7-23-2014

Height and Calories in Early Childhood

Andrew S. Griffen
The University of Tokyo, griffen@e.u-tokyo.ac.jp

Height and Calories in Early Childhood

Abstract
This paper estimates a height production function using data from a randomized nutrition intervention conducted in rural Guatemala from 1969 - 1977. Using the experimental intervention as an instrument, the IV estimates of the effect of calories on height are an order of magnitude larger than the OLS estimates. Information from a unique measurement error process in the calorie data, counterfactuals results from the estimated model and external evidence from migration studies suggest that the divergence between the OLS and IV estimates is driven by the LATE interpretation of IV. Attenuation bias corrected OLS estimates of the height production function imply that calories gaps in early childhood can explain at most 16% of the height gap between Guatemalan children and the US born children of Guatemalan immigrants.

Keywords
Health production, Health and Economic Development, Human resources, Human development, Migration

Disciplines
Biological and Physical Anthropology | Community Health and Preventive Medicine | Demography, Population, and Ecology | Economics | Health Economics | Nutrition | Public Health Education and Promotion

Comments
Griffen, Andrew S. 2013. "Height and Calories in Early Childhood." GCC Working Paper Series, GCC 13-09.

Height and Calories in Early Childhood

Andrew S. Griffen[1]

This paper estimates a height production function using data from a randomized nutrition intervention conducted in rural Guatemala from 1969 - 1977. Using the experimental intervention as an instrument, the IV estimates of the effect of calories on height are an order of magnitude larger than the OLS estimates. Information from a unique measurement error process in the calorie data, counterfactuals results from the estimated model and external evidence from migration studies suggest that the divergence between the OLS and IV estimates is driven by the LATE interpretation of IV. Attenuation bias corrected OLS estimates of the height production function imply that calories gaps in early childhood can explain at most 16% of the height gap between Guatemalan children and the US born children of Guatemalan immigrants. (*JEL:* I12, I15, O15)

Current version: July 23rd, 2014

Grand Challenges Canada Grant
Working Paper Number 13-09

[1]I would like to thank Jere Behrman, Cecilia Calderon, Flavio Cunha, Hide Ichimura, Yasuyuki Sawada, Petra Todd and Eric Weese for helpful comments as well as seminar participants at the University of Tokyo, the Institute of Developing Economies (IDE-JETRO), Autonoma Universitat de Barcelona and the University of Hawaii. I gratefully acknowledge financial support from the Grand Challenges Canada Grant 0072-03 to the Grantee, The Trustees of University of Pennsylvania. Contact: Economics Research Building, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan, Email: griffen@e.u-tokyo.ac.jp

I. Introduction

The formation of human height has been the subject of interest by economists in recent years because of its importance as a predictor of wages (Persico et al., 2004; Case and Paxson, 2008), health (Deaton, 2007), cognitive skills (Case and Paxson, 2008), and longevity (Fogel and Costa, 1997). The causal mechanisms behind these associations are complex but the economics profession has adopted the "nutritionist view" on human height as an easily measured indicator of general health status (Steckel, 1995). Understanding the formation of height can be an important guide for programs that attempt to improve health, especially in developing countries. Although sibling and twin research suggests that environmental factors explain only 20% of the variation in human height (Silventoinen, 2003), the role of the potentially policy specific environmental factors is disputed.[2] Some research finds strong correlations between per capita GDP during childhood (as a proxy for nutrition) and subsequent average adult population heights (Steckel, 1995) while others argue that cross population differences in averages heights are driven more by selective survival of taller children (Bozzoli et al., 2009; Gørgens et al., 2012) than by nutritional differences in early childhood (Deaton, 2007). Silventoinen (2003) states that one difficultly with much auxology (human growth) research is that it tends to use aggregate or cross-sectional data, which makes causal conclusions difficult.[3]

The formation of human height involves a complex interplay between polygenetic markers for height, the use of energy inputs by cells to divide, and how disease and deprivation causes feedback from the endocrine system to actually pause or suspend cellular division and growth (Tanner, 1990). Essentially theory points to genetics, nutrition and disease as all playing roles in the formation of height (Silventoinen, 2003) and a succinct characterization given by Deaton (2007), who states that "[h]eight is determined by genetic potential and by net nutrition," is a useful starting point for empirical work. A variety of observational, quasi-experimental and experimental evidence are all consistent with this biological theory of height formation and further suggest the importance of early environmental factors. Studies have examined the co-movement of secular trends in height and living standards (Fogel and Costa, 1997) and used heights to measure variation in standards of living by socioeconomic class (Komlos, 1994). Quasi-experimental evidence has

[2] Silventoinen (2003) also speculates that the effect of environment explain more of the variation in height in developming countries.

[3] An exception is work in public health studying inputs into stunting and child mortality that uses meta analyses of cross country RCTs as inputs into a simulation model to predict child mortality in developing countries (Bhutta et al., 2008, 2013). Although less attention is paid to the behavioral interpretation of parameters and the focus of that work is more on child mortality, the conclusions from my preferred OLS model are, broadly speaking, of a similar magnitude to their findings and the IV models presented in my paper are inconsistent with their work.

found impact on average population heights from famines (Meng and Qian, 2009), neonatal mortality (Bozzoli et al., 2009), and income shocks (Banerjee et al., 2010), and nutrition experiments, such as the Institute of Nutrition of Central America and Panama (INCAP) Longitudinal Study 1969 - 1977, provide strong evidence that nutrient intake during early childhood has both short and long term impacts on many anthropometric measures and other kinds of human capital (Martorell, 1995b; Hoddinott et al., 2008; Behrman et al., 2009). Heckman (2007) argues that, while the health economics has focused on documenting sensitive periods of development, more work should investigate the possibility of remediation following the production function literature on cognitive and noncognitive skills. Some studies strongly imply the possibility of such remediation; a study of the Dutch Hunger Winter 1944-45 (Stein et al., 1975) did not find impacts on adult height from famine exposure. A production function is the appropriate tool to answer many questions about hypothetical inputs interventions (Rosenzweig and Schultz, 1983). Furthermore, the parallels between the production process for skills and for height are actually quite strong and a common framework (Todd and Wolpin, 2003, 2007; Cunha and Heckman, 2007) can be brought to bear on both problems. A drawback relative to quasi-experimental methods for investigating height formation is the often strong modeling assumptions and the extensive panel data requirements on height, nutrition and disease.

To that end, this paper estimates a height production function using data from the INCAP Longitudinal Study, a randomized nutrition intervention conducted in rural Guatemala from 1969 - 1977, that collected extensive longitudinal data on calorie intake, disease measures and height.[4] I begin by presenting a general theoretical framework motivated by the theory of human growth and then proceed to impose a series of assumptions on the model to derive an empirically estimable function. This approach both clarifies the assumptions needed to estimate econometric models of height production and suggests a series of specification tests for the nested models. Using the INCAP experimental intervention and distance to the calorie distribution center in the INCAP experiment as instrumental variables for calorie inputs, the IV estimates of the production function suggest that the effect of calories on height is an order of magnitude larger than the OLS estimates. To explore the divergence between the OLS and IV estimates, I investigate the role of measurement error in attenuating the OLS estimates and the plausibility of the model counterfactuals us-

[4]In addition to height, the INCAP data have extensive measurement of other anthropometric outcomes. The basic pattern of regression results in the paper also hold for weight, head circumference, arm circumference, calf circumference, tricep skinfold, bicondylar breadth, arm length, subscapular skinfold and calf skinfold as outcome measures. The consistency across outcomes is perhaps not surprising given their correlation with height. Of the different anthropometric outcomes, I chose height as the outcome of interest given the focus on height in the economics literature as well as height information from Guatemalan migration studies.

ing both the OLS and IV estimates. Information from a unique measurement error process in the collected INCAP calorie data collection allows me to bound the effects of attenuation bias. The size of this attenuation bias cannot explain the divergence between the OLS and IV estimates. Furthermore, comparing calorie differences between the US and Guatemala during the same period as the INCAP experiment, I use the estimated production function to investigate the role of the US-Guatemala "calorie gap" in increasing the height of Guatemalan children. The difference between the OLS and IV estimates implies large differences in the counterfactual height gaps. Comparing these counterfactual height gap estimates from the production function to height gap estimates from migration studies of Guatemalan families to the US suggests that the IV estimates produce extremely unrealistic counterfactual estimates. Failing this out of sample validation test, the failure of traditional over-identification tests, and the variation in the IV estimates across instruments, despite the prima facie validity of the instruments, suggest that the IV estimates are not identifying policy relevant production function parameters (Heckman, 1997; Heckman et al., 2006). The attenuation adjusted, change in height OLS estimates imply that the average difference in calories between the US and Guatemala can explain at most 16% of the average difference in height between the INCAP children and height of the US born children from a sample of Guatemalan immigrants to the US. These results are consistent with (Deaton, 2007) who questions the specific role of nutrition in explaining population differences in height. In the conclusion, I discuss several other policy relevant inputs of interest and the interpretation of these findings.

The related literature on height is truly towering and spans the biomedical, public health and social sciences with sometimes seemingly little interaction between the fields. In economics, Behrman and Deolalikar (1988) and Strauss and Thomas (1998) both review a large literature on nutrition and health. Early height research in economics focused on whether there were causal links between maternal education and child height (Wolfe and Behrman, 1987; Behrman and Wolfe, 1987) and unpacking those mechanisms (Thomas et al., 1991). Martorell and Habicht (1986) provide a review of growth in childhood specifically in developing countries from an auxology perspective and Bhutta et al. (2008, 2013) provide an extensive meta-analyses from RCT public health interventions in developing countries. Even the literature on the INCAP experiment itself is quite extensive with much early work on the experimental impacts and subsequent analyses examining longer term impacts for each of the follow-up data collections during adolescence and adulthood (Martorell, 1995b). This paper builds on earlier work estimating models of height formation using these data by Schroeder et al. (1992) through incorporating instrumental variables estimation into the estimation of the effects of calories on height, developing the notion

of a height production function more closely with the estimation equation, addressing the role of measurement error in calories, conducting extensive specification testing and using and validating the estimated models with out of sample forecasts. An additional broader contribution is the use of instrumental variables in auxology research, which is apparently rare. The instrumental variable strategy in this paper follows several second generation papers using the INCAP data that have used the initial experimental variation as an instrumental variable to understand how early childhood nutrition effects educational outcomes (Maluccio et al., 2009), cognitive skills (Behrman et al., 2014) and wages (Calderon, 2009).

The remainder of the paper proceeds as follows. Section II develops a model of height production and section III describes the INCAP experiment and data collection. Section IV presents the results about the INCAP measurement error process and discusses the LATE interpretation of the IV estimates. Finally, section V presents the estimation results, section VI describes the counterfactuals and section VII concludes.

II. A Production Function for Height

Following theoretical notions about height formation, I model height as a function of genetic potential, disease and nutrition. In particular, I focus on diarrhea as a measure of the disease input and calories as a measure of the nutrition input. The use of both measures is motivated by theoretical and empirical evidence. Martorell and Habicht (1986) state that "[o]f all the problems of infection affecting young children in developing countries none is as important as diarrheal diseases."[5] Furthermore, cells actually use caloric energy during division and will suspend or halt growth during periods of low energy intake (Tanner, 1990). Energy, measured in the form of calories, are actually an aggregated measure of protein, carbohydrates and fats all of which are potentially individually important for growth.[6] However, as Pitt et al. (1990) point out, calories in developing countries often

[5]In addition to number of days sick with diarrhea, the data also contain parental reported number of days their child was sick with infectious disease symptoms, had ears/eyes/nose disease symptoms, had disease skin/hair symptoms, had a lack of appetite, was apathetic or irritable, spent in bed with illness, experienced fever and was with serious illness. I tried combining the disease measures using principal-components analysis and treating the first predict component as the disease input in the production function. None of the empirical patterns in the estimated model were affected.

[6]The question of whether protein deficiencies are important for height growth in developing countries has a long history and the effect of closing the "protein gap" was an important research and policy issue in nutrition field in the 1950s, 60s and 70s (Campbell et al., 2007). Interestingly, the original INCAP experiment was designed and funded to contribute exactly to that debate (Read and Habicht, 1993) although the INCAP researchers subsequently concluded from the experimental effects that energy intake (calories) and not its particular form (e.g., protein) mattered for growth in the INCAP context. However, it seems the literature has not entirely reached a consensus. Tanner (1990) writes that "[p]rotein seems less important than was once thought" while Silventoinen (2003) states that "[i]n developing countries ... [t]he lack of

come from one or two food sources so that even if data existed with individual consumption measures of protein, carbohydrates and fats, the measures would be highly collinear and it would be difficult to identify the effects of each component separately. The same argument also applies to questions about identifying the effects of diet diversity and micronutrients. More direct evidence on micronutrients from an extensive meta analysis by Bhutta et al. (2008) suggests only small impacts on height growth from micronutrient interventions. Essentially the subsequent model and empirical results can be interpreted as the effect of calories conditional on a particular aggregation of calories, diet diversity and micronutrients.

Consider the following general model of height production:

$$H_{ia} = H_a(C_{ia}, ..., C_{i0}, D_{ia}, ..., D_{i0}, X_i, \mu_i, \varepsilon_{ia}) \tag{1}$$

where H_{ia} is the height of child i at age a. $C_{ia}, ..., C_{i0}$ are the complete history of calories chosen for that child in periods before the measurement of height, $D_{ia}, ..., D_{i0}$ is the history of disease, X_i is an individual specific, time invariant determinant of height such as gender, μ_i is an unobserved (to the researcher) child specific determinant of height, and ε_{ia} represents a shock to height production. In this most general specification the entire history of inputs can affect current period height. The subscript a in (1) indicates that the effects of inputs on height can vary by age. For example, the marginal impact of calories on child height could be larger when the child is younger. There are many econometric challenges to estimating (1). The functional form is unknown, the inputs are chosen by children or family members with knowledge of the unobserved heterogeneity μ_i, reverse causality is potentially important (growing children consume more calories but more calories do not make children grow), and calorie data are noisy. Moreover, optimizing behavior can make the problem even more difficult. Parents can respond to previously low inputs if a child is "lagging behind" or can give more calories to make-up for a period of poor growth. That type of behavior would cause correlation between any observed current inputs and unobserved lagged inputs. To estimate (1), the researcher needs to make a series of assumptions about its functional form and about the information available to parents and children when choosing inputs. Because the estimated parameter from a production function gives a ceteris paribus impact of an input, one can, conditional on the maintained assumptions of the

protein is a particularly important factor explaining the slow (height) growth rate." The latter statement seems to be not supported by the INCAP data itself. However, Read and Habicht (1993) propose that protein may be limiting in certain environments but not in others, which in effect calls for a production function with complementaries in the sources of calories as inputs. Although in this case researchers would likely need cross country or cross region data to have sufficient variation in calorie composition and even then there would arise questions of omitted variables.

model, use the estimates to perform counterfactual analyses, such as what would be the impact of increases in specific inputs on height outcomes.

To derive an empirically implementable strategy, I make the following assumptions. The spirit of the exercise follows (Todd and Wolpin, 2003, 2007) in making explicit the assumptions needed for different estimators to be consistent.[7]

Assumption I: *The height production function is linear*

$$\begin{aligned} H_{ia} =& \beta_a^a C_{ia} + \beta_{a-1}^a C_{ia-1} + ... + \beta_0^a C_{i0} \\ &+ \alpha_a^a D_{ai} + \alpha_{a-1}^a D_{ia-1} + ... + \alpha_0^a D_{i0} + \\ &+ \gamma_a X_i + \lambda_a \mu_i + \varepsilon_{ai} \end{aligned} \tag{2}$$

The parameters in (2) have both subscripts and superscripts. This allows the marginal effect of inputs applied in the period indicated by the subscript to have different effects on height in the period denoted by the superscript. Because the data will be measured in discrete time intervals, I assume that an input with subscript a refers to inputs applied between a and $a-1$. In addition, both the time invariant inputs and the child-specific endowment can have different effects on height by age as γ_a and λ_a have age-specific subscripts. Without further assumptions, it would be necessary to estimate a different equation by each age, which both limits the options to control for endogeneity and reduces statistical power. Consider the following additional assumption:

Assumption II: *Lagged inputs have constant effects by age*

In terms of the coefficients, assumption II can be expressed as $\beta_{a-t}^{a'} = \beta_{a-t}^{a'-1}$, $\alpha_{a-t}^{a'} = \alpha_{a-t}^{a'-1} \forall a', t \geq 1$ and helps to simplify the model by canceling the effects of lagged inputs. To see this, consider differencing (2) by age under assumption II:

$$\begin{aligned} \Delta H_{ia} = H_{ia} - H_{ia-1} =& \beta_a^a C_{ia} + (\beta_{a-1}^a - \beta_{a-1}^{a-1}) C_{a-1i} + ... + (\beta_0^a - \beta_0^{a-1}) C_{i0} \\ & \alpha_a^a D_{ia} + (\alpha_{a-1}^a - \alpha_{a-1}^{a-1}) D_{ia-1} + ... + (\alpha_0^a - \alpha_0^{a-1}) D_{i0} + \\ & (\gamma_a^a - \gamma_{a-1}^a) X_i + (\lambda_a - \lambda_{a-1}) \mu_i + \varepsilon_{ia} - \varepsilon_{ia-1} \\ \Delta H_{ia} =& \beta_a^a C_{ia} + \alpha_a^a D_{ia} + \tilde{\gamma}_a^a X_i + \underbrace{(\lambda_a - \lambda_{a-1}) \mu_i + \varepsilon_{ia} - \varepsilon_{a-1i}}_{\text{Unobserved}} \end{aligned} \tag{3}$$

where the second equality follows from assumption II. Notice that X_i does not cancel because the effect of the parameter is time varying in the full model. Now the coefficient

[7] De Cao (2011) also adapts the Todd-Wolpin framework to study height using data from the Cebu Longitudinal Health and Nutrition Survey. Her results are not directly comparable to mine given that the Cebu data is missing data on children from age 2 to 8.

has an added tilde and a different interpretation as the effect of the time invariant variable on changes in height. Although the model is much simplified in terms of parameters to estimate, there are still several difficulties. First, changes in the productivity of the endowment, $\lambda_a - \lambda_{a-1}$, representing growth, could be correlated with inputs. Second, parents or children may respond to lagged shocks, which would produce correlation between inputs and ε_{a-1i}. Third, the model would need to be estimated separately by each age. Instrumental variables can in principle solve problems one and two while the ability to estimate age specific heterogeneity becomes a statistical power issue. The following assumption allows the model to be simplified further even further:

Assumption III: *Contemporary inputs have constant effects*

In terms of the coefficients, assumption IV assumes that $\beta_a^a = \beta_{a-1}^{a-1}$, $\alpha_a^a = \alpha_{a-1}^{a-1}$ and $\gamma_a^a = \gamma_{a-1}^{a-1}$ $\forall a$. One important consequence of III is that the data can be pooled over ages so that the change in height model becomes simply:

$$\Delta H_{ia} = \beta C_{ia} + \alpha D_{ia} + \tilde{\gamma} X_i + (\lambda_a - \lambda_{a-1})\mu_i + \varepsilon_{ia} - \varepsilon_{a-1i} \tag{4}$$

An additional consequences of assumptions II and III is that the height model (2) becomes:

$$H_{ia} = \beta \sum_{t=1}^{a} C_{ti} + \alpha \sum_{t=1}^{a} D_{ti} + \gamma X_t + \lambda_a \mu_i + \varepsilon_{ia} \tag{5}$$

so that only total inputs matter up age a, which is a cumulative specification for inputs. Several of the assumptions can be tested directly with the data. Assumption I can be tested by including squared terms and interactions. Assumption II can be tested by including lagged inputs into the change in height production function and testing their joint significance. Assumption III can be tested by estimating models separately by age. Finally, comparing estimates from the height model and the change in height model and comparing estimates with and without instruments can inform about the likelihood of correlation between inputs and unobserved determinants of growth.

III. Data

To estimate the different variants of the height production function, I use data from the INCAP Longitudinal Study 1969 - 1977, a village-wide randomized nutrition intervention and data collection in rural Guatemala.[8] Researchers selected four villages to participate

[8]See Habicht and Martorell (1992) for an in-depth description and discussion of the experimental design. One important feature of the experiment is that some children were treated only for parts of their life because

in the study. In both treatment and control villages, a centralized feeding center was set up that offered supplements that were fortified with micronutrients and offered twice-daily. All villagers could consume the supplements and more supplement was given upon request. Following a matched pair design, two treatment villages received the offer to consume *atole*, which was a high-calorie, high protein vegetable porridge, and two control villages received the offer to consume *fresco*, which was a slightly sweet, cool drink with less calories and no protein per serving.

Daily supplement intake data was recorded for all children and pregnant and lactating mothers. In addition, researchers collected anthropometric measures, such as height and weight, on all children under the age of seven, and data on home calorie consumption at three, six and twelve month intervals with the frequency of data collecting depending on the child's age. Data on number of days sick with diarrhea and other illnesses was also collected in three, six and twelve months intervals. Given the theoretical discussion given above that height depends on net nutrient intake (calories net of disease) and genetic endowments, all of the ingredients are available to estimate the height production function.[9]

There are three sources of variability in the frequency of data collection. The first source comes from the design of the data collection for anthropometric measures and home calories intake. Children between ages 0 and 2 were purposively measured at 3 month intervals, children between 2 to 4 years old at 6 month intervals and children from 4 to 7 years old at 12 month intervals. The second source comes from the different frequency of measurement of home and supplement calories. The third source of variability arises because the actual age of the children at each data collection did not correspond exactly to

of the beginning and ending of the experiment and data collection. For example, children born prior to 1969 only entered the sample after 1969 and children born after 1971 and onwards only have data collection until at most age 6. This created some natural variation in exposure to the experiment that has been examined in previous work using longer term follow-up data. In the current work, because the data collection ends when children exit the sample I cannot make use of the timing of exposure. The data collection also changes the samples used to estimate the height and the change in height models. In the height model, the necessity of an entire history of calorie inputs limits the sample because of the sample truncation from the beginning of the experiment and data collection. The change in height model, like other value-added models, essentially only needs two data points so the model is less demanding of the data and the estimation sample is subsequently larger. In addition, under the maintained modeling assumptions, the truncation of the data collection as a result of the beginning and ending of the experiment becomes less important in the change in height model.

[9]This is not the first estimate of the effect of calories on height using these data. The closest paper is Schroeder et al. (1992). However, in their paper many specification tests are not performed and reverse causality is a source of bias. Moreover, the focus in this paper is different because I develop the assumptions embedded in the estimating equation closely with the notion of a height production function and use the production function to perform out of sample forecasts. In addition, their estimates condition on a measure of socioeconomic status, which potentially changes the interpretation of the coefficients on the nutrition inputs (Wolpin, 1997). It is also unclear what the socioeconomic status variable proxies because under the assumptions listed above the INCAP data contains all of the necessary inputs. Finally, this paper also adds the use of instrumental variables, which was not performed in the initial analyses.

the 3, 6 or 12 month intervals. For example, at the 24 month data collection, some children might be 23 months old and some children might be 25 months old.

To address these forms of data frequency variability, I treat the data in the following way. First, to keep the timing consistent across data collection rounds, I use the data at a twelve month frequency. That is, I use the height and home calories measures at 1 year old, 2 years old, etc. In this way, when taking the model to the data, the interpretation of the model coefficient of interest becomes the effect on height of a one calorie increase in average daily caloric intake over the course of a year. Second, for the supplement calorie inputs, I use the exact amount of average supplement intake recording from the previous year, and for the home calorie inputs, I use the measured amount of home calories at each yearly data collection round as the measure of home calories for the previous year. I simply sum the measures to get a variable for average daily total calorie intake for the year. Combining the measures in this way obviously induces difference sources of measurement error given the way calories are measured. I address the role of measurement error in the next section and incorporate the theoretical results into the estimation. Third, to address the variation in the age at measurement conditional on the data collection round, I linearly interpolate between data rounds so that I have an estimate of the child's height at the exact age corresponding to 1 year old, 2 years old, etc. In practice, the correlation coefficient between interpolated height and the measured height is above 0.99 at all ages so using the interpolated height or the measured height makes no quantitative difference for the magnitude of the model estimates.

The data contain two sources of potentially exogenous variation. The first source is the experimental variation across villages itself. Although clustering bias is a potential issue, Martorell (1995a) report that "the anthropometric characteristics of both children and adults were strikingly similar in the different villages before 1969 [when the experiment began]," which reflects the matched pair nature of the experimental design. The second source of exogenous variation is the measured distance from each house to the feeding center. In contrast to the experimental variation, the distance instrument utilizes child specific, within village variation. As I will show, both the experimental and distance instruments are very strong predictors of individual total calorie intake and are both plausibly exogenous to unobserved determinants of child growth.[10]

[10]Maluccio et al. (2009) also use the experimental variation as instrument for supplement intake to examine the effect on educational outcomes but elect not to use the distance variation. This, however, is because schools were actually also located in the center of the villages so that there may have been a direct impact of distance on educational outcomes. In my case, the distance instrument may be inappropriate in levels because perhaps the location of the feeding center was not exogenous to unobserved child characteristics however it seems that in differences that the criticism would be less likely to hold. In addition, the IV and OLS divergence is greatest when using the experimental variation only so the point holds *a fortiori* using the

Table 1 displays summary statistics from the data. The data are disaggregated by *atole* and *fresco* villages and pooled over age and time. The children in the atole villages were significantly taller and also consumed significantly more total calories. For each kind of village, the increase in supplement calories outweighed the decrease in home calories, which is evidence of the so-called "fly-paper" effect where a transfer "sticks" to the child.[11] Because the modeling assumptions impose different requirements on the data and the sample size changes depending on the model being estimated I display atole and fresco differences both by the height estimation sample and by the change in height estimation sample. The basic pattern of differences holds across both samples, however, notably, the children in the height sample are shorter on average than children in the change in height sample.

IV. Measurement Error, IV and LATE

Data on calories are typically collected from 24-hour recall surveys. Families or individuals are asked to recall everything they ate in the previous 24 hours. Researchers then use conversion factors to convert the types and quantities of food reported into calories, proteins, and other nutrients. 24-hour food recall surveys have test-retest reliability on the order of 0.5 (Lechtig et al., 1976), where the test-retest reliability is the correlation coefficient between two repeated measures.[12] If calories are measured with classical measurement error, then the OLS coefficient is attenuated according to:[13]

$$E[\hat{\beta}_{OLS}] = \beta \frac{\sigma_c^2}{\sigma_c^2 + \sigma_u^2}$$

more plausibly exogenous instrument.

[11] Using the INCAP data Islam and Hoddinott (2009) rigorously document this effect.

[12] Test-retest reliability may be affected by conversion error in converting recorded calories into actual calories consumed, recall error on the part of individuals and actual variability in consumption by individuals across days. In my model, I use average calorie intake over a period, which implicitly assumes perfect substitutability across time within a period. Clearly variation in calorie intake may be an important issue in how calories affect growth. However, given the difficulty in collecting accurate individual level calorie intake at a specific point in time, the possibility of collecting an accurate, high frequency panel of calorie intake is very remote.

[13] The formula for attenuation bias is more complicated in the presence of multiple regressors but the pattern of empirical results in this paper are not sensitive to excluding the additional regressors. Another source of concern may be measurement error in diarrhea measures, which are self-reported. In addition, days sick with diarrhea may also be correlated with unobserved healthiness or tendency to grow. Unfortunately neither the experimental nor the distance instruments are strong predictors of days sick with diarrhea, which mitigates the possibility to control for endogeneity or measurement error using IV. However, the results from calories in this paper suggest some caution in estimating a biological production function with IV. In addition, the calorie results are similar when excluding days sick with diarrhea which suggests linearly independent effects.

where σ_u^2 is the variance of measurement error in calories and σ_c^2 is the variance in calories. The formula for the test-retest reliability is also exactly equal to the attenuation bias, which shows an attenuation adjusted estimate can be backed out by dividing the OLS estimate by the a priori known test-retest reliability of 0.5 according to:

$$\hat{\beta}_{\text{Adjusted OLS}} = \frac{\hat{\beta}_{OLS}}{.5}$$

This analysis suggests that OLS estimates of the marginal impact of calories will be approximately half of the true coefficient. As is well-known, instrumental variables can remove attenuating effects of measurement error so one estimation strategy available in the INCAP data is to use the experimental and distance variables to purge the attenuation bias. The expected finding would be for the OLS coefficient to double in magnitude. However, one unique feature of the INCAP data is that the calorie intake data collected at the feeding center do not contain measurement error because the amount of supplement calories that each particular child consumed was measured precisely.[14] This means that the degree of measurement error for each child is related to the proportion of calories consumed at the feeding center. If the child ate entirely at the feeding center then their calorie input would be measured without error whereas if the child ate entirely at home then their calorie input would be measured with the traditional amount of error.

One way to model this type of measurement error is to consider that the observed total calories is a mixture of a correctly measured calorie input and an incorrectly measured calorie input with the mixing proportion determined by the amount consumed at the feeding center. Let C^h be the amount of calories consumed if the child ate only at home and let C^c be the amount of calories that would be consumed if the child ate only at the feeding center. Assume also that observed home calories in the absence of a feeding center would still be measured with classical measurement error according to: $C^{h,o} = C^h + U$. Both $C^{h,o}$ and C^h are hypothetical quantities because in reality the child will eat some amount at both home and at the feeding center. Now the observed total calories $C^{T,o}$ is:

$$C^{T,o} = PC^{h,o} + (1-P)C^c$$

where P is the fraction of calories consumed at home. In the data, I observe $PC^{h,o}$, $(1-P)C^c$ and $C^{T,o}$ but not P, $C^{h,o}$ or C^c. Under this kind of measurement error, it can be shown

[14] For each child and each visit to the feeding center, the staff measured both the precise amount of atole and fresco dispensed and the amount leftover. Daily leftovers were totaled and cross-checked against the recorded total amount consumed and dispensed.

(Appendix B) that the OLS estimate is attenuated according to:

$$E[\hat{\beta}_{OLS}] = \beta \frac{\sigma_c^2}{\sigma_c^2 + E[P^2]\sigma_u^2}$$

Because the percentage of calories consumed at home $P \in [0,1]$ then $E[P^2]\sigma_u^2 < \sigma_u^2$ and the attenuation bias is smaller than in the case of classical measurement error. Therefore, the estimated value of β from the INCAP measurement error lies between the OLS estimate and the attenuated adjusted OLS estimate. The intuition behind the result is that observing some calorie data without measurement error lowers the variance of the measurement error relative to the variance in calories, which decreases the attenuation bias. Notice that as the P becomes degenerate at 1 (calories are measured with error) then $E[P^2] = 1$ and the standard attenuation bias result holds. Whereas if P becomes degenerate at 0 (calories are measured without error) then $E[P^2] = 0$ and the OLS estimate correctly identifies β under the maintained assumptions. This analysis suggests that correction for the standard OLS attenuation bias result can serve as an upper bound on attenuation bias in the INCAP data. Another interesting result is this kind of mixture of a mismeasured and correctly measured variable can also be removed with an instrumental variable provided that the instrument is uncorrelated with the measurement error U (see Appendix B).

A final issue that arises in using IV to estimate a production function is the interpretation of IV when there are random coefficients and individuals select into treatment based on the value of their random coefficient. For example, suppose that equation (3) is enriched so that each child has a different marginal productivity of calories, which is denoted by adding an i subscript to $\beta_{i,a}^a$:

$$\Delta H_{ia} = \beta_{i,a}^a C_{ia} + \alpha_a^a D_{ia} + \tilde{\gamma}_a^a X_i + (\lambda_a - \lambda_{a-1})\mu_i + \varepsilon_{ia} - \varepsilon_{a-1i}$$

This is a fairly realistic assumption and says that for certain children at certain time periods that increasing calories is more productive than at other time periods and for other children. The question that arises is what does IV estimate in such a model? A series of papers (Heckman, 1997; Heckman et al., 2006) argues that in this case of "essential heterogeneity" that IV can estimate a parameter that is uninterpretable for economic policy. The basic concern is that even if the instrumental variable has strong first stage explanatory power and even if the instrument is plausibly exogenous to with respect to unobserved determinants of height, which seems reasonable in the current context given the experimental and distance variation, if the children still select into treatments (calorie intake) based on the value of their child specific $\beta_{i,a}$, then Heckman et al. (2006) demonstrate that

IV in this case can be larger, smaller or even the opposite sign to the true causal effect. Although Heckman et al. (2006) develop statistical tests to test for the case of essential heterogeneity, a perhaps simpler way to test the validity of the models is simply to test their predictive ability, which is one of the aims of the current paper. One interesting indicator of essentially heterogeneity is the failure of overidentification tests, which in their model does not indicate a lack of exogeneity but rather picks up effect heterogeneity coming from different instrument sets identifying and weighting individuals with different effects and different margins of choice (Heckman et al., 2006).

IV. Results

I first present the empirical results without age specific heterogeneity and then proceed to the model with age specific effects. The baseline empirical results are given in Tables 2 and 3. Tables 2a and 2b report the first and second stage regression estimates of the cumulative height production function, treating calories as an endogenous variable and using a dummy variable for whether the child lived in an atole village and the distance from the child's house to the feeding center within each village as instruments. I consider three different instrument sets: IV-1 uses only the atole dummy, IV-2 uses the atole dummy and distance as instruments and IV-3 uses only distance as an instrument. In Table 2a, the instrument sets all have strong explanatory power with the first stage F-tests ranging from 22.9 to 27.8, which exceeds the recommended thresholds (Staiger and Stock, 1997). Moreover the signs are all in the expected direction with the treatment village dummy associated with higher total calorie intake and being farther from the feeding center associated with lower total calorie intake.

In Table 2b, I present both the OLS results and the second stage IV regressions. Regression (1) displays estimates of the height production function using OLS. The coefficients have the expected signs with cumulative calories having a positive association with height and cumulative diarrhea a negative association. Regressions (2), (3) and (4) in Table 2b show model estimates with the different instrument sets treating calories as an endogenous variable. The most important and salient difference between the OLS and IV estimates is that the magnitude of the coefficient of calories increases by 2.7 to 11-fold. The difference between the size of OLS and IV coefficients is statistical significant in regressions (2) and (3), reflected in failing to reject the hypothesis of equality of the OLS and IV estimators using the Durbin-Wu-Hausman (DWH) test in regressions (2) and (3). Regression (4) does not reject the equality of the OLS and IV estimates. Also in Table 2b, regression (3) has an overidentified instrument set and the regression fails the Sargan overidentification test, which suggests either a lack of exogeneity of the instruments or that the different

instrument sets are picking up effect heterogeneity.

Table 3a and 3b report the first and second stages estimates of equation (5) with change in height as the dependent variable, calories as the endogenous variable of interest and using the same instrument sets as previously. The instruments are again powerful with the first stage F-tests in Table 3a ranging from 10.7 to 22.1 and the signs of the coefficients both intuitive and seemingly large in magnitude. In Table 3b, regression (1) estimates the technology by OLS. The coefficients again have the expected sign. Calories have a positive effect, diarrhea has a negative effect, and age has a negative effect, which is picking up the falling velocity of growth throughout childhood (Tanner, 1990). Columns (2), (3) and (4) in Table 3b estimate the change in height specification using the same instrument sets described above. The pattern is similar to Table 2b. Compared to the OLS estimates in (1), the IV-estimates in (2), (3) and (4) show a large increase in the magnitude of the coefficient on calories by 4 to 19 fold. Given that maximum attenuation factor is 0.5, the adjusted OLS estimates would be twice the estimated effect in column (1), which suggests that the OLS attenuation adjusted estimate is at the lower end of the range of IV estimates. A DWH test rejects the equality of the OLS and IV estimates in specification (2) and (3) but again not in specification (4).

Comparing the estimates across Table 2b and 3b, the change in height specification lowers the marginal effect of calories on height in all specifications. This is consistent with the idea that children who are taller receive more calories. A formal statistical test of this hypothesis rejects the null of the equality of height and change in height specifications, which indicates perhaps bias from not addressing unobserved heterogeneity in growth across children.[15]

In Table 4a and 4b, I use the same estimation strategy as above except I allow for parameter heterogeneity in calories by age by interacting age dummies with the calorie inputs. Table 4a displays the OLS and second stage IV regression results for the height model and Table 4b shows the OLS and second stage IV regression results for the change in height model. In both of these models, because of the added heterogeneity, the F-stats are much lower and below the recommended thresholds, ranging between 1.88 and 4.48 in

[15]Two other potential sources of bias are (1) a non-linear calorie input relationship and (2) reverse causality between increases in calories and change in height. To check bias (1), in results not presented here, I considered a quadratic calorie input and used both instruments in the IV estimation. The large OLS vs. IV patterns continue to hold. Although I find some evidence for a non-linear calorie effect in the t-test for the squared calorie term, the effect does not appear quantitatively important in the counterfactuals. For bias (2), reverse causality between increases in calories and change in height would actually suggest the OLS estimates would be biased *upwards* because faster growing children begin to eat more but eating more does not cause faster growth. In that case an instrument that exogenously varies calories should find a smaller effect than OLS estimates. The opposite pattern is found in the empirical results.

the height model and between 0.0438 and 4.42 in the change in height model. Although it seems important to understand heterogeneity in effects by age, evidently considering this amount of heterogeneity stretches the statistical power of the data so the added realism of the model should be balanced against any potential statistical issues.

Reflecting previous work (Schroeder et al., 1992), I replicate that the OLS results show decreasing marginal productivity of calories by age.[16] This decline by age is much more pronounced in the change in height model with essentially zero marginal productivity of calories after 48 months in the change in height model. The IV results again display the same pattern being much larger in magnitude than the OLS results. In addition, they also display the same decreasing productivity by age. The effect is especially pronounced for the change in height model. A Durbin-Wu-Hausman test rejects the equality of the OLS and IV models with age heterogeneity in most specifications and a statistical test of the equality of calorie productivity by age also rejects the null in most specifications. Although the falling productivity by age is well-known and consistent with much other work (Schroeder et al., 1992), the main finding of the large difference between the OLS and IV results continues to hold to the model with additional heterogeneity.

V. Counterfactual

To both understand the validity of the estimated models and to predict increases in height from policy relevant, out of sample increases in calorie intake, this paper next considers a counterfactual policy where the Guatemalan children are fed on "first-world" diet. To implement such a counterfactual, I need an estimate of the "calorie gap" between children in the U.S. and Guatamalan children in the INCAP data. Data on calories in the U.S. comes from the National Health and Nutrition Examination Survey I (NHANES I). The NHANES I is a nationally representative survey of 32,000 people age 1-74 from 1971 to 1975. In addition to height and age information, the survey collects information on calorie intake from a 24-hour food consumption intake survey. I chose the NHANES I study because the sampling period 1971 to 1975 overlaps the INCAP study from 1969 to 1977. In addition, there were less problems with overeating and obesity in the 1970s. I restrict the sample to children aged 1-7 in the NHANES I data to match the ages of children in the INCAP data.

Pooling the INCAP and NHANES I data, Table 5 displays a series of regressions with calories as the dependent variable. "US" is a variable that equals one if the observation comes from the NHANES I survey. Regression (1) shows the unadjusted calorie-gap;

[16]This is conceptually different than decreasing velocity of growth although decreasing productivity by age and their co-movement by child's age certainly suggests a close relationship.

the average difference in calories between a child in the INCAP data and the NHANES I study. The difference is approximately 600 calories. However, the difference overstates the magnitude of the calorie gap because the US children are taller and taller children receive more calories, even in Guatemala. Adding controls for height and age lowers the calorie gap to, averaging across the estimates, approximately 430 calories. One interesting feature is that the interaction between the US indicator and age is statistically insignificant, which implies that the calorie gap is constant at 430 calories across ages.

Prior to exploring the predicted impacts of closing the calorie gap using the model estimates, I first need a benchmark for the counterfactual results. A interesting study by Bogin et al. (2002) provides exactly such a benchmark. Their paper compares the heights of two populations of children. The first is a group of Maya children living in Guatemala and the second is a group of American children born to Maya immigrants to the US. In Table 6, I reproduce the relevant comparison between the American Maya children and the Guatemalan Maya children in the Bogin et al. (2002) study. I also add average height by age from the Guatemalan INCAP children and average height by age from the NHANES I. Averaging across the Bogin et al. (2002) results, the average differences between Guatemalan Maya children and US Maya children is 9.8 cm at 7 years old, which suggests a sizable gap in average height between these two groups. Also note that despite the large height differences between the two groups of Maya children, the US Maya children are still substantially shorter than children in the NHANES I; the gap between an average US child and the US Maya children of Guatemalan immigrants is 6.4 centimeters at age 7.

Before comparing the results of Bogin to the current study, several caveats are in order. First, the children in the INCAP study were ladino and not Maya, where the ladinos are a Spanish speaking population of mixed European and Maya descent, which potentially implies a different genetic makeup for height potential in the two populations. Second, both the Bogin samples and the INCAP experiment are samples of convenience and potentially non-representative. However, interestingly for both caveats is that the heights of the Guatemalan children in the Bogin et al. (2002) paper and INCAP study are not quantitatively different. Third, there is the possibility that migration to the US is related to the height of either potential migrants or the height of their children. However, Bogin et al. (2002) argue that the migration decisions were exogenous and driven by factors such as "civil war, economic crisis and a cholera epidemic."[17] Setting aside the issues regarding

[17]It is also not clear that even if those events are random with respect to parents' height that they would not affect the height of their children through health and income shocks. It would have been interesting to compare the heights of the *parents* of the children in the study who decided to immigrate or not. It seems they did not collect those data. However, a previous study (Bogin, 1988) did not find evidence of "phenotypic

the validity of the height gap as benchmark, I use the 9.8 cm average difference in heights to gauge the plausibility of the magnitude of the subsequent counterfactuals.

Table 7 displays results from increasing the diet of the Guatemala children by 430 calories over a period of 6 years.[18] Because the unit of time in the production function is one year, the interpretation on the calorie coefficient is the marginal impact on height in centimeters of a one calorie increase in average daily calories over the course of a year. Then for each possible coefficient the counterfactual is computed according to $6 years$ x $430 \frac{calories}{day}$ x $\beta \frac{\frac{cm}{year}}{\frac{calories}{day}}$ where the units of the model coefficient β are centimeters per year per average daily calorie. The 'Height' column in Table 7 uses estimates of the technology in Table 2b, where each row corresponds to the appropriate estimate in Table 2b. The 'Change in Height' column uses the parameter estimates from Table 3b.

The differences in the height and change in height specifications estimates and the difference in OLS and IV estimates translates into large differences in the counterfactuals. For the height production function specification, the counterfactual results in an increase in between 2.69 cm to 28.27 cm of height. This within specification variation is driven both by the divergence between the OLS and IV estimates and variation among the IV estimates.

The height and change in height specification also generate substantial differences in the counterfactuals, with the change in height counterfactuals all below the height counterfactuals. For the change in height specification, OLS estimate gives a counterfactual of is 0.77 cm while the IV estimates range from 3.34 to 14.71. Finally, I add an "Adjusted OLS" counterfactual, which uses as the parameter estimate the attenuation adjusted OLS estimate. Recall that this estimate serves as an upper bound for the attenuating effects of measurement error in the INCAP data.[19] Those counterfactual results are 5.37 cm for the height specification and 1.53 cm for the change in height specification.

Also, in Table 7, the column '% of Height Gap' shows the computed counterfactual as a percentage of 9.8 cm difference report in Bogin et al. (2002). Again, the differences in the OLS and IV estimates translate into large differences in magnitude between the counterfactual height impacts of the calorie intervention. As the preferred model, I focus on the counterfactuals from the change in height specification. While the OLS estimate suggests that only 8% of the height gap would be closed by closing the calorie gap, the IV

selective migration" from Central America to the US, which suggests that this may not be a source of bias.

[18]The INCAP data do not have "home calories" for children less than one year old because the children are primarily breastfeeding. Therefore the production function is estimated from 1 - 7 years of age in the INCAP data so I consider a counterfactual of 6 years.

[19]The "Adjusted OLS" is actually a soft upper bound because OLS estimate is attenuated according to the INCAP measurement error and not classical measurement error.

estimates range from closing 34% to 150% of the height gap between Guatemalan children and Guatemalan children born in the US.

In Tables 8a and 8b, I perform the same counterfactual exercise using the production functions age specific heterogeneity in the marginal impact of calories and again benchmarking with the 9.8cm difference.[20] This time, instead of closing the calorie gap for all of early childhood, I consider closing the calorie gap for different durations in early childhood: between ages 1 - 2, between ages 1 - 3, etc. The results parallel the baseline model with very large differences in counterfactuals between the OLS and IV and substantial differences across IV counterfactuals. The main difference from the baseline model is that the falling impact of calories by age indicates that the height gap counterfactuals are finished by age 3. Finally, the unrealistic predictions of the counterfactuals using the IV estimates continue to hold or are exacerbated in the age heterogeneity models.

The large differences between the IV estimates and the implication of one of the IV estimates that the height gap would not only be overturned, but that the height of Guatemalans would increase by 50% of the height gap only through calories, suggests indirectly that the IV estimates are not identifying the policy relevant average marginal impact of calories. The OLS estimate counterfactual is an order of magnitude smaller. In fact, using the attenuation bias result, an upper bound on the impacts of measurement error in the INCAP data is given by the attenuation adjusted change in height counterfactual, which suggests that at most 16% of the height gap can be explained by closing the US-Guatemalan calorie gap during early childhood. Age specific effects give essentially the same pattern of results but that the effects would occur at earlier ages.

VI. Conclusion

This paper uses an extremely rich longitudinal data set on calorie intake, disease and height to estimate a height production function on a sample of rural Guatemalan children born between 1962 to 1977. The econometric model is consistent with theoretical notions of human growth and the richness of the micro data allows me to perform several specification tests for different biases. The results suggest biases from unobserved difference in growth, substantial measurement error in calories and policy biased estimates arising from the LATE characterization of IV. Using the estimated height production functions, I compute a counterfactual effect on height of closing the "calorie gap" and giving Guatemalan children the average diet of a child in the United States during the 1970s. The idea of the counterfactual is to understand how important calories in early childhood are in explaining

[20]Because of noise in the estimates, I impose in the counterfactuals that statistically insignificant and negative calorie coefficients have zero impact on height or changes in height.

differences in average heights across populations. The attenuation adjusted OLS estimate shows that equalizing the calorie gap would close at most 16% of the height gap. Although this estimate is approximately twice the effect implied by estimates from OLS, the results still suggest that 84% of the height gap is unexplained by calorie inputs. These results are consistent with recent evidence by Deaton (2007) questioning the role of nutrition in explaining average differences in heights across populations. Although the results show that increased calorie consumption will clearly improve height outcomes, other potential inputs such as the composition of calories, micronutrients, diet diversity, prenatal care, the disease environment, medical inputs, intergenerational linkages in height (Behrman et al., 2009), and work requirements that affect net nutrition are all potentially important channels, either directly as inputs or indirectly by affecting the productivity of other inputs, in explaining group level differences in average heights. Finally, the results suggest caution in using instrumental variables to estimate biological processes. Even in the current ideal case of an extensive panel data collection with exogenous, experimental variation in availability of calories, the IV estimates produce unrealistic and misleading estimates of the productivity of calorie inputs.

References

Abhijit Banerjee, Esther Duflo, Gilles Postel-Vinay, and Tim Watts. Long-run health impacts of income shocks: wine and phylloxera in nineteenth-century france. *The Review of Economics and Statistics*, 92(4):714–728, 2010.

Jere R Behrman and Barbara L Wolfe. How does mother's schooling affect family health, nutrition, medical care usage, and household sanitation? *Journal of Econometrics*, 36 (1):185–204, 1987.

Jere R Behrman, Maria C Calderon, Samuel H Preston, John Hoddinott, Reynaldo Martorell, and Aryeh D Stein. Nutritional supplementation in girls influences the growth of their children: prospective study in guatemala. *The American Journal of Clinical Nutrition*, 90(5):1372–1379, 2009.

Jere R Behrman, John Hoddinott, John A. Maluccio, Erica Soler-Hampejsek, Emily L. Behrman, Reynaldo Martorell, Agnes Quisumbing, Manuel Ramirez, and Aryeh D. Stein. What determines adult skills? impacts of pre-school, school and post-school experiences in guatemala. *Latin American Economic Review*, page forthcoming, 2014.

J.R. Behrman and A.B. Deolalikar. Health and nutrition. *Handbook of Development Economics*, 1:631–711, 1988.

Zulfiqar A Bhutta, Tahmeed Ahmed, Robert E Black, Simon Cousens, Kathryn Dewey, Elsa Giugliani, Batool A Haider, Betty Kirkwood, Saul S Morris, HPS Sachdev, et al. What works? interventions for maternal and child undernutrition and survival. *The Lancet*, 371(9610):417–440, 2008.

Zulfiqar A Bhutta, Jai K Das, Arjumand Rizvi, Michelle F Gaffey, Neff Walker, Susan Horton, Patrick Webb, Anna Lartey, and Robert E Black. Evidence-based interventions for improvement of maternal and child nutrition: what can be done and at what cost? *The Lancet*, 2013.

B. Bogin, P. Smith, AB Orden, MI Varela Silva, and J. Loucky. Rapid change in height and body proportions of Maya American children. *American Journal of Human Biology*, 14 (6):753, 2002.

Barry Bogin. Rural-to-urban migration. *Biological aspects of human migration*, page 90, 1988.

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

Maria Cecilia Calderon. Three essays on gender inequalities, human capital and development in guatemala. *Dissertation, University of Pennsylvania*, 2009.

T Colin Campbell, Thomas M Campbell, et al. *The China study: the most comprehensive study of nutrition ever conducted and the startling implications for diet, weight loss and long-term health*. Wakefield Press, 2007.

A. Case and C. Paxson. Stature and Status: Height, Ability, and Labor Market Outcomes. *Journal of Political Economy*, 116(3):499–532, 2008.

Flavio Cunha and James Heckman. The technology of skill formation. *American Economic Review*, 97(2):31–47, 2007.

Elisabetta De Cao. The height production function from birth to early adulthood. Technical report, " Carlo F. Dondena" Centre for Research on Social Dynamics (DONDENA), Università Commerciale Luigi Bocconi, 2011.

A. Deaton. From the Cover: Economics of Health and Mortality Special Feature: Height, health, and development. *Proceedings of the National Academy of Sciences*, 104(33): 13232, 2007.

R.W. Fogel and D.L. Costa. A theory of technophysio evolution, with some implications for forecasting population, health care costs, and pension costs. *Demography*, pages 49–66, 1997.

Tue Gørgens, Xin Meng, and Rhema Vaithianathan. Stunting and selection effects of famine: A case study of the great chinese famine. *Journal of Development Economics*, 97(1):99–111, 2012.

JP Habicht and R. Martorell. Objectives, research design, and implementation of the INCAP [Institute of Nutrition of Central America and Panama] longitudinal study. *Food and Nutrition Bulletin*, 14(3):176–190, 1992.

J. Heckman. Instrumental variables: A study of implicit behavioral assumptions used in making program evaluations. *Journal of Human Resources*, pages 441–462, 1997.

James J Heckman, Sergio Urzua, and Edward Vytlacil. Understanding instrumental variables in models with essential heterogeneity. *The Review of Economics and Statistics*, 88(3):389–432, 2006.

J.J. Heckman. The technology and neuroscience of capacity formation. *Proceedings of the National Academy of Sciences*, 104(33):13250–55, 2007.

J. Hoddinott, J.A. Maluccio, J.R. Behrman, R. Flores, and R. Martorell. Effect of a nutrition intervention during early childhood on economic productivity in Guatemalan adults. *The Lancet*, 371(9610):411–416, 2008.

Mahnaz Islam and John Hoddinott. Evidence of intrahousehold flypaper effects from a nutrition intervention in rural guatemala. *Economic Development and Cultural Change*, 57(2):215–238, 2009.

J. Komlos. *Stature, living standards, and economic development: essays in anthropometric history*. University of Chicago Press, 1994.

A. Lechtig, C. Yarbrough, R. Martorell, H. Delgado, and RE Klein. The one-day recall dietary survey: a review of its usefulness to estimate protein and calorie intake. *Archivos Latinoamericanos de Nutrición*, 26(3):243, 1976.

John A Maluccio, John Hoddinott, Jere R Behrman, Reynaldo Martorell, Agnes R Quisumbing, and Aryeh D Stein. The impact of improving nutrition during early childhood on education among guatemalan adults*. *The Economic Journal*, 119(537):734–763, 2009.

R. Martorell. The INCAP longitudinal study (1969-1977) and its follow-up (1988-1989); an overview of results. *Community-Based Longitudinal Nutrition and Health Studies: Classical Examples from Guatemala, Haiti and Mexico*, 1995:124–142, 1995a.

R. Martorell. Results and implications of the INCAP follow-up study. *Journal of Nutrition*, 125(4 Suppl):1127S, 1995b.

Reynaldo Martorell and Jean-Pierre Habicht. *Growth in early childhood in developing countries*. 1986.

Xin Meng and Nancy Qian. The long term consequences of famine on survivors: evidence from a unique natural experiment using china's great famine. Technical report, National Bureau of Economic Research, 2009.

N. Persico, A. Postlewaite, and D. Silverman. The Effect of Adolescent Experience on Labor Market Outcomes: The Case of Height. *Journal of Political Economy*, 112(5): 1019–1053, 2004.

Mark M Pitt, Mark R Rosenzweig, and Md Nazmul Hassan. Productivity, health, and inequality in the intrahousehold distribution of food in low-income countries. *The American Economic Review*, pages 1139–1156, 1990.

Merrill S Read and Jean-Pierre Habicht. History of the incap longitudinal study on the effects of early nutrition supplementation in child growth and development. *FOOD AND NUTRITION BULLETIN-UNITED NATIONS UNIVERSITY-*, 14:169–169, 1993.

Mark R Rosenzweig and T Paul Schultz. Estimating a household production function: Heterogeneity, the demand for health inputs, and their effects on birth weight. *The Journal of Political Economy*, pages 723–746, 1983.

D.G. Schroeder, H. Kaplowitz, and R. Martorell. Patterns and predictors of participation and consumption of supplement in an intervention study in rural Guatemala. *Food and Nutrition Bulletin*, 14:191–200, 1992.

K. Silventoinen. Determinants of variation in adult body height. *Journal of Biosocial Science*, 35(02):263–285, 2003.

Doug Staiger and James H Stock. Instrumental variables regression with weak instruments. *Econometrica*, 65(3):557–586, 1997.

R.H. Steckel. Stature and the Standard of Living. *Journal of Economic Literature*, pages 1903–1940, 1995.

Z. Stein, M. Susser, G. Saenger, and F. Marolla. Famine and human development: The Dutch hunger winter of 1944-45, 1975.

J. Strauss and D. Thomas. Health, nutrition, and economic development. *Journal of Economic Literature*, 36(2):766–817, 1998.

J.M. Tanner. *Foetus into man: physical growth from conception to maturity*. Harvard Univ Pr, 1990.

Duncan Thomas, John Strauss, and Maria-Helena Henriques. How does mother's education affect child height? *Journal of human resources*, pages 183–211, 1991.

P.E. Todd and K.I. Wolpin. On the Specification and Estimation of the Production Function for Cognitive Achievement. *The Economic Journal*, 113(485):3–33, 2003.

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

Barbara L Wolfe and Jere R Behrman. Women's schooling and children's health: Are the effects robust with adult sibling control for the women's childhood background? *Journal of Health Economics*, 6(3):239–254, 1987.

Kenneth I Wolpin. Determinants and consequences of the mortality and health of infants and children. *Handbook of Population and Family Economics*, 1:483–557, 1997.

Appendix A: Tables

Table 1
Descriptive Statistics By Treatment (*atole*) vs. Control (*fresco*)

	Height Sample			Change in Height Sample		
	atole	*fresco*	Difference	*atole*	*fresco*	Difference
Children's Height (cm)	90.74	88.66	2.08‡	94.02	92.19	1.83‡
	(10.34)	(10.2)		(10.52)	(10.47)	
Average Daily Total Calories	1013.84	970.69	43.15‡	1032.48	993.78	38.7‡
	(348.12)	(387.66)		(346.92)	(361.22)	
Average Daily Home Calories	860.83	925.82	-64.99‡	884.29	941.25	-56.96‡
	326.37	372.08		331.14	347.66	
Average Daily Supplement Calories	153.01	44.86	108.15‡	148.18	52.53	95.65‡
	(120.16)	(40.15)		(114.06)	(43.97)	
Days Sick with Diarrhea (per month)	6.56	6.43	0.13	5.46	6.46	-1.0†
	(9.74)	(9.57)		(10.4)	(11.03)	
Child's Age (months)	44.27	44.08	0.19‡	49.92	49.85	.07
	(16.72)	(16.26)		(17.13)	(16.87)	
Percent male	.52	.57	-.05	.53	.57	-.04†
	(.5)	(.5)		(.5)	(.5)	
N	842	777		1,404	1,301	

Notes: Standard deviations are reported in parentheses. "Height Sample" are the observations used for estimation in Table 2. "Change in Height Sample" are the observations used for estimation in Table 3. The models estimated using instrument sets IV-2 and IV-3 (see below) have slightly less observations because of missing distance data but the patterns of differences across the samples is similar in the slightly smaller sample.
‡ $p<0.01$, † $p<0.05$, * $p<0.1$

Table 2a
First-stage regression
Dependent Variable: Cumulative calories

	IV-1	IV-2	IV-3
	(2)	(3)	(4)
Atole	184.1‡	159.8‡	-
	(34.9)	(35.6)	
Distance to Feeding Center	-	-70.3‡	-80.8‡
		(16.1)	(16.0)
Cumulative Days with Diarrhea	-2.71‡	-2.78‡	-2.78‡
	(0.71)	(0.71)	(0.72)
Age	91.9‡	91.9‡	91.9‡
	(1.1)	(1.1)	(1.1)
Sex	210.3‡	219.9‡	212.2‡
	(35.5)	(35.8)	(35.9)
Constant	-1751.7‡	-1572.6‡	-1460.3‡
	(57.0)	(72.0)	(67.9)
N	1,619	1,590	1,590
R^2	0.82	0.82	0.82
First stage F-test	27.8	22.9	25.36

Notes: First-stage regressions corresponding to the height IV regressions in Table 2b. For the excluded exogenous variables, IV-1 uses exposure to *atole*, IV-2 uses both distance to the feeding center and *atole* as instruments and IV-3 uses distance to the feeding center as an instrument.
‡ $p<0.01$, † $p<0.05$, * $p<0.1$

Table 2b
Dependent Variable: Height

	OLS	IV-1	IV-2	IV-3
	(1)	(2)	(3)	(4)
Cumulative calories	0.0010‡	0.011‡	0.0071‡	0.0027†
	(0.000)	(0.002)	(0.001)	(0.001)
Cumulative days with diarrhea	-0.0166‡	0.0101	-0.0021	-0.0135†
	(0.004)	(0.010)	(0.007)	(0.005)
Age	0.4858‡	-0.4254†	-0.0673	0.3339‡
	(0.016)	(0.202)	(0.114)	(0.109)
Sex	0.7216‡	-1.2741†	-0.6116	0.3164
	(0.205)	(0.601)	(0.397)	(0.330)
Constant	65.82‡	82.21‡	75.93‡	68.67‡
	(0.379)	(3.669)	(2.102)	(1.992)
N	1,619	1,619	1,590	1,590
R^2	0.85	0.39	0.68	0.84
Durbin-Wu-Hausman	-	0.00	0.00	0.15
Sargan	-	-	0.0001	-
IV/OLS	-	11	7.1	2.7

Notes: Huber-White standard errors are reported in parentheses. IV-1 uses exposure to *atole* as an instrument, IV-2 uses distance to the feeding center and *atole* as instruments and IV-3 uses distance as an instrument. First stage F-test is the p-value from test of the joint significance of the first stage regressors. Durbin-Wu-Hausman is the p-value from a test of the null hypothesis of the equality of the calorie OLS and IV estimates. Sargan test reports the p-value from an overidentification test.
‡ $p<0.01$, † $p<0.05$, * $p<0.1$

Table 3a
First-stage regressions
Dependent Variable: Calories

	IV-1	IV-2	IV-3
	(2)	(3)	(4)
Atole	39.9‡	29.1†	
	(12.2)	(14.8)	-
Distance to Feeding Center	-	-29.3‡	-31.1‡
		(6.7)	(6.6)
Days with Diarrhea	-0.99*	-1.10	-1.12‡
	(0.58)	(0.67)	(0.67)
Age	9.0 ‡	11.0‡	11.0‡
	(0.36)	(0.44)	(0.44)
Sex	75.3‡	98.4‡	96.7‡
	(12.3)	(14.8)	(14.8)
Constant	508.1‡	499.6‡	520.3‡
	(22.2)	(31.2)	(29.4)
N	2,705	1,811	1,811
R^2	0.20	0.28	0.28
First stage F-test	10.7	13.0	22.1

Notes: First-stage regressions corresponding to the change in height IV regressions in Table 3b. For the excluded exogenous variables, IV-1 uses exposure to *atole*, IV-2 uses both distance to the feeding center and *atole* as instruments and IV-3 uses distance to the feeding center as an instrument.
‡ $p<0.01$, † $p<0.05$, * $p<0.1$

Table 3b
Dependent Variable: Change in Height

	OLS	IV-1	IV-2	IV-3
	(1)	(2)	(3)	(4)
Calories	0.0003‡	0.0057†	0.0026†	0.0013
	(0.000)	(0.002)	(0.001)	(0.001)
Days with Diarrhea	-0.0127‡	-0.0069	-0.0124‡	-0.0137‡
	(0.003)	(0.005)	(0.004)	(0.004)
Age	-0.0682‡	-0.1169‡	-0.0971‡	-0.0825‡
	(0.002)	(0.021)	(0.013)	(0.013)
Sex	-0.26‡	-0.66‡	-0.50‡	-0.38‡
	(0.067)	(0.196)	(0.139)	(0.139)
Constant	10.55‡	7.69‡	9.65‡	10.24‡
	(0.14)	(1.25)	(0.53)	(0.53)
N	2,705	2,705	1,811	1,811
R^2	0.298	-0.416	0.218	0.303
Durbin-Wu-Hausman	-	0.001	0.054	0.53
Sargan	-	-	0.006	-
Test $\beta_h = \beta_{ch}$	0.000	0.025	0.001	0.240
IV/OLS		19	8.66	4.33

Notes: Huber-White standard errors are reported in parentheses. IV-1 uses exposure to *atole* as an instrument, IV-2 uses distance to the feeding center and *atole* as instruments and IV-3 uses distance as an instrument. First stage F-test is an omnibus test of the joint significance of the first stage regressors. Test $\beta_h = \beta_{ch}$ gives the p-values for a null hypothesis of the equality of the coefficient on calories in the height specification versus the change in height specification for each of OLS, IV1, IV2 and IV3. Durbin-Wu-Hausman is a test of the difference between the IV estimates and the OLS estimate. Sargan tests the exogeneity of overidenifying instruments in regression (3) only.
‡ $p<0.01$, † $p<0.05$, * $p<0.1$

Table 4a
Dependent variable: Height, parameter heterogeneity by age

	OLS	IV-1	IV-2	IV-3
	(1)	(2)	(3)	(4)
Cumulative calories 0-24 month	0.0030***	0.0120	0.0204***	0.0106
	(0.001)	(0.009)	(0.005)	(0.008)
Cumulative calories 0-36 months	0.0024***	0.0123***	0.0099***	0.0052**
	(0.000)	(0.002)	(0.001)	(0.002)
Cumulative calories 0-48 months	0.0017***	0.0112***	0.0062***	0.0031***
	(0.000)	(0.003)	(0.001)	(0.001)
Cumulative calories 0-60 months	0.0011***	0.0104***	0.0043***	0.0019*
	(0.000)	(0.003)	(0.001)	(0.001)
Cumulative calories 0-72 months	0.0006***	0.0099***	0.0031***	0.0011
	(0.000)	(0.004)	(0.001)	(0.001)
Cumulative calories 0-84 months	0.0002	0.0094**	0.0021*	0.0005
	(0.000)	(0.004)	(0.001)	(0.001)
Cummulative Days with Diarrhea	-0.0161***	0.0088	-0.0067	-0.0142***
	(0.004)	(0.011)	(0.006)	(0.005)
Age	0.5598***	-0.2884	0.5526***	0.6186***
	(0.023)	(0.479)	(0.157)	(0.218)
Sex	0.6181***	-1.2943**	-0.4544	0.2272
	(0.197)	(0.644)	(0.331)	(0.320)
Constant		76.3845***	50.0723***	55.1441***
		(17.061)	(6.523)	(10.347)
Observations	1,619	1,619	1,590	1,590
R-squared	0.859	0.448	0.750	0.846
First stage F-test		1.88	4.48	2.10
Durbin-Wu-Hausman		0.00	0.00	0.90
$\beta_a = \beta$	0.00	0.00	0.00	0.00

Notes: IV-1 uses exposure to *atole* as an instrument, IV-2 uses distance to the feeding center and *atole* as instruments and IV-3 uses distance as an instrument. First stage F-test is an omnibus test of the joint significance of the first stage regressors. Durbin-Wu-Hausman is a test of the difference between the IV estimates and the OLS estimate. Test $\beta_a = \beta$ reports the p-values for a joint test of the equality of the marginal impacts of calories by age.

Table 4b
Change in height, parameter heterogeneity by age

	OLS	IV-1	IV-2	IV-3
	(1)	(2)	(3)	(4)
Calories 12-24 months	0.0013***	0.0170	0.0098***	0.0077**
	(0.000)	(0.015)	(0.002)	(0.003)
Calories 24-36 months	0.0010***	0.0060**	0.0055***	0.0039**
	(0.000)	(0.003)	(0.001)	(0.002)
Calories 36-48 months	0.0003**	-0.0013	0.0024***	0.0010
	(0.000)	(0.012)	(0.001)	(0.001)
Calories 48-60 months	0.0000	-0.0073	0.0003	-0.0011
	(0.000)	(0.022)	(0.001)	(0.001)
Calories 60-72 months	-0.0002	-0.0126	-0.0015	-0.0028*
	(0.000)	(0.031)	(0.001)	(0.002)
Calories 72-84 months	-0.0003	-0.0171	-0.0027*	-0.0040*
	(0.000)	(0.038)	(0.002)	(0.002)
Days with Diarrhea	-0.0126***	-0.0023	-0.0081*	-0.0091*
	(0.003)	(0.014)	(0.005)	(0.005)
Age	-0.0386***	0.4947	0.1138**	0.1182
	(0.007)	(0.967)	(0.050)	(0.074)
Sex	-0.2667***	-0.1390	-0.5168***	-0.3772**
	(0.066)	(0.913)	(0.130)	(0.155)
Constant		-14.9357	-0.4300	0.8000
		(34.553)	(2.403)	(3.906)
Observations	2,705	2,705	1,811	1,811
R-squared	0.308	-1.988	-0.000	0.114
First stage F-test		0.0438	4.4197	2.7454
Durbin-Wu-Hausman		0.00	0.00	0.091
$\beta_a = \beta$	0.0000	0.14	0.0001	0.0027

Notes: IV-1 uses exposure to *atole* as an instrument, IV-2 uses distance to the feeding center and *atole* as instruments and IV-3 uses distance as an instrument. First stage F-test is an omnibus test of the joint significance of the first stage regressors. Test $\beta_h = \beta_{ch}$ gives the p-values for a null hypothesis of the equality of the coefficient on calories in the height specification versus the change in height specification for each of OLS, IV1, IV2 and IV3. Durbin-Wu-Hausman is a test of the difference between the IV estimates and the OLS estimates. Test $\beta_a = \beta$ reports the p-values for a joint test of the equality of the marginal impacts of calories by age.

Table 5
Calorie Gap: US vs. Guatemala

	Dependent Variable: Daily Calorie Intake					
	(1)	(2)	(3)	(4)	(5)	(6)
US	592‡	430‡	442‡	462‡	471‡	348†
	(12)	(12)	(19)	(21)	(32)	(169)
Height		16‡	15‡	15‡	15‡	13‡
		(0)	(1)	(1)	(1)	(1)
Age			7	66‡	68‡	75‡
			(10)	(21)	(22)	(22)
Age-Squared				-7‡	-7‡	-7‡
				(2)	(2)	(2)
US x Age					-2	-15
					(6)	(18)
US x Height						2
						(2)
N	6,838	6,838	6,838	6,838	6,838	6,838
R-squared	0.25	0.38	0.38	0.38	0.38	0.38

Notes: Robust standard errors are reported in parentheses. US is an indicator variable that equals 1 if the observation comes from the NHANES-I and 0 from the INCAP.
‡ $p<0.01$, † $p<0.05$, * $p<0.1$

Table 6
Cross Study Average Heights (cm): USA vs. Guatamala

	Bogin et al. (2002)				
	USA-1992	USA-2000	Guate-1998	INCAP 1969-77	NHANES I 1971-75
Age 5	111.3	111.5	102.2	99.1	112.8
Age 6	113.2	115.8	105.0	104.9	118.9
Age 7	118.5	119.7	109.3	110.5	124.9

Note: Average population height reported in centimeters. The Bogin et al. (2002) samples are USA-1992 and USA-2000, from the children of Maya immigrants born and living in the U.S., and Guate-1998, a sample of Maya Guatemala school children. The Bogin samples were collected in 1992, 2000 and 1998, respectively. INCAP 1969-77 is from a sample of four ladino villages in the INCAP Longitudinal Study 1969 - 1977. NHANES I 1971-1975 is a nationally representative sample from the USA.

Table 7
Production Function Counterfactuals: Baseline Model

	Height		Change in Height	
	Δ Average Height	% Height Gap	Δ Average Height	% Height Gap
OLS	2.69	27%	0.77	8%
Adjusted OLS	5.37	55%	1.53	16%
IV1	28.27	288%	14.71	150%
IV2	18.20	186%	6.75	69%
IV3	6.94	71%	3.34	34%

Note: Statistics reported in centimeters. Height refers to the height technology estimate using height as a dependent variable. Change in height refers to using change in height as a dependent variable. Adjusted OLS refers to the OLS estimates adjusted for attenuation bias. Δ Average Height refers to the change in average height at 7 years old and % Height Gap refers to the percent of the 9.8cm height gap.

Table 8a
Production Function Counterfactuals: Height model with parameter heterogeneity by age

	OLS		**IV1**			**IV2**			**IV3**		
Intervention timing	Δ Average Height	% Height Gap	Δ Average Height	% Height Gap	IV/OLS	Δ Average Height	% Height Gap	IV/OLS	Δ Average Height	% Height Gap	IV/OLS
1-2 years	1.3	13.2%	5.18	52.9%	4	8.75	89.3%	6.7	4.54	46.4%	3.5
1-3 years	2.32	23.6%	10.48	107%	4.5	12.99	132.5%	5.6	6.76	69%	2.9
1-4 years	3.04	31%	15.3	156.1%	5	15.65	159.7%	5.1	8.09	82.6%	2.7
1-5 years	3.5	35.8%	19.79	201.9%	5.6	17.49	178.4%	5	8.91	90.9%	2.5
1-6 years	3.77	38.5%	24.04	245.3%	6.4	18.8	191.9%	5	9.37	95.6%	2.5
1-7 years	3.86	39.4%	28.06	286.3%	7.3	19.7	201%	5.1	9.58	97.7%	2.5

Notes: Δ Average Height refers to the change in average height in centimeters given the length of the intervention in the far left column. % Height Gap refers to the percent of the 9.8cm height gap and IV/OLS divides each IV counterfactual by the OLS counterfactual result.

Table 8b
Production Function Counterfactuals: Change in height model with parameter heterogeneity by age

	OLS		**IV1**			**IV2**			**IV3**		
Intervention timing	Δ Average Height	% Height Gap	Δ Average Height	% of 9.8cm	IV/OLS	Δ Average Height	% Height Gap	IV/OLS	Δ Average Height	% Height Gap	IV/OLS
1-2 years	.56	5.7%	7.29	74.4%	13.1	4.21	43%	7.6	3.33	33.9%	6
1-3 years	.99	10.1%	9.88	100.8%	10	6.56	67%	5	5	51.1%	5
1-4 years	1.12	11.5%	9.88	100.8%	8.8	7.62	77.7%	6.8	5.45	55.7%	4.9
1-5 years	1.13	11.5%	9.88	100.8%	8.7	7.74	79%	6.8	5.45	55.7%	4.8
1-6 years	1.13	11.5%	9.88	100.8%	8.7	7.74	79%	6.8	5.45	55.7%	4.8
1-7 years	1.13	11.5%	9.88	100.8%	8.7	7.74	79%	6.8	5.45	55.7%	4.8

Notes: Δ Average Height refers to the change in average height in centimeters given the length of the intervention in the far left column. % Height Gap refers to the percent of the 9.8cm height gap and IV/OLS divides each IV counterfactual by the OLS counterfactual result.

Appendix B

Let Y be the outcome of interest (either height, H, or change in height, ΔH, depending on the maintained modeling assumptions). For calories, following the notation in the paper, let total observed calories be given by $C^{T,o} = C^T + U$ where C^T is total calories and U is a random variable with $E[U] = E[(Y - E[Y])U] = E[(C^T - E[C^T])U] = 0$ and $E[U^2] = \sigma_u^2$. The OLS estimate of the true marginal impact of calories, β_0, is attenuated according to:

$$\begin{aligned} E[\hat{\beta}_{OLS}] &= \frac{Cov(Y, C^{T,o})}{Var(C^{T,o})} \\ &= \frac{Cov(Y, C^T + U)}{Var(C^T + U)} \\ &= \frac{Cov(Y, C^T) + Cov(H, U)}{Var(C^T) + Var(U)} \\ &= \frac{Cov(Y, C^T)}{Var(C^T)} \frac{Var(C^T)}{Var(C^T) + Var(U)} \\ &= \beta_0 \frac{\sigma_c^2}{\sigma_c^2 + \sigma_u^2} \end{aligned}$$

Instrumental variables can remove the attenuating effects of measurement error. If the instrument Z is exogenous, correlated with total calories and uncorrelated with the measurement error U according to $E[(Z - E[Z])U] = 0$, then:

$$\begin{aligned} E[\hat{\beta}_{IV}] &= \frac{Cov(Y, Z)}{Cov(C^{T,o}, Z)} \\ &= \frac{Cov(Y, Z)}{Cov(C^T + U, Z)} \\ &= \frac{Cov(Y, Z)}{Cov(C^T, Z) + Cov(U, Z)} \\ &= \frac{Cov(Y, Z)}{Cov(C^T, Z)} \\ &= \beta_0 \end{aligned}$$

In the INCAP experiment, instead of observing $C^{T,o} = C^T + U$, we actually observe $C^{T,o} = PC^{h,o} + (1 - P)C^c$ with potential observed home calories given by $C^{h,o} = C^h + U$, where I assume U has the same properties as previously. P is a random variable that takes values in $[0, 1]$ and if I assume that U and P are independent, then the attenuation bias is given

by:

$$\begin{aligned}
E[\hat{\beta}_{OLS}] &= \frac{Cov(H, C^{T,o})}{Var(C^{T,o})} \\
&= \frac{Cov(H, PC^{h,o} + (1-P)C^c)}{Var(PC^{h,o} + (1-P)C^c)} \\
&= \frac{Cov(H, P(C^h + U) + (1-P)C^c)}{Var(P(C^h + U) + (1-P)C^c)} \\
&= \frac{Cov(H, C^T + PU)}{Var(C^T + PU)} \\
&= \frac{Cov(H, C^T) + Cov(H, PU)}{Var(C^T) + 2Cov(C^T, PU) + E[P^2U^2]} \\
&= \frac{Cov(H, C^T)}{Var(C^T) + E[P^2]\sigma_u^2} \\
&= \frac{Cov(H, C^T)}{Var(C^T)} \frac{Var(C^T)}{Var(C^T) + E[P^2]\sigma_u^2} \\
&= \beta_0 \frac{\sigma_c^2}{\sigma_c^2 + E[P^2]\sigma_u^2}
\end{aligned}$$

Because the range of P is $[0,1]$ then $0 \leq P^2 \leq 1$ and so:

$$0 \leq \int P^2 \phi(P) dP = E[P^2] \leq 1$$

In the INCAP measurement error, the attenuation bias is less severe than the attenuation bias in the classical measurement error case:

$$\begin{aligned}
E[P^2]\sigma_u^2 &\leq \sigma_u^2 \\
\sigma_c^2 + E[P^2]\sigma_u^2 &\leq \sigma_c^2 + \sigma_u^2 \\
\frac{1}{\sigma_c^2 + \sigma_u^2} &\leq \frac{1}{\sigma_c^2 + E[P^2]\sigma_u^2} \\
\frac{\sigma_c^2}{\sigma_c^2 + \sigma_u^2} &\leq \frac{\sigma_c^2}{\sigma_c^2 + E[P^2]\sigma_u^2}
\end{aligned}$$

The inequality holds with equality if P is distributed degenerately at 1, which is exactly the case where all calories are measured with error. Notice also that if P is distributed degenerately at 0, which is the case of no measurement error in calories, then the right-hand side equals 1 so that there is no attenuation bias. Essentially the INCAP attenuation

bias is bounded between the traditional measurement error case and the true coefficient:

$$\beta_0 \frac{\sigma_c^2}{\sigma_c^2 + \sigma_u^2} \leq \beta_0 \frac{\sigma_c^2}{\sigma_c^2 + E[P^2]\sigma_u^2} \leq \beta_0$$

In the measurement error process given above, an instrumental variable that decontaminates classical measurement error also decontaminates the INCAP measurement error:

$$\begin{aligned}
E[\hat{\beta}_{IV}] &= \frac{Cov(H,Z)}{Cov(C^{T,o},Z)} \\
&= \frac{Cov(H,Z)}{Cov(PC^{h,o} + (1-P)C^c,Z)} \\
&= \frac{Cov(H,Z)}{Cov(P(C^h+U) + (1-P)C^c,Z)} \\
&= \frac{Cov(H,Z)}{Cov(C^T + PU,Z)} \\
&= \frac{Cov(H,Z)}{Cov(C^T,Z) + Cov(PU,Z)} \\
&= \frac{Cov(H,Z)}{Cov(C^T,Z)} \\
&= \beta_0
\end{aligned}$$