942	942	942	942
R-squared	0.467	0.845	0.425	0.468	0.847	0.423	0.467	0.846	0.422	0.466	0.846	0.422

[a] † indicates household characteristics averaged at village level.
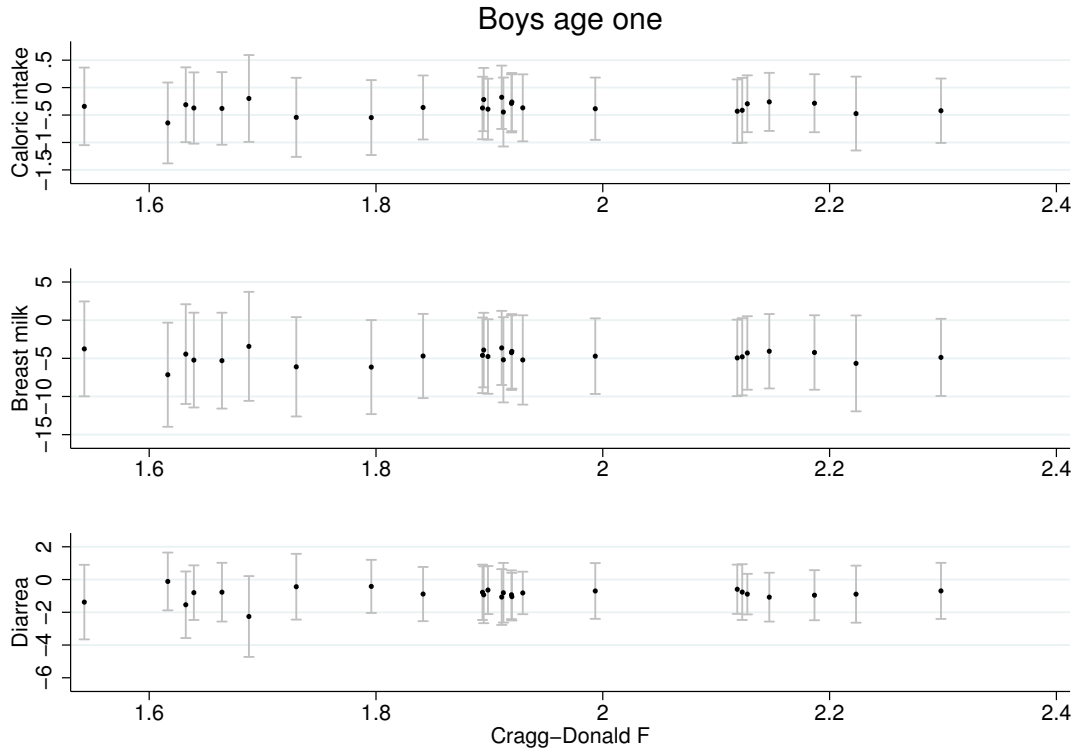[b] Note: Every model includes age and age squared.

Figure A1: Boys' age one coefficients. The x-axis corresponds to the Cragg-Donald F test, while the y-axis corresponds to the coefficient values in centimetres. Each result corresponds to an IVFE model where a different IV set is considered. All the IV sets satisfy the under identification (Kleibergen-Paap test ≤ 0.05) and over identification tests (J Hansen > 0.05) for both boys and girls of age one. All regressions include age and age squared between wave 6 and 0. 8% of the regressions is based on a set of 4 IVs, 4% on a set of 5 IVs, 20% on a set of 6 IVs, 32% on a set of 7 IVs, 32% on a set of 8 IVs, and 4% on a set of 9 IVs.
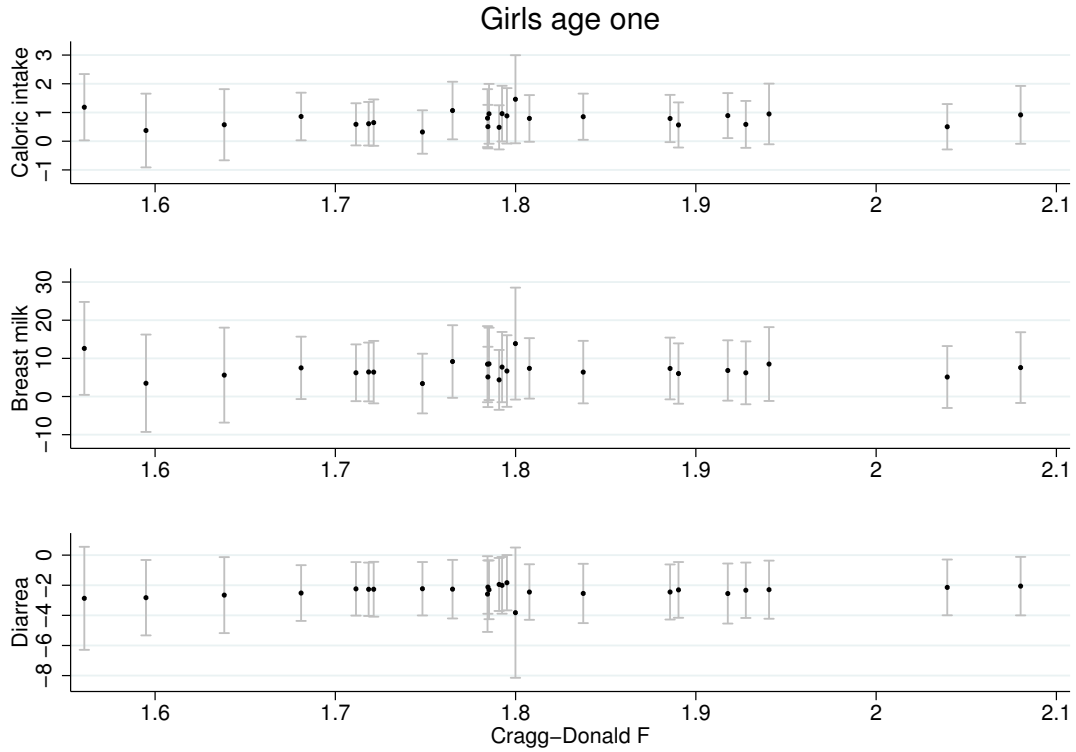
Figure A2: Girls' age one coefficients. The x-axis corresponds to the Cragg-Donald F test, while the y-axis corresponds to the coefficient values in centimetres. Each result corresponds to an IVFE model where a different IV set is considered. All the IV sets satisfy the under identification (Kleibergen-Paap test ≤ 0.05) and over identification tests (J Hansen > 0.05) for both boys and girls of age one. All regressions include age and age squared between wave 6 and 0. 8% of the regressions is based on a set of 4 IVs, 4% on a set of 5 IVs, 20% on a set of 6 IVs, 32% on a set of 7 IVs, 32% on a set of 8 IVs, and 4% on a set of 9 IVs.
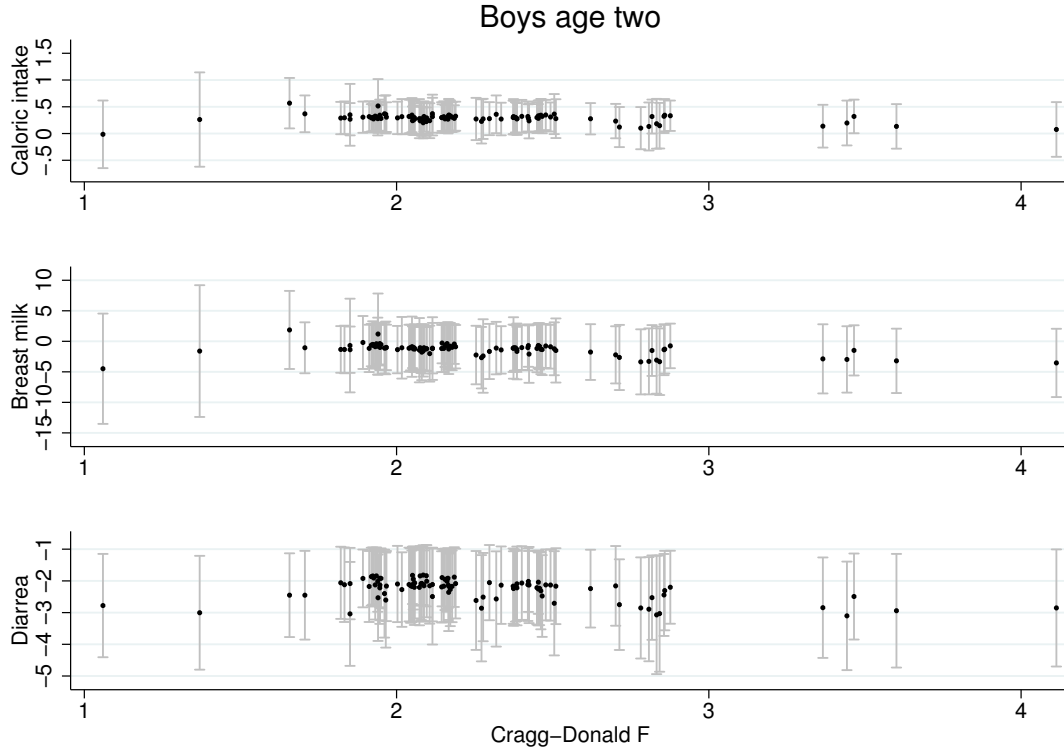
Figure A3: Boys' age two coefficients. The x-axis corresponds to the Cragg-Donald F test, while the y-axis corresponds to the coefficient values in centimetres. Each result corresponds to an IVFE model where a different IV set is considered. All the IV sets satisfy the under identification (Kleibergen-Paap test ≤ 0.05) and over identification tests (J Hansen > 0.05) for both boys and girls of age two. All regressions include age and age squared between wave 12 and 6. 2.91% of the regressions is based on a set of 3 IVs, 4.85% 4 IVs, 13.59% 5 IVs, 17.48% 6 IVs, 31.07% 7 IVs, 18.45% 8 IVs, 11.65% 9 IVs.
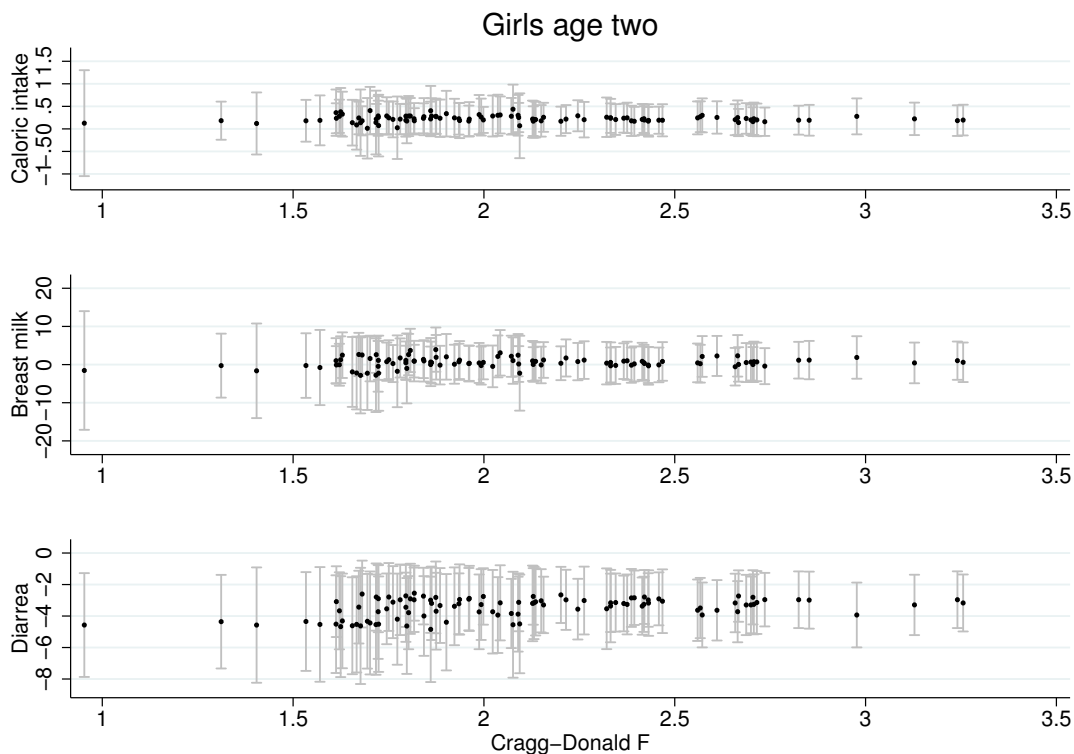
Figure A4: Girls' age two coefficients. The x-axis corresponds to the Cragg-Donald F test, while the y-axis corresponds to the coefficient values in centimetres. Each result corresponds to an IVFE model where a different IV set is considered. All the IV sets satisfy the under identification (Kleibergen-Paap test ≤ 0.05) and over identification tests (J Hansen > 0.05) for both boys and girls of age two. All regressions include age and age squared between wave 12 and 6. 2.91% of the regressions is based on a set of 3 IVs, 4.85% 4 IVs, 13.59% 5 IVs, 17.48% 6 IVs, 31.07% 7 IVs, 18.45% 8 IVs, 11.65% 9 IVs.

Table A6: Boys' hybrid height production function at age one. Dependent variable: change in height ($\Delta_{Height}1$)

	FE	IVFE	IVFE	IVFE	IVFE
		IVSET1	IVSET2	IVSET3	IVSET4
Family income age one	0.00215***	0.00198*	0.00144*	0.00128	0.00140*
	[0.000523]	[0.000852]	[0.000708]	[0.000801]	[0.000709]
Breast milk age one	0.426	-5.892*	-4.637	-4.265	-4.364
	[0.399]	[2.901]	[2.655]	[2.672]	[2.618]
Caloric intake age one	0.0624	-0.532	-0.298	-0.224	-0.265
	[0.0493]	[0.335]	[0.280]	[0.315]	[0.278]
Diarrhea age one	-0.0758	-0.317	-0.694	-0.836	-0.818
	[0.0966]	[0.983]	[0.854]	[0.944]	[0.827]
Observations	1095	957	957	940	957
R^2	0.986	0.981	0.982	0.982	0.982
Kleibergen-Paap Underif. Test		11.893*	15.113**	13.046*	16.981**
p-value		0.0182	0.0099	0.0423	0.0094
Hansen J Overid. Test		1.808	3.573	2.148	4.105
p-value		0.6132	0.4668	0.8283	0.5344
Kleibergen-Paap Wald statistic		2.015	2.153	1.823	2.131
F-test statistic					
Breast milk		9.01***	8.97***	9.86***	8.73***
Caloric intake		19.99***	24.38***	16.19***	22.67***
Diarrhea		3.93***	3.37**	3.45***	3.24**

[a] Every model includes age and age squared between two consecutive waves.

[b] The kcal is exclusive of breast milk.

[c] A change of one unity in caloric intake corresponds to 100 kcal.

[d] A change of one unity in income corresponds to 1000 pesos.

[e] Robust standard error in parenthesis.

[f] * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

[g] Signif. codes: (*) if p < .05, (**) if p < .01, (***) if p < .001.

[h] The Kleibergen-Paap Wald statistic is the robust version of the Cragg-Donald Wald F statistic. It is always below the Stock-Yogo critical values.

[g] IVset1 contains: rainy season+evaporated milk price+distance road+time to infant store+typhoon +powder milk price; IVset2 is IVset1+refrigerator; IVset3 is IVset1+piped water+banana price; IVset4 is IVset1+refrigerator+piped water.

Table A7: Girls' hybrid height production function at age one. Dependent variable: change in height ($\Delta_{Height}1$)

	FE	IVFE	IVFE	IVFE	IVFE
		IVSET1	IVSET2	IVSET3	IVSET4
Family income age one	0.000682*	0.000750	0.000996	-0.000623	0.00101
	[0.000336]	[0.00128]	[0.00115]	[0.00160]	[0.00114]
Breast milk age one	1.037**	3.838	1.034	4.003	1.246
	[0.365]	[4.590]	[3.566]	[4.437]	[3.289]
Caloric intake age one	0.138**	0.536	0.195	0.811	0.215
	[0.0532]	[0.477]	[0.330]	[0.447]	[0.303]
Diarrhea age one	-0.257*	-2.465**	-2.344**	-3.258**	-2.347**
	[0.102]	[0.815]	[0.788]	[1.021]	[0.790]
Observations	1003	882	882	867	882
R^2	0.985	0.977	0.979	0.971	0.979
Kleibergen-Paap Underif. Test		8.588	13.666*	12.513	17.411**
p-value		0.0723	0.0179	0.0515	0.0079
Hansen J Overid. Test		6.46	7.438	3.619	7.45
p-value		0.0913	0.1145	0.6055	0.1892
Kleibergen-Paap Wald statistic		1.574	2.081	1.978	2.47
F-test statistic					
Breast milk		11.20***	9.64***	4.78***	8.43***
Caloric intake		18.13***	19.47***	11.09***	17.15***
Diarrhea		5.46***	4.68***	4.26***	4.11***

[a] Every model includes age and age squared between two consecutive waves.

[b] The kcal is exclusive of breast milk.

[c] A change of one unity in caloric intake corresponds to 100 kcal.

[d] A change of one unity in income corresponds to 1000 pesos.

[e] Robust standard error in parenthesis.

[f] * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

[g] Signif. codes: (*) if p < .05, (**) if p < .01, (***) if p < .001.

[h] The Kleibergen-Paap Wald statistic is the robust version of the Cragg-Donald Wald F statistic. It is always below the Stock-Yogo critical values.

[g] IVset1 contains: rainy season+evaporated milk price+distance road+time to infant store+typhoon +powder milk price; IVset2 is IVset1+refrigerator; IVset3 is IVset1+piped water+banana price; IVset4 is IVset1+refrigerator+piped water.

Table A8: Boys' hybrid height production function at age two. Dependent variable: change in height ($\Delta_{Height}2$)

	FE	IVFE	IVFE	IVFE	IVFE
		IVSET1	IVSET2	IVSET3	IVSET4
Family income age two	0.00101*	-0.00215	-0.00189	-0.00198	-0.00207
	[0.000416]	[0.00147]	[0.00118]	[0.00137]	[0.00146]
Breast milk age two	0.905***	0.534	0.484	0.363	0.688
	[0.197]	[1.891]	[1.834]	[1.643]	[1.791]
Caloric intake age two	0.169***	0.440*	0.412**	0.436**	0.442*
	[0.0234]	[0.182]	[0.159]	[0.164]	[0.177]
Diarrhea age two	-0.270***	-2.442**	-2.331***	-2.015**	-2.350**
	[0.0567]	[0.750]	[0.630]	[0.648]	[0.726]
Observations	1106	923	923	923	923
R^2	0.947	0.865	0.874	0.890	0.871
Kleibergen-Paap Underif. Test		15.683**	17.383**	16.815**	15.366**
p-value		0.0078	0.0038	0.0049	0.0089
Hansen J Overid. Test		6.07	5.805	10.200*	6.364
p-value		0.194	0.2142	0.0372	0.1735
Kleibergen-Paap Wald statistic		1.935	2.091	2.096	2.042
F-test statistic					
Breast milk		6.07***	7.25***	7.33***	6.10***
Caloric intake		16.57***	17.13***	20.07***	15.78***
Diarrhea		4.45***	4.43***	6.30***	4.89***

[a] Every model includes age and age squared between two consecutive waves.

[b] The kcal is exclusive of breast milk.

[c] A change of one unity in caloric intake corresponds to 100 kcal.

[d] A change of one unity in income corresponds to 1000 pesos.

[e] Robust standard error in parenthesis.

[f] * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

[g] Signif. codes: (*) if p < .05, (**) if p < .01, (***) if p < .001.

[h] The Kleibergen-Paap Wald statistic is the robust version of the Cragg-Donald Wald F statistic. It is always below the Stock-Yogo critical values.

[g] IVset1 contains: distance road+time to infant store+kerosene price+powder milk price+egg price +infant food store+piped water; IVset2 is IVset1-kerosene price+refrigerator; IVset3 is IVset1-time to infant store+banana price; IVset4 is IVset1-time to infant store+evaporated milk price.

Table A9: Girls' hybrid height production function at age two. Dependent variable: change in height ($\Delta_{Height}2$)

	FE	IVFE	IVFE	IVFE	IVFE
		IVSET1	IVSET2	IVSET3	IVSET4
Family income age two	0.000230	-0.000885	-0.0000857	-0.000166	-0.00120
	[0.000499]	[0.00170]	[0.00137]	[0.00170]	[0.00198]
Breast milk age two	1.038***	-0.638	-0.698	0.926	0.572
	[0.224]	[2.904]	[2.635]	[2.871]	[3.111]
Caloric intake age two	0.177***	0.225	0.108	0.211	0.338
	[0.0294]	[0.251]	[0.212]	[0.242]	[0.270]
Diarrhea age two	-0.242***	-3.438**	-2.854**	-3.711**	-4.200**
	[0.0727]	[1.228]	[1.036]	[1.171]	[1.353]
Observations	989	824	824	824	824
R^2	0.937	0.812	0.853	0.794	0.749
Kleibergen-Paap Underif. Test		13.362*	14.631*	12.479*	11.378*
p-value		0.0202	0.0121	0.0288	0.0444
Hansen J Overid. Test		3.222	6.147	3.744	1.914
p-value		0.5214	0.1884	0.4417	0.7516
Kleibergen-Paap Wald statistic		2.021	2.469	1.782	1.608
F-test statistic					
Breast milk		3.80***	4.00***	3.71***	3.34**
Caloric intake		15.90***	17.01***	15.73***	14.91***
Diarrhea		3.60***	3.92***	3.24**	3.34***

[a] Every model includes age and age squared between two consecutive waves.

[b] The kcal is exclusive of breast milk.

[c] A change of one unity in caloric intake corresponds to 100 kcal.

[d] A change of one unity in income corresponds to 1000 pesos.

[e] Robust standard error in parenthesis.

[f] * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

[g] Signif. codes: (*) if p < .05, (**) if p < .01, (***) if p < .001.

[h] The Kleibergen-Paap Wald statistic is the robust version of the Cragg-Donald Wald F statistic. It is always below the Stock-Yogo critical values.

[g] IVset1 contains: distance road+time to infant store+kerosene price+powder milk price+egg price +infant food store+piped water; IVset2 is IVset1-kerosene price+refrigerator; IVset3 is IVset1-time to infant store+banana price; IVset4 is IVset1-time to infant store+evaporated milk price.