Growing and Learning when Consumption is Seasonal: Long-term Evidence from Tanzania*†

Paul Christian*‡ Brian Dillon‡

February 9, 2017

Abstract

In this paper we show that the seasonality of food consumption during childhood, conditional on average food consumption, impacts long-run human capital development. We develop a model that distinguishes between differences in average consumption levels, seasonal fluctuations, and idiosyncratic shocks, and estimate the model using high frequency panel data from early 1990s Tanzania. We then test whether the average level and the seasonality of a child's consumption profile affect height and educational attainment in a 2010 follow-up survey with the same children. Results show that the negative effects of greater seasonality are 30-60% of the magnitudes of the positive effects of greater average consumption (in the same units). Put differently, children expected to have identical human capital based on annualized consumption measures will have substantially different outcomes if one child's consumption is more seasonal than the other's. We discuss implications for measurement and policy.

Keywords: seasonality; consumption; human capital; height; education; catch-up growth.

JEL codes: O12, J13, I13, I31

*†We are grateful to the African Development Bank, World Bank, and Cornell University for funding through the STAARS program. We thank Chris Barrett, Anirban Basu, Jere Behrman, Leah Bevis, Joshua Blumenstock, Michael Carter, Arun Chandraskehar, Norma Coe, Alison Cullen, Joachim De Weerdt, Andrew Foster, Rachel Heath, Heather Hill, Kalle Hirvonen, Martijn Huysmans, Nathan Jensen, Arianna Legovini, Mark Long, Linden McBride, Ellen McCullough, Tyler McCormick, Jenny Romich, Jacob Vigdor, and seminar participants at Cornell, KU Leuven, the University of Washington, the Second Annual Global Food Security Conference, and the STAARS conference in Addis Ababa, for helpful discussions and comments. We are joint first authors on the paper. Any errors are our responsibility.

*‡World Bank. Email: `pchristian@worldbank.org`.

‡Corresponding author. University of Washington, Evans School of Public Policy and Governance, 109 Parrington Hall, Seattle, WA, 98195, USA. Email: `bdillon2@uw.edu`. Telephone: 206.221.4601. Fax: n/a.

1 Introduction

In this paper we address the following question: Does a consistently seasonal diet during childhood have long-run effects on human capital formation? Consider two children who have the same level of consumption when measured over the course of a year. These children are indistinguishable using standard measures of poverty, which are based on annualized consumption. Suppose the first child has relatively smooth consumption over all 12 months, while the second consumes substantially more in the months after harvest (the harvest season) than in the months prior to the harvest (the lean season). The goal of this paper is to empirically examine whether the second child is worse off, and if so, by how much. That is, our aim is to determine whether the seasonality of consumption impacts child development, conditional on the average level.

This question is motivated by two lines of prior research. The first is the large literature in economics, public health, and demography on the connection between nutrition during childhood and later life outcomes (Glewwe, Jacoby and King, 2001; Case, Fertig and Paxson, 2005; Cunha and Heckman, 2007; Victora et al., 2008; Bleakley, 2010; Currie and Almond, 2011; Behrman, 2016). Many of the papers in that literature use shocks to consumption from droughts, wars, nutrition programs, or other exogenous events to study the long-run impacts of changes in consumption levels on health or educational attainment (Hoddinott and Kinsey, 2001; Alderman, Hoddinott and Kinsey, 2006; Bundervoet, Verwimp and Akresh, 2009; Maluccio et al., 2009; Akresh, Lucchetti and Thirumurthy, 2012; Singh, Park and Dercon, 2013; Dercon and Porter, 2014; Hoyne, Schanzenbach and Almond, 2016). The emphasis in this line of research is on long-term impacts of variation in consumption levels, but not on seasonality.

A second body of work studies the link between seasonal fluctuations in consumption and short-term measures of health and wellbeing (Behrman, 1993; Bhagowalia, Chen and Masters, 2011; Egata, Berhane and Worku, 2013; Hirvonen, Taffesse and Hassen, 2015). This topic is especially important in low-income, agrarian countries, where the composition and level of consumption typically fluctuate with the agricultural cycle (Branca et al., 1993; Mitchikpe et al., 2009; Miller et al., 2013; Chikhungu and Madise, 2014; Arsenault et al., 2014; Abay and Hirvonen, 2016). In many agricultural countries, households use coping strategies to survive the lean season that involve meal skipping, selectively reducing some household members' food intake at each meal, or increasing dependence on starchy staple foods (Behrman, 1988*a*,*b*; Sahn, 1989; Adams, 1995; Shetty, 1999). These can lead to measurable lean season deficits in weight, mid-upper arm circumference (MUAC), and other measures of wasting and stunting (Maleta et al., 2003).

It is not *ex ante* obvious that seasonality will have long-run impacts on child development. If lean season deprivations are compensated for by catch-up growth in the post-harvest period, potential deficits may be mitigated or eliminated (Martorell, Khan and Schroeder, 1994; Alderman, Hoddinott and Kinsey, 2006). Current debates in epidemiology and nutrition center on the extent to which full catch-up is possible (Behrman, 2016). Furthermore, co-insurance can help mitigate the effects of annually recurring negative shocks to food availability, potentially dampening their long-run impacts (Coate and Ravallion, 1993; Townsend, 1994; Morduch, 1995; Dercon and Krishnan, 2000; Fafchamps and Lund, 2003; Kinnan, 2011). Perhaps most importantly, because our analysis will condition on average consumption, two children with different degrees of seasonality will not be consistently ranked in terms of daily consumption; in some months one child will consume more, in other months, the other child will.

We study the effects of seasonality on two human capital outcomes: height and educational attainment. These outcomes have long precedent in the literature on nutrition and child development. In Section 2 we describe the theoretical rationale for why we might expect height and educational attainment to respond to the seasonality of consumption (conditional on the average level). We discuss both biological pathways, due to seasonal interruptions in physical or cognitive development, and behavioral pathways, e.g., pulling older children from school during the lean season to help at home or in the labor market.

In Section 3 we describe the data. Data availability is a challenge for this topic, because one needs to observe both seasonal food consumption during childhood and realizations of height and educational attainment many years later. We use the Kagera Health and Development Survey (KHDS), a 19-year panel study from Tanzania. The KHDS includes a high frequency panel survey from 1991-1994, in which nearly 800 households were surveyed

4 times, roughly 6 months apart. These data include household-level consumption measures from which we can estimate the seasonality of consumption (but as we describe in Section 4, we do more than just use the 4 observations per household to measure seasonality). Our outcome variables come from a 2010 follow-up study, during which the KHDS team collected anthropometry measures and information on educational attainment for all respondents who could be tracked (see Section 5.3 for a discussion of attrition).

In Section 4 we develop the empirical approach and identification strategy. There is a challenging identification problem inherent to this question. The issue is that any exogenous shock to the seasonality of consumption will also impact the average level. It is difficult to think of an instrument or experimental design that would hold average consumption constant while randomly varying seasonality (without taking food away from some participants in some periods). Hence, our approach to identification is partly structural, with a series of robustness checks to validate the identifying assumptions. First, we develop a theoretical model of daily consumption that separates average consumption, seasonal fluctuations, and idiosyncratic shocks. We then use the 1991-1994 data to estimate the model, allowing consumption on day d, $d \in \{1 \ldots 365\}$, to be a function of observable household characteristics. This aspect of the approach is fully general, in that it nests the case with no seasonality. We then predict consumption for every household on every day of the year, and show that the standard deviation of the predicted consumption sequence is an estimate of the seasonal component of consumption, purged of the idiosyncratic deviations. Likewise, the mean of the sequence is an estimate of average consumption.

Our final empirical step is to use the predicted consumption statistics from 1991-1994 as explanatory variables in regressions with 2010 height or educational attainment as dependent variables. These regressions address our main research question: conditional on the average level of consumption during childhood, how does human capital vary with the seasonality of consumption? Because the key independent variables are predicted, we bootstrap standard errors over the two steps of the estimation procedure. We also estimate a series of robustness checks in which we systematically drop variables from the consumption model and include them with their squares in the 2010 regressions. These checks verify that the seasonality measure is not mistakenly capturing other non-linear relationships between household characteristics and long-run outcomes.

In Section 5 we present our findings. We find a robust, negative relationship between consumption seasonality and human capital formation. Across specifications, the negative relationship between seasonality and human capital is 30-60% of the magnitude of the positive relationship between average consumption and human capital (in the same units). When we allow for heterogeneous effects by age, we find a pattern consistent with the predictions of the theory in Section 2. The effects of seasonality on height is greatest for children in utero and during infancy, during the critical first 1,000 days of life. Effects on education are most pronounced for older children, suggesting that behavioral channels such as dropping out of school to help on the farm are more important in this sample than early life impacts on cognitive performance. When we further allow for heterogeneity by both age and gender, we see that the height effects during infancy are concentrated among girls, while the education effects during adolescence are largely driven by boys. One interpretation of those results is that households favor infant boys over infant girls when making difficult consumption choices during the lean season; yet, boys appear to be more likely to drop out of school to work during periods of lean season deprivation. These results are also consistent with potentially gender-specific biological channels linking early life circumstances and developmental outcomes. We cannot distinguish the biological channel from the gender-specific investment channel, but our findings are consistent with others that have found girls to be particularly responsive to early life shocks (Mancini and Yang, 2009). Taken as a whole, the robustness of the results and the alignment with theory gives us confidence that the estimated relationships are causal.

Finally, in Section 6 we conclude with a brief discussion. We first discuss the interpretation of the magnitudes of estimated effects in this paper, relative to the literature. We then suggest ways that current consumption and poverty measures could be augmented to incorporate information about seasonality, and conclude with a note about policies to provide counter-cyclical consumption support.

2 Background and theoretical motivation

In this section we describe how prior work suggests a number of possible connections between consumption seasonality and human capital. The outcomes studied here, height and education, share an important common feature: both are stock variables accumulated gradually throughout childhood and adolescence. There are also some key differences. For instance, height is directly influenced by a range of factors unobserved in socioeconomic data (such as genetic endowment), while educational attainment (over the support relevant to this paper) is the consequence of a sequence of choices over which agents have control. Hence, we develop the theoretical motivations separately.

2.1 Seasonality and height

Linear growth leading to the accumulation of height depends on a complex set of environmental, genetic, health, and nutritional factors. A large literature has demonstrated that chronic undernutrition or negative nutrition shocks during childhood can lead to height deficits that persist into adulthood (Victora et al., 2008; Almond and Currie, 2011; Attanasio, 2015). There is some evidence of "catch-up growth", i.e., higher-than-average growth rates by previously disadvantaged children when they experience a shift to receive adequate nutritional support (Martorell, Khan and Schroeder, 1994; Alderman, Hoddinott and Kinsey, 2006). The literature has not settled on a definitive answer for whether full catch-up is possible (Behrman, 2016). Yet, even if full catch-up from negative nutritional shocks is biologically possible, negative impacts will persist into the future for any individuals who do not access the nutritional and medical resources required for catch-up.

There are two ways to think about the link between persistent seasonal variation in consumption and the accumulation of height. The first is direct. Seasonal deprivations in caloric and nutritional intake can be thought of as "mini-shocks" that occur annually. In stylized terms, receipt of a sufficient diet during any period enables a child to grow in accordance to her or his genetic potential, other things equal. Deficits in the amount or composition of nutrient intake undercut that growth. A child whose consumption falls below critical thresholds for a regular period each year will then experience repeated, minor interruptions in growth. For children experiencing acute malnutrition, lean season deprivations are obvious. For other children, each seasonal episode of under-nourishment and slow growth may be small enough to elude detection, while the cumulative effect leads to a measurable height deficiency in adulthood.

A second set of pathways arises through the possibility of dynamic complementarities (Foster, 1995; Cunha and Heckman, 2007). There is evidence that for some aspects of cognitive and physical development, early deprivations make a person less responsive to future investments. For example, early life malnutrition can delay the onset and shorten the duration of the adolescent growth spurt, resulting in lower adult height (Martorell, 1999; Limony, Kozieł and Friger, 2015). However, dynamic forces may also mitigate the effect of seasonality on height. Catch-up growth may occur through biological mechanisms that increase the efficiency with which consumption is translated into growth, or through adaptation by the household to provide more nutrients to an under-fed child (Maleta et al., 2003; Abay and Hirvonen, 2016). If such dynamic complementarities are present, the question of whether lean-season deprivations aggregate into measurable differences during adulthood rests on the net effects.

Although children can grow continuously from infancy through late adolescence, there are key periods of faster growth. These are from birth through age 2, and during the adolescent growth spurt. The latter typically occurs from age 11-15 for girls and 13-17 for boys in well nourished populations (Rogol, 2000), but can be delayed to persist beyond age 17 (Simondon et al., 1998). To accommodate this, in the empirical analysis we allow for heterogeneity by age group. We also estimate specifications in which we allow for heterogeneity by age and gender, as prior work has shown both biological and behavioral channels by which nutrition variation can affect boys and girls differently (Mancini and Yang, 2009; Maluccio et al., 2009).

2.2 Seasonality and educational attainment

The mechanisms linking seasonal consumption to educational attainment are more numerous than those for height. The additional complexity arises from the nature of the education production function, which takes both cognitive ability and child time as inputs.

Effects on education that occur through cognitive ability are similar in many respects to the mechanisms described in the previous subsection, for height. Cognition develops in stages. Basic functions such as the regulation of sight and hearing are established in infancy, while language processing is mostly developed by age 5, and the capacity for higher order cognition develops into adolescence (Grantham-McGregor et al., 2007). Annually recurring periods of low nutrition can impede or slow the pace of cognitive development. Evidence from other settings suggests that dynamic complementarities are particularly important in this domain (Heckman, 2010). Interventions to improve children's nutritional circumstances – in this setting, analogous changes would be the return to harvest season levels of consumption, or changes in household circumstances that reduce consumption seasonality – are not as effective if delivered later in childhood, after the accumulation of previous deficits. Because cognitive ability affects success in school and the probability of continuation, consumption seasonality may reduce educational attainment through the accumulated negative impacts of lean season deprivation on cognitive development (Currie and Thomas, 1999; Peet et al., 2015; Grantham-McGregor et al., 2007).

There are also two behavioral channels by which consumption seasonality can impact educational attainment. The first is through the experience of hunger during school. There is evidence that children who are hungry at school have difficulties paying attention and completing tasks, and score lower on standardized tests than non-hungry children (Pollitt et al., 1982-1983; Powell et al., 1998; Rampersaud et al., 2005; Kristjansson et al., 2007; Wisniewski, 2010). It is plausible that children who experience periods of acute hunger during the lean season are less likely to succeed in school during that period, resulting in poorer performance, reduced attendance, and earlier drop-out.

A second behavioral channel stems from the competing demands on child time. In Tanzania, even young children may assist with household tasks, such as fetching water, collecting firewood, preparing meals, or tending to animals (Beegle, Dehejia and Gatti, 2006). During adolescence and teenage years, they begin caring for smaller children, assisting on the farm, or working off-farm for a family enterprise or in the market. In bad agricultural years, lean season deprivation may be severe enough that the marginal value of child time outside of school exceeds the expected marginal return to schooling, at least in the short run. If so, students may drop out to assist with household chores, freeing up adult time to search for remunerative work, or may seek work themselves in order to help the household buy food. Evidence from the KHDS suggests that households do view child labor as a means to smooth consumption in the face of unexpected shocks (Beegle, Dehejia and Gatti, 2006). Effects operating through this mechanism are most likely to be present for the oldest children, for whom the opportunity cost of time in school is greatest.

In sum, the existing literature suggests a number of mechanisms through which seasonal variation in consumption might affect height and educational attainment in the long run. Of course, the same mechanisms also indicate an important role for the *average* level of consumption in determining these outcomes. Hence, any empirical model linking seasonal consumption variation with human capital outcomes must also account for differences in the average level of consumption.

3 Data and descriptive statistics

In this section we describe the setting and the data. The first subsection gives an overview of the data set. We then explain the variables from 1991-1994 that are used to estimate the consumption model of Section 4. In the final subsection we describe the human capital outcome variables from 2010.

3.1 The Kagera Health and Development Survey

The data for this project are from the Kagera Health and Development Survey (KHDS). The KHDS consists of three survey efforts resulting in a panel data set that spans nearly 20 years. The original respondents were selected from 51 communities in the Kagera region. This is a hilly, rainy region, in the far northwest corner of Tanzania. It is bordered on the east by Lake Victoria, the north by Uganda, and the west by Rwanda. One of the primary goals of the first survey was to understand the severity and impacts of the HIV/AIDS crisis, which was in its first decade when the project was conceived.

The first set of surveys, KHDS 1, consists of four panel rounds collected from late 1991 to early 1994. The sample includes 6,353 individuals in 915 households. Researchers interviewed 759 households in all four rounds.[1] The survey covered a wide range of topics, including demographics, consumption, health, child anthropometry, agriculture, time use, labor supply, shocks, credit, and other topics common to the Living Standards Measurement Surveys (LSMS) format. The KHDS survey team worked continuously, spacing each household's interviews by roughly six months.

Follow-up surveys with KHDS 1 respondents were conducted in 2004 (KHDS 2) and 2010 (KHDS 3). The survey team interviewed the initial respondents, their new family members, and any outmigrants or split-off households. In 2004, researchers re-interviewed at least one person from 713 of the 759 households that had completed all 4 waves in KHDS 1, for a panel re-survey rate of 94%. In 2010, the team interviewed someone from 706 of the initial 759 households (93%). At the individual level, the re-interview rates among the non-deceased were 82% in 2004 and 85% in 2010. The remarkably low attrition rate for a long term sample in this setting is attributable to the team's extensive efforts to track respondents, including those who had moved to other parts of Tanzania or to neighboring countries.

To link the childhood experience of sample individuals with their later life outcomes, we use consumption measures from KHDS 1, and measures of height and educational attainment from KHDS 3.

3.2 Consumption and household characteristics from 1991-1994

Our first empirical goal is to estimate from the KHDS 1 data a consumption function that allows for the possibility of seasonal variation. This requires a measure of consumption, and a set of household characteristics to serve as arguments.

The KHDS 1 consumption module is of the "representative consumption" format. For a wide range of food goods, respondents were asked whether household members consumed the good during each month of the year. If the answer was "Yes", they were then asked to indicate the number of times each good was consumed in a typical non-zero month, and to

[1] In the first round, 840 households were surveyed, with 81 of these not appearing in the last round, the vast majority of which had moved outside the study area. Seventy-five households were selected as replacements for the 81 who dropped out, leading to 915 total unique households included across rounds.

estimate the market value (for home produced goods) or expenditure (for purchased goods) for a typical instance of consumption. Respondents were able to give separate answers for typical consumption during a rainy season month and a dry season month, with the seasons defined by the respondent. In the first round of KHDS 1, respondents were asked to consider consumption over the previous 12 months. In rounds 2-4, they were asked to use a 6-month recall period.

This way of collecting consumption data is no longer standard, in part because of the significant measurement error associated with 6–12 month recall periods. To minimize the effects of recall bias, we treat the consumption responses as relevant only for the month preceding each interview. That is, for interviews conducted on or after the 15th day of the month, we assign the consumption response to the interview month. For interviews on days 1-14, we assign the consumption response to the previous month. The assumption underlying this approach is that the most recent month is the most salient, and therefore the most likely to influence the response. In choosing whether to use the dry or rainy season response to characterize the most recent month, we use the village-level modal response to questions about the rainy season.[2]

With this approach, for each household we observe consumption in one specific month, for each of the four KHDS 1 rounds. We then calculate the daily total food expenditure of household i, on item j, interviewed in month m, as: $c_{ijm} = \frac{frequency_{ijm} * expenditure_{ijm}}{days_m}$, where $frequency$ is the number of times per month that the household consumes that item, $expenditure$ is the typical expenditure per consumption instance, and $days$ is the number of days in the month. We deflate expenditure using a Laspeyres price index that takes account of both spatial and inter-temporal price variation (see De Weerdt et al. (2012) for details).

Figure 1 shows the histogram of KHDS 1 survey dates in the lower panel, and a local polynomial regression of observed daily household consumption (assigned to the interview day) in the upper panel. In the lower panel, the broad temporal coverage of the survey is clear. The survey team worked continuously during the study. Although this is not reflected in the figure, the team also distributed its work evenly across districts, so that survey timing

[2] In practice, it makes little difference whether we use the household-specific response or the village-level mode, as they are closely aligned.

and location do not co-vary. Because the team worked continuously in all districts, did not systemically favor surveying one type of household before another, and re-visited households on a roughly 6-month cycle, it turns out that interview dates are quasi-randomly assigned. In the analysis to follow, the quasi-random assignment of interview dates will be helpful for power, though not technically necessary for identification (see Section 4).

In the upper panel of Figure 1, the seasonality of consumption is clear. Average consumption rises gradually after harvest in August and September, peaks in January, and drops off substantially through the following months. In real terms, peak consumption is greater in the first full survey year than in the second. The trough reaches roughly the same minimum across years. In the aggregate it is clear both that consumption follows an annual cycle, and that there are other factors that influence the level and variability of consumption during any given year.

Table 1 shows summary statistics for the 758 households interviewed in all four rounds of KHDS 1, pooled across rounds.[3] Over 70% of households are headed by a male. The average household head is approximately 50 years old, and has 4 years of formal education. Households are larger than the typical household in Tanzania today, averaging nearly 7 members and 5 adult equivalent members. The average household owns over 5 acres and the livestock equivalent of nearly 1.5 head of cattle (in tropical livestock units). The asset index is a relative measure of wealth, normalized to have a mean of zero and a standard deviation of 1, constructed from the first principal component of survey data on ownership of durable assets and household dwelling characteristics (Filmer and Pritchett, 2001; Sahn and Stifel, 2003). While agriculture represents the primary livelihood strategy in this region, 40% of households engage in some form of non-farm enterprise. To the extent that diversification of income sources is a consumption smoothing strategy, this is a possible source of between-household variation in the degree of consumption seasonality.

[3]Of the 758 households interviewed in all four rounds of KHDS 1, 23 did not complete the full food consumption and expenditure modules in all 4 rounds. The 736 with complete consumption data constitute the main analysis sample.

3.3 Sample and outcome variables in 2010

We include in the estimation sample all respondents who were aged 17 to 36 in 2010, which is equivalent to age -2 to 17 in 1991. These criteria restrict the sample to children who were 20 years of age or younger at the end of KHDS 1. We define age groups and eligibility by working backwards from age reported in 2010, because we expect less misreporting from adults stating their own age than from proxy respondents answering for children in KHDS 1. In total, 2,859 respondents who were children during KHDS 1 were successfully tracked during KHDS 3.[4]

The first outcome variable, height, was directly measured and recorded by KHDS 3 enumerators. The height measure is available for 2,578 individuals, or 90.2% of the sample. The other primary outcome variable, total years of education completed through 2010, is available for 2,633 of the age-eligible respondents (92.1%). Nine respondents indicated that they had obtained only "Adult education" or "Koranic school", both of which we coded as zero years of formal education.

Table 2 provides summary statistics for the outcome variables. The table shows the overall statistics and the breakdown by age subgroups. The age group categories are based on the theoretical discussion in Sections 2.1 and 2.2. Age groups 2 and 3, representing children aged 3-10 and 11-17 in 1991, respectively, each account for roughly 40% of the sample. The youngest group is smaller, at 22% of the sample. This is due in part to the smaller age range, and possibly also to the higher mortality rate for infants than for older children.

The sample is balanced on gender, at 49% male. The mean height is 162.5 cm, and the mean years of education is 7.6. There are slight differences between age groups in average height, likely due to variation in the number of remaining growth years as of 2010. There are also slight differences in educational attainment, with younger children acquiring more education. This reflects secular trends in Tanzania toward increased enrollment and attachment to primary and secondary education, similar to trends in much of sub-Saharan Africa.

The KHDS sample is drawn from a single, relatively homogeneous region. Hence, there is less variation in the explanatory and outcome variables than in the country as a

[4]In Section 5.3 we discuss attrition.

whole. For example, in a different data set from Tanzania – the nationally representative 2010-2011 LSMS-ISA data – the average height of adults of comparable ages is 160.2 cm, but the standard deviation is 10.3, nearly 25% higher than in Table 2.[5] This will have implications for how we interpret results.

4 Identification and estimation

In this section we explain the identification and estimation of a model linking consumption seasonality during childhood to human capital development. The section proceeds in three steps. We first develop a model of seasonally varying consumption and explain how we estimate a household-level measure of seasonality. We then explain the empirical model linking seasonality in KHDS 1 (1991-1994) to height and educational attainment in KHDS 3 (2010). We conclude with a discussion of identification and robustness.

4.1 Modeling consumption

Consider the following model of daily household consumption. Let $c(d, X_{ydh})$ be food consumption by household h on day d in year y, where $d = 1, \ldots, 365$, $y = 1, \ldots, Y$, and the matrix X_{ydh} consists of household characteristics that affect consumption. We can write a simple linear model of consumption as:

$$c(d, X_{ydh}) = X_{ydh}\phi + \nu_{ydh} \tag{1}$$

where ϕ is a coefficient vector and ν_{ydh} is a stochastic component. X_{ydh} can include a trend term or time dummy variables if necessary. When there is a seasonal component to consumption that recurs annually, ν_{ydh} can be decomposed into that component and a statistical error term:

$$c(d, X_{ydh}, Z_{ydh}) = X_{ydh}\phi + \Gamma(d, Z_{ydh}) + \psi_{ydh} \tag{2}$$

[5] The LSMS-ISA is a multi-country panel survey effort combining nationally representative Living Standard Measurement Surveys with an agriculturally focused Integrated Survey on Agriculture. The Tanzania survey is collected by the National Bureau of Statistics, with assistance from the World Bank. See http://go.worldbank.org/BCLXW38HY0 for more details.

where Γ is a function that maps household characteristics on day d into a day-specific seasonal innovation, and ψ_{ydh} is an i.i.d. error term with mean zero and variance σ^2_ψ. The matrix Z_{ydh} consists of household characteristics, and may contain some, all, or none of the same elements as X_{ydh}. Assume for now that components of X_{ydh} and Z_{ydh} are constant over all d and y with positive support (we will relax this assumption momentarily). The approach in (2) nests the case with no seasonality, which would have $\Gamma(d, Z_{ydh}) = 0$ for all d. Suppose that $\Gamma(d, Z_{ydh})$ can be represented as the product of two components, $\Gamma(d, Z_{ydh}) = \gamma(d)f(Z_{ydh})$. The term $\gamma(d)$ represents a sequence of day-specific innovations that are common to all households. This sequence has a mean $\overline{\gamma}$ and variance σ^2_γ. The effect of the daily innovation is attenuated or exacerbated by household characteristics through the function $f(Z_{ydh})$. The conditional variance of $\Gamma(d, Z_{ydh})$ can be written as $\sigma^2_{\gamma h} \equiv \sigma^2_\gamma f(Z_{ydh})^2$. If we assume $Cov(\gamma, \psi) = 0$, the conditional variance of the household consumption sequence $c_{ydh} = c(d, X_{ydh}, Z_{ydh})$ is given by:

$$\begin{aligned} Var(c_{ydh} \,|\, X,\, Z) &= \mathbb{E}\left[(c_{ydh} - \mathbb{E}\left[c_{ydh}\right])^2 \,|\, X,\, Z\right] \\ &= \mathbb{E}\left[(X_{ydh}\phi + \gamma(d)f(Z_{ydh}) + \psi_{ydh} - (X_{ydh}\phi + \overline{\gamma}f(Z_{ydh}))^2 \,|\, X,\, Z\right] \\ &= \sigma^2_{\gamma h} + \sigma^2_\psi \end{aligned} \tag{3}$$

Expression (3) makes an intuitive point. As long as stochastic consumption shocks are orthogonal to regularly occurring seasonal fluctuations, the variance of the consumption sequence is the sum of the conditional variance due to seasonality and the common variance due to unexpected shocks. It follows immediately that:

$$Var(\mathbb{E}[c_{ydh} \,|\, X,\, Z]) = \sigma^2_{\gamma h} \tag{4}$$

i.e., the variance of the sequence of expected consumption realizations is simply the conditional variance of the seasonal component. This suggests a straightforward empirical strategy for isolating and estimating a household-specific measure of the seasonal component of consumption variation: estimate the model, project consumption on every day of the year as a function of household observables, and then take the variance of that sequence.

To complete the model for estimation we must structurally represent the seasonal component $\gamma(d)f(Z_{ydh})$. We write the household-specific term as a linear function of household characteristics, $f(Z_{ydh}) = Z_{ydh}\rho$, where ρ is a coefficient vector. This term allows the amplitude of seasonality to vary across households. In deciding how to model the common seasonal component $\gamma(d)$, we use the wave-like shape of the aggregate consumption path from Figure 1 as our guide, and adopt a sine function representation. We make one modification, generalizing the sine function to allow for asymmetric periods above and below zero. This is accomplished by the addition of a parameter τ for the day d on which consumption returns to its day 1 level (equivalent to the standard sine function crossing the horizontal axis at $x = \pi$). We make this generalization so that the data, rather than the functional form, can divide the calendar year into periods above/below day 1 consumption. The full model then becomes:

$$
\begin{aligned}
c_{ydh} &= X_{ydh}\phi + \gamma(d)Z_{ydh}\rho + \psi_{ydh} \\
&= X_{ydh}\phi + \left\{ \sin\left(\frac{\pi d}{\tau}\right) \mathbb{I}[d \leq \tau] + \sin\left(\pi + \pi\frac{d-\tau}{365-\tau}\right) \mathbb{I}[d > \tau] \right\} Z_{ydh}\rho + \psi_{ydh} \\
&= X_{ydh}\phi + w(d,\tau)Z_{ydh}\rho + \psi_{ydh} \qquad (5)
\end{aligned}
$$

where $w(d,\tau)$ is a weight determined by d and τ, and the full set of parameters consists of τ, ϕ, and ρ.

We estimate the model in (5) using maximum likelihood with Gaussian errors. The data for estimation are from KHDS 1. For each KHDS 1 household we observe c_{ydh} on 4 randomly selected days over a 28-month period. Using all observations for all households in the analysis provides balanced coverage of the entire study period (see Figure 1). In our main specifications we use the variables reported in Table 1, with the addition of district effects, for both X_{ydh} and Z_{ydh}. Because the components of Z_{ydh} are weighted by a function of the interview date, no exclusion restriction is required for identification. We discuss this further in Section 4.3. The quasi-random assignment of interview dates, which we discussed in Section 3.1, is also not necessary for identification. However, quasi-random assignment is useful for power, as it ensures that the variation in the components of X_{ydh} and Z_{ydh} is well-represented across months of the year.

With an estimate in hand of $\{\hat{\phi}, \hat{\rho}, \hat{\tau}\}$, projected consumption on day d of year y for household with characteristics X_{ydh} and Z_{ydh} is given by $\hat{c}_{ydh} = X_{ydh}\hat{\phi} + w(d, \hat{\tau})Z_{ydh}\hat{\rho}$. This projection requires values for X_{ydh} and Z_{ydh} on every day. Here, we must relax the assumption that the variables in X_{ydh} are Z_{ydh} are fixed over the study period. While many of the variables in Table 1 are constant, some are not. To smoothly accommodate inter-temporal variation in household characteristics, we interpolate to non-interview days using the weighted average of the variables from the nearest surveys before and after each date d.[6] The variance of the resulting sequence of household-specific consumption predictions, $\hat{s}_h^2$, is our empirical estimate of (4). The mean of the sequence, $\hat{m}_h$, is an estimate of the average level of consumption by household members.

Because we allow Z_{ydh} to evolve over time, $\hat{s}_h^2$ is not exactly analogous to $\sigma_{\gamma h}^2$ from expression (4). We could fully align the two if we took the additional steps of modeling the evolution of the variables in Z_{ydh}, and added the appropriate adjustment to (4). We elected not to do this, because it would introduce complexity with little value. The important point is that $\sigma_{\gamma h}^2$ need not be constant over time, because it is a function of some time-varying household characteristics. Hence, we estimate $\hat{s}_h^2$ as the average seasonal component for the household over the study period, allowing for the possibility that the degree of seasonality might change as household characteristics change.

4.2 Linking seasonality to human capital outcomes

The discussion in Section 2 described a number of channels by which consumption seasonality may negatively impact child development, conditional on the average level of consumption. To test the reduced form hypotheses suggested by that discussion, we estimate regressions of the following form:

$$\ln Outcome_{ih}^{2010} = \alpha + \beta_1 \ln \hat{m}_h + \beta_2 \ln \hat{s}_h + \gamma W_{ih} + \delta + \epsilon_{ih} \tag{6}$$

where i indexes individuals, h indexes households from KHDS 1, $Outcome_{ih}^{2010}$ is either height or educational attainment, W_{ih} is a vector of individual controls for age, gender, and other

[6] For this reason, we only use data from households that were surveyed all four rounds, so as not to extrapolate to dates more than a year before or after from the most recent interview.

variables used in robustness checks, δ is a vector of district effects, and ϵ_{ih} is a stochastic error term. The variables $\hat{m}_h$ and $\hat{s}_h$ are as defined in the previous subsection; all other variables are from KHDS 3 (in 2010).[7] We adopt a double log specification for the estimated consumption statistics so that we can interpret coefficient estimates as elasticities, and so that results do not depend on whether we use the estimated variance ($\hat{s}_h^2$) or the estimated standard deviation ($\hat{s}_h$) to characterize consumption seasonality.

Because (6) includes predicted values, we bootstrap the full estimation procedure across equations (5) and (6). See Appendix section B for a description of the bootstrap procedure. The point estimates and pattern of statistical significance from the bootstrapped estimates are the same as those from estimation in two separate steps with OLS standard errors clustered at the level of the 1991-1994 household.

The theoretical discussion in Section 2 suggests clear predictions for the signs of the key parameters in (6). We expect to find $\beta_1 > 0$, because higher average consumption is generally associated with better outcomes in an undernourished population. The predication that seasonality reduces human capital, conditional on the mean level of consumption, is implemented by testing whether β_2 is negative (i.e., we test whether we can reject the hypothesis $H_0 : \beta_2 \geq 0$).

We also estimate specifications that include interactions of age or age $\times$ gender with the consumption statistics. In these specifications we assign children to the three age groups indicated in Table 2. This allows us to test specific mechanisms laid in the theoretical discussion of Section 2. For example, a negative relationship between seasonality and educational attainment for children in utero or during infancy (age group 1) would indicate a causal pathway through cognitive development, whereas a similar finding only for children in age group 3 would more likely indicate a behavioral mechanism through reduced school attendance.

An attractive feature of this approach is that by predicting consumption on every day of the year and then calculating statistics from the predicted values, the estimated

[7] We also estimated specifications using the observed values of m_h and s_h, calculated directly from the 4 observations in KHDS 1. The pattern of results from those specification is similar to what we report in Section 5. However, this alternative approach is clearly misspecified, because the raw statistics m_h and s_h contain both seasonal and idiosyncratic variation. Results are available upon request.

moments are conditionally independent of unobservable household characteristics. Hence, in the human capital regressions represented by equation (6), measurement error in $\hat{m}_h$ and $\hat{s}_h$ is not driven by unobserved household-specific factors that might lead a household to have higher- or lower-than-expected consumption. Unobservables can only lead to bias in (6) indirectly, through correlation with elements of X_{ydh} and Z_{ydh}. That is the rationale for a series of robustness checks that we describe in the next subsection.

4.3 Identification

The empirical approach developed in the previous two subsections partly relies on the structure of the consumption model to identify the effects of seasonality. How comfortable should we be with that aspect of the identification strategy? Clearly, it would be preferable if seasonality were randomly assigned. In this section we discuss whether even seemingly exogenous assignment of seasonal variation could identify the effect of interest, without the benefit of a model or additional controls. We then explain some robustness checks and components of the design that give us confidence in the identification strategy.

The central identification challenge stems from our interest in seasonality, which is related to the second moment of consumption. Yet, the first moment of consumption – the average level – is also an input to child development. We think it unlikely that there is an instrument or experimental design that could generate exogenous variation in seasonality without also affecting average consumption. The approach most likely to succeed would be in a lab, where researchers could assign participants to 2×2 factorial subgroups with {High, Low} mean consumption and {High, Low} seasonality of consumption. This is clearly infeasible in practice, not least because of the ethical barriers to running a food restriction study of this nature with children. Other alternatives – randomly assigning a treated group of households to receive counter-cyclical consumption support at key times of year, or providing all households with the same annual food transfer, but randomizing the seasonality of transfers – would either generate confounding variation in average consumption, or only be relevant for populations receiving regular consumption support, or both.

A related identification challenge stems from the conceptualization of seasonality as a treatment. Households are not assigned a consumption path; they choose it, at least in part. And while food consumption might be exogenous from the child's perspective, it is possible that parents who exhibit a strong preference for consumption smoothing, or who have greater capacity to smooth, differ in important ways from parents who do not. This could lead to biased estimation of (5) and mis-attribution of variation in long-run human capital outcomes to seasonality.

A consequence of these inherent challenges is that the structure of the consumption model plays a role in identification. It is for this reason that we use a functional form that is reflective of the wave-like shape of the aggregate consumption cycle. Also, because the model allows predicted consumption to be a smooth function of a large set of household characteristics, all observations contribute to the estimation of all parameters of (5), and thereby to the consumption predictions for all other households on all days. And, as mentioned, it is critical that we condition on average consumption, as this controls for the first order effects of consumption variation on human capital outcomes.

Otherwise, these identification concerns amount to a concern that $\hat{s}_h$ could pick up the effects of other seasonally varying characteristics that are not accounted for by $\hat{m}_h$, or, relatedly, that we are loading all non-linear relationships between variables in X_{ydh} and human capital outcomes into our estimates of $\hat{s}_h$. The latter is not, in fact, likely to be a problem, as it requires an omitted variable to have positively correlated non-linear effects on the amplitude of intra-annual consumption seasonality and the accumulation of long-run human capital. We think this is unlikely, because characteristics that improve long-run outcomes would be expected to facilitate, rather than impede, consumption smoothing. Nevertheless, to check the robustness of our results to these concerns, we sequentially re-estimate the model in (5) after dropping one variable at a time from X_{ydh} and Z_{ydh} and including the dropped variable and its square as control variables in (6). That is, for each variable x^j_{ydh} in X_{ydh} and Z_{ydh} we estimate

$$c_{ydh} = X^{-j}_{ydh}\phi + w(d,\tau)Z^{-j}_{ydh}\rho + \psi_{ydh} \tag{7}$$

where the superscript $^{-j}$ indicates that variable x^j_{ydh} is excluded, and

$$\ln Outcome^{2010}_{ih} = \alpha + \beta_1 \ln \hat{m}^{-j}_h + \beta_2 \ln \hat{s}^{-j}_h + \gamma_1 \overline{x}_{jh} + \gamma_2 \overline{x}^2_{jh} + \delta + \epsilon_{ih} \tag{8}$$

where $\overline{x}_{jh}$ is the household h mean of x^j_{ydh} over the four rounds of KHDS 1. In these specifications, any unobserved variable that is correlated with x^j_{ydh} and with seasonality, but that also impacts human capital directly, may attenuate the coefficient on $\hat{s}_h$ in model (8).[8]

Note that if we were to continue this procedure indefinitely by removing multiple variables from the consumption model and including them in the human capital regressions, we would eventually undermine the consumption model entirely. Even the test with one variable at a time carries the risk of over-penalizing the coefficients $\hat{\beta}_1$ and $\hat{\beta}_2$ in (8), if the contribution of x_{jh} to human capital is mediated primarily by consumption during childhood. Hence, our aims with these robustness checks are only to demonstrate that no single variable is driving the results, and that $\hat{s}_h$ is not picking up other non-linear effects that are outside the scope of the model.

Despite these efforts, we cannot definitively rule out that some part of the association we find between seasonality and human capital could be due to an omitted variable. The nature of this problem may be such that no single paper can establish beyond doubt a causal link between consumption seasonality and child development. In that case, our contribution is to document and measure an important risk factor for low human capital outcomes. However, for reasons described in this subsection, we believe both that omitted variables bias is unlikely, given our design, and that our approach rivals other feasible solutions to the identification challenges that are inherent to this research question.

5 Results

We now turn to the empirical results. In the first subsection we present the estimates of model (5) and the associated estimates of $\hat{m}_h$ and $\hat{s}_h$. We then show the estimates of human capital models based on equation (6). We conclude this section with an examination of robustness.

[8]We do not bootstrap the standard errors for these robustness checks, because it would take roughly 4 months of continuous estimation to complete them all (the interpolation and prediction is time intensive). However, in our main specifications the bootstrap results lead to standard errors that are so similar to those from OLS with clustering that there are no implications for inference.

5.1 Modeling consumption: estimation results

The complete estimates of the parameter vector for equation (5) are not of specific interest. The model is for prediction purposes only, and no attempt was made to exclude highly correlated variables.[9] Hence, there is limited value in interpreting the individual coefficient estimates. However, to demonstrate the search process for one key parameter, the left panel of Figure 2 plots the log likelihood against the 365 possible choices of τ. The likelihood is a smooth function of τ, and is maximized at $\hat{\tau} = 151$. Because we set the start date of the annual cycle to a point in late November that corresponded visually to $x = 0$ for a standard sine function, this indicates that average consumption rises at the end of the year, falls back to its November level at the end of April, and remains below the November reference level for 214 days.

To give a visual example of the model's performance, the right panel of Figure 2 shows observed and predicted consumption for two households. The $\times$s and dots show the observed consumption values, and the lines show the predicted consumption path (the $\times$s are associated with the dashed line, the dots with the solid line). There are three things to note in the figure. The first is that the predicted values preserve the ordering of both average consumption and the degree of seasonality (though this need not always be the case): the $\times$s have a higher mean and higher variance than the dots, and this is also clearly true for the dashed line relative to the solid line. Second, the predicted values for the $\times$ household are all below the observed values, while the predicted values for the dot household are all above the observed values. The implication is that based on household characteristics, we would expect the $\times$ (dot) household to have slightly lower (higher) consumption than observed. In this way the model is purging idiosyncratic variations in observed consumption, related perhaps to the fact that the dot household interviews lag the $\times$ household interviews by a couple of months. Third, the predicted sequence for the $\times$ household has a noticeably lower variance than the sample variance of the four $\times$ observations. This suggests that the especially high first value for the $\times$ household, which lies far above expectations, is due in part to measurement error or to an unexpected positive shock.

We can further assess model performance by comparing the means of $\hat{m}_h$ and $\hat{s}_h$ to

[9]Parameter estimates are provided in Appendix Table A1.

the mean and standard deviation of observed consumption. The empirical mean of daily consumption per adult equivalent is 121.8 TZS, while the mean of $\hat{m}_h$ is 121.2 TZS. This close match is reassuring, as the model is intended to center average consumption around its empirical mean. The empirical standard deviation of daily consumption per adult equivalent is 62.1 TZS, while the mean of $\hat{s}_h$ is 23.7 TZS. This is again in line with expectations. The mean of $\hat{s}_h$ is lower than the raw standard deviation because $\hat{s}_h$ isolates the seasonal variation, while the observed standard deviation includes both seasonality and idiosyncratic shocks. The relative magnitudes suggest that seasonal variation accounts for approximately 36% of variation around mean consumption.[10]

5.2 Main results

We now turn to the main results. The first set of estimates are from the specification in equation (6) without age–gender interactions. Table 3 shows the estimated coefficients. In column 1, a 10% increase in the mean of consumption is associated with .21% greater height. This is the expected positive association between the level of consumption during childhood and adult height. By contrast, an increase of 10% in the seasonality term is associated with 0.07% lower height, conditional on the mean. Both results are highly statistically significant, particularly given that the main hypotheses are one-sided. The magnitudes are small relative to some other studies, such as those that use droughts or wars to identify shocks to consumption (Hoddinott and Kinsey, 2001; Akresh, Lucchetti and Thirumurthy, 2012). This is not surprising given that the sample is relatively homogeneous and that identification is not based on a catastrophic change in child circumstances (see next section for more discussion of magnitudes). It is the relative magnitude of the coefficients that is most striking. The seasonality coefficient is a third of that of the mean coefficient. In other words, reducing the standard deviation of the seasonal component of consumption by 10% would have equivalent effects on height to those of a 3.3% increase in mean consumption.

Column 2 of Table 3 shows the estimates with the log of years of education as the dependent variable. Once again, the coefficients on the level and seasonality terms are

[10]This is an approximation because some allowance must be made for suppressed variation in the predicted variable $\hat{s}_h$. Hence the rationale for bootstrapping standard errors in the human capital regressions.

statistically significant and in the expected direction. A 10% increase in $\hat{m}_h$ is associated with a 4% increase in educational attainment, while a 10% increase in $\hat{s}_h$ is associated with a 1% decrease in educational attainment. The absolute magnitudes of both elasticities are substantially greater than those for height, though the relative magnitude is similar. In the case of educational attainment, a 10% decrease in the standard deviation of consumption is equivalent to a 2.5% increase in the mean.

In Section 2 we discussed the possibility of critical windows for human capital development, particularly during the first 1,000 days of life and during adolescence. Table 4 shows estimates of (6) with interactions between the consumption statistics and dummy variables for age groups. Results are based on regressions with the same controls as in Table 3, but Table 4 presents only the combined effects for each age group and the p-values from F-tests of significance.

Here, an even more striking pattern emerges. Column 1 shows that the seasonality term is associated with the most significant decreases in height for infants and children in utero, i.e., during the critical first 1,000 days of life. This suggests that at least in this population, a seasonally varying diet has the greatest effect on linear growth when it occurs during the formative period at the start of life. In contrast, the effect of seasonality on educational attainment is greatest for children aged 11-17, and only statistically significant for the two older age groups. One interpretation of this finding is that seasonality impacts education less by impeding cognitive development during infancy than by raising the opportunity cost of school for older children, or by making it harder for students to focus and apply effort in school.

Finally, in Table 5 we show a similar set of results disaggregated by age group and gender. The aim of this specification is to examine whether there are important differences for boys and girls that suggest biological or behavioral channels by which seasonality has gender-differentiated effects. In column 1, we see that the effect of seasonality on height, in utero and during infancy, is present only for girls. This is consistent with parents favoring male infants more than female infants during the lean season, or with gender-specific biological channels regulating development. In contrast, the effects on educational attainment for school-aged children are present only for boys. This suggests either that seasonal hunger has a greater impact on boys' ability to pay attention and perform in school, or that boys are more likely to be pulled from school during the lean season to assist in the household.

The only result in Table 5 that does not lend itself to clear interpretation is the statistically significant coefficient on the Log of $\hat{s}_h$ term for boys in age group 2. One possibility is that lean season deprivations have a greater impact on boys than girls during these years. Another possibility is that some of the older boys in age group 2 are experiencing negative impacts on growth during the adolescent growth spurt. However, in that case it is still puzzling that the effects are not present for girls, who usually experience puberty and adolescent growth sooner than boys.

5.3 Robustness

In Section 4.3 we described the rationale for a series of robustness checks in which we systematically drop variables from the consumption model and include their level and square from 1991-1994 in the 2010 regressions. To focus on key results without displaying hundreds of additional coefficients, we report the robustness results for the age groups of primary interest from specifications that include age group interactions (similar to those in Table 4).[11]

Figure 3 shows the point estimates and confidence intervals for regressions with height as the dependent variable. The coefficients shown are for age group 1. The clear pattern is that the coefficient on Log of $\hat{s}_h$ is robust to dropping variables and including them directly in the second stage. The coefficients on Log of $\hat{s}_h$ shown in the figure are statistically significant at conventional levels in all but two cases, and in both of those cases the p-values are below 0.14. Furthermore, because consumption is scaled by adult equivalent units, the attenuation of the Log of $\hat{s}_h$ coefficient from including this variable is partly mechanical, through the correlation between the adult equivalents variable and the denominator of the dependent variable.

Figure 4 shows a similar pattern for the education results. The coefficients shown

[11]Note that these are not chosen because they are our "strongest" results. The significant results from Table 5 are even more robust than the ones shown here, in that magnitudes are stable and statistical significance is unchanged when subject to the same procedure.

are for age group 3. Once again the coefficient on Log of $\hat{s}_h$ is attenuated most when the "Adult equivalents" variable is dropped from the consumption model and included in the human capital regressions. In all other cases the estimated coefficient lies between -0.1 and -0.2, and is statistically significant at conventional levels.

Taken as a whole, these additional regressions provide broad support for the robustness of the main findings.

A separate robustness concern relates to the possibility of non-random attrition in the long-term tracking of respondents from the 1991-1994 sample. While differential attrition would not invalidate our main findings, it would raise questions about their generalizability. To examine this, Table 6 shows a balance table for characteristics in 1991 between attritors (those not tracked in 2010) and non-attritors (those who were tracked). The table includes the predicted consumption statistics Log of $\hat{m}_h$ and Log of $\hat{s}_h$, individual age, individual gender, and all of the household-level variables from Table 1.

There are a number of variables in Table 6 that are statistically significantly different for attritors and non-attritors at conventional levels. For most of those, the differences are too small in magnitude to be of interest. Two sets of variables bear mention. The first are the predicted consumption statistics. Compared to non-attritors, attritors have on average 8% higher $\hat{m}_h$ and 17% higher $\hat{s}_h$. If our main regression results are correct, then these differences should be roughly offsetting, as far as height and educational attainment in 2010 are concerned. More importantly, the higher average $\hat{s}_h$ among the non-tracked group indicates that our results are not based on a subsample that happens to have greater-than-average consumption seasonality. The second group of variables to discuss are those with differences that are relatively large in magnitude. These are acres owned, livestock owned (measured as Tropical Livestock Units), and household size measured either as number of people or as number of adult equivalents. Differences in these variables are almost surely due to tracking in a long-term panel. Households with more land and livestock are less likely to move and therefore easier to find; and, the more people there are in the household, the more likely it is that at least one of them will be tracked (which facilitates tracking the others).

Overall, this attrition analysis gives us confidence that the main findings are based on a generally representative subsample of children. If anything, because tracked individuals come from households with less variable consumption and slightly higher land and livestock holdings than average, our results may understate the true population effect of consumption variability on human capital.

6 Discussion

In this concluding section we discuss two aspects of the findings. First, we provide additional context for interpreting the magnitudes of the estimated effects. We then discuss the implications for poverty measurement and for policy.

6.1 Magnitudes

Across Tables 3–5 our findings indicate that a 10% increase in average consumption during childhood leads to a 0.15–0.36% increase in height and a 3.2–5.0% increase in educational attainment, while a 10% increase in consumption seasonality during childhood leads to a .05–0.16% decrease in height and a 1.0–1.9% decrease in educational attainment. These magnitudes are smaller than those in some other studies of the link between consumption or nutrition and human capital. For example, Maluccio et al. (2009) report that children receiving a 10% increase in caloric support through the INCAP supplementation program in Guatemala experienced increases of 1.2 years of schooling attainment (for girls) and 2.5 additional cm of height at age 3 (for all children).[12] Those effects are 4-10 times greater than the effects of $\hat{m}_h$ measured here, and an order of magnitude larger than the effects of $\hat{s}_h$ (roughly 12-20 times the magnitude), although directly comparing the impacts on height is problematic because this paper looks at much longer-term effects. Regardless, it is generally the case that the effects measured here are smaller in magnitude than many in the literature.

We think there are two main reasons for this difference. The first is that our findings are based on a relatively short window of observation during childhood. Children acquire height and cognitive skills / education over two decades, while we use consumption statistics from just over two years. It is well established in the literature on transitory poverty that household consumption profiles can and do change, both gradually over the medium term

[12]For details on the INCAP study, see Stein et al. (2008).

and sharply from one year to the next (Baulch and Hoddinott, 2000; Dercon and Krishnan, 2000; Barrett, 2005). In this regard, our measures of $\hat{m}_h$ and $\hat{s}_h$ are noisy, and subject to attenuation bias due to measurement error. This also explains why we see different effects by age group. If consumption profiles were fixed throughout childhood, then the age at which we observe a child would not matter for estimating the link between seasonality and human capital (any sufficiently long snapshot of consumption during childhood would suffice).

The second reason for the difference in magnitudes is that the KHDS sample is more homogeneous than the population of Tanzania as a whole. The original 1991-1994 households were selected from a single region with a single major ethnic group, a single agro-climatic zone, a common set of primary crops, and broadly similar access to farm and non-farm employment opportunities (which facilitate income diversification and consumption smoothing). Indeed, we noted in Section 3.3 that the standard deviation of height in the nationally representative LSMS-ISA data from Tanzania is nearly 25% greater than the comparable statistic from the KHDS sample. The effects in a more heterogeneous population may well be larger than those estimated in the KHDS, where there is less variation to explain.

Given these limitations, it is possible that our results greatly understate the true relationship between seasonality and human capital. We noted above that the effects of a level increase in consumption from INCAP were 4-10 times the magnitude of increases in $\hat{m}_h$ in this paper. If we assume that the seasonality effects are attenuated by the same degree as the average consumption effects, then scaling up the magnitude of the seasonality effects by 7 (the midpoint of the 4-10 range) gives an indication of how large the negative consequences of seasonal consumption might be. With this re-scaling, a 10% increase in seasonality would lead to a .35–1.1% decrease in height (0.6–1.8 centimeters) and a 7.0–13.3% decrease in educational attainment (0.5–1.0 years). At these levels, elimination of chronic seasonality has impacts on human capital similar to those from avoiding droughts or major conflicts. This is only a back-of-the-envelope exercise, as it involves numerous assumptions about the comparability between our findings and those from INCAP. Nevertheless, it provides useful context for understanding how detrimental seasonal consumption might be to human capital development.

6.2 Implications for measurement and policy

If consumption seasonality matters for child development, what does that suggest for research and policy? Currently, poverty lines are set with reference to average food consumption. Resources are often targeted to those with the lowest average food consumption or lowest predicted average food consumption. We have shown that two children who have similar levels of consumption when averaged over the year will have markedly different human capital outcomes if one child's consumption profile is more seasonal than the other. One implication of this finding is that current approaches to measuring poverty would have greater relevance to long-term child wellbeing if augmented with a measure of seasonality. This is an important finding regardless of whether or not the association we document is causal.

There are (at least) two ways that current consumption measures could be augmented to account for seasonality. First, if a consumption survey is to be implemented on a geographically stratified, 12-month calendar (as some are), then analysts could estimate a model of seasonal consumption similar to that used in stage 1 of our analysis. In doing so, the emphasis should be on projecting daily consumption and estimating $\hat{m}_h$ and $\hat{s}_h$, not on interpreting the magnitudes of the coefficients in the model. In targeting scarce resources in a food security intervention, the case could then be made that two households with similar values of $\hat{m}_h$, but different levels of $\hat{s}_h$, should receive different levels of intervention.

A second approach would involve adapting the Food Consumption Score, or some other widely used method to analyze food security, to include questions about seasonality. Most such measures do not take into account seasonal consumption fluctuations (though they may be implemented during the lean season precisely because of this concern). Proposing a specific re-design of such measures would take us beyond the scope of this paper. Yet, our findings suggest a new and important rationale for incorporating some information about seasonality into food security measurement tools.

In broader policy terms, it is already the case that some interventions are designed to provide counter-cyclical consumption support. Targeted food aid is often implicitly counter-cyclical, whether through the strategic release of grain reserves to slow the pace of food price increases during the lean season, or through intervention in the face of a drought or other large covariant shock. Yet, the findings here suggest that governments and development practitioners could go a step further. Because seasonality has an effect on longterm outcomes that is separate from that of average consumption, there is an argument for seasonally varying any program that involves consumption support (such as a subsidy for food purchases or a food transfer through a conditional program). Depending on the cost of moving assets across time, seasonally varying support could have a substantially larger effect on human capital than providing uniform transfers throughout the year. This is the case even if total resources are held constant.

Finally, our results demonstrate a potentially important channel by which structural transformation influences human development. When the shares of manufacturing and services in the economy surpass that of the agriculture sector, incomes becomes decidedly less seasonal. This reduces intra-annual variability in consumption. Our results suggest that this reduction in seasonality will have a meaningful effect on child development, independent of that which is due to income growth.

Appendix - intended for online publication only

A Consumption model: parameter estimates

Table A1 shows estimates of parameter vectors $\hat{\phi}$ and $\hat{\rho}$ from the consumption model in equation (5).

B Bootstrap procedure for standard errors

To address the concerns that (*a*) variation in household food consumption may be suppressed in the prediction stage, and (*b*) errors may be clustered at levels we do not account for, we use a bootstrap procedure to compute standard errors in tables (4), (5), and (6). The steps of this procedure are as follows:

1. Draw a sample with replacement of individuals in 2010 for which we observe height or education and can link to a 1991-1994 household.

2. Keep all 1991-1994 households that have at least one offspring who was selected in iteration j.

3. Re-estimate equation (5) on this set of 1991-1994 households with each household weighted by $\hat{W} = n_h/n_{hj}$ where h is the number of tracked children in 2010 for a given household from 91-94, n_h is the number of 94-94 households with h children in the original 91-94 and 2010 samples, and n_{hj} is the number of unique 91-94 households who have h children appear in the 2010 sample from iteration j.[13]

[13]Weights are needed to account for the fact that because we sample from 2010 individuals, households with more children in the 2010 sample have a greater probability of being selected at least once than households with fewer children appearing in the 2010 sample. Because each household contributes one observation in 2010, failing to weight would mean selection too few 91-94 households on average and over-representing households who have more children in the 2010 sample. Estimating weights within iteration ensures that the weighted sample in each iteration is exactly the size of the original 91-94 sample of households. Across iterations, these estimated weights converge to 1/(probability that a 91-94 household with h children in 2010 is selected in iteration j) $= 1/[1-(1-(h/N))^N]$ where h is the number of children from a given 91-94 household that appear in the 2010 data and N is the total number of individuals in the 2010 data.

4. Using the parameters estimated in step 3, project daily consumption $\hat{c}$ for each household and compute $\hat{m}$ and $\hat{s}$, the mean and standard deviation of the $\hat{c}$ path for each household.

5. Re-estimate equation (6) using 2010 outcomes and $\hat{m}$ and $\hat{s}$ estimated in step 4.

6. Repeat steps 1-5 10,000 times, storing the coefficients from step 5 for each iteration. The p-values for the null hypothesis that the effect of $\hat{m}$ or $\hat{s}$ on 2010 outcomes is 0 are the percentiles of these 10,000 which fall above (below) 0 when the mean of coefficients on $\hat{m}$ or $\hat{s}$ are greater than (less than) 0.

Tables A2, A3, and A4 below show the coefficients from estimating the coefficients and p-values on the original sample and using the bootstrap procedure. In all cases, the point estimates of the coefficients are virtually identical to the ones conventionally computed on the original sample. The p-values associated with the coefficients in the two methods are also qualitatively very similar, with very few cases where the choice of method for computing p-values influences whether a coefficient is significant at standard levels.

References

Abay, Kibrewossen, and Kalle Hirvonen. 2016. "Does market access mitigate the impact of seasonality on child growth?: Panel data evidence from northern Ethiopia." *IFPRI working paper.*

Adams, AM. 1995. "Seasonal variations in energy balance among agriculturalists in central Mali: compromise or adaptation?" *European Journal of Clinical Nutrition*, 49(11): 809–823.

Akresh, Richard, Leonardo Lucchetti, and Harsha Thirumurthy. 2012. "Wars and child health: Evidence from the Eritrean–Ethiopian conflict." *Journal of Development Economics*, 99(2): 330–340.

Alderman, Harold, John Hoddinott, and Bill Kinsey. 2006. "Long Term Consequences of Early Childhood Malnutrition." *Oxford Economic Papers*, 58(3): 450–474.

Almond, Douglas, and Janet Currie. 2011. "Killing Me Softly: The Fetal Origins Hypothesis." *Journal of Economic Perspectives*, 25(3): 153–172.

Arsenault, Joanne E, Laetitia Nikiema, Pauline Allemand, Kossiwavi A Ayassou, Hermann Lanou, Mourad Moursi, Fabiana F De Moura, and Yves Martin-Prevel. 2014. "Seasonal differences in food and nutrient intakes among young children and their mothers in rural Burkina Faso." *Journal of Nutritional Science*, 3: e55.

Attanasio, Orazio P. 2015. "The Determinants of Human Capital Formation During the Early Years of Life: Theory, Measurement, and Policies." *Journal of the European Economic Association*, 6: 949–997.

Barrett, Christopher B. 2005. "Rural poverty dynamics: development policy implications." *Agricultural Economics*, 32(s1): 45–60.

Baulch, Bob, and John Hoddinott. 2000. "Economic mobility and poverty dynamics in developing countries." *The Journal of Development Studies*, 36(6): 1–24.

Beegle, Kathleen, Rajeev H Dehejia, and Roberta Gatti. 2006. "Child labor and agricultural shocks." *Journal of development Economics*, 81(1): 80–96.

Behrman, Jere R. 1988*a*. "Intrahousehold allocation of nutrients in rural India: Are boys favored? Do parents exhibit inequality aversion?" *Oxford Economic Papers*, 40(1): 32–54.

Behrman, Jere R. 1988*b*. "Nutrition, health, birth order and seasonality: Intrahousehold allocation among children in rural India." *Journal of Development Economics*, 28(1): 43–62.

Behrman, Jere R. 1993. "The economic rationale for investing in nutrition in developing countries." *World Development*, 21(11): 1749–1771.

Behrman, Jere R. 2016. "Growth Faltering in the First Thousand Days After Conception and Catch-Up Growth." In *The Oxford Handbook of Economics and Human Biology.* , ed. John Komlos and Inas R. Kelly, Chapter 1, 266–290. Oxford:Oxford University Press.

Bhagowalia, Priya, Susan E Chen, and William A Masters. 2011. "Effects and determinants of mild underweight among preschool children across countries and over time." *Economics & Human Biology*, 9(1): 66–77.

Bleakley, Hoyt. 2010. "Health, human capital, and development." *Annual Review of Economics*, 2: 283.

Branca, F, G Pastore, Tsegaye Demissie, and A Ferro-Luzzi. 1993. "The nutritional impact of seasonality in children and adults of rural Ethiopia." *European Journal of Clinical Nutrition*, 47(12): 840–850.

Bundervoet, Tom, Philip Verwimp, and Richard Akresh. 2009. "Health and civil war in rural Burundi." *Journal of Human Resources*, 44(2): 536–563.

Case, Anne, Angela Fertig, and Christina Paxson. 2005. "The lasting impact of childhood health and circumstance." *Journal of health economics*, 24(2): 365–389.

Chikhungu, Lana Clara, and Nyovani Janet Madise. 2014. "Seasonal variation of child under nutrition in Malawi: is seasonal food availability an important factor? Findings from a national level cross-sectional study." *BMC Public Health*, 14(1): 1.

Coate, Stephen, and Martin Ravallion. 1993. "Reciprocity without commitment: Characterization and performance of informal insurance arrangements." *Journal of Development Economics*, 40(1): 1–24.

Cunha, Flavio, and James Heckman. 2007. "Technology of Skill Formation." *American Economic Review*, 97(2): 31–47.

Currie, Janet, and Douglas Almond. 2011. "Human capital development before age five." *Handbook of Labor Economics*, 4: 1315–1486.

Currie, Janet, and Duncan Thomas. 1999. "Early Test Scores, Socioeconomic Status and Future Outcomes." *NBER Working Paper Series*, , (6943).

Dercon, Stefan, and Catherine Porter. 2014. "Live Aid Revisited: Long-Term Impacts of the 1984 Ethiopian Famine on Children." *Journal of the European Economic Association*, 12(4): 927–948.

Dercon, Stefan, and Pramila Krishnan. 2000. "Vulnerability, seasonality and poverty in Ethiopia." *The Journal of Development Studies*, 36(6): 25–53.

De Weerdt, Joachim, Kathleen Beegle, Helene Bie Lilleor, Stefan Dercon, Kalle Hirvonen, Martina Kirchberger, and Sofya Krutikova. 2012. "Kagera Health and Development Survey 2010: Basic Information Document." *Rockwool Foundation Working Paper Series*, 46.

Egata, Gudina, Yemane Berhane, and Alemayehu Worku. 2013. "Seasonal variation in the prevalence of acute undernutrition among children under five years of age in east rural Ethiopia: a longitudinal study." *BMC Public Health*, 13(1): 1.

Fafchamps, Marcel, and Susan Lund. 2003. "Risk-sharing networks in rural Philippines." *Journal of Development Economics*, 71(2): 261–287.

Filmer, Deon, and Lant H Pritchett. 2001. "Estimating wealth effects without expenditure data—or tears: An application to educational enrollments in states of India." *Demography*, 38(1): 115–132.

Foster, Andrew. 1995. "Prices, Credit Markets and Child Growth in Low-Income Rural Areas." *The Economic Journal*, 105: 551–570.

Glewwe, Paul, Hanan G Jacoby, and Elizabeth M King. 2001. "Early childhood nutrition and academic achievement: a longitudinal analysis." *Journal of Public Economics*, 81(3): 345–368.

Grantham-McGregor, Sally, Yin Bun Cheung, Santiago Cueto, Paul Glewwe, Linda Richter, Barbara Strupp, and the International Child Development Steering Group. 2007. "Developmental potential in the first 5 years for children in developing countries." *The Lancet*, 369: 60–70.

Heckman, James J. 2010. "Schools, Skills, and Synapses." *Economic Inquiry*, 46(3): 289.

Hirvonen, Kalle, Alemayehu Seyoum Taffesse, and Ibrahim Worku Hassen. 2015. "Seasonality and household diets in Ethiopia." *Public Health Nutrition*, 1–8.

Hoddinott, John, and Bill Kinsey. 2001. "Child growth in the time of drought." *Oxford Bulletin of Economics and Statistics*, 63(4): 409–436.

Hoyne, Hilary, Diane Whitmore Schanzenbach, and Douglas Almond. 2016. "Long-Run Impacts of Childhood Access to the Safety Net." *The American Economic Review*, 106(4): 903–934.

Kinnan, Cynthia. 2011. "Distinguishing barriers to insurance in Thai villages." Northwestern University working paper.

Kristjansson, EA, V Robinson, M Petticrew, B MacDonald, J Krasevec, L Janzen, T Greenhalgh, G Wells, J MacGowan, A Farmer, BJ Shea, A Mayhew, and P Tugwell. 2007. "School Feeding for improving the physical and psychosocial health of disadvantaged elementary school children." *Cochrane Database Systematic Review*, 24(1).

Limony, Yehuda, Sławomir Kozieł, and Michael Friger. 2015. "Age of onset of a normally timed pubertal growth spurt affects the final height of children." *Pediatric research*, 78(3): 351–355.

Maleta, K, SM Virtanen, M Espo, T Kulmala, and P Ashorn. 2003. "Seasonality of growth and the relationship between weight and height gain in children under three years of age in rural Malawi." *Acta Paediatrica*, 92(4): 491–497.

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

Mancini, Sharon, and Dean Yang. 2009. "Under the Weather: Health Schooling, and Economic Consequences of Early-Life Rainfall." *American Economic Review*, 99(3).

Martorell, Reynaldo. 1999. "The nature of child malnutrition and its long-term implications." *Food and Nutrition Bulletin*, 20(3): 288–292.

Martorell, R, LK Khan, and DG Schroeder. 1994. "Reversibility of stunting: epidemiological findings in children from developing countries." *European Journal of Clinical Nutrition*, 48: S45–57.

Miller, Jacqueline, Brett Ritchie, Cuong Tran, Sean Beggs, Christina Olly Lada, Kathryn Whetter, and Lynne Cobiac. 2013. "Seasonal variation in the nutritional status of children aged 6 to 60 months in a resettlement village in West Timor." *Asia Pacific Journal of Clinical Nutrition*, 22(3): 449.

Mitchikpe, Comlan ES, Romain AM Dossa, Eric AD Ategbo, Joop MA Van Raaij, and Frans J Kok. 2009. "Seasonal variation in food pattern but not in energy and nutrient intakes of rural Beninese school-aged children." *Public health nutrition*, 12(03): 414–422.

Morduch, Jonathan. 1995. "Income smoothing and consumption smoothing." *The Journal of Economic Perspectives*, 9(3): 103–114.

Peet, Evan D., Dana C. McCoy, Goodarz Danaei, Majid Ezzati, Wafaie Fawzi, Marjo-Riitta Jarvelin, Demetris Pillas, and Gunther Fink. 2015. "Early Childhood Development and Schooling Attainment: Longitudinal Evidence from British, Finnish and Philippine Birth Cohorts." *PLOS One*, 10(9).

Pollitt, E, NL Lewis, C Garza, and RJ Shulman. 1982-1983. "Fasting and Cognitive Function." *Journal of Psychiatry Research*, 17(2): 169=74.

Powell, Christine A, Susan P Walker, Susan M Chang, and Sally M Grantham-McGregor. 1998. "Nutrition and education: a randomized trial of the effects of breakfast in rural primary school children." *The American Journal of Clinical Nutrition*, 68(4): 873–879.

Rampersaud, Gail C, Mark A Pereira, Beverly L Girard, Judi Adams, and Jordan D Metzl. 2005. "Breakfast habits, nutritional status, body weight, and academic performance in children and adolescents." *Journal of the American Dietetic Association*, 105(5): 743–760.

Rogol, Alan D. 2000. "Growth and pubertal development in children and adolescents: effects of diet and physical activity." *American Journal of Clinical Nutrition*, 72(2): 521s–528s.

Sahn, David E., ed. 1989. *Seasonal Variability in Third World Agriculture: The Consequences for Food Security.* Johns Hopkins University Press.

Sahn, David E, and David Stifel. 2003. "Exploring alternative measures of welfare in the absence of expenditure data." *Review of Income and Wealth*, 49(4): 463–489.

Shetty, PS. 1999. "Adaptation to low energy intakes: the responses and limits to low intakes in infants, children and adults." *European Journal of Clinical Nutrition*, 53: S14–33.

Simondon, KB, F Simondon, I Simon, A Diallo, E Benefice, P Traissac, and B Maire. 1998. "Preschool stunting, age at menarche and adolescent height: a longitudinal study in rural Senegal." *European Journal of Clinical Nutrition*, 52: 412–418.

Singh, Abhijeet, Albert Park, and Stefan Dercon. 2013. "School Meals as a Safety Net: An Evaluation of the Midday Meal Scheme in India." *Economic Development and Cultural Change*, 2: 275–306.

Stein, Aryeh D, Paul Melgar, John Hoddinott, and Reynaldo Martorell. 2008. "Cohort profile: The Institute of Nutrition of Central America and Panama (INCAP) nutrition trial cohort study." *International Journal of Epidemiology*, 37(4): 716–720.

Townsend, Robert M. 1994. "Risk and insurance in village India." *Econometrica*, 539–591.

Victora, Cesar G, Linda Adair, Caroline Fall, Pedro C Hallal, Reynaldo Martorell, Linda Richter, Harshpal Singh Sachdev, Maternal, Child Undernutrition Study Group, et al. 2008. "Maternal and child undernutrition: consequences for adult health and human capital." *The Lancet*, 371(9609): 340–357.

Wisniewski, Suzanne LW. 2010. "Child nutrition, health problems, and school achievement in Sri Lanka." *World Development*, 38(3): 315–332.

Tables

Table 1: Household summary statistics from 4 rounds of KHDS 1, 1991-1994, pooled

Variable	Mean	Standard deviation
Head is male (=1)	0.72	0.45
Head age (years)	49.46	16.96
Head education (years)	4.07	3.20
Head can read (=1)	0.72	0.45
Head basic math (=1)	0.72	0.45
Asset index	-0.01	0.97
Acres owned	5.02	27.64
Value of agricultural capital (TZS)	5928.01	67034.83
Has non-farm business (=1)	0.40	0.49
Tropical livestock units	1.45	7.29
Household size, number	6.81	3.53
Household size, adult equivalents	4.66	2.45
Muslim (=1)	0.12	0.32
Catholic (=1)	0.60	0.49
Christian (=1)	0.24	0.43
Haya ethnicity (=1)	0.63	0.48
Hangaza ethnicity (=1)	0.13	0.33
Speaks Swahili (=1)	0.91	0.29

Notes: Authors' calculations from KHDS data. "Tropical livestock units" is an index of livestock holdings, scaled so that 1 unit is equivalent to 1 cow. Adult equivalent units are the sum of household members where members have the following weights: age 0 to 5 years: weight = 0.5; age 5 to 14 or age 50+: weight = 0.7; age 15 to 50: weight =1. Statistics are for 758 households that were interviewed in all rounds of KHDS 1.

Table 2: Summary statistics in 2010

Variable	N	Mean	Standard deviation	Median
Age group 1 (17–21 in 2010; -2–2 in 1991)	2859	0.22		
Age group 2 (22–29 in 2010; 3–10 in 1991)	2859	0.40		
Age group 3 (30–36 in 2010; 11–17 in 1991)	2859	0.38		
Male (=1)	2859	0.49		
Height (cm)	2578	162.5	8.3	162.2
Education (years)	2633	7.6	2.9	7.0
Height (cm) - age group 1	517	160.5	8.1	160.3
Height (cm) - age group 2	1039	162.4	8.4	162.0
Height (cm) - age group 3	1022	163.6	8.2	163.3
Education (years) - age group 1	593	7.9	2.8	7.0
Education (years) - age group 2	1022	7.7	3.3	7.0
Education (years) - age group 3	1018	7.3	2.6	7.0

Notes: Authors' calculations from KHDS data. Data are summary statistics from KHDS 3, collected in 2010.

Table 3: Height and education regressed on logs of $\hat{m}_h$ and $\hat{s}_h$

Dependent variable:	Log height (1)	Log years educ. (2)
Log of $\hat{m}_h$	0.021^{***}	0.402^{***}
	(0.000)	(0.000)
Log of $\hat{s}_h$	-0.007^{***}	-0.097^{***}
	(0.004)	(0.002)
Male(=1)	0.060^{***}	0.020
	(0.000)	(0.141)
Age group 2 (=1)	0.014^{***}	-0.056^{**}
	(0.000)	(0.013)
Age group 3 (=1)	0.021^{***}	-0.078^{***}
	(0.000)	(0.001)
Observations	2309	2333
R^2	0.39	0.06

Notes: Authors' calculations from KHDS data. All parameters and p-values are computed through 10,000 bootstrap iterations as described in the appendix. All regressions include district fixed effects. Observations are the number of 2010 individuals used to estimate equation (6). ***sig. at 1%, **sig. at 5%, *sig. at 10%.

Table 4: Height and education regressed on logs of $\hat{m}_h$ and $\hat{s}_h$, by age group

Dependent variable:	Log height (1)	Log years educ. (2)
Log of $\hat{m}_h$, age group 1	0.021^{**}	0.401^{***}
	(0.020)	(0.000)
Log of $\hat{m}_h$, age group 2	0.025^{***}	0.409^{***}
	(0.000)	(0.000)
Log of $\hat{m}_h$, age group 3	0.017^{***}	0.402^{***}
	(0.004)	(0.000)
Log of $\hat{s}_h$, age group 1	-0.011^{***}	-0.044
	(0.009)	(0.216)
Log of $\hat{s}_h$, age group 2	-0.006^{**}	-0.095^{**}
	(0.044)	(0.031)
Log of $\hat{s}_h$, age group 3	-0.005^{*}	-0.130^{***}
	(0.083)	(0.001)
Observations	2309	2333
R^2	0.39	0.06

Notes: Authors' calculations from KHDS data. Estimates are the combined effects of interactions between binary variables for age group membership and the predicted consumption statistics. All parameters and p-values are computed through 10,000 bootstrap iterations as described in Appendix section B. All regressions include district fixed effects. Observations are the number of 2010 individuals used to estimate equation (6). ***sig. at 1%, **sig. at 5%, *sig. at 10%.

Table 5: Height and education regressed on logs of $\hat{m}_h$ and $\hat{s}_h$, by age and gender

Dependent variable:	Log height (1)	Log years educ. (2)
Log of $\hat{m}_h$, age group 1, women	0.015	0.467***
	(0.136)	(0.000)
Log of $\hat{m}_h$, age group 2, women	0.020**	0.451***
	(0.020)	(0.000)
Log of $\hat{m}_h$, age group 3, women	0.016**	0.318***
	(0.040)	(0.001)
Log of $\hat{s}_h$, age group 1, women	−0.016***	-0.002
	(0.006)	(0.503)
Log of $\hat{s}_h$, age group 2, women	-0.001	-0.050
	(0.409)	(0.213)
Log of $\hat{s}_h$, age group 3, women	-0.005	-0.077
	(0.168)	(0.101)
Log of $\hat{m}_h$, age group 1, men	0.024*	0.334***
	(0.058)	(0.007)
Log of $\hat{m}_h$, age group 2, men	0.032***	0.372**
	(0.002)	(0.026)
Log of $\hat{m}_h$, age group 3, men	0.019**	0.495***
	(0.019)	(0.000)
Log of $\hat{s}_h$, age group 1, men	-0.005	-0.100
	(0.257)	(0.123)
Log of $\hat{s}_h$, age group 2, men	−0.012***	−0.135**
	(0.006)	(0.035)
Log of $\hat{s}_h$, age group 3, men	-0.006	−0.191***
	(0.142)	(0.000)
Observations	2309	2333
R^2	0.39	0.06

Notes: Authors' calculations from KHDS data. Estimates are the combined effects of interactions between binary variables for age group membership, gender, and the predicted consumption statistics. All parameters and p-values are computed through 10,000 bootstrap iterations as described in Appendix section B. All regressions include district fixed effects. Observations are the number of 2010 individuals used to estimate equation (6). ***sig. at 1%, **sig. at 5%, *sig. at 10%.

Table 6: Tests for Differential Attrition

Variable	Mean (attritors) (1)	Mean (non-attritors) (2)	Difference (3)	P-value (4)
Log of $\hat{m}_h$	4.72	4.64	0.08	0.00
Log of $\hat{s}_h$	2.82	2.65	0.17	0.00
Individual is male (=1)	1.49	1.50	-0.00	0.91
Individual age (years)	9.05	8.62	0.43	0.08
Head is male (=1)	0.73	0.76	-0.03	0.12
Head age (years)	48.67	48.49	0.18	0.80
Head education (years)	4.24	4.18	0.06	0.68
Head can read (=1)	0.72	0.77	-0.05	0.01
Head basic math (=1)	0.73	0.75	-0.03	0.20
Asset index	0.15	0.05	0.09	0.03
Acres owned	4.15	5.48	-1.33	0.00
Value of agricultural capital (TZS)	1791	1682	109	0.88
Has non-farm business (=1)	0.30	0.29	0.01	0.78
Tropical livestock units	0.94	2.11	-1.18	0.00
Household size, number	7.14	8.08	-0.94	0.00
Household size, adult equivalents	5.60	6.29	-0.69	0.00
Muslim (=1)	0.12	0.13	-0.01	0.61
Catholic (=1)	0.58	0.59	-0.02	0.47
Christian (=1)	0.27	0.24	0.03	0.12
Haya ethnicity (=1)	0.63	0.60	0.03	0.15
Hangaza ethnicity (=1)	0.13	0.12	0.02	0.30
Speaks Swahili (=1)	0.91	0.93	-0.02	0.07

Notes: Authors' calculations from KHDS data. Table includes all individuals aged -2 to 17 in the first round of the 1991-1994 surveys. Attritors are those individuals who do not appear in the 2010 surveys; non-attritors are those who do. P-values are for the difference in mean value of variables in the first round of the 1991-1994 surveys.

Appendix Tables – for online publication only

Table A1: Estimates of $\hat{\phi}$ and $\hat{\rho}$ from model (5)

	$\hat{\phi}$		$\hat{\rho}$	
	Coefficient	st. error	Coefficient	st. error
	(1)	(2)	(3)	(4)
Head is male (=1)	35.07**	(16.79)	-10.24	(14.32)
Head age (years)	0.71*	(0.39)	-0.48	(0.59)
Head education (years)	-3.13	(4.46)	-4.45	(4.02)
Head can read (=1)	25.76	(29.86)	-0.64	(42.39)
Head basic math (=1)	26.64	(29.06)	7.00	(38.73)
Adult equivalents	55.56***	(9.74)	4.60	(12.48)
Acres owned	4.72***	(1.54)	4.97**	(2.00)
Livestock (TLU)	7.71**	(3.42)	2.18	(3.30)
Haya ethnicity (=1)	-8.92	(27.42)	24.64	(33.95)
Hangaza ethnicity (=1)	11.58	(51.32)	30.33	(47.29)
Muslim (=1)	20.84	(30.66)	47.87**	(22.97)
Catholic (=1)	-5.27	(19.32)	2.51	(23.93)
Speaks Swahili (=1)	27.71	(24.06)	-21.75	(32.86)
Household size	8.69	(7.51)	3.92	(8.24)
Has non-farm business (=1)	-5.22	(15.97)	-15.13	(25.63)
Asset index	92.51***	(14.64)	13.22	(14.21)
Value of agricultural capital (TZS)	0.0004	(0.0004)	0.0002	(0.0008)
Observations	2943			
R^2	0.35			

Notes: Authors' calculations from KHDS data. Model also includes district and year effects. ***sig. at 1%, **sig. at 5%, *sig. at 10%.

Table A2: Height and education regressed on $\hat{m}_h$ and $\hat{s}_h$

Dependent variable:	Log height [BS] (1)	Log height [Conv.] (2)	Log years educ [BS] (3)	Log years educ [Conv.] (4)
Log of $\hat{m}_h$	0.021***	0.021***	0.402***	0.401***
	(0.000)	(0.001)	(0.000)	(0.000)
Log of $\hat{s}_h$	−0.007***	−0.007**	−0.097***	−0.097**
	(0.004)	(0.041)	(0.002)	(0.018)
Male(=1)	0.060***	0.060***	0.020	0.020
	(0.000)	(0.000)	(0.141)	(0.295)
Age group 2 (=1)	0.014***	0.014***	−0.056**	−0.056**
	(0.000)	(0.000)	(0.013)	(0.025)
Age group 3 (=1)	0.021***	0.021***	−0.078***	−0.078***
	(0.000)	(0.000)	(0.001)	(0.001)
Observations	2309	2309	2333	2333
R^2	0.39	0.39	0.06	0.06

Notes: Authors' calculations from KHDS data. In columns 2 and 4, p-values in parentheses are computed conventionally. In columns 1 and 3 coefficients are calcuated from the mean coefficient over 10,000 bootstrap iterations and p-values in parentheses are computed from distributions of these bootstrapped coefficients. All regressions include district fixed effects. Observations are the mean number of observations used to estimate equation (6) across bootstrap samples. ***sig. at 1%, **sig. at 5%, *sig. at 10%.

Table A3: Height and education regressed on $\hat{m}_h$ and $\hat{s}_h$, by age group

Dependent variable:	Log height [BS] (1)	Log height [Conv.] (2)	Log years educ [BS] (3)	Log years educ [Conv.] (4)
Log of $\hat{m}_h$, age group 1	0.021^{**}	0.021^{*}	0.401^{***}	0.399^{***}
	(0.020)	(0.056)	(0.000)	(0.000)
Log of $\hat{m}_h$, age group 2	0.025^{***}	0.025^{***}	0.409^{***}	0.405^{***}
	(0.000)	(0.002)	(0.000)	(0.001)
Log of $\hat{m}_h$, age group 3	0.017^{***}	0.017^{**}	0.402^{***}	0.401^{***}
	(0.004)	(0.026)	(0.000)	(0.000)
Log of $\hat{s}_h$, age group 1	-0.011^{***}	-0.012^{**}	-0.044	-0.043
	(0.009)	(0.020)	(0.216)	(0.502)
Log of $\hat{s}_h$, age group 2	-0.006^{**}	-0.006	-0.095^{**}	-0.095^{*}
	(0.044)	(0.147)	(0.031)	(0.086)
Log of $\hat{s}_h$, age group 3	-0.005^{*}	-0.005	-0.130^{***}	-0.129^{***}
	(0.083)	(0.216)	(0.001)	(0.008)
Observations	2309	2309	2333	2333
R^2	0.39	0.38	0.06	0.05

Notes: Authors' calculations from KHDS data. Estimates are the combined effects of interactions between binary variables for age group membership, and the predicted consumption statistics from equation (6). In columns 2 and 4, p-values in parentheses are computed conventionally. In columns 1 and 3 coefficients are calcuated from the mean coefficient over 10,000 bootstrap iterations and p-values in parentheses are computed from distributions of these bootstrapped coefficients. All regressions include district fixed effects. Observations are the mean number of observations used to estimate equation (6) across bootstrap samples. ***sig. at 1%, **sig. at 5%, *sig. at 10%.

Table A4: Height and education regressed on $\hat{m}_h$ and $\hat{s}_h$, by age and gender

Dependent variable:	Log height [BS] (1)	Log height [Conv.] (2)	Log years educ [BS] (3)	Log years educ [Conv.] (4)
Log of $\hat{m}_h$, age group 1, women	0.015	0.015	0.467***	0.463***
	(0.136)	(0.269)	(0.000)	(0.001)
Log of $\hat{m}_h$, age group 2, women	0.020**	0.020*	0.451***	0.448***
	(0.020)	(0.054)	(0.000)	(0.001)
Log of $\hat{m}_h$, age group 3, women	0.016**	0.017	0.318***	0.316***
	(0.040)	(0.106)	(0.001)	(0.007)
Log of $\hat{s}_h$, age group 1, women	−0.016***	−0.015**	-0.002	0.002
	(0.006)	(0.010)	(0.503)	(0.979)
Log of $\hat{s}_h$, age group 2, women	-0.001	-0.001	-0.050	-0.049
	(0.409)	(0.834)	(0.213)	(0.473)
Log of $\hat{s}_h$, age group 3, women	-0.005	-0.005	-0.077	-0.076
	(0.168)	(0.363)	(0.101)	(0.233)
Log of $\hat{m}_h$, age group 1, men	0.024*	0.024	0.334***	0.333**
	(0.058)	(0.122)	(0.007)	(0.026)
Log of $\hat{m}_h$, age group 2, men	0.032***	0.032***	0.372**	0.365*
	(0.002)	(0.005)	(0.026)	(0.058)
Log of $\hat{m}_h$, age group 3, men	0.019**	0.019*	0.495***	0.494***
	(0.019)	(0.068)	(0.000)	(0.000)
Log of $\hat{s}_h$, age group 1, men	-0.005	-0.005	-0.100	-0.098
	(0.257)	(0.509)	(0.123)	(0.264)
Log of $\hat{s}_h$, age group 2, men	−0.012***	−0.012**	−0.135**	−0.134*
	(0.006)	(0.028)	(0.035)	(0.072)
Log of $\hat{s}_h$, age group 3, men	-0.006	-0.006	−0.191***	−0.190***
	(0.142)	(0.335)	(0.000)	(0.000)
Observations	2309	2309	2333	2333
R^2	0.39	0.39	0.06	0.05

Notes: Authors' calculations from KHDS data. Estimates are the combined effects of interactions between binary variables for age group membership, gender, and the predicted consumption statistics from equation (6). In columns 2 and 4, p-values in parentheses are computed conventionally. In columns 1 and 3 coefficients are calcuated from the mean coefficient over 10,000 bootstrap iterations and p-values in parentheses are computed from distributions of these bootstrapped coefficients. All regressions include district fixed effects. Observations are the mean number of observations used to estimate equation (6) across bootstrap samples. ***sig. at 1%, **sig. at 5%, *sig. at 10%.

Figures

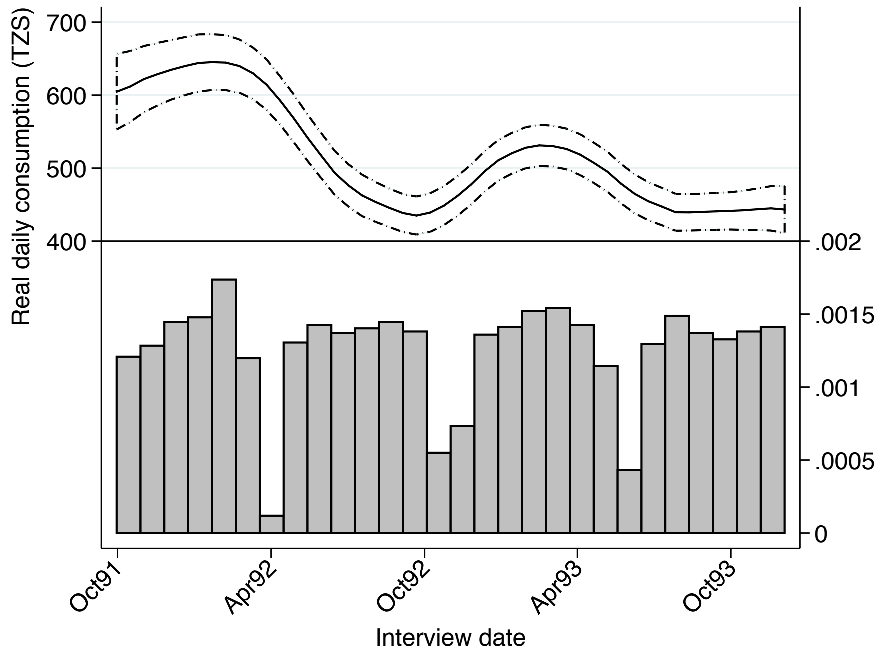
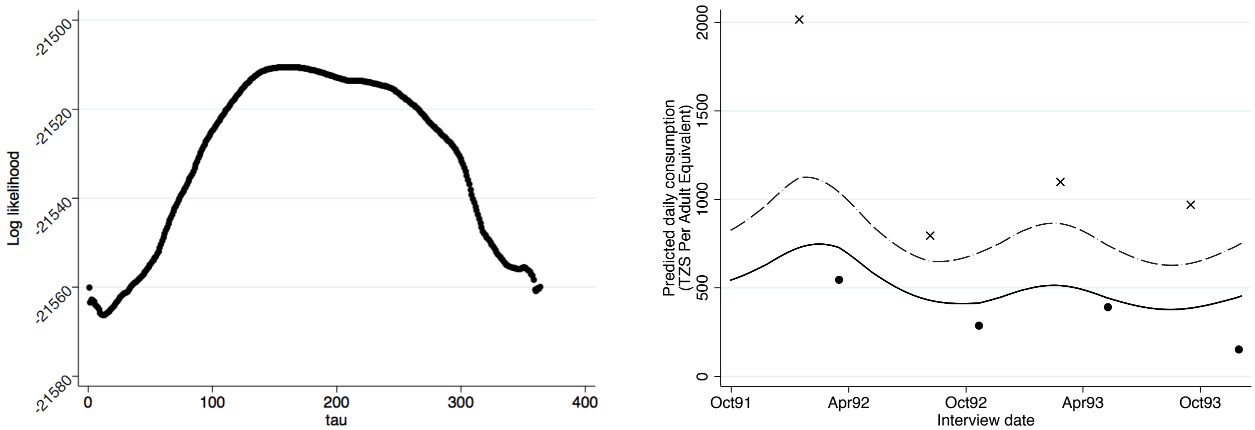


Figure 1: Histogram of survey dates and time path of daily consumption estimates.
Source: Authors' calculations from KHDS 1 data.



A. Log likelihood of (5) plotted against $\hat{\tau}$

B. Predicted and observed food expenditure

Figure 2: Consumption model: estimation results

Source: Authors' calculations from KHDS 1 data. In Panel B, $\times$s and dots show expenditure measured by the survey for two sample households; the lines show predicted daily consumption. The $\times$s are associated with the dashed line, the dots with the solid line.

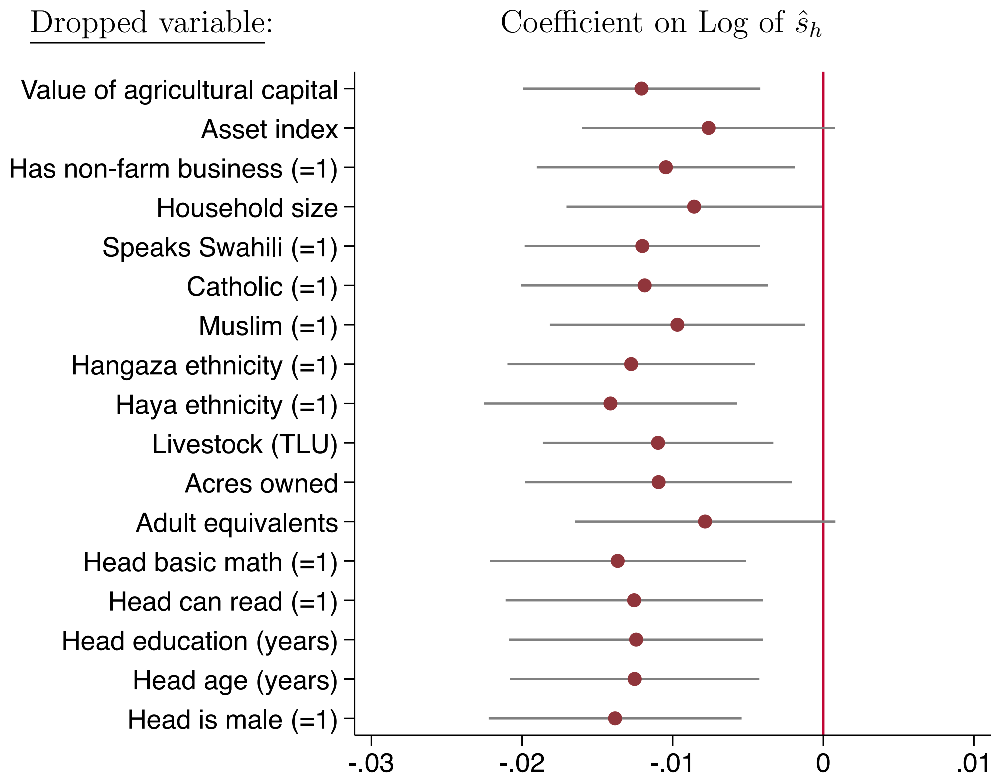


Figure 3: Estimates of $\hat{\beta}_2$ in equation (8), with height as dependent variable

Source: Authors' calculations from KHDS data. Figure shows the point estimates and 90% confidence intervals for the coefficient on Log of $\hat{s}$ for children in age group 1 (-2 to 2 years of age in 1991). Standard errors clustered at the level of the 1991-1994 village.

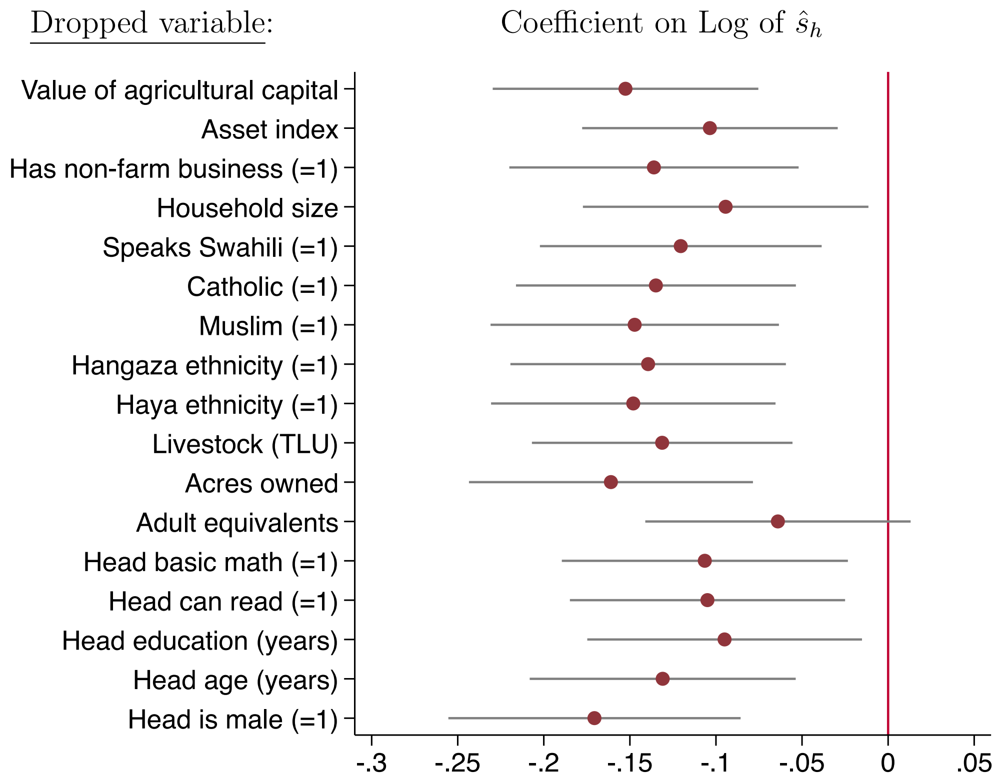


Figure 4: Estimates of $\hat{\beta}_2$ in equation (8), with years of education as dependent variable

Source: Authors' calculations from KHDS data. Figure shows the point estimates and 90% confidence intervals for the coefficient on Log of $\hat{s}$ for children in age group 3 (11 to 17 years of age in 1991). Standard errors clustered at the level of the 1991-1994 village.