Is there Complete, Partial, or No Recovery from Childhood Malnutrition? - Empirical Evidence from Indonesia*

Subha Mani†

JOB MARKET PAPER

November 14, 2007

Abstract: In Indonesia, more than 30% of children under the age of 5 years suffer from chronic malnourishment. The long-term consequences of childhood malnutrition are well established in the literature. Yet, little is known about the extent to which these children are able to recover from some of the long-term deficits in health outcomes caused by childhood undernourishment. To capture the association between nutritional deficiency at young ages and subsequent health status, a panel data is constructed using observations on children between the age of 3 and 59 months in 1993 who are followed through the 1997 and 2000 waves of the Indonesian Family Life Survey. A dynamic conditional health demand function is estimated, where the coefficient on the one-period lagged health status captures the extent of recovery, if any, from childhood malnutrition. This coefficient is also known as the 'catch-up' term. Variants of the IV/GMM estimation strategy are used here to obtain an unbiased and consistent coefficient estimate on the lagged dependent variable. While the OLS coefficient estimate on the one-period lagged health status is 0.53, it is only 0.23 in a first-difference GMM framework, indicating an upward bias in the OLS parameter estimate. A coefficient of 0.23 on the one-period lagged health status indicates that poor nutrition at young ages will cause some, but not severe, retardation in the growth of future height indicating partial catch-up effects. In the absence of any catch-up, by adolescence, a malnourished child will grow to be 4.15 cm shorter than a well-nourished child. However, a coefficient of 0.23 estimated here indicates that by adolescence, a malnourished child will grow to be only 0.95 cm shorter than a well-nourished child. The first-difference GMM estimation strategy used here is especially attractive as it relies on much weaker stochastic assumptions than earlier papers and addresses both omitted variables bias and measurement error bias in data.

Keywords: Child health, Lagged dependent variable, First-difference, Indonesia

JEL Classification: I 10, 0 12, R 20, D 10

* **Acknowledgments:** I would like to thank John Ham, Duncan Thomas, and seminar participants at Bureau for Research in Economic Analysis and Development (BREAD) summer school 2005, Asia Pacific Rim University 2006, University of Melbourne 2007, Monash University 2007, Northeast Universities Development Consortium (NEUDC) 2007, and University of Southern California for helpful comments and suggestions. I am thankful to John Hoddinott and Jeffrey B. Nugent for their comments, advice, and support. I am especially indebted to John Strauss for his guidance, continuous encouragement, and support. Financial support was provided by United Nations University-World Institute for Development Economics Research (UNU-WIDER) and College of Letters, Arts, and Sciences, University of Southern California. All remaining errors are mine.

† Contact details: Department of Economics, University of Southern California; smani@usc.edu; 323-252-3670

1 Introduction

Social scientists from diverse fields such as Economics, Nutrition, and Epidemiology have come to agree that childhood malnutrition affects future well-being by decreasing the total human capital accumulated over an individual's life course[1]. For example: Alderman et. al (2006) show using data from Zimbabwe that undernourishment at young ages lowers both attained height and completed grades of schooling measured in adolescence, of which the decline in educational outcomes is estimated to translate into a 14% reduction in lifetime earnings[2]. However, if individuals are able to recover from some of the deficits in health outcomes caused by nutritional deficiencies at young ages, then some of the negative consequences associated with poor nutrition can be mitigated.

The main objective of this paper is to identify the extent to which individuals are subsequently able to compensate for some of the poor nutritional outcomes from the past. This paper finds that malnutrition during childhood will cause only some permanent growth retardation in an individual's physical well-being as measured by height attainments. This implies that at least some of the negative consequences associated with childhood malnutrition can be mitigated at an early age.

In this paper, height-for-age z-scores[3] (HAZ) and height in cm are used as indicators of nutritional status. These measures are particularly advantageous as they have been identified as indicators of chronic malnutrition and long-run physical well-being[4]. In addition, these measures are not confounded by systematic measurement error in data[5].

[1]See Glewwe and Miguel (2008) for review on the role played by child health in determining schooling outcomes. See Strauss and Thomas (2008) for a most recent review on the association between child health and future health status.

[2]Poor nutrition during childhood affects future health status thereby affecting future earnings. See Thomas and Strauss (1997) for the role played by adult height attainments in determining wage earnings using data from Brazil.

[3]Height-for-age z-sores are standardized height's calculated using the 1977 NCHS (National Center for Health Services) tables drawn from the United States population conditional upon age (in months) and sex.

[4]Waterlow (1988); Tanner (1981); Strauss and Thomas (1995); Martorell (1999); Martorell and Habicht (1986) have all discussed that height related measures capture cumulative investments in child health. Height related measures are affected by only long-term health shocks and nutritional deficiencies such as vitamin A deficiency and not short-term illnesses such as diarrhea that lasts 2-3 days.

[5]An example of systematic measurement error, Thomas and Frankenberg (2002) point out that men in general tend to self-report themselves as being taller than they actually are and women tend to report themselves as being lighter than they are.

The existing literature classifies children with HAZ$<$-2 as undernourished and or stunted [Waterlow (1988); Onis et. al (2000)]. Stunting in young children remains a serious source of concern in several developing countries, including Indonesia, as poor nutrition during childhood has long lasting impact on an individual's overall well-being. Table 1 in the appendix indicates that, in 2000, 34.8% of children from Indonesia under the age of 5 years were classified as stunted[6]. This number is large and comparable to many poor countries of the world (Onis et. al, 2000). The degree to which this stunting actually causes severe retardation in the future physical well-being of these children from Indonesia is an empirical question - unknown to policy makers and researchers in the field.

There exists a vast literature[7] that estimates the extent to which undernourishment at young ages affects subsequent health status [Adair (1999); Fedorov and Sahn (2005); Hoddinott and Kinsey (2001); Alderman et. al (2006)]. The major difficulty in estimating such a relationship arises due to the presence of unobservables such as child's innate ability to fight diseases, parental preferences, and community connections; all of which are likely to be correlated with an individual's past nutritional status. In addition, random measurement error in anthropometric outcomes makes it difficult to obtain an unbiased estimate on the child's past health status. Hoddinott and Kinsey (2001), Alderman et. al (2006), and Fedorov and Sahn (2005) are some exceptions who have successfully addressed some of the aforementioned econometric concerns.

The main contribution of this paper is to ascertain the extent to which children in Indonesia are able to recover from some of the long-run deficits in health outcomes caused by childhood malnutrition. The innovation comes in the use of time-varying community level characteristics from 1993 to identify the changes in lagged health status between 1997 and 1993 in a first-difference framework. This paper also attempts to isolate the impact of some of the key socioeconomic determinants of nutritional status among children.

A panel data set is constructed using observations on children between the age of 3 and 59 months in 1993, who are followed through the 1997 and 2000 waves of the Indonesian Family Life

[6]Source: Indonesian Family Life Survey (IFLS).
[7]Section 2 of this paper and Strauss and Thomas (2008) for a more detailed review.

Survey (IFLS). This paper first, estimates a static conditional health demand function to capture the impact of current socioeconomic factors in determining current health status. Second, a dynamic conditional health demand function is estimated which captures the extent of recovery, if any, from childhood malnutrition. The extent of recovery from poor nutrition is determined by the coefficient on the one-period lagged health status, also known as the 'catch-up' term. A coefficient of zero on the one-period lagged nutritional status in the dynamic function indicates 'complete catch-up'. A coefficient of one on the one-period lagged health status indicates 'no catch-up'. A coefficient between zero and one on the one-period lagged health status indicates 'partial catch-up' (Hoddinott and Kinsey, 2001). Finally the paper introduces an interaction term between the one-period lagged health status and lagged age in months in the dynamic specification to determine if and to what extent recovery from poor nutritional outcomes varies by age.

The parameter estimates obtained from the static conditional health demand regression indicates parental schooling, parental height, household income, and community infrastructure as some important determinants of child health. In the dynamic specification, a first-difference GMM estimation strategy is used which yields a coefficient estimate of 0.23 on the one-period lagged health status. A coefficient of 0.23 suggests partial catch-up effects; that is, malnutrition during childhood will cause only some, permanent retardation in growth in height. Using the same first-difference GMM strategy, we find that younger children have marginally larger catch-up potential than older children.

The above findings suggests that by adolescence, a malnourished child in the absence of any catch-up, that is, a coefficient of 1 on the lagged health status, would grow to be 4.15 cm shorter than a well-nourished child. However, in the presence of partial catch-up effects, such as, a coefficient of 0.23 estimated here, indicates that a malnourished child will grow to be only 0.95 cm shorter than a well-nourished child.

These results have further implications on schooling attainments. For example: Maccini and Yang (2005) have examined the impact of improvements in health status as measured by height in cm on schooling attainments using data from the IFLS. Using their predictions, I find that the decline in stature by 0.95 cm here will result in individual's accumulating 0.60 less grades of schooling. This estimate will be four times larger if there was no catch-up, that is, childhood malnutrition would lower attained height in adolescence by 4.15 cm and schooling attainment by 2.4 completed grades of schooling[8].

This paper contributes to the extant literature in two ways - First, the paper overall contributes to the larger literature in economic development addressing concerns regarding child health outcomes. It establishes the relationship between current health status and lagged health status bringing out the permanent effects of childhood malnutrition on individual's future physical well-being which is further correlated with his/her overall economic and social well-being. The paper also identifies the key socioeconomic factors that must be appropriately targeted towards improving nutritional status among children.

Second, the paper addresses a number of methodological issues that in principle can be applied to any dynamic model. The paper clearly identifies a range of IV/GMM estimation strategies that can be used to address the endogeneity issues (omitted variables and or measurement error) and discusses how the estimation strategy adopted depends upon the main source of concern related to the endogeneity problem. The first-difference GMM strategy adopted here - (a) addresses biases arising from time-invariant child-specific (genetic ability), household-specific (parental preferences), and community-specific (political connections) unobservables that are likely to affect both current and lagged health status, (b) corrects for potential biases arising from random measurement error in anthropometric data, (c) uses instruments that neither rely on lack of serial correlation in the error terms, nor on the lack of correlation between the instruments and the time-invariant unobservables (example: genetic endowments) from the empirical specification. The paper also contributes to the growing discussion in academia about instrument relevance and uses test statistics and hypothesis tests to support the relevance of instruments used in the first-difference GMM framework. Finally, the results obtained here are also robust to sample attrition, a common problem that arises due to the use of longitudinal data.

[8]The methodology used for calculating these predictions is drawn from Alderman et. al (2006).

The rest of the paper is organized as follows. Section 2 provides a brief review of the related literature. Section 3 outlines the theoretical model specified to derive the dynamic conditional health demand function. The empirical specification and identification strategy used here are described in section 4. Survey instruments, sample composition, summary statistics, and attrition rates are provided in section 5. The main regression results are discussed in section 6. Concluding remarks follow in section 7.

2 Literature review on catch-up effects

The definition of catch-up effects[9] varies significantly in the literature. Growth retardation and subsequent catching-up in health outcomes depends on whether the shocks that result in growth retardation are transitory or permanent. Transitory factors are likely to inhibit growth in short-run indicators of health outcomes such as weight and hemoglobin. Whereas permanent shocks inhibit growth retardation in height attainments. The focus of this paper is to investigate the extent of catch-up potential in the more long-run determinants of health, such as height.

The term ‘catch-up’ here signifies the extent to which childhood malnutrition causes permanent retardation in the growth of future health status. ‘Complete catch-up’ implies that childhood malnutrition will not permanently lower the child’s future growth potential and that the child can potentially also follow a higher growth path compared to his/her genetically predetermined growth path. ‘No catch-up’ implies that a child once classified as under nourished, will be permanently locked into a lower growth trajectory. ‘Partial catch-up’ implies that childhood malnutrition will cause some, but not severe, retardation in the child’s predetermined growth path.

As noted earlier, growth retardation in height attainments, particularly during childhood, if not recuperated at an earlier age can significantly lower an individual’s total human capital accumulated, affecting his/her overall well-being. Hence, social scientists have made an attempt to examine the magnitude to which individual’s can recover from some of the deficits in health

[9] See Boersma and Wit (1997) for a whole range of possible definitions to define ‘catch-up’ growth in health outcomes.

outcomes caused by childhood malnutrition.

Different lines of inquiry are used to examine the relationship between health during childhood and future health status. The review of this literature begins with the discussion of the important INCAP (Institute of Nutrition of Central America and Panama) study, a nutrition supplementation program started during the late 1960's in rural Guatemala. The main finding of the INCAP study indicates that nutrition during pregnancy and the first few years of life improved health status during childhood and reduced stunting at age 3 [Martorell (1999); Martorell (1995); Habicht et. al (1995)]. The experimental design followed in the INCAP study not only shows that there exists catch-up potential in health outcomes but also suggests that nutritional interventions at early ages contributes towards the improvements in child health.

In the absence of an experimental design, Foster (1995) using data from Bangladesh use prior period exogenous changes in weather outcomes to identify the changes in subsequent health, as measured by weight. The study finds that it is the better-off households that were able to reduce the impact of the weather shock (flood) on child health and finds that access to credit is one of the important factors that enabled children to overcome some of the adverse economic conditions created by the flood.[10]

Some of the other studies in the literature have used longitudinal data to estimate dynamic models which are used to identify the extent to which childhood malnutrition affects subsequent health status [Adair (1999); Hoddinott and Kinsey (2001); Fedorov and Sahn (2005); Alderman et. al (2006); Johnston and Macvean (1995)]. Among these, Adair (1999) and Johnston and Macvean (1995) fail to address attrition bias and omitted variables bias. In particular, lagged health status is not treated as endogenous[11].

Three other closely related studies that are much more sound are Fedorov and Sahn (2005);

[10] See more on this literature in Strauss and Thomas (1998)

[11] Johnston and Macvean (1995) use type of fuel used and number of electrical appliances as right hand side covariates, both of which are likely to be correlated with household's socio-economic status. Adair (1999) use low birth weight, early menarche, height in the baseline year; all of which are correlated with household and individual-specific time-invariant unobservables. Almost 50% of the observations are attrited over time [Johnston and Macvean (1995)]. Selection problems are magnified by running regressions on stunted and non-stunted children as classified from the baseline year [Johnston and Macvean (1995); Adair (1999)]

Hoddinott and Kinsey (2001); Alderman et. al (2006). These studies not only examine the actual extent to which catch-up growth exists but also employ estimation techniques that address econometric concerns such as attrition bias and endogeneity of the lagged dependent. The three aforementioned papers discussed here estimate a dynamic conditional health demand functions to estimate the coefficient on the lagged dependent variable, that is, the catch-up term.

Fedorov and Sahn (2005) specify a dynamic conditional child health demand function in levels[12] and Hoddinott and Kinsey (2001) and Alderman et. al (2006) use a child growth specification[13].

Fedorov and Sahn (2005) use an Arellano-Bond (1991) and alternatively an Arellano-Bover (1995) type estimation strategy yielding coefficients of 0.19 and 0.21 on lagged height, respectively. Their results indicate reasonable catch-up potential. The main limitation of their paper is that the results rely on a very strong assumption, that is, lack of serial correlation in the error terms, which is usually not satisfied in dynamic panel data models [Deaton (1997); Blundell and Bond (1998); Blundell et. al (2000)][14].

Hoddinott and Kinsey (2001) use both two-stage least squares (2SLS) and maternal fixed-effects estimation techniques. In a levels models, yielding a coefficient of 0.56 and 0.18 respectively on the catch-up term reflecting partial catch-up effects. The 2SLS method adopted in Hoddinott and Kinsey (2001) addresses problems arising from random measurement error but may not address omitted variable bias arising from the potential correlation between the instruments and the individual and household-specific time-invariant unobservables[15]. The maternal fixed-effects esti-

[12]Fedorov and Sahn (2005) specify a levels specification; $H_{it} = \beta_0 + \beta_1 H_{it-1} + \sum_{j=1}^{R} \beta_j^X X_{jit} + \sum_{j=1}^{S} \beta_j^Z Z_{ji} + \epsilon_i + \epsilon_h + \epsilon_c + \epsilon_{it}$. Where ϵ_i is child specific time-invariant unobservable, ϵ_h is household specific time-invariant unobservable, ϵ_c is community specific time-invariant unobservable, and ϵ_{it} is the random time-varying i.i.d. term.

[13]$H_{it} - H_{it-1} = \beta_0 + \beta_G H_{it-1} + \sum_{j=1}^{R} \beta_j^X X_{jit} + \sum_{j=1}^{S} \beta_j^Z Z_{ji} + \epsilon_i + \epsilon_h + \epsilon_m + \epsilon_{it}$. The coefficient on β_1 from a dynamic levels specification on footnote 12 is equal to 1 + β_G from the growth specification here. Where ϵ_i is child specific time-invariant unobservable, ϵ_h is a household specific time-invariant unobservable, ϵ_m is the mother specific time-invariant unobservable, and ϵ_{it} is the random time-varying i.i.d term.

[14]It is shown later in the paper using a Hausman (1978) type specification test that the assumption of zero first-order and second-order serial correlation in the error terms is in fact not valid for the data in hand and may not necessarily be valid for other papers with a short time dimension (say less than 5 periods) as well.

[15]For example: birth weight (instrument used in Hoddinott and Kinsey, 2001) itself can also be endogenous on two accounts - One, children with higher birth weight reflect higher unobserved healthiness/innate ability and hence potentially correlated with other child specific unobservables in the model. Two, birth weight is usually known for births taken in a health facility reflecting household's socioeconomic status (Strauss and Thomas, 2008). This makes

mation strategy adopted by Hoddinott and Kinsey (2001) addresses omitted variable bias problem but cannot address measurement error bias.

Alderman et. al (2006) use a maternal fixed-effects instrumental variable (MFE-IV) estimation strategy which results in a catch-up coefficient of 0.43 in levels, reflecting partial catch-up effects. Their paper addresses biases coming from measurement error in data and other household and community specific time-invariant unobservables, addressing almost all sources of omitted variables bias and measurement error bias in data. However, individual-specific time-invariant unobservables such as the child's innate ability to fight diseases is treated as random. The individual-specific time-invariant unobservables such as the child's genetic ability to fight diseases and absorb nutrients could potentially be correlated with the instruments used in the first-stage regressions (no. of days the child was living prior to August 1980). The estimation strategy adopted by Alderman et. al (2006) though addresses biases coming from the correlation between household-specific unobservables and child's lagged health status, individual specific time-invariant unobservables remain a potential source of concern.

In addition, both Hoddinott and Kinsey (2001) and Alderman et. al (2006) estimate a growth specification that is likely to magnify the measurement error in height attainments and bias the estimated coefficient on the lagged dependent variable towards -1 which is equivalent to 0 in levels specification.

As discussed above, the following three papers - Fedorov and Sahn (2005), Hoddinott and Kinsey (2001), and Alderman et. al (2006) cannot completely address for both omitted variable bias and measurement error bias in data. It is the ability of the first-difference GMM strategy used in this paper that makes it especially attractive.

birth weight correlated with other household-specific unobservables in the model as well.

3 Model

Parents make investments in their children's health with the aim of improving the child's overall well-being. Following Fedorov and Sahn (2005), Strauss and Thomas (1998, 2008), health status in period t can be specified as a function of health inputs, environmental factors, individual demographic characteristics, household background characteristics, genetic endowments, time-varying health shocks, and time-invariant health endowments.

$$H_t = h(M_t, M_{t-1}, ..., M_0, I_t, I_{t-1}, ..., I_0, D_\sigma, \theta_{c\sigma}, \theta_c, \mu_{h\sigma}, \mu_h, G) \qquad \sigma = 0, 1, ...t \tag{1}$$

H_t is current health status measured by height-for-age z-score or height in cm. M_t is health input at time t which includes food and non-food consumption goods used towards the maintenance and or improvement of child health. It is assumed that households do not derive any direct utility from the consumption of health inputs except from its indirect use in the accumulation of child health output. I_t characterizes the environment where the child lives capturing infrastructure availability and disease environment in the community. D_σ reflects all time-varying demographic characteristics such as the child's age. $\theta_{c\sigma}$ includes all time-varying health shocks like fever and diarrhea. θ_c summarizes information about all time-invariant characteristics such as the child's gender and time-invariant health endowments like the child's innate ability to absorb nutrients and fight diseases. $\mu_{h\sigma}$ and μ_h capture household specific time-varying and time-invariant demographics and background characteristics such as parents rearing and caring practices. G summarizes information about all genetic endowments capturing genotype[16] and phenotype[17] influences that affect child health.

Following Strauss and Thomas (1992, 1995), the one-period lagged health status is assumed to be a sufficient statistic that captures the impact of all health inputs, environmental factors, and other

[16]Genotype influences include genetic endowments that are passed from the parents to the child via their DNA.

[17]Phenotype influences capture all observable characteristics of an individual, such as shape, size, color, and behavior that result from the interaction of genotype influences with the environment.

time-varying characteristics starting from birth up until the last observed period in the sample. By making this assumption we can substitute for all past period's determinants of child health by the one-period lagged health status in equation (1)[18]. Redefining equation (1), the dynamic child health production function can be re-written as:

$$H_t = f(H_{t-1}, M_t, I_t, D_t, \theta_{ct}, \theta_c, \mu_{ht}, \mu_h, G) \tag{2}$$

Where, health status in the current period is a function of the one-period lagged health status, current period health inputs, environmental factors, demographics, genetic endowments, health shocks, and household characteristics. The optimal choice of health inputs is determined by the household's utility maximization problem described below.

The household maximizes expected lifetime utility - U (3), subject to a lifetime budget constraint (4) where assets at end of period T must be equal to the difference between lifetime earnings and lifetime expenditure, and a period specific dynamic child health production function (5).

$$Max : U = E_t \sum_{t=0}^{T} \beta^t u_t[C_t, H_t, L_t; \theta_{pt}] \tag{3}$$

$$A_T = (\prod_{t=0}^{T}(1+r_t))A_0 + \sum_{t=0}^{T}(\prod_{\tau=t}^{T}(1+r_\tau))(w_t(T_t - L_t) + \pi_t - P_t^c C_t - P_t^m M_t) \tag{4}$$

$$H_t = f(H_{t-1}, M_t, I_t, D_t, \theta_{ct}, \theta_c, \mu_{ht}, \mu_h, G) \tag{5}$$

The sub-utility function (u_t) in each period depends upon consumption goods that include food and non-food consumption commodities, C_t, leisure, L_t, health status of the child, H_t, and certain unobserved preference shocks, θ_{pt}. β is the subjective discount factor which captures household preferences for higher utility today *vis-a-vis* the future. P_t^c is a vector of prices of food and non-

[18]We acknowledge that this assumption is strong but testing this assumption is beyond the scope of this paper.

food consumption goods. P_t^m is a vector of price of health inputs. w_t is the wage rate (price of leisure). T_t is parents total time endowment and A_0 is assets the households owns at the beginning of period 0. Profit income from farm and non-farm activities and all other sources of non-labor income is captured by π_t.

The solution for the above optimization problem relies on the following assumptions: (1) the household's lifetime utility function is assumed to be additively separable over time [Deaton and Meullbauer (1980); Fedorov and Sahn (2005); Strauss and Thomas (2008)]. (2) The sub-utility functions are quasi-concave and twice differentiable. (3) The one-period lagged health status is a sufficient statistic capturing the impact of all past health inputs and resources in determining current health status. (4) The household can potentially borrow and or lend against its future in each period tF. Combining all the above assumptions, using dynamic programming, the optimal dynamic conditional health input demand function (M_t^*)[19] can be written as:

$$M_t^* = m(H_{t-1}, P_t^c, P_t^m, w_t, I_t, \lambda, D_t, \theta_{ct}, \theta_c, \mu_{ht}, \mu_h, G, E_t(M_{t+1}^*)) \tag{6}$$

M_t^* is a function of the one-period lagged health status, current period prices of consumption goods, prices of other health inputs, wage rates, environmental factors, λ, a set of time-varying and time-invariant child level and household level characteristics, and $E_t(M_{t+1}^*)$. Where λ is the marginal utility of wealth in period 0. M_t^* today is not only a function of past and current period factors but is also affected by the household's expectations at date t about all future period's prices, incomes, environmental characteristics, and other factors which enter M_t^* through $E_t(M_{t+1}^*)$.

The dynamic conditional health demand function (7) can be obtained by replacing M_t in equation (5) by M_t^* in equation (6):

$$H_t^* = h(H_{t-1}, P_t^c, P_t^m, w_t, I_t, \lambda, D_t, \theta_{ct}, \theta_c, \mu_{ht}, \mu_h, G, E_t(M_{t+1}^*)) \tag{7}$$

[19] See Strauss and Thomas (2008) for a similar, yet even more general model with clear exposition of the solution method and assumptions needed to derive such a dynamic model.

A lot of the assumptions outlined earlier in this section are acknowledged to be strong, but testing these assumptions is not the aim or contribution of this paper. In addition, it is shown later how relaxing some of these very strong assumptions does not empirically change our results.

4 Empirical specification and identification

The main aim of this paper is to establish using a dynamic model the extent to which childhood malnutrition can permanently alter height attainments in the future. In addition, this paper also characterizes the determinants of child health outcomes in a static environment[20]. The static (8) and dynamic (9) conditional health demand functions estimated in this paper can be written as follows:

$$H_{it} = \beta_0 + \sum_{j=1}^{R} \beta_j^X X_{jit} + \sum_{j=1}^{S} \beta_j^Z Z_{ji} + \epsilon_c + \upsilon_{it}; \quad \upsilon_{it} = \epsilon_i + \epsilon_h + \epsilon_{it} \tag{8}$$

$$H_{it} = \beta_0 + \beta_1 H_{it-1} + \sum_{j=1}^{R} \beta_j^X X_{jit} + \sum_{j=1}^{S} \beta_j^Z Z_{ji} + \epsilon_i + \epsilon_h + \epsilon_c + \epsilon_{it} \tag{9}$$

H_{it} and H_{it-1} are the child's height-for-age z-score or height measured in centimeters at time t and t-1 respectively, where subscript i refers to the individual. X's are time-varying regressors which include child's age, household income, and community characteristics such as prices of food consumption goods, prices of health inputs, and community infrastructure variables. Z's include time-invariant regressors such as parental schooling variables and parental height.

It is in general difficult to obtain a composite measure of full income which captures wage and non-wage income from all sources. Along with the difficulty in obtaining a composite measure of household income, there also lies great deal of measurement error in incomes reported. Hence, similar to papers in the existing literature, this paper uses log of real per capita household con-

[20] See Mani (2007) for the derivation of the static conditional health demand function.

sumption expenditure [log(PCE)] as a proxy for household's measure of full income in the static specification (8) [Thomas et. al (1990); Thomas and Strauss (1992)].

In the dynamic model, λ is known as the marginal utility of wealth in period 0. λ is a function of both retrospective information (period 0 to period t-1) and prospective information (period t+1 to period T) on prices, incomes, child characteristics, and household characteristics, that enter the demand function through the lifetime budget constraint. Empirically, treating marginal utility of wealth as a constant would be a strong assumption since it relies on the existence of complete markets. However, households in most developing countries are credit constrained. Therefore, the assumption of complete markets is empirically relaxed here by using some sort of a proxy for household's access to credit. I use lagged measure of log of household's real per capita consumption expenditure in the right hand side to control for household's access to credit. The lagged measure of log(PCE) is assumed to capture the household's access to credit. The sequence of expected future household characteristics, prices, incomes, and other factors affecting current health through $E_t(M^*_{t+1})$ empirically enters either through the time-invariant household specific unobservables (ϵ_h) or the time-varying i.i.d term (ϵ_{it}) given in equation (9). Whether the sequence of factors that affect current health status through $E_t(M^*_{t+1})$ enter the empirical specification via (ϵ_h) and or (ϵ_{it}) depends upon whether the household assumes some of these expectations to be time-invariant or not. No specific assumptions are needed in this case since it does not affect our empirical work any differently. However, we do need to assume that the impact of $E_t(M^*_{t+1})$ on H^*_t enters the demand for current health only additively.

There are four sources of unobservable in the dynamic specification (equation 9) - ϵ_i, ϵ_h, ϵ_c, and ϵ_{it}. ϵ_i captures the time-invariant individual-specific unobservables reflecting the child's inherent healthiness. ϵ_h captures all time-invariant household-specific unobservables reflecting parental preferences toward child health and parents time discount rate. ϵ_c captures all time-invariant community-specific unobservables like community time-invariant endowments and political associations. ϵ_{it} includes child specific time-varying unobservables such as expected future health shocks, current health shocks, and expected future prices of - consumption goods, health inputs, wage rates, and other household characteristics, some of which are unknown to the child and all of which are unknown to the econometricians at date t. In addition to the unobservables in the dynamic specification, there are also two sources of unobservables in the static specification - υ_{it} and ϵ_c where, υ_{it} in equation (8) is assumed to be a time-varying i.i.d term[21] and ϵ_c is time-invariant community specific unobservable which is removed through community fixed-effects.

The condition of zero correlation between the error term and explanatory variables may never be satisfied with the inclusion of the lagged dependent variable in the right hand side [Deaton (1997); Blundell and Bond (1998); Wooldridge (2002)]. Hence with H_{it-1} endogenous, standard OLS estimate of β_1 is likely to be biased and inconsistent. The sources of endogeneity in H_{it-1} deserve careful explanation.

The one-period lagged health status, H_{it-1}, is likely to be correlated with the time-invariant individual-specific unobservables like the child's ability to fight diseases, which creates an upward bias in the estimated coefficient on the one-period lagged health status - β_1. The one-period lagged health status is also likely to be positively correlated with the time-invariant household-specific unobservables like parental preferences towards child health and time discount rate, again creating an upward bias in the estimated coefficient on the one-period lagged health status - β_1. Parents could also invest more in children who had lower health status in the last period making the coefficient on β_1 biased downwards. The time-invariant community-specific unobservables like political connections of a community are also likely to be positively correlated with the lagged dependent variable creating an upward bias in the estimated coefficient on β_1. At the same time, pro-poor policies at the community level can bias the estimated coefficient of β_1 downwards. In addition, β_1 is likely to be biased downwards, towards zero due to the presence of classical measurement error in height attainments.

Given the different sources of the potential biases in H_{it-1}, it is difficult to assign the net

[21] In the static specification, there are not enough observations with at least two children from the same mother or household to be able to separately control for household specific time-invariant unobservables and hence we must treat the time-invariant unobservables at the individual and the household level as random. An alternate method would be to estimate the static demand function in first-differences to remove all time-invariant unobservables from the specification. However, this comes at the great cost of losing the impact of all time-invariant parental characteristics in determining child health.

direction of bias on the estimated coefficient on the one-period lagged health status - β_1. However, one can broadly classify the main sources of the endogeneity in the estimated coefficient on the one-period lagged health status as omitted variables and or random measurement error in data.

It is empirically a difficult challenge to correct for both omitted variable bias and measurement error in data. This paper discusses variants of the IV/GMM estimation strategies that can be used to address either omitted variables bias and or random measurement error in data.

The first IV strategy followed here is a simple two-stage least-square (2SLS) with province fixed-effects, where the dynamic levels specification (9) is estimated using two-period lagged (1993) community characteristics as instruments for lagged height under the assumption that the community characteristics are exogenous, and that the time-invariant individual-specific, location-specific, and household-specific unobservables are random. The 2SLS estimation strategy followed here addresses random measurement error bias in H_{it-1} as the lagged community characteristics are likely to be uncorrelated with the time-varying individual-specific time-varying unobservables (ϵ_{it}). However, one cannot rule out for the correlation between the time-invariant unobservables (ϵ_i, ϵ_c, and ϵ_h) and the instruments used due to the presence of potential non-random program placement effects, and hence the estimated coefficient on H_{it-1} is likely to be biased upwards (if ϵ_c captures political association) or downwards (if ϵ_c captures pro-poor policies)[22].

A simple solution for removing all sources of unobserved heterogeneity (ϵ_i, ϵ_c, and ϵ_h) would be to estimate the dynamic specification (equation 9) in first-differences. The advantage of first-differencing is that it takes away all time-invariant unobservables from the estimation equation there by taking care of one of the potential sources of endogeneity in β_1, omitted variable bias. The disadvantage of first-differencing being that it also takes away the impact of certain important right hand side variables such as mother's height, father's height, mother's schooling, and father's

[22] Community level infrastructure variables are likely to be correlated with community specific time-invariant unobservables (see Rosenzweig and Wolpin, 1986). Ghuman et. al (2005) also discusses how the community level time-invariant unobservables could also be potentially correlated with other household specific observables creating an upward bias in the estimated coefficient on the household characteristics. In the two-stage least square estimates applied to the dynamic specification, we include for province fixed-effects and not location fixed-effects due to problems of multicollinearity that arise from little over-time variation in the location varying characteristics. Province fixed-effects address some of the concerns regarding endogenous program placement effects, but not all.

schooling that have an independent effect on H_{it}. A lot of the potential variation among the right hand side variables is also lost due to first-differencing. First-differencing alone cannot address biases coming from the correlation between $\delta(H_{it-1})$ and $\delta(\epsilon_{it})$ which stems from measurement error in data, and is still to be addressed. A simple first-difference method in the presence of random measurement error will create an even larger downward bias in the estimated coefficient on the first-differenced lagged height (see Griliches and Hausman, 1986).

The second estimation strategy adopted here follows an Arellano-Bond (1991) framework where the first-differences in lagged height is instrumented with community characteristics from 1993, and height from 1993, maintaining the assumption of lack of serial correlation in the error terms, and exogeneity of the community characteristics. It is shown in the results section (section 6.4) that the assumption of lack of serial correlation in the error terms is not satisfied here, and hence an Arellano-Bond (1991) type estimator cannot be used to obtain an unbiased and consistent parameter estimate on the catch-up term.

Certain other variants of the GMM estimation strategy like the Arellano-Bover (1995) and the System GMM (Blundell and Bond, 1998) estimators can potentially address both sources of endogeneity - omitted variables and measurement error. However, these two estimators also rely on the lack of serial correlation in the error terms, which is not satisfied in this paper, and hence cannot be used to obtain an unbiased and consistent parameter estimate on the catch-up term[23].

Third, the preferred first-difference GMM strategy adopted here uses only community characteristics from 1993 and it's interactions with child's age and mother's schooling to identify the changes in height between 1997 and 1993 (first-differenced lagged height). The first-difference GMM strategy only relies on the assumption of exogeneity of the community characteristics and provides us with an unbiased and consistent coefficient estimate on the catch-up term. Two-period lagged (1993) community characteristics like number of health posts in a community and other measures of community infrastructure are used to identify the changes in height between 1997 and 1993.

[23]See Blundell et. al (2000) for an outline on the additional restrictions needed for obtaining unbiased and consistent coefficient on the lagged dependent variable using the Arellano-Bover (1995) and System GMM estimators.

Health posts also locally known as *posyandus* which are located in almost all communities in Indonesia. These posyandus are community-sponsored sub-village health posts which provide basic maternal and child health care to neighborhood groups. They are primarily targeted towards meeting the health care needs of younger children in the age of 0 and 5 years - who are most vulnerable to health shocks. Health posts provide immunization services, oral rehydration solution packets, and vitamin supplements on a monthly basis. On occasional instances it also provides food supplements to young children. Health posts in a community actively contribute towards meeting the health care needs of children and hence the number of health posts present in a community during 1993 can be used as a good identifying variable to explain the subsequent changes in child health between 1993 and 1997[24]. Additionally, interactions between mother's schooling and the number of health posts in 1993; interactions between child's age in months in 1993 and the number of health posts in 1993 capture for the age and mother specific returns to availability of health post in the community. Electricity in the community reflects infrastructure availability and the disease environment, both of which affect subsequent changes in child height. Taken together, these instruments capture access to preventive measures of health and to some extent curative measures of health, both of which affect subsequent changes in child height. Recall that under the assumption that the community characteristics are exogenous, all the above mentioned instruments are valid for identifying the subsequent changes in height attainments among young children[25].

So far the potential pros and cons of following the different IV/GMM estimation strategies have been discussed. The results section outlines the actual coefficient estimates on the lagged dependent variable and the direction of bias in the estimated coefficient on the catch-up term. This paper attempts to choose the estimator that addresses both omitted variables bias and measurement error bias.

[24]The statistical relevance of these instruments used is discussed in section 6.3

[25]1993 measures of all community characteristics can be potentially used as instruments to identify the changes in health status between 1997 and 1993 (first-differenced lagged height). However, there is a severe weak instrument problem associated with using all the community characteristics from 1993 to identify the changes in lagged health status between 1997 and 1993

5 Data and variables

5.1 Indonesian Family Life Survey

The data used in this paper comes from the 1993, 1997 and 2000 waves of the Indonesian Family Life Survey (IFLS), a large-scale socio-economic survey conducted in Indonesia. The IFLS collects extensive information at the individual, the household, and the community level. The survey includes modules on measures of health, household composition, labor and non-labor income, farm and non-farm assets, pregnancy, schooling, consumption expenditure, contraceptive use, sibling information, and immunization [see Frankenberg et. al (1995, 2000) and Strauss et. al (2004) for more details on sample selection and survey instruments].

The IFLS is an ongoing longitudinal survey, the first wave of which was fielded during late 1993 and early 1994 (IFLS1). In IFLS1, 7224 households were interviewed. The first follow-up wave was surveyed during the second half of 1997 (IFLS2) just before the major economic and financial crisis in Indonesia. In IFLS2, 7629 households were interviewed of which 6752 were original IFLS1 households and 877 were split-off households. The third wave (IFLS2+) was a special follow-up survey fielded during the late 1998. A 25% sub-sample of the original IFLS1 households were contacted in late 1998 with the aim of analyzing the immediate impact of the 1997-98 economic and financial crisis. The fourth wave of the IFLS was fielded in 2000 (IFLS3). A total of 10435 households were interviewed in 2000. Of these, 6661 were original IFLS1 households and 3774 households were split-off households. The sample surveyed in 1993-94 represented 83% of the Indonesian population living in 13 of Indonesia's 27 provinces at the time. The 13 provinces are spread across the islands of Java, Bali, Kalimantan, Sumatra, West Nusa Tenggara, and Sulawesi. Provinces were selected to maximize representation of the population, capture the cultural socio-economic diversity of Indonesia, and yet be cost-effective given the size and the terrain of the country. A total of 321 enumeration areas (EAs)/communities were selected from these 13 provinces for final survey purposes.

The IFLS is unique in a number of ways - (1) it links individual, household and community level data bringing together an enormous amount of information that enables us to better understand the impact of household characteristics on individual level observables controlling for community infrastructure availability. (2) IFLS interviews members from different age groups (0-14 years interviewed by proxy, 15-49 years, and 50 years and older) capturing the overall demographic composition in a household. (3) Few other surveys collect health related measures, in particular, height in centimeters is not commonly collected in all household surveys. (4) The IFLS is particularly useful in estimating a dynamic panel data model as estimating such a model requires data from atleast two time periods and a lot of exogenous variables that can be used as potential instruments to address the endogeneity issues in the lagged dependent variable. (5) The IFLS data quality is excellent as numerous checks were done at the field level and at the data entry level. For example: IFLS provides best guessed age in years, date of birth year, date of birth month, and date of birth day information for all panel and new respondents from all three waves of the survey. Numerous variables are double-checked across waves and across books within the same wave to provide correct information to the user.

Location/geographic information for all respondents is available at four administrative unit levels in Indonesia (from smallest to the largest): community, kecamatan (subdistrict), kabupatan (municipality) and province. One would ideally like to use the community level code as the location variable to remove any location-specific time-invariant unobservables from the model and also control for community level time-varying characteristics in the right hand side of the empirical specification. There are two challenges in using the original community codes as the location variable in this study: First, community level data is only available for respondents residing in the 321 original IFLS communities. The IFLS does not provide detailed community level information for mover households except for some communities in 2000 [see details in the mini-CFS questionnaire from Strauss et. al (2004)]. Second, to do any location-specific fixed-effects, data must be available on at least 2 children residing in the same community from each of the three waves of the IFLS. In order to be able to match households with community level information in all three waves of the survey, and estimate fixed-effects models to remove time-invariant community level unobservables, the following decision rule is used to create the “location” variable used in this paper that is aimed at overcoming the two above mentioned constraints.

The “location” variable created here is assigned with the community code if there are 5 or more children residing in the same community[26]. In cases where this criterion fails, the “location” variable is assigned the code corresponding to the next level of aggregation, i.e., the kecamatan[27] code following the same rules. Similarly the kabupatan and lastly the province codes are assigned to the location variable in order to obtain at least 5 children from each of the newly created location variable. This new aggregation of the geographic units helps us combine household level and community level information and also allows the use of fixed-effects estimation techniques at the location level. It is this “location” variable which captures geographic information corresponding to each household in all three waves of the IFLS. All community level characteristics reported in the tables vary at the location level created here and not at the original community id level.

5.2 Attrition rates

Sample attrition primarily occurs at two levels - the individual level and the household level. Attrition at the individual level occurs when an individual from the original wave either cannot be followed in the subsequent waves or information on the dependent variable is missing due to measurement error in data or due to other restrictions imposed by the author. Attrition can be a problem only if, firstly, observable factors that result in attrition are correlated with the error term in the specification of interest (9), and secondly, if unobservables in the attrition equation are correlated with the unobservables in the empirical specification of interest (Fitzgerald et. al, 1998). This section provides details on household level and individual level attrition rates using the IFLS and addresses concerns regarding attrition bias.

[26]It is usually the case that less than 5 children are found only in communities which were not the original IFLS1 communities and are communities where mover households resided.

[27]The kecamatan and kabupatan codes are based on BPS (Indonesian central bureau of statistics) codification that can be easily linked to other nationally representation data like the SUSENAS. The definition of a kecamatan and a kabupatan continues to change over time. In order to use systematic codes of the kecamatan and kabupatans over time, I use the 1999 BPS codes that define the kecamatan and kabuptan codes for all IFLS communities from all three years of the survey.

In IFLS1, 7224 households were interviewed. In IFLS2, 94.3% of all original IFLS1 households were re-contacted. In IFLS3, 90.9% of all original IFLS1 households were interviewed (Strauss et. al, 2004). The follow-up surveys were only designed to target the original IFLS1 households. Household level attrition is at about 1.4% per year between 1993 and 1997 and at about 1.3% per year between 1993 and 2000. The IFLS follows households that move out of the community in which they are interviewed in the baseline year keeping household level attrition low [see Thomas et. al (2001) for more details on sample attrition in IFLS]. In addition details about attrition rates at the individual level are provided below.

From IFLS1 complete information on age in months, sex, and height in cm is available for 2203 children between the age of 3 and 59 months. Of these 2203 children, 1966 were followed in 1997, and 2051 of the original sample was re-contacted in 2000. A total of 1819 children between the age of 3 and 59 months in 1993 can be followed through the 1997 and 2000 waves of the IFLS - this sample excludes observations deleted due to measurement error in height attainments or age in months. There was an overall rate of 10.76% attrition between 1993 and 1997 and 6.90% between 1997 and 2000. Re-contact rates were much lower in 1997 as compared to 2000[28]. A simple mean test on the difference in height attainments between all children in 1993 and children who were lost over time is -0.76 with standard error of 0.80. The mean difference in height attainments between all children present in 1993 and a sub-sample of those who were present in 1997 and 2000 is not statistically different. This indicates that attrition rates are not related to differences in initial period health status[29], suggesting that attrition is more likely to be random.

[28]In analyzing household level attrition rates, Thomas et. al (2001) also find that attrition rates are higher between 1993 and 1997 as compared to 1997 and 1998. They attribute this decline in attrition rate to be associated with learning by doing in running a large-scale household level survey.

[29]Additionally a linear probability model on attrition is also estimated where the dependent variable, attrition is defined equal to 1 if the individual can be followed through the 1993, 1997 and 2000 waves of the IFLS, and zero otherwise. The right hand side regressors include height-for-age z-score, mother's schooling, father's schooling, mother's height, father's height, gender, age in months, measure of household income, mother's age, father's age, rural dummy, and location indicators. All the right hand side regressors belong to the baseline survey year, 1993. The coefficient on HAZ from 1993 is 0.002 with a standard error of 0.004, indicating an insignificant impact in determining attrition. Among the other regressors mentioned above, it is only the rural dummy which has a significant impact on attrition apart from the location indicators. Children residing in rural areas are more likely to be followed as compared to children residing in urban areas in the baseline year. This is similar to the findings by Thomas et. al (2001), who find that household level attrition rates are higher in urban areas compared to rural areas. In summary, the liner probability model also verifies that attrition is unrelated to endogenous observables like the child's health status from 1993 and

In addition, individual level attrition is not a real concern in this paper, given the estimation strategy adopted here. First-differencing removes all potential sources of unobservables like the child's genetic endowments which is likely to be correlated with the observables or unobservables that result in attrition, thereby creating attrition bias. In the presence of a first-difference estimation strategy, the only possible remaining source of attrition is that arising from the presence of random health shocks, such as infectious diseases that may affect health status in 1993. But, these health shocks from 1993 are also likely to be uncorrelated with the health shocks in 1997 and/or 2000. Hence, attrition arising from the existence of random, time-varying health shocks is not likely to contaminate the parameter estimate on the lagged dependent variable.

5.3 Sample size, variables, and descriptives

Martorell and Habicht (1986) and Satyanarayana et. al (1980) point out that decline in growth in height during the first few years of life largely determines the small stature exhibited by adults in developing countries. In addition height measured at young ages is also strongly correlated with attained body size as an adult [Spurr (1988), Martorell (1995)]. Hence, in this paper the initial sample is restricted to children less than 5 years of age in 1993[30]. In addition, the sample in this paper is restricted to include children who are less than 12 years of age in 2000 in order to keep the child health production function time-invariant for the complete sample here[31]. This additional restriction does not result in the loss of several observations because the initial sample includes children who are between the age of 3 and 59 months in 1993 and hence, by 2000, over 99% of the sample is still under 144 months of age. The final sample includes 1819 children for whom there

measure of household income. Hence the parameter estimates reported in this paper are not likely to be confounded by selection issues. See table 10 in appendix for complete results of the attrition regression.

[30] Although some amount of catch-up growth occurs during adolescence, it is not sufficient to overcome the initial loss in the growth in height (Martorell, 1999). Additionally, the catch-up potential in adolescence is limited by maturation. Early maturation also hinders catch-up potential. Almost all children mature somewhere between 11-14 years, thereby restricting growth potential. Hence, catch-up growth estimated using the sample of children less than 12 years, reflects a possible lower bound on the extent of actual long-run catch-up possible by the time the child is an adult and stature becomes predetermined for life. However, at the same time maturation during adolescence suggests that this catch-up coefficient is not likely to be a lot smaller that the true lifetime catch-up estimate.

[31] The child health production function varies between young children and teenagers going through pubescent growth spurts (Waterlow, 1988).

exist complete anthropometric details from all three waves of the survey.

The outcome variables of interest in this paper are: height-for-age z-score (HAZ) and height in centimeters. HAZ score is used as the dependent variable in estimating the static conditional child health demand function as specified by equation (8). Height in centimeters is used as the dependent variable in estimating a dynamic conditional health demand function as specified by equation (9).

Height-for-age z-score and height in cm are both well established long-run indicators of individual health status. Figure 1 in the appendix shows that z-scores flatten out by 48 months of age. Also the majority of children in the dynamic specification are older than 48 months, by which z-scores flatten out leaving little scope for any dynamics[32] However, height attained in centimeters is not only a long-run indicator of health status but also captures the dynamic effects in health outcomes. Figure 2 in the appendix highlights the strong linear relationship between height in centimeters and age in months, depicting continuous changes in height attainments.

The right hand side variables in the regression estimates include - age of the child, male dummy, male dummy interacted with age in months, logarithm of real per capita household consumption expenditure, mother's height in centimeters, father's height in centimeters, mother's completed grades of schooling, and father's completed grades of schooling. In addition to the aforementioned child level and household level characteristics, the regression estimates also include a series of location level time-varying characteristics such as an indicator for whether the individual lives in a rural area, log of real price of rice, log of real price of condensed milk, log of real price of cooking oil, distance to health center in km, dummy for presence of paved road, percentage of households with electricity, log of real hourly male wage rates, log of real hourly female wage rates, and number of health posts in a community. Information on age of the child, gender, and per capita consumption expenditure is obtained from the household questionnaires. Age and sex variables were checked to be consistent over time across all three waves of the survey. Information on household residence is also obtained from the household questionnaire. Prices of food

[32]It is growth faltering at young ages among children from developing countries that results in the decline in the z-scores. Most of this growth faltering occurs due to poor nutrition and diseases. See Shrimpton et. al (2001) for more discussion on growth faltering in young children.

consumption goods such as price of rice, price of cooking oil, and price of condensed milk are obtained from the community questionnaires. All prices are converted in real terms and expressed in logs. Hourly male and female wage rates are also converted in real terms and expressed in logs. Information on whether the community has a paved road or not, number of health posts located in a community, distance to the health center in km, and percentage of households with electricity in a community are also obtained from the community questionnaire.

Tables 1 and 2 show trends in mean height-for-age z-scores and percentage of children classified as stunted over the three waves of the IFLS. There exists significant improvement in mean height-for-age z-scores over time for children using both repeated cross-sectional[33] and panel data[34]. The statistics indicate that mean height-for-age z-scores worsen until 1997 and then improve during 1997-2000. The percentage of children classified as stunted also increases between 1993 and 1997 and then declines between 1997 and 2000. In summary, trends in child health status as measured by height-for-age z-scores have improved by the year 2000.

Table 3 depicts the relationship between levels of stunting during childhood (as measured in 1993) and height attained in centimeters during later stages of life (as measured in 2000). Male children initially classified as stunted in 1993 grow to be 4.65 cm shorter than their counterparts in 2000, who did not suffer from any evidence of long-run malnutrition during childhood. Similarly, female children initially classified as stunted in 1993 grow to be 3.81 cm shorter than their female counterparts who did not suffer from any malnutrition during childhood. There is no evidence of gender-differences in height attainments among stunted and non-stunted children. The pattern of no gender-differentials is also found in another important aspect of human capital accumulation, education as measured by primary school enrollment rates (Deolalikar, 1993). Also in examining mortality rates, Kevane and Levine (2001) find no evidence of "missing girls", i.e., daughters are not likely to suffer from higher rates of mortality than sons. Levine and Anes (2003) show that even in the aftermath of the crisis, girls did not fare worse than boys. Most of the literature from

[33] Cross-section data includes data for children between the ages 3 and 59 months in 1993, 1997, and 2000 waves of the IFLS.

[34] Panel data includes data for children initially between the ages 3 and 59 months in 1993 who are followed through the 1997, and 2000 waves of the IFLS.

Indonesia, suggests that there is no evidence of gender bias in favor of male children.

In this paper, pooling tests on gender in the first-differenced dynamic instrument variable specification gives an overall chi-square of 55.74 (0.00), which favors separating the sample for boys from girls and then estimating the first-difference equation. However, a chi-square test on all right hand side variables except the age and gender interacted coefficients is 10.61 (0.64). This suggests that the differences in height between boys and girls occurs only due to the age and sex specific differences in growth of height attainments and not due to differential catch-up effects between boys and girls[35] or any other socioeconomic characteristics. Hence, in this paper only coefficient estimates from the pooled regressions are reported controlling for interactions between the male dummy and age in months variables to capture the gender specific growth patterns in height attainments.

Table 4 gives information on the mean and standard deviation of all variables used in the regression specification.

6 Results

6.1 Results from estimating a static health demand function

Columns 1-4 in table 5, report coefficient estimates from regressing height-for-age z-score on child characteristics, household characteristics, and location interacted time dummies. In column 5 of table 5 these location-interacted time dummies are replaced with actual location time-varying characteristics. Columns 1-4 in table 5 capture the independent impact of household level and child level characteristics in determining current health status after controlling for all community level characteristics by using location/community interacted time dummies on the right hand side of the empirical specification. The regression coefficients reported in table 5 follow an OLS/IV estimation strategy with location fixed-effects. The standard errors reported in table 5 are not only

[35] A chi-square on the interaction between the first-differenced lagged height and the male dummy from the pooled first-difference GMM specification is 1.00 (0.31) which indicates that there are no gender differential catch-up effects in health outcomes.

adjusted for clustering at the individual level, but are also robust to the presence of any arbitrary form of heteroskedasticity.

The coefficient estimates obtained on the child and household characteristics from columns 4 and 5 of table 5 are not statistically different from each other, indicating that the choice of using location interacted time dummies vs. location time-varying characteristics is not likely to bias the coefficient estimates on the household characteristics and child characteristics reported in columns 4 and 5 of table 5.

The coefficient on the male dummy from table 5 has a negative sign, suggesting that females have better health than male children. This result is striking when compared to other Asian countries like India and Bangladesh which exhibit comparable levels of stunting, where one finds large significant gender differentials in favor of boys *vis-a-vis* girls. For Indonesia this is not very surprising, since the country does not traditionally suffer from large gender differences in human capital accumulation outcomes (see pg 25-26).

The relationship between height-for-age z-score and age in months is non-linear and the coefficient on the spline variables captures this non-linearity; indicating that z-scores decline till the age of 24 months and then improve and remain steady and or unchanged after 48 months. The interaction terms between the spline variables and male dummy captures the gender specific changes in health outcomes. Overall, females have higher z-scores as compared to their male counterparts.

Household characteristics included in the regression estimates are parent's completed grades of schooling, parental height in centimeters, and measure of household income. Parents schooling variable captures for the efficiency with which health inputs are transformed into health output. The coefficient estimates on mother's completed grades of schooling and father's completed grades of schooling reported in table 5 shows an expected positive relationship between parental schooling and child health. Every additional year of mother's schooling increases z-scores by 0.015 (column 1, table 5) standard deviations. Father's schooling has a positive though insignificant impact on z-scores. The IV estimates reported in column 4, table 5 also our preferred estimates indicate that neither of the parental characteristics have a statistically significant impact in determining child health. The positive correlation between household per capita consumption expenditure and mother's schooling is likely to have biased the coefficient estimate on mother's schooling upwards in column 1, table 5. This is contrary to much of the evidence in the literature (see Strauss and Thomas, 1998 for review). However, Behrman and Rosenzweig (2002) in a separate treatment, examining the impact of mother's schooling on child schooling (another measure of human capital) show that mother's schooling has little role in determining child schooling once the regression estimates are appropriately corrected for endogeneity in the mother's schooling variable.

Parental height variables capture the impact of genetic endowments in determining current health. Mother's height in centimeters and father's height in centimeters both capture the impact of different genetic endowments in ascertaining the child's current health status[36]. Every 1 centimeter increase in mother's height improves z-scores by 0.04 standard deviations and every 1 centimeter increase in father's height improves z-scores by 0.03 standard deviations (column 4, table 5). Mothers' height is likely to have a higher impact in determining child health as compared to fathers' height. This is similar to the results found by Ghuman et. al (2005) and Thomas and Strauss (1992).

The final household characteristic included in the regression specification is that of household income. Logarithm of real per capita household consumption expenditure is used to capture the household's complete resource availability. OLS estimates of log(PCE) from column 1, table 5 can be both biased upwards due to its correlation with time-invariant household-specific unobservables and biased downwards due to measurement error in data. Assets are exogenously determined in a static model and hence, log(PCE) is replaced with productive assets and total assets respectively in columns 2 and 3 of table 5. The results indicate that children residing in households with higher income enjoy better health. IV estimates of log (PCE) are reported in column 4 of table 5 where log(PCE) is instrumented with the sum of household productive assets, unproductive assets, and unearned income, which are assumed to be exogenous in a static model. The coefficient estimate on log(PCE) increases from 0.08 (column 1, table 5) to 0.24 (column 4, table 5) showing

[36]See Thomas and Strauss (1992) for discussion on the role played by parent-specific genetic endowments in explaining current health status.

that IV estimates of income have much larger impact on current health status. The increase in the coefficient estimate of log(PCE) from OLS to IV regressions indicates that OLS estimates of log(PCE) is likely to be biased downward due to measurement error and not biased upwards due to omitted variables[37]. The role of income is largely consistent with most related work examining the determinants of child health[38].

The role of community/location time-varying characteristics is also important in determining child health. In the light of endogenous program placement effects, not accounting for the correlation between community infrastructure variables and community level unobservables can bias coefficient estimates on the community characteristics [Rosenzweig and Wolpin (1986)]. To address this issue in its entirety, the original community codes must be used for estimating community fixed-effects models. However, the data restrictions discussed in section 5.1 of the paper suggest that the smallest level of geographic aggregation must be re-defined to a 'location' in order to remove any geographic time-invariant unobservables from the model. The coefficient estimates from an IV location fixed-effects model is reported in column 5, table 5 which captures the impact of the community/location level time-varying characteristics in determining child health, addressing only some potential concerns related to endogenous program placement effects.

Among the community level time-varying characteristics, the paper controls for prices of consumption goods, health inputs, wage rates, and community infrastructure variables. Prices of consumption goods included are - price of rice, price of cooking oil, and price of condensed milk[39]. The increase in the price of rice is associated with improvements in child health in both urban and rural areas (column 5, table 5). A priori one would think that increase in the price of the staple food consumption commodity must be associated with a decline in child health. However, a positive coefficient on rice prices in rural areas indicates that households in these areas are net producers of rice and not net consumers of rice. As for urban areas, the positive coefficient on the price of rice is

[37]The F statistic on the excluded instruments and the Hansen J statistic from the first-stage regression for the IV estimates reported in table 5 are appended at the end of table 5 and the complete first-stage regression estimates are summarized in table 11 of the appendix.

[38]Thomas et. al (1991); Thomas and Strauss (1992); Haddad et. al (2003); Glick and Shan (1998); all find a strong positive effect of per capita consumption expenditure in determining child health.

[39]prices are converted in real terms and expressed in logs throughout the paper

still surprising as residents in urban areas are likely to be net consumers of rice and not net producers of rice. Rice in itself has little nutritional component, it is only a source of carbohydrates for the body which provides energy. In addition, if households had access to other food consumption goods (excluding rice and including better substitutes for rice) like cassava, milk, vegetables and meat; then the prices of those food consumption goods would be more important in determining child health as compared to price of rice.

Increase in the price of cooking oil is associated with decline in child health (column 5, table 5). Spending on cooking oil may not be a large proportion of household per capita consumption expenditure but reflects spending on essential consumption goods. One important consumption good aimed only for children is condensed milk, also included in the regression results. The advantage of using condensed milk is that it does not need refrigeration, an important advantage in a country where not all households own a refrigerator. The price of condensed milk has a positive but insignificant impact in determining child health. Due to a lot of the missing variables in the price data for other consumption goods, this paper can only control for the price of rice, price of cooking oil, and price of condensed milk among our right hand side variables. It is acknowledged that ideally a range of consumption goods must be included in the right hand side. However data constraints do not allow us to control for prices of more consumption goods.

Also included in the regressions are prices of health inputs as captured by distance to health center, and price of parents time as captured by male and female specific hourly wage rates in a community[40]. None of these have a statistically significant impact on child health.

Measures of community infrastructure availability such as number of health posts (access to health care), presence of paved road (access to bigger cities), and availability of electricity (storage facility) are used as additional control variables. Number of health posts in a community is also shown to have a negative impact on child health. This is contrary to economic intuition but, similar to coefficient estimates found on the community infrastructure variables in the presence of endogenous program placement effects (Rosenzweig and Wolpin, 1986)[41]. Community infrastruc-

[40]Hourly wage rates are converted in real terms and expressed in logs

[41]Some of the biases associated with endogenous program placement effects are addressed by the use of location

ture availability like presence of paved road in the community and availability of electricity in the community are both positively associated with improvements in child health. Children residing in communities with a paved road have 0.10 standard deviation higher z-scores as compared to their counterparts from other communities. Similarly children residing in communities with greater prevalence of electricity have 0.0026 standard deviation higher z-scores.

6.2 Catch-up effects - complete, partial, or none?

The results from estimating a dynamic conditional health demand function using variants of the IV/GMM estimation strategy are reported in table 6. OLS estimate on the one-period lagged height is 0.53 (see column 1, table 6), this indicates less than partial catch-up in attained height. The OLS estimate is likely to be biased and inconsistent as it suffers from omitted variable bias and measurement error bias - as previously discussed in section 4.

The coefficient estimate on the one-period lagged height using a simple 2SLS estimation strategy is 0.83 (column 2, table 6), which is even larger than the OLS parameter estimate. The 2SLS estimation strategy uses community characteristics from 1993 as instruments for the lagged dependent variable, addressing the downward bias in the catch-up term caused by random measurement error. But cannot address biases arising from the correlation between time-invariant unobservables (ϵ_i, ϵ_c, and ϵ_h) and lagged height due to endogenous program placement effects, and hence, the parameter estimate obtained on the catch-up term using this strategy continues to be biased and inconsistent.

The coefficient estimate on the catch-up term reported in column 3, table 6 is -0.18 and is biased downwards as compared to the OLS estimate, 0.53 (column 1, table 6). An OLS method applied to a first-difference specification creates an even larger downward bias compared to an OLS method applied to a levels specification, magnifying the measurement error problem (see Griliches and

fixed-effects. However, the limitation of not being able to use the actual community code for fixed-effects is still likely to contaminate the parameter estimates on the community infrastructure variables. Especially the number of health posts in a community is endogenously determined and correlated with a lot of the time-invariant community level unobservables. The presence of an imperfect measure to remove the community time-invariant unobservables can continue to bias the estimated coefficient on the community infrastructure variables.

Hausman, 1986 for a discussion on this).

Parameter estimate from an Arellano-Bond (1991) type first-difference GMM strategy uses community characteristics from 1993 and height in cm from 1993 as instruments for the first-differenced one-period lagged height. The coefficient estimate on the first-differenced lagged height for this specification is reported in column 4, table 6 which produces a coefficient estimate of -0.07 on the catch-up term. The Arellano-Bond (1991) strategy does not address measurement error bias due to the correlation between the time-varying error terms and the two-period lagged height in the instrument set. The Hausman (1978) type specification test reported in section 6.4 shows that the assumption of lack of serial correlation in the time-varying error terms is not valid for this paper and hence the Arellno-Bond (1991) estimation strategy will also produce a biased and inconsistent coefficient estimate on the catch-up term.

The first-differenced GMM specification uses community characteristics from 1993 as instruments for the first-differenced one-period lagged height. This results in a coefficient estimate of 0.23 (column 5, table 6) on the catch-up term. The coefficient on the catch-up term from the first-difference GMM specification indicates larger catch-up effects compared to the coefficient estimate reported in the OLS specification, suggesting an upward bias in the OLS parameter estimate of the catch-up term. The catch-up term of 0.23 indicates more than partial catch-up in height attainments, that is, children with less than average height in 1993 will not continue to obtain less than average height attainments in 2000. This indicates that malnutrition during childhood is not likely to lock these children into lower health status as measured by height in centimeters in the future. The catch-up coefficient obtained from following a first-difference GMM strategy provides us with our preferred estimate on the catch-up term as it addresses both omitted variables bias (via first-differencing) and measurement error bias (via instrumental-variable techniques) in data.

In column 6, table 6, an alternate measure of household's long-run resource availability is used where one-period lagged assets (productive and non-productive assets included) are used to replace the one-period lagged log(PCE). This specification is to verify the robustness of the catch-up estimate, i.e., to see if the use of the two different measures of household resource availability alters the coefficient estimates on the catch-up term[42]. The coefficient estimate on the catch-up term reported in column 6, table 6 is 0.23 and uses a first-difference GMM strategy with the same instruments as those used in column 5, table 6. The coefficient estimates reported on the first-differenced lagged height in columns 5 and 6 of table 6 are statistically different from both zero and the ordinary least square parameter estimate[43]. In addition, the coefficient estimate on the catch-up term obtained from columns 5 and 6 of table 6 are not statistically different form each other which suggests that coefficient estimates on the first-differenced lagged height in columns 5 and 6 of table 6 are robust to the variables used to capture household's long-run resource availability and or household's access to credit.

Even if we were to assume that there are complete markets, that is, households can freely borrow and lend in each period. The assumption of complete markets would then imply that there should be no measure of household resource availability in the right hand side of the dynamic empirical specification. Estimating the dynamic specification using a first-difference GMM estimation strategy dropping lagged log(PCE) (our measure of household's access to credit) from the RHS using the same instruments as in column 5, table 6, yields a coefficient estimate of 0.24 on the catch-up term which is statistically significant at 10%. This result suggests that the coefficient estimate on the catch-up term is robust to the assumption of complete markets made in the model section of the paper.

Now comparing the coefficient estimate on the catch-up term obtained in this paper to some of the earlier literature. Hoddinott and Kinsey (2001) find a catch-up coefficient of 0.56 using data on children from Zimbabwe. Fedorov and Sahn (2005) report a coefficient of 0.19 on the catch-up term using data on children from Russia. Alderman et. al (2006) estimate a catch-up coefficient of 0.43 again using data on children from Zimbabwe. Children from Russia exhibit higher levels of catch-up potential as compared to children from Zimbabwe. Children from Indonesia too exhibit

[42]Even if we were to treat lagged log(PCE) as endogenous in the first-difference specification, then also the estimated coefficient on the catch-up term remains unchanged.

[43]A simple chi-square on the coefficient on the first-difference GMM being different from the OLS coefficient estimate are 4.82 (0.02) for estimates reported in column 5, table 6 and 4.65 (0.03) for estimates reported in column 6, table 6 with p-values in the bracket.

higher levels of catch-up potential compared to children from Zimbabwe.

6.3 A test of weak instruments in the dynamic panel specification

The preferred IV estimates reported here are also additionally robust to an important econometric concern - instrument validity. An instrument is defined to be valid only if it satisfies the following two conditions - (1) the excluded instruments must be strongly correlated with the endogenous regressor and (2) the instrument must be uncorrelated with the error term in the second stage regression. In the presence of weak correlation between the instruments and the endogenous regressors, the IV estimates reported here are likely to suffer from a higher bias and inconsistency compared to the bias obtained on the OLS parameter estimate (Blundell, 2005). It is hence important to verify that the IV estimates reported here satisfy the two above mentioned conditions.

Stock et. al (2002) and Staiger and Stock (1997) have discussed some test statistic that can be used to test the relevance of the instrument used in an IV estimation framework. Stock et. al. (2002) and Stock and Yogo (2005) define an instrument to be weak based on two criteria - First, based on the relative two-stage least squares (TSLS) bias where the instrument is deemed to be strong if the Cragg-Donald F statistic is large such that the TSLS bias with respect to the OLS bias is say at most x% (5, 10, 15 depending the extent of bias the author wants to allow). The second criterion is based on size, i.e., the instruments are defined to be strong if the Cragg-Donald F statistic is large enough that a 5% hypothesis test is rejected no more than say $x\%$ of the time, otherwise the instruments are weak. The Cragg-Donald F statistic is however based on the assumption of lack of first-order and second-order serial correlation in the error terms which is not valid in the current setting and hence the Cragg-Donald F statistic is not an appropriate test statistic for the dynamic panel data model estimated in this paper.

The bias in an IV coefficient estimate relative to an OLS estimate can also be approximated with the inverse of the F statistic on the excluded instruments obtained from the first-stage regressions (Murray, 2006). Based on the above definition of relative bias, the larger the F the smaller the relative bias from following an IV strategy compared to an OLS estimation approach. If F =1 the bias in 2SLS can be approximated to the bias in OLS estimates. If $F<1$ then the bias in 2SLS is even larger than the bias in OLS estimate. Staiger and Stock (1997) suggest a simple rule of thumb to test for instrument relevance. They suggest that in the presence of a single endogenous regressor, instruments are deemed to be weak if the first-stage F statistic on the excluded instruments is less than 10. However, the number 10 itself is quite arbitrary in its choice. In general, weak instruments cause two problems: (1) it brings the bias in the 2SLS/IV estimate closer or even larger than the OLS estimate. (2) It reduces the standard errors in IV estimates thereby producing incorrect inferences.

Since there does not exist a precise test statistic to check for instrument relevance of the instruments used in the first-difference GMM estimates reported in columns 5 and 6 of table 6. A combination of factors jointly help to support that the true coefficient estimate on the catch-up term is close to 0.23 and is statistically different from both zero and the OLS parameter estimate. The first-stage F statistic reported in columns 5 and 6 of table 6 are 3.06 and 3.14 respectively. The F statistics reported here if compared to the Staiger and Stock (1997) rule of thumb would identify the instruments as weak. However, using a different set of lagged community characteristics to identify the exogenous variation in the first-differences in lagged height maintaining the same stochastic assumptions as for the estimates reported in columns 5 and 6 of table 6 gives a coefficient estimate of 0.25 on the catch-up term with a first stage F statistic of 7.63, which is closer to 10. This clearly indicates no problem of weak instruments. The standard weak instrument problem does not seem to apply to this case since neither the significance of the parameter estimates changes and nor does the actual magnitude obtained changes under the presence of a smaller first-stage F statistic.

In addition to the test of strong correlation between the endogenous regressor and the instrument, it must also be the case that the instrument is uncorrelated with the error term in the second stage regression. The Hansen J statistic (1996) of 2.31 with a p-value of 0.51 (column5, table 6) and 2.12 with a p-value of 0.54 (column6, table 6) suggests that we cannot reject the null of instrument validity for the instruments specified in columns 5 and 6 of table 6. The coefficient estimate on the Hansen J statistic and the first-stage F test statistic on the excluded instruments are all appended at the end of the regression tables.

The two conditions of instrument relevance discussed in this section provide additional support for the reliability of the preferred estimates obtained using the first-difference GMM strategy.

6.4 A test of serial correlation in the error terms

In this section an attempt is made to determine whether or not there is serial correlation in the error terms of a dynamic panel model. An Arellano-Bond (1991) estimation strategy may not be suitable for the dynamic specification because of the presence of serial correlation in the time varying error terms, however this must be tested. A Hausman (1978) type test is incorporated to the Arellano-Bond (1991) and the first-difference GMM strategies specified in columns 5 and 6 of table 6. Under the null that there is no serial correlation in the error terms, the Arellano-Bond (1991) strategy must yield consistent and efficient parameter estimates on the first-differenced lagged height. However, if this assumption fails, then the first-difference GMM estimate must be chosen which is consistent and efficient under the alternative but not under the null.

The first-difference GMM (in column 5, table 6) estimator is tested against the Arellano-Bond (1991) (in column 4, table 6) estimator, where two-period lagged height is used as an instrument for the first-difference in lagged height in addition to all the instruments specified in the first-difference GMM specification reported in column 5 of table 6. The Hausman (1978) test statistic yields a coefficient of 0.30 with 0.12 as the standard error, rejecting the null. The coefficient estimates on the first-differenced lagged height are statistically significant and different under the two estimation strategies suggesting that the null of zero first-order and second-order serial correlation in the error terms is rejected. This section provides additional support in favor of the first-difference GMM strategy as the most preferred estimation strategy to be followed in a dynamic model especially, where serial correlation between the error terms is inevitable.

6.5 Role of child, household, and community characteristics in the dynamic conditional health demand function

Table 6 reports coefficient estimates from the regression of the dynamic conditional child health demand function specified in equation (9). Column 1 in table 6 reports coefficient estimates from following a simple OLS estimation strategy. The preferred first-difference GMM estimates are reported in column 5, table 6.

The coefficient on lag age in months from column 5, table 6 captures the positive relationship between age in months and attained height. The interaction term between lag age in months and the male dummy suggest that with age, improvements in height are slightly higher among females. This is similar to the patterns found in the static regression results.

In addition to the age and sex variables as controls in our right hand side, duration, i.e., the length of period measured in months between the two consecutive survey rounds controls for the uneven gap between the three survey rounds (1993, 1997 and 2000). For every additional month between survey rounds, there is a 0.49 centimeter increase in attained height between 1997 and 2000 (column 5, table 6). The coefficient on the interaction of lag age in months and duration captures the age differential growth patterns in height. The longer the duration between survey rounds, the slower the changes in height attainments among older cohorts. The interaction terms between lag age in months and both duration and the male dummy capture the age and sex differential patterns in growth of height attainments. The longer the duration and older the child, the larger will be growth in height for male children relative to their female counterparts. Child characteristics capture the biological process of growth in height that differs by age and sex. The coefficient estimates from the child characteristics are largely consistent with that found in the literature.

The coefficient estimates on the parental schooling variables and parental height variables reported in columns 1 and 2 of table 6 are inconsistent and hence, little can be drawn from these estimates. In our preferred specification, due to first-differencing the impact of parental characteristics is lost. This is one limitation of using any first-difference estimation approach.

Another household characteristic included in the regression estimates is the one-period lagged household consumption expenditure. Regression estimates from table 6 show that the one-period lagged log(PCE) in the dynamic function has a positive effect on current health status. The coefficient on lagged log(PCE) is 0.51 (column 1, table 6) in the OLS specification indicating a large positive impact of income on current health even after controlling for the one-period lagged health status. The coefficient on lagged log(PCE) in the first-difference GMM specification reduces to 0.22 (column 5, table 6) indicating the presence of a possible upward bias in the OLS coefficient estimate of lagged log(PCE) resulting from the correlation between time-invariant household specific unobservables and lagged log(PCE) reported in column 1, table 6. Income and child health exhibit a strong positive and significant relationship.

Community characteristics play an important role in determining child health outcomes in static models. Little is known about their influence in dynamic settings. Fedorov and Sahn (2005) report coefficient estimates on a series of community characteristics from the estimation of a static and dynamic conditional child health demand function and find that community characteristics have a larger role to play in determining current health in dynamic settings. There exists some evidence for the important role played by price of rice, price of cooking oil, measure of electricity, and measure of paved road in determining the child's current health status in a static framework as reported in table 5.

At the same time there is no impact of these community characteristics in the dynamic specifications especially for the preferred estimates reported in columns 5, table 6. After controlling for the one-period lagged health status, the effect of past community characteristics in determining current health largely diminishes. First-differencing removes all time-invariant variation among the right hand side regressors and additional instrumenting of the first-difference specification, results in a loss of over time variation in the right hand side variables. Both these factors explain for the little role played by health inputs in determining current health status in the dynamic specification.

6.6 Do catch-up effects differ with age?

It is usually hypothesized that younger children will experience larger catch-up effects as compared to older children [Martorell and Habicht (1986); Habicht et. al (1995)]. For example: Schroeder et. al (1995); Habicht et. al (1995) show that the impact of the nutritional intervention program in rural Guatemala had the most significant impact on improving the stature of children less than 3 years of age. This paper attempts to find similar support by adding an interaction term between the one-period lagged health status and lag age in months in the dynamic specification. A positive and significant coefficient estimate on the interaction term will indicate lower catch-up potential among older children. However, adding the interaction term in the empirical specification increases the endogeneity problem.

Columns 1 and 2 of table 7 report coefficient estimates on the one-period lagged health status and the interaction term between lagged health status and lag age in months using OLS and first-difference GMM estimation strategies. The first-difference GMM estimates reported in column 2, table 7 indicates a coefficient of 0.0010 on the interaction term indicating age differential catch-up effects, i.e., older children experience lower catch-up as compared to younger children. The F statistics on the excluded instruments are also valid and appended at the end of table 7. The Hansen J statistic testing the null of zero correlation between the error and the instrument set is also satisfied. Figure 3 plots the catch-up effects against age in months based on the regression estimates from column 2, table 7. Figure 3 indicates that, even though little, younger children exhibit higher catch-up potential than older children.

6.7 Further implications

Stunting during early childhood has long-term effects on an individual's future economic and social well-being. This paper captures the extent to which stunting in childhood manifests into poor health status in the future. In the absence of strong causal effects between childhood malnutrition and subsequent health status, some of the poor consequences associated with childhood malnutrition can be mitigated. In this paper, I find that childhood malnutrition causes some but not significant growth retardation in an individual's future physical well-being as measured by height attainments. I find that a malnourished child in the absence of any catch-up potential would by adolescence, grow to be 4.15 cm shorter than a well-nourished child. However, in the presence of partial catch-up effects, i.e., a coefficient of 0.23 as estimated in this paper indicates that a malnourished child will by adolescence grow to be only 0.95 cm shorter than a well-nourished child. This recovery from childhood stunting also has impact on an individual's schooling attainments and other socioeconomic characteristics. For example: Maccini and Yang (2005) examine the association between adult height attainments and schooling attainments using data from the IFLS. Using their estimates on the causal effects of adult height attainments on schooling attainments, and combining the methodology outlined in Alderman at. al (2006), I compute the magnitude to which the presence of partial catch-up effects affects schooling attainments. I find that a malnourished child, in the presence of partial catch-up effects (0.23) as predicted in this paper, by adolescence, is likely to complete 0.6 less grades of schooling compared to a well-nourished child from the same population. In the absence of any catch-up potential this coefficient estimate is likely to be four times larger.

7 Conclusion

In view of the ever growing concern among development economists for child health, this paper identifies the determinants of nutritional outcomes among Indonesian children. The findings suggests that it is mother's height, father's height, log of real per capita consumption expenditure, price of consumption goods, and measures of community infrastructure that are important for improving nutritional outcomes among children. The results outlined call attention to programs and policies that focus on community level infrastructure development, regulating prices of essential consumption goods, and providing access to credit.

This paper also captures the extent to which childhood malnutrition affects subsequent health status. A dynamic conditional health demand function is estimated where the coefficient on the lagged dependent variable captures the extent of recovery, if any, from childhood malnutrition. A coefficient of 0.23 on the one-period lagged health status indicates reasonable catch-up in height attainments. Recall from the introduction section, in the presence of partial catch-up potential, by adolescence, a malnourished child will grow to be 0.95 cm shorter than a well-nourished child. In the absence of any catch-up, by adolescence, a malnourished child will grow to be 4.15 cm shorter than a well-nourished child. Using the coefficient estimates reported in Maccini and Yang (2005) on the impact of height on various socioeconomic outcomes, I calculate that a decline in stature by 0.95 cm lowers schooling attainments by 0.6 grades of schooling. There is also some evidence showing that catch-up effects are marginally higher among younger children than older cohorts.

From a practical standpoint, the presence of partial catch-up effects and age-differential catch-up effects suggests that continued efforts must be made on the part of households and policy makers towards improving children's nutritional status at all ages. However, special emphasis must be on younger age groups as their catch-up potential is still the highest.

It is important that policy prescription is drawn from good empirical work. The first-difference GMM estimation strategy used here relies on much weaker stochastic assumptions than earlier work and addresses omitted variable bias and measurement error bias in data. The results reported here are in addition robust to econometric concerns such as sample attrition and weak instruments.

This paper and other papers from the earlier literature can be criticized due to the presence of potential regression to the mean effects [Cameron et. al (2005), and Coly et. al (2006)]. In this paper, I have mitigated some of this problem by addressing issues related to measurement error and sample selection in data (Cameron et. al, 2005). However, the individual-specific time-varying growth spurts in stature also result in regression to the mean effects. Therefore the presence of regression to the mean effects can never be completely ruled out.

To summarize, this paper uses both static and dynamic frameworks to outline the determinants of child health. The static results characterize the factors that must be targeted to improve nutritional status among children in Indonesia. On the other hand the dynamic results indicate that there exists catch-up potential in health outcomes, that is, children who suffered from chronic malnutrition during childhood are not likely to remain as undernourished forever. The presence of catch-up potential suggests that focused attempts must be made towards improving nutritional outcomes of children at all ages with special emphasis on the very young.

References

1. Adair, L. S. 1999. Filipino children exhibit catch-up growth from age 2 to 12 years. *Journal of Nutrition*, 129: 1140-48.
2. Alderman, H., Hoddinott, J., and Kinsey, W. 2006. Long-term consequences of early childhood malnutrition. *Oxford Economic Papers*, 58(3): 450-474.
3. Arellano, M. and Bover, O. 1995. Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics*, 68: 29-51.
4. Arellano, M. and Bond, S. 1991. Some tests of specification for panel data: Mote carlo evidence and an application of employment equations. *Review of Economic Studies*, 58: 277-297.
5. Barrera, A. 1990. The role of maternal schooling and it's interaction with public health programs in child health production function. *Journal of Development Economics*, 32: 69-91.
6. Behrman, J. R. and Rosenzweig, M. R. 2002. Does increasing women's schooling raise the schooling of the next generation? *American Economic Review*, 92: 323-334.
7. Blundell, R. and S. Bond. 1998. Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87: 115-143.
8. Blundell, R., Bond, S., and Windmeijer, F. 2000. Estimation in dynamic panel data models: Improving on the performance of the standard GMM estimators. Working Paper 00/12, *The Institute for Fiscal Studies*.
9. Blundell, R. and MaCurdy, T. 1999. Labor supply: A review of alternative approaches. *Handbook of Labor Economics*, III: 1560-169.
10. Blundell, R. 2005. Dynamic panel data methods. *Micro econometric lecture notes*.
11. Boersma, B. and Wit, J. M. 1997. Catch-up growth. *Endocrine reviews*, 18 (5): 646-661.
12. Cameron, N., Preece, M. A., and Cole T. J. 2005. Catch-up growth or regression to the mean? Recovery from stunting revisited. *American Journal of Human Biology*, 17: 412-417.
13. Cebu study team. 1992. A child health production function estimated from longitudinal data. *Journal of Development Economics*, 38: 323-51.
14. Coly, A. N., Miler, J., Diallo, A., Ndiaye, T., Benefice, E., Simondon, F., Wade, S., and Simondon, K.B. 2006. Preschool stunting, adolescent migration, catch-up growth, and adult height in young Senegalese men and women of rural origin. *Journal of Nutrition*, pp 2412-2420.
15. Deaton, A. 1997. The analysis of household surveys: A microeconometric approach in development policy. Baltimore: *Johns Hopkins University press*.
16. Deaton, A. and Muellbauer, J. 1980. Economics and consumer behavior. *Cambridge university press*.
17. Deolalikar, A. 1993. Gender differences in the returns to schooling and in schooling enrollments in Indonesia. *Journal of Human Resources*, 28(4): 899-932.
18. Deolalikar, A. 1996. Child nutritional status and child growth in Kenya: Socioeconomic determinants. *Journal of International Development*, 8: 375-94.

19. Fedorov, L. and Sahn, D. 2005. Socioeconomic determinants of children's health in Russia: A longitudinal survey. *Economic Development and Cultural Change*, pp 479-500.

20. Fitzgerald, J., Gottschalk, P., and Moffitt, R. 1998. An analysis of sample attrition in panel data. *Journal of Human Resources*, 33(2): 251-299.

21. Foster, A. 1995. Prices, credit markets, and child growth in low-income rural areas. *Economic Journal*, pp 551-70.

22. Frankenberg, E. and Thomas, D. 2000. The Indonesia family life survey (IFLS): Study design and results from Waves 1 and 2. *DRU-2238/1-NIA/NICHD*.

23. Frankenberg, E. and Karoly, L. 1995. The 1993 Indonesian family life survey: Overview and field report. RAND. *DRU-1195/1-NICHD/AID*

24. Frankenberg, E., Thomas, D., and Suriastini, W. 2005. Can expanding access to basic health care improve children's health status? Lessons from Indonesia's midwife in the village program. *Population Studies*, vol 59.

25. Ghuman, S., Behrman, J., Borja, J., Gultiano, S., and King, E. 2005. Family background, service providers, and early childhood development in Philippines: proxies and interactions. *Economic Development and Cultural Change*, pp 129-164.

26. Glewwe, P., and Miguel, E. 2008. The impact of child health and nutrition on education in less developed countries. forthcoming in the *Handbook of Development Economics*, vol. 4, edited by T. Paul Schultz and John Strauss eds., Amsterdam: North Holland Press.

27. Glick, P. and Sahn, D. E. 1998. Maternal labor supply and child nutrition in West Africa. *Oxford Bulletin of Economics and Statistics*, 60 (August): 325-55.

28. Griliches, Z. and Hausman, J. A.. 1986. Errors-in-Variables in panel data. *Journal of Econometrics*, 31: 93-118.

29. Grossman, M. 1972. On the concept of human capital and demand for health. *Journal of Political Economy*, 80(2): 223-255.

30. Haddad, L., Alderman, H., Appleton, S., Song, L., and Yohannes, Y. 2003. Reducing child malnutrition: How far does income growth take us? *World Bank Economic Review*, 17(June): 10731.

31. Habicht, J. P., Martorell, R., and Rivera, J. 1995. Nutritional impact of supplementation in the INCAP longitudinal study: Analytic strategies and inferences. *Journal of Nutrition*, 125(4S): 1042S-1050S.

32. Hansen, L.P., J. Heaton, and A. Yaron. 1996. Finite sample properties of some alternative GMM estimators. *Journal of Business and Economic Statistics*, 14(3): 262-280.

33. Hausman, J. A. 1978. Specification tests in econometrics. *Econometrica*, 46: 1251-1271.

34. Heller, P. and Drake, W. 1979. Malnutrition, child morbidity, and the family decision process. *Journal of Development Economics*, 6: 203-235.

35. Horton, S. 1986. Child nutrition and family size in the Philippines. *Journal of Development Economics*, 23: 161-176.

36. Hoddinott, J. and Kinsey, W. 2001. Child growth in the time of drought. *Oxford Bulletin of Economics and Statistics*, 63(4): 409-436.

37. Jacoby, H. and Skoufias, E. 1997. Risk, financial markets, and human capital in a developing country. *Review of Economic Studies*, 64: 311-335.
38. Johnston, F. E. and Macvean, R. B. 1995. Growth-faltering and catch-up growth in relation to environmental change in children of disadvantaged community from Guatemala city. *American Journal of Human Biology*, 7: 731-740.
39. Kevane, M. and Levine, D. 2001. The changing status of daughters in Indonesia, *mimeo*.
40. Levine, D. and Anes, M. 2003. Gender bias and the Indonesian financial crisis: Were girls hit hardest? *Center for International and Development Economics Research*.
41. Maccini, S. and Yang, D. 2005. Returns to health: Evidence from exogenous height variation in Indonesia, *mimeo*
42. MaCurgy, T. 1981. An empirical model of labor supply in a life-cycle setting. *Journal of Political Economy*, 89(6): 1059-1085.
43. Mani, S. 2007. Role of the household and the community in determining child health, *WIDER Research paper No. 2007/X*, Helsinki: UNU-WIDER.
44. Martorell, R. 1995. Promoting healthy growth: Rationale and benefits. In P. Pinstrup-Andersen, D. Pelletier, and H. Alderman (Eds.), *Child Growth and Nutrition in Developing Countries: Priorities for Action*, pp. 15-31, Ithaca: Cornell University Press.
45. Martorell, R. 1999. The nature of child malnutrition and its long-term implications. *Food and Nutrition Bulletin*, 20: 288-92.
46. Martorell, R. and Habicht, J. 1986. Growth in early childhood in developing countries. In F. Falkner and J.M. Tanner, eds., *Human Growth, Methodology, Ecological, Genetic, and Nutritional Effects on Growth*, second edition, New York: Plenum, 241-262.
47. Martorell, R., Habicht, J. P., and Rivera, J. 1995. History and design of the INCAP Longitudinal Study (1969-77) and its follow-up (1988-89). *Journal of Nutrition*, 125(4S): 1027S-1041S.
48. Murray, M. P. 2006. Avoiding invalid instruments and coping with weak instruments. *Journal of Economic Perspectives*, 20(4): 111-132.
49. Onis, de M., Frongillo, E., and Blossner, M. 2000. Is malnutrition declining? An analysis of changes in levels of child malnutrition since 1980. *Bulletin of the World Health Organization*, 78 (10).
50. Rosenzweig, M.R. and Wolpin, K.J. 1986. Evaluating the effects of optimally distributed public programs. *American Economic Review*, 76(3): 470-87.
51. Sahn, D. 1994. The contribution of income to improved nutrition in Cote de Ivoire. *Journal of African Economics*, 3: 29-61.
52. Satyanarayana, K., Radhaiah, G., Murali, M. R., Thimmayama, B.V.S., Pralhad, R. N., and Narasinga, R.B.S. 1989. The adolescent growth spurt of height among rural Indian boys in relation to childhood nutritional background: An 18 year longitudinal study. *Annals Human Biology*, 16: 289-300.
53. Shrimpton, S., Victora, C.G., Onis, M. de, Lima, R.C., Blossner, M., and Clugston, G. 2001. Worldwide timing of growth faltering: Implications for nutritional interventions. Pediatrics, 107 (5): e75.

54. Spurr, G. B. 1988. Body size, physical work capacity, and productivity in hard work: Is bigger better? In *Linear Growth Retardation in Less Developed Countries* (Waterlow, J. C., ed.), Nestle Nutrition Workshop Series Volume 14. Raven Press, New York, NY.

55. Staiger, D. and Stock, J. H. 1997. Instrumental variables regression with weak instruments. *Econometrica*, 65(May): 557-86.

56. Stock, J. and Yogo, M. 2004. Testing for weak instruments in linear IV regression, *mimeo*, Harvard University.

57. Stock, J., Wright, J., and Yogo, M. 2002. A survey of weak instruments and weak identification in generalized method of moments. *Journal of Business and Economic Statistics*, 20(4): 518-529.

58. Stock, J.H. and Yogo, M. 2005. Testing for weak instruments in linear IV regression. In D.W.K. Andrews and J.H. Stock, eds. *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*. Cambridge: Cambridge University Press, pp. 80-108.

59. Strauss, J., Lavy, V., and Thomas, D. 1996. Public policy and anthropometric outcomes in the Cote d'Ivoire. *Journal of Public Economics*, 61: 155-192.

60. Strauss, J. 1986. Does better nutrition raise farm productivity, *Journal of Political Economy*, 94(2): 297-320.

61. Strauss, J. and Thomas, D. 1995. Human resources: empirical modeling of household and family decisions. *Handbook of Development Economics*, vol. 3, edited by Jere R. Behrman and T.N. Srinivasan.

62. Strauss, J. and Thomas, D. 1998. Health, Nutrition and Economic Development, *Journal of Economic Literature*, 36(2): 766-817.

63. Strauss, J., Beegle, K., Sikoki, B., Dwiyanto, A., Herawati, Y. and Witoelar, F. 2004. The third wave of the Indonesia Family Life Survey (IFLS3): Overview and field report, *WR-144/1-NIA/NICHD*.

64. Strauss, J., Beegle, K., Dwiyanto, A., Herawati, Y., Pattinasarany, D., Satriawan, E., Sikoki, B., Sukamdi and Witoelar, F. 2004. Indonesian living standards - before and after the financial crisis, *Rand Corporation*.

65. Strauss, J. and Thomas, D. 2008. Health over the life course, forthcoming in the *Handbook of Development Economics*, vol. 4, edited by T. Paul Schultz and John Strauss eds., Amsterdam: North Holland Press.

66. Tanner, J. M. 1981. A history of the study of human growth, *Cambridge university press*.

67. Thomas, D., Strauss, J., and Henriques, M. 1990. Child survival, height for age, and household characteristics in Brazil. *Journal of Development Economics*, 33: 197-234.

68. Thomas, D., Strauss, J., and Henriques, M. 1991. How does mother's education affect child height. *Journal of Human Resources*, 26(2): 183-211.

69. Thomas, D. and Strauss, J. 1992. Prices, infrastructure, household characteristics, and child height. *Journal of Development Economics*, 32: 301-331.

70. Thomas, D. and Strauss, J. 1997. Health and wages: Evidence on men and women in urban Brazil. *Journal of Econometrics*, 77(1): 159-185.

71. Thomas, D., Frankenberg, E., and Smith, J. 2001. Lost but not Forgotten: Attrition and follow up in the Indonesian family life survey. *Journal of Human Resources*, 36 (3): 556-592

72. Thomas, D. and Frankenberg, E. 2002. The measurement and interpretation of health in social surveys, *mimeo*.

73. Thomas, D., Frankenberg, E., Beegle, K. and Teruel, G. Household budgets, household composition, and the crisis in Indonesia: Evidence at a longitudinal survey data, *mimeo*.

74. Waterlow, J. 1988. Linear growth retardation in less developed countries. *Nestle Nutrition workshop series vol. 14, chapters 1 and 2*.

75. WHO. 2000. Global Data base on Child Growth and Malnutrition. *Geneva: WHO*.

76. Wolfe, B. and Behrman, J.. 1982. Determinants of child mortality, health, and nutrition in a developing Country. *Journal of Development Economics*, 11: 163-193.

77. Wooldridge, J. 2002. Econometric Analysis of Cross Section and Panel Data, *Cambridge: MIT Press*.

Table 1: Summary statistics on Height-for-age z-score for children between the age of 3 and 59 months in 1993, 1997, and 2000

Years	Observations	% HAZ <-2	Mean	Mean difference (years)
1993	2203	40.26	-1.578	-0.127*** (1997-1993)
		(0.010)	(0.038)	(0.051)
1997	2356	41.38	-1.705	0.272*** (2000-1997)
		(0.010)	(0.036)	(0.044)
2000	3537	34.88	-1.432	0.145*** (2000-1993)
		(0.008)	(0.028)	(0.046)

Notes:
Source: IFLS - 1993, 1997, and 2000
Standard errors reported in parenthesis are robust to clustering at the household level
*** significant at 1%, ** significant at 5%, * significant at 10%

Table 2: Summary statistics on Height-for-age z-score for children between the age of 3 and 59 months in 1993, who are followed through the 1997 and 2000 waves of the IFLS

Years	Observations	% HAZ <-2	Mean	Mean difference (years)
1993	1819	40.626	-1.625	-0.134*** (1997-1993)
		(0.011)	(0.039)	(0.036)
1997	1819	42.056	-1.758	0.077*** (2000-1997)
		(0.012)	(0.027)	(0.019)
2000	1819	38.647	-1.681	-0.055*** (2000-1993)
		(0.011)	(0.025)	(0.037)

Notes:
Source: IFLS - 1993, 1997, and 2000
Standard errors reported in parenthesis are robust to clustering at the household level
*** significant at 1%, ** significant at 5%, * significant at 10%

Figure 1: **Lowess plot on height-for-age z-score against age in months for all panel children**

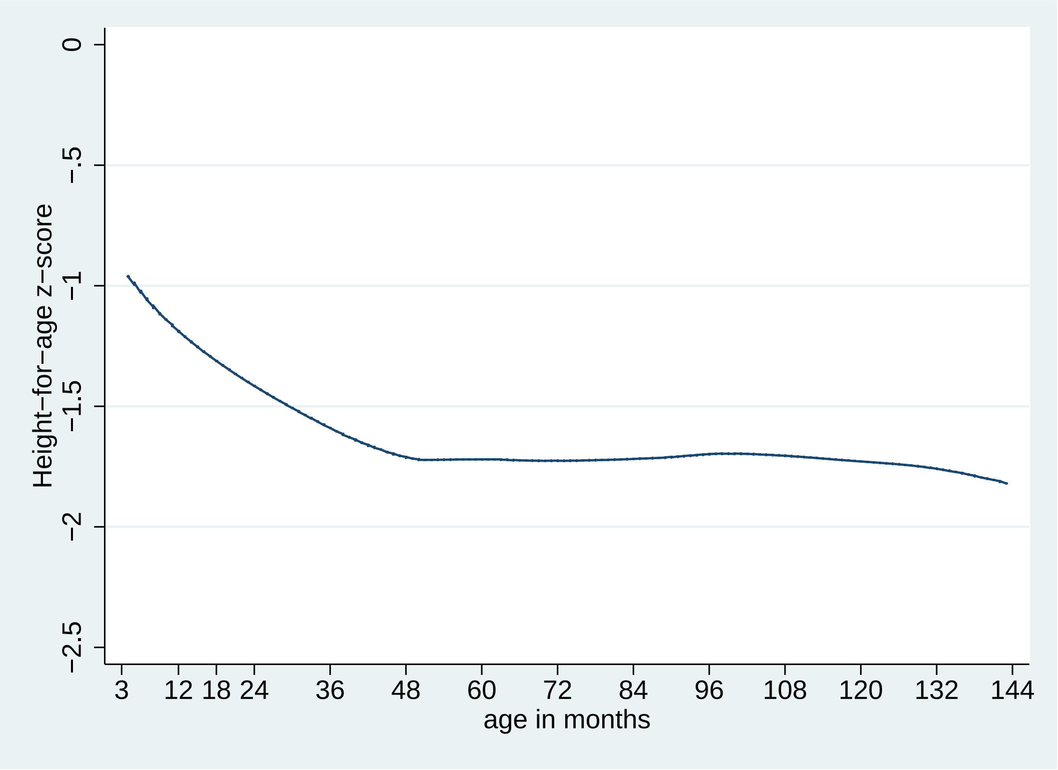


Figure 2: **Lowess plot on height in cms against age in months for all panel children**

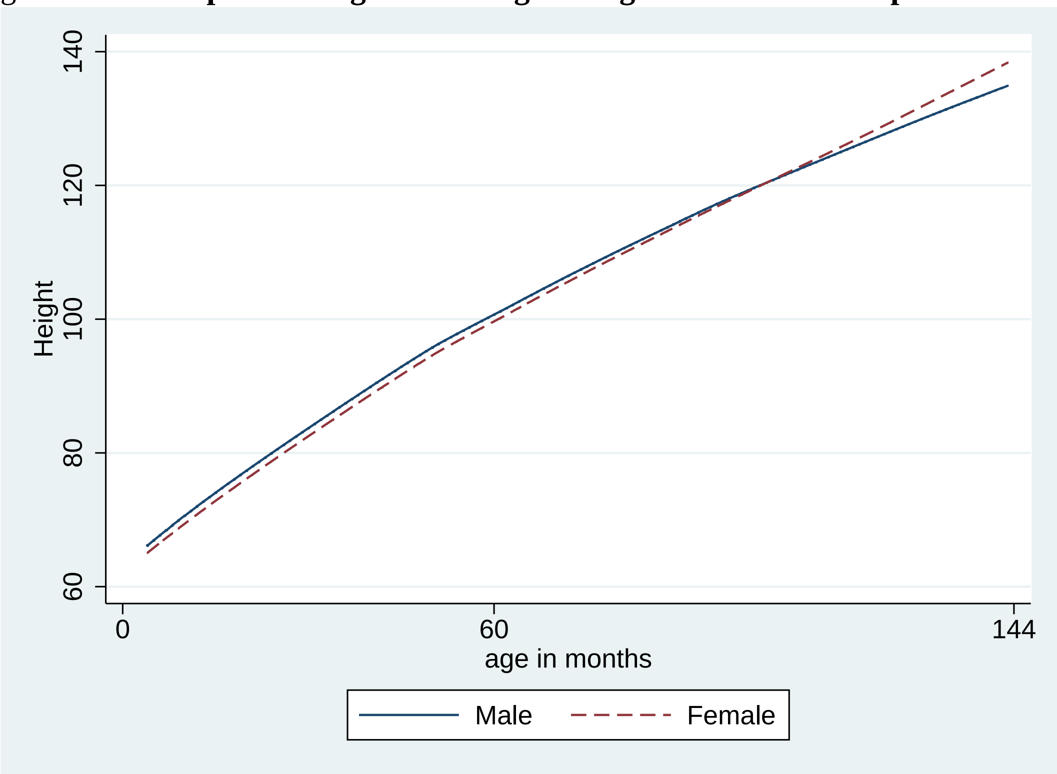


Table 3: Mean height attained in 2000 for all panel children between the age of 3 and 59 months in 1993

	Male(966)	Female (853)	Difference
Stunted (739)	121.35	122.06	-0.71
	(0.36)	(0.42)	(0.55)
Non-Stunted (1080)	126.00	125.87	0.12
	(0.46)	(0.37)	(0.59)

Source: IFLS - 1993, 1997, and 2000

Children with HAZ<-2 in 1993 were classified as stunted

Children with HAZ>=-2 in 1993 were classified as non-stunted

*** significant at 1%, ** significant at 5%, * significant at 10%

Table 4: Summary statistics of all variables used in the empirical specification

Variable	Observations	Mean	Std. dev
Height-for-age z-score (HAZ)	5457	-1.68	1.30
Height in cm	5457	105.86	19.42
Mother's height in cm	5457	150.54	5.11
Father's height in cm	5457	161.38	5.36
Mother's completed grades of schooling	5457	5.96	3.93
Father's completed grades of schooling	5457	6.90	4.33
Log of real per capita household consumption expenditure	5457	9.87	0.76
Square root of real per capita household productive assets	5457	1.51	2.61
Square root of real per capita household total assets	5457	4.48	3.79
Distance to the community health center in km	5457	5.08	4.57
Percentage of households with electricity	5457	76.68	26.92
Log of real male wage rate	5457	6.56	0.52
Log of real female wage rate	5457	6.19	0.85
Log of real price of rice	5457	0.86	0.20
Log of real price of condensed milk	5457	5.17	1.52
Log of real price of cooking oil	5457	1.74	0.43
Dummy=1 if the community has paved road	5457	0.74	0.44
Number of health posts in a community	5457	7.23	6.51

Source: IFLS - 1993, 1997, and 2000

Community here is the same as the location variable defined in the paper

Table 5: Regression Estimates of a Static Health Demand Function

Covariates	(1) **OLS** **HAZ**	(2) **OLS** **HAZ**	(3) **OLS** **HAZ**	(4) **IV** **HAZ**$^{+}$	(5) **IV** **HAZ**$^{++}$
Male dummy	-0.7659***	-0.7528***	-0.7647***	-0.7890***	-0.6991**
	(0.28)	(0.28)	(0.28)	(0.28)	(0.27)
Spline in age in months ($<$ 24 months)	-0.0780***	-0.0773***	-0.0778***	-0.0793***	-0.0779***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Spline in age in months ($>=$ 24 months)	-0.0013	-0.0012	-0.0013	-0.0015	0.0017*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.0008)
Spline in age in months (< 24)	0.0340***	0.0333***	0.0338***	0.0352***	0.0304**
*male dummy	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Spline in age in months ($>= 24$)	-0.0029***	-0.0028***	-0.0028***	-0.0030***	-0.0028*
*male dummy	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Mother's height	0.0480***	0.0482***	0.0480***	0.0475***	0.0475***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Father's height	0.0359***	0.0364***	0.0357***	0.0351***	0.0347***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Mother's schooling	0.0154**	0.0187***	0.0161**	0.0094	0.0082
	(0.007)	(0.006)	(0.007)	(0.007)	(0.007)
Father's schooling	0.0026	0.0051	0.0024	-0.0019	-0.0026
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
log(PCE)	0.0886***			0.2478***	0.2414***
	(0.03)			(0.09)	(0.08)
Productive assets		-0.0012			
		(0.007)			
Total assets			0.0158***		
			(0.005)		
Price of rice					0.3277**
					(0.13)
Price of cooking oil					-0.0992***
					(0.03)

Price of condensed milk					0.0023
					(0.01)
Rural dummy					0.0399
					(0.17)
Rural dummy*price of rice					-0.3113**
					(0.15)
Number of health posts					-0.0151*
					(0.008)
Distance to health center					0.0052
					(0.004)
Electricity					0.0029***
					(0.001)
Dummy for paved road					0.1021*
					(0.05)
Male wage rate					0.0152
					(0.04)
Female wage rate					0.0069
					(0.02)
observations	5457	5457	5457	5457	5457
Location fixed-effects	Yes	Yes	Yes	Yes	Yes

Notes:

- Source: IFLS - 1993, 1997, and 2000; *** significant at 1%, ** significant at 5%, * significant at 10%

- Robust standard errors adjusted clustering at the individual level are reported in parenthesis

- +, preferred estimates with community interacted time dummies; ++, preferred estimates with actual community level time-varying characteristics

- In (4), log(PCE) is instrumented with household productive assets. The F statistic on the excluded instruments is 197

- In (5), log(PCE) is instrumented with total household assets. The F on the excluded instruments is 196.42

- The first-stage regression estimates for column 5 are reported in table 11 of the appendix

- Also included in the regressions are dummy variables capturing missing observations on mother's schooling, father's schooling, mother's height, and father's height, where the missing observation was imputed by the sample mean

- Prices of consumption goods and hourly wage rates are converted in real terms and expressed in logs

Table 6: Regression Estimates of a Dynamic Health Demand Function

Covariates	(1) **OLS Height**	(2) **Two-Stage least-square Height**	(3) **OLS First-difference without IV's Height**	(4) **Arellano-Bond Height**	(5) **First-difference GMM Height**	(6) **First-difference GMM Height**
Lagged height or	0.5305***	0.8396***	-0.1820***	-0.0714**	0.2339*	0.2375*
catch-up coefficient	(0.02)	(0.21)	(0.03)	(0.03)	(0.13)	(0.13)
Male dummy	9.5326***	4.1735				
	(3.23)	(26.09)				
Lag age in months	0.4556***	0.0225	0.4172***	0.4044***	0.4290***	0.4276***
	(0.03)	(0.21)	(0.13)	(0.03)	(0.03)	(0.03)
Lag age in months*male dummy	-0.1533***	-0.0209	-0.1803***	-0.1725***	-0.1717***	-0.1692***
	(0.04)	(0.32)	(0.04)	(0.04)	(0.04)	(0.04)
Duration	0.7897***	0.2242	0.1737**	0.2312***	0.4950***	0.4937***
	(0.05)	(0.53)	(0.08)	(0.08)	(0.13)	(0.13)
Duration*male dummy	-0.1929***	-0.0202	-0.1583*	-0.1709	-0.1846	-0.1788
	(0.07)	(0.76)	(0.10)	(0.11)	(0.12)	(0.12)
Duration * lag age in months	-0.0075***	0.0015	0.0021***	0.0008	-0.0036***	-0.0036***
	(0.00)	(0.007)	(0.00)	(0.0007)	(0.002)	(0.002)
Duration*lag age in months	0.0030**	-0.0011	0.0043***	0.0039	0.0037***	0.0037***
*male dummy	(0.00)	(0.009)	(0.00)	(0.0008)	(0.000)	(0.000)
Mother's height	0.1798***	0.1167**				

	(0.01)	(0.05)				
Father’s height	0.1342***	0.0759*				
	(0.01)	(0.04)				
Mother’s schooling	0.0211	-0.0031				
	(0.02)	(0.03)				
Father’s schooling	0.0215	-0.0034				
	(0.02)	(0.02)				
Lagged log(PCE)	0.5161***	0.2448	-0.0114	0.1156	0.2240*	
	(0.12)	(0.16)	(0.11)	(0.11)	(0.13)	
Lagged household						0.0237
assets						(0.03)
Price of rice	0.8965	1.1670	-0.4556	-0.5379	-0.1291	-0.0420
	(1.22)	(1.01)	(0.67)	(0.65)	(0.74)	(0.74)
Price of cooking oil	-0.1884	-0.2140	0.0924	0.0377	-0.0398	-0.0693
	(0.26)	(0.31)	(0.16)	(0.16)	(0.19)	(0.19)
Price of condensed milk	0.0103	-0.1003	-0.0385	-0.0028	-0.0233	-0.0167
	(0.09)	(0.13)	(0.06)	(0.06)	(0.07)	(0.07)
Rural dummy	-0.3136	1.6042	-0.8436	-1.1385	0.0323	-0.0167
	(1.17)	(1.010)	(1.08)	(1.10)	(1.31)	(1.31)
Rural dummy*price of rice	-0.6590	-2.5500**	0.2103	0.3830	-0.1465	-0.1216
	(1.18)	(1.21)	(0.80)	(0.79)	(0.91)	(0.91)
Number of health posts	0.0017	0.0188	-0.0101	-0.0009	0.0016	0.0004
	(0.03)	(0.015)	(0.01)	(0.01)	(0.02)	(0.01)

Distance to health center	-0.0311	0.0415**	0.0132	0.0149	-0.0094	-0.0095
	(0.02)	(0.02)	(0.01)	(0.018)	(0.02)	(0.02)
Electricity	-0.0049	-0.0029	-0.0070	-0.0048	-0.0017	-0.0017
	(0.0075)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Dummy for paved road	0.0125	0.2032	-0.0349	0.0065	-0.0375	-0.0492
	(0.28)	(0.25)	(0.26)	(0.25)	(0.27)	(0.27)
Male wage rate	0.3662	0.3178	-0.1311	-0.0947	0.0169	0.0231
	(0.27)	(0.30)	(0.16)	(0.16)	(0.21)	(0.21)
Female wage rate	0.2610	-0.0898	0.0432	0.1568	0.1495	0.1466
	(0.17)	(0.15)	(0.12)	(0.11)	(0.13)	(0.12)
observations	5457	3638	1819	1819	1819	1819
Location fixed-effects	Yes	No	No	No	No	No
Province fixed-effects	No	Yes	No	No	No	No
F statistic on the excluded instruments from the first-stage regressions		3.60		31.90	3.06	3.14
		(0.02)		(0.00)	(0.01)	(0.01)
Hansen J statistic		0.022		9.86	2.31	2.12
		(0.88)		(0.04)	(0.51)	(0.54)
Hausman Specification Test					0.30	

specification 4 against specification 5		(0.12)	
Hausman Specification Test specification 4 against specification 6			0.30 (0.12)
C statistic testing the orthogonality of height in 1993 used as instruments for specification (4)	6.24 (0.01)		
C statistic testing the orthogonality of the first-differenced lagged log(PCE) in specification (5)		0.16 (0.68)	

Notes:

- Source: IFLS - 1993, 1997, and 2000

- *** significant at 1%, ** significant at 5%, * significant at 10%

- In (1), robust standard errors adjusted for clustering at the individual level are reported in parenthesis

- In (2)-(7) robust standard errors adjusted for clustering at the community level are reported in parenthesis

- In (2), Instruments used - two-period lagged measure of prevalence of electricity in the community, two-period lagged dummy=1 if the road in the community is paved.

- In (4), two-period lagged measure of electricity in the community, two-period lagged no. of health posts in the community, two-period lagged no. of health posts interacted with two-period lagged age in months, two period lagged no. of health posts interacted with mother's schooling, and two-period lagged height in cm

- In (5), two-period lagged measure of electricity in the community, two-period lagged no. of health posts in the community, two-period lagged no. of health posts interacted with two-period lagged age in months, and two-period lagged no. of health posts interacted with mother's schooling

- In (6), two-period lagged measure of electricity in the community, two-period lagged no. of health posts in the community, two-period lagged no. of health posts interacted with two-period lagged age in months, and two-period lagged no. of health posts interacted with mother’s schooling.

- Also included in the regressions are dummy variables capturing missing observations for each of the following variables - mothers schooling, fathers schooling, mothers height and fathers height, where the missing observation was imputed by the sample mean.

- p-values are reported for the F statistic on the excluded instruments and the Hansen J statistic.

- Prices of consumption goods and hourly wage rates are converted in real terms and expressed in logs

- Two-period lagged corresponds to information from the year 1993

Table 7: Regression Estimates of a Dynamic health demand function with interaction between lagged height and lag age in months

Covariates	(1) **OLS Height**	(2) **First-difference GMM Height preferred estimates**
Lagged height	0.3670***	0.2408**
	(0.02)	(0.09)
Lagged height*lag age in months	0.0033***	0.0010*
	(0.0003)	(0.0006)
Male dummy	9.5034****	
	(3.23)	
Lag age in months	-0.3189***	0.1884
	(0.03)	(0.15)
Lag age in months*male dummy	-0.1564****	-0.1660***
	(0.05)	(0.05)
Duration	0.1911**	0.3501**
	(0.09)	(0.13)
Duration*male dummy	-0.1893***	-0.1633
	(0.07)	(0.12)
Duration*lag age in months	0.0036**	-0.0009
	(0.007)	(0.002)
Duration*lag age in months*	0.0030**	0.0037***
male dummy	(0.001)	(0.001)
Mother's height	0.1707***	
	(0.01)	
Father's height	0.1263***	
	(0.01)	
Mother's schooling	0.0178	
	(0.02)	
Father's schooling	0.0208	

	(0.02)	
Lagged log(PCE)	0.5240***	0.2074
	(0.12)	(0.14)
Price of rice	1.5203	0.1964
	(1.24)	(0.79)
Price of cooking oil	-0.1403	-0.0165
	(0.26)	(0.20)
Price of condensed milk	0.0396	-0.0096
	(0.09)	(0.07)
Rural dummy	0.3524	0.2002
	(1.16)	(1.39)
Rural dummy*price of rice	-1.1019	-0.4095
	(1.18)	(0.97)
Number of health posts	-0.0094	-0.0021
	(0.02)	(0.02)
Distance to health center	-0.0305	-0.0175
	(0.02)	(0.02)
Electricity	-0.0035	-0.0018
	(0.007)	(0.006)
Dummy for paved road	0.0596	-0.0359
	(0.28)	(0.28)
Male wage rate	0.2707	0.0341
	(0.28)	(0.23)
Female wage rate	0.2324	0.1259
	(0.17)	(0.13)
observations	5457	1819
Location fixed-effects	Yes	No
F statistic on the excluded instruments from the first-stage		19.69 (0.00)

regressions	
Hansen J statistic	0.42
	(0.81)
C statistic testing the orthogonality of the two-period lagged log(PCE) in specification (2)	0.24 (0.61)

Notes:

- Source: IFLS - 1993, 1997, and 2000

- *** significant at 1%, ** significant at 5%, * significant at 10%

- In (1), Robust standard errors adjusted for clustering at the individual level reported in parenthesis

- In (2), instruments used - two-period lagged log(PCE), two-period lagged number of health posts in the community, two-period lag age in months, two-period lag age in months interacted with two-period lagged no. of health posts in the community.

- Also included in the OLS regression are dummy variables capturing missing observations for each of the following variables - mother's schooling, father's schooling, mother's height, and father's height where the missing observation was imputed by the sample mean.

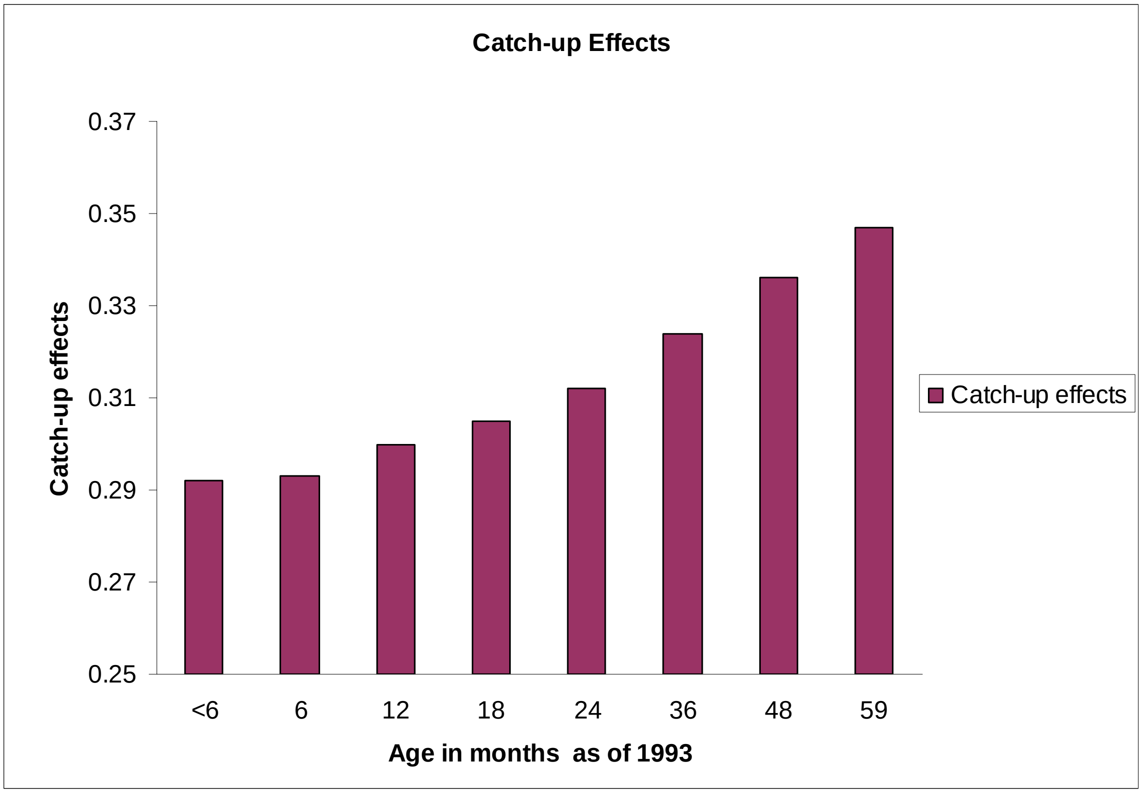
- P-values are reported for the F statistic on the excluded instrument and the Hansen J statistic.

- The F on the excluded instruments from the lagged height*lagged age in months - 64.50

- Prices of consumption goods and hourly wage rates are converted in real terms and expressed in logs

- Two-period lagged corresponds to information from the year 1993

Figure 3: Catch-up Effects



Appendix

Table 8: First-stage regression Results for the preferred estimates reported in columns 5 and 6 of table 6

Excluded and included instruments from the first-stage regressions	**coefficient estimates on the first-stage regressions variables reported in column 5, table 6**	**coefficient estimates on the first-stage regressions variables reported in column 6, table 6**
excluded instruments		
Two-period lagged electricity	0.03	0.004
	(0.005)	(0.05)
Two-period lagged no. of health posts	0.12*	0.12*
	(0.06)	(0.06)
Two-period lagged no. of health posts*	-0.003**	-0.003**
two-period lagged age in months	(0.001)	(0.001)
Two-period lagged no. of health posts*	0.007**	0.007**
mothers schooling	(0.003)	(0.003)
Included instruments		
First-difference in Lag age in months	0.04	0.04
	(0.06)	(0.06)
First-difference in Lag age in months	-0.05	-0.05
*male dummy	(0.05)	(0.05)
First-difference in Duration	-0.60***	-0.59***
	(0.14)	(0.14)
First-difference in Duration*	-0.04	-0.05
male dummy	(0.14)	(0.15)
First-difference in Duration	0.01***	0.01***
*lag age in months	(0.00)	(0.00)
First-difference in Duration*	0.001	0.001
lag age in months*male dummy	(0.001)	(0.001)
First-difference in Lagged log(PCE)	-0.44**	
	(0.17)	

First-difference in Lagged total assets		-0.03
		(0.04)
First-difference in Price of rice	-0.70	-0.85
	(0.86)	(0.85)
First-difference in Price of cooking oil	0.34	0.39
	(0.28)	(0.28)
First-difference in Price of condensed milk	-0.04	-0.05
	(0.09)	(0.08)
First-difference in Rural dummy	-1.81	-1.75
	(1.06)	(1.04)
First-difference in Rural dummy price of rice	0.52	0.45
	(0.70)	(0.71)
First-difference in Number of health posts	-0.01	-0.01
	(0.03)	(0.03)
First-difference in Male wage rate	-0.29	-0.29
	(0.34)	(0.33)
First-difference in Female wage rate	-0.09	-0.08
	(0.21)	(0.21)
First-difference in Distance to health center	0.05**	0.05**
	(0.02)	(0.02)
First-difference in Electricity	-0.009	-0.009
	(0.006)	(0.006)
First-difference in Dummy for paved road	0.07	0.08
	(0.37)	(0.37)
F statistic on the excluded instruments from the first-stage regressions	3.06	3.14
Hansen J statistic	2.31	2.12
	(0.51)	(0.54)

Notes:

- Source: IFLS - 1993, 1997, and 2000

- Two-period lagged corresponds to information from the year 1993

Table 9: Dynamic child health demand function estimated in first-differences using community interacted time dummies in the RHS

Covariates	**First-difference GMM**
	Height
Lagged height or catch-up coefficient	0.28**
	(0.11)
Lag age in months	0.42***
	(0.03)
Lag age in months*male dummy	-0.17***
	(0.05)
Duration	0.50***
	(0.12)
Duration*male dummy	-0.18
	(0.12)
Duration*lag age in months	-0.004**
	(0.001)
Duration*lag age in months*male dummy	0.003***
	(0.00)
Lag log(PCE)	0.21
	(0.15)
F statistic on the excluded instruments from the first-stage regressions	17.54 (0.00)
Hansen J statistic	2.69
	(0.26)

Notes:

- Source: IFLS - 1993, 1997, and 2000

- *** significant at 1%, ** significant at 5%, * significant at 10%

- Instruments used are log(PCE) from 1993, mother schooling interacted with no. of health posts from 1993, no. of heath posts in 1993 interacted with child's age in 1993

- A chi-square test on the catch-up term being statistically different from the OLS estimate gives a chi-square of 4.93 with 0.02 as standard error. This indicates that the catch-up term is significantly different from the OLS estimate.

- A chi-square test on the catch-up term being statistically different from zero gives a chi-square of 6.50 with 0.01 as standard error. This indicates that the catch-up term is significantly different zero.

Table 10: Determinants of sample attrition

Covariates	**OLS**
	attrition
Height-for-age z-score	0.002
	(0.004)
Male dummy	-0.0181
	(0.013)
Age in months	-0.0006
	(0.0004)
Mother's schooling	0.0027
	(0.002)
Father's schooling	-0.0020
	(0.002)
Mother's height	0.0006
	(0.001)
Father's height	0.002
	(0.001)
log(PCE)	-0.0002
	(0.01)
Mother's age	-0.0007
	(0.001)
Father's age	-0.0008
	(0.001)
Rural dummy	0.1428**
	(0.06)
Location fixed-effects	Yes
observations	2203

- Source: IFLS - 1993; robust standard errors reported in the parenthesis

- *** significant at 1%, ** significant at 5%, * significant at 10%

- Attrition =1 if the individual can be followed through the 1993, 1997, and 2000 waves of the IFLS and zero otherwise

Table 11: First-stage regression Results for the preferred estimates reported in columns 5 of table 5

Excluded and included instruments from the first-stage regressions	**coefficient estimates on the first-stage regressions variables reported in column 5, table 5**
excluded instruments	
Total assets	0.06***
	(0.004)
included instruments	
Male dummy	0.05
	(0.08)
Spline in age in months ($<$ 24 months)	0.007***
	(0.002)
Spline in age in months ($>=$ 24 months)	-0.001***
	(0.0003)
Spline in age in months (< 24)*male dummy	-0.003
	(0.003)
Spline in age in months ($>= 24$)*male dummy	0.0005
	(0.0004)
Mother's height	0.002
	(0.001)
Father's height	0.002
	(0.001)
Mother's schooling	0.02***
	(0.003)
Father's schooling	0.01***
	(0.003)
Price of rice	- 0.22***
	(0.07)
Price of cooking oil	0.14***
	(0.02)

Price of condensed milk	0.003
	(0.007)
Rural dummy	-0.32***
	(0.08)
Rural dummy*price of rice	0.15*
	(0.08)
Number of health posts	-0.0003
	(0.002)
Distance to health center	-0.007
	(0.003)
Electricity	-0.0002
	(0.0004)
Dummy for paved road	0.004
	(0.02)
Male wage rate	0.06**
	(0.02)
Female wage rate	0.03**
	(0.01)
observations	5457
Location fixed-effects	Yes
F statsitic on the excluded instruments from the first-stage regressions	196.42

- Source: IFLS - 1993

- *** significant at 1%, ** significant at 5%, * significant at 10%