How responsive is body weight to transitory income changes? Evidence from rural Tanzania

Niklas Bengtsson*

First draft, Nov 20 2006. This draft, June 1, 2007.

Abstract

We use time-series of rainfall along with individual fixed effects to estimate the response of body weight to transitory changes in household income. Our data consists of a longitudinal sample of subsistence farmers in rural Tazania, representing one of the poorest populations in the world. We find that the response is positive for all household members, but highest among female children. For female children (Age 10), a ten-percent increase in household expenditure induces an increase in body weight with about 1.2 percent, about 0.2 kilo. The results suggest that malnutrition to a non-trivial extent is subject to binding income constraints in this region.

Work in progress - comments welcome!

*Department of Economics, Uppsala University. niklas.bengtsson@nek.uu.se

1 Introduction

Throughout rural Africa, household expenditure is very sensitive to weather-induced changes in rural production and income. In this paper, we ask to what extent such fluctuations are reflected in the physical capacity of the population. Is nutritional status sensitive to transitory changes in household expenditure? Are female children particularly vulnerable?

Although widespread malnutrition typically occurs after series of weather failure, it has been recognized for some time that malnutrition is not simply determined by the regional supply of food. In many sub-Saharan countries, endemic malnutrition is present also during normal periods, when the total amount of food production plus imports is enough to feed everyone. Amartya Sen was among the first to note that severe outbreaks of malnutrition not seldom co-exist with regional prosperity, along with little or no decrease in the total level of regional food production (Sen, 1981). Rather than simply being a function of regional food availability, Sen argues, hunger catastrophes occur when the household's purchasing power with respect to food (or "food entitlement") decreases.

Sen's entitlement approach has had a large influence on development policy, with clear policy guidance: if you want to reduce malnutrition, you need to reduce poverty. However, a second wave of literature has come to question weather poverty is the main, or even most important, cause of malnutrition. This literature builds on two empirical observations. A first one is that regional malnutrition seems to occur also when the total cost of the calories needed to assure "light physical capacity" – FAO's criterion for nutritional well-being – is so small that even the poorest households should be able to secure an adequate level of nutrient intake. Secondly, estimates of the demand for calories and other nutrients suggest that households seem unwilling to channel positive changes in income towards calories. Although the income elasticity of food is typically close to unity in developing countries, the income elasticity of calories seems to be closer to zero, even among households that would not be considered well-nourished by international standards. In fact, a handful of papers suggest that households' demand for calories and nutrients is completely irresponsive to income changes (Behrman and Deolalikar, 1987, 1990; Bouis and Haddad, 1994; a recent paper is Aromolaran, 2004). Behrman and Dealolikar (1989) propose an economic rationale for this finding, namely that households have a strong "taste for variety" or more luxury calories, even at the lowest levels of income.

This so-called "revisionist" literature has been contested (see Subramanian and Deaton, 1996), but the question whether the nutritional status of individuals in sub-Saharan Africa will improve as income increases is still relevant for development policy. If the demand for nutrients and calories is largely invariant to income changes, policymakers mostly concerned with improving nutrition may find "standard" economic reforms – trade liberalization, micro-credits, income tax policies, etc. – less relevant. On the other hand, if malnutrition is subject to binding income constraints, economic reforms will be aligned with more health-oriented policy initiatives. Today, the former attutude is probably the most prevalent in policy documents. In a paper in the World Bank Policy Review, Haddad et al. (2003) concluded that child malnutrition will not be eradicated by economic growth alone, if it is not accompanied by reforms aimed at improving nutrition directly.

In this paper, we test whether large transitory swings in household income have no effect on nutritional outcome. Our data consist of a longitudinal sample of subsistence farmers in rural Tanzania, collected by the World Bank. The Kagera Health and Development Survey (KHDS) dataset represents one of the poorest populations in the world and one whose nutritional status, defined as BMI and weight-for-age, is below international standards. Rather than focusing on nutrient intake, such as the quantity of calories consumed, we use body weight as a proxy for nutritional outcome, following the praxis of Haddad et al. (2003) and Strauss and Thomas (1998). This approach allows us to compare the response of nutritional outcome to expenditure across family members. In order to capture the transitory component in household expenditure, we use rainfall as an instrument along with individual fixed effects. As most of the households in our sample rely on crop yields as the main source of income, rainfall can explain a non-trivial share of the intertemporal variation in household expenditure.

An instrumental variable strategy should improve the estimates obtained by ordinary least squares for three main reasons. First, a typical problem in survey-based econometric exercises is the attenuation bias stemming from measurement error; people are unlikely to be able to recall household expenditure or income for the last six months perfectly. We argue that a rainfall-induced variation in household income will be less tainted by measurement error. The second objective is to evade the simultaneity between nutritional status and income. The reciprocal relationship between nutrition and income has been a central feature in some important work in development microeconomics (following the the work by Harvey Leibenstein in 1957). Third, rainfall fluctuations arguably captures a transitory and exogenous component in household's income and expenditure, uncorrelated to life-cycle decisions, knowledge or other variables that may enter the households' preferences over nutrition.[1]

Despite the large body of research on malnutrition – and the popularity of the "weather-instrument" – there is only a handful papers that studies the impact of transitory income shocks on nutritional outcomes.[2] This literature is mostly concerned with child nutrition. Hoddinott and Kinsey (2001) and Alderman et al (2005) study child growth in Zimbabwe and Ethiopia, respectively, and report that children born during drought experienced a slowdown in human growth. Rose (1999) studies the interaction between gender, "favorable weather shocks" during early childhood and the probability of survival, and finds that the excessive female child mortality is increased during periods of adverse weather shocks. Foster (1995) compares the impact of flooding on child weight in Bangladesh across land-owners and landless households, in order to test for liquidity constraints, and finds that the variation in child weight was more pronounced among landless households. Dercon and Krishnan (2000) use nutritional status for adult males and females in order to study risk-sharing within households in Ethiopia, and report assymetric intrahousehold responses to various "shock"-variables.

We add to this literature, and contribute in some important methodological respects. First, we use a time-varying instrument along with individual fixed effects. By doing so we can disentangle the random element in year-to-year rainfall fluctuations, keeping the mean levels of rainfall and expenditure constant.[3] Second, earlier estimates of the causal impact of income shocks have typically been based on reduced-form equations – i.e., studied the impact of the shock-variable in itself (the exception is Foster, 1995). IV-estimates are, however, more useful for comparisons across datasets and samples, and will provide

[1] In a related study on the same dataset, Alderman et al. (2006) study the effect of long-term economic growth and policy interventions on child health outcomes in a random effects model. Assuming away omitted variable bias and simultaneity, they use (self-reported) income and roof quality as an instruments for household consumption. We argue that such instruments are unlikely to be helpful even to remedy measurement error bias, since, as pointed out by the same authors, self-reported income figures are typically more unreliable than consumption data.

[2] More broadly related papers have used weather variation in order to test hypothesis about the income-consumption nexus. This strand of research includes Wolpin (1982), Jacoby and Skoufias (1998), Paxson (1992) and Duflo and Udry (2004). Kochar (1998) studies whether weather-induced income shocks increases labor supply. Miguel (2004) uses rainfall variation to estimate the causal impact of economic shocks on ethnic conflict.

[3] None of the abovementioned papers are able to control for unobservable individual-specific effects. Of the paper that uses observed calorie-intake, however, Behrman and Deolalikar (1990) uses a fixed-effects approach.

us with an economically relevant measure of the severeness of income fluctuations. Such estimates are becoming more and more important as empirical contributions in development economics tend to focus on the internal validity of their estimates (as opposed to producing externally commensurable estimates). A third contribution is that we will not exclusively focus on child weight, but compare our estimates across household members. Despite the theoretical and documented importance of intrahoueshold allocation of calories and nutrients, direct evidence on intrahousehold mechanisms is quite rare. Fourth, the completeness of our dataset allow us to control for incidences of malaria and other diseases that are known to be triggered by rainfall in order to study the stability of our IV-estimates.

The paper is organized as follows. In Section 2 we outline a basic framework for studying the response of rainfall-induced changes in household income. We put special focus on the measurement of income (and expenditure) when rainfall is thought to enter the budget constraints. In Section 3, we address some empirical considerations regarding the definition of rainfall. The data and key variables are presented in Section 4. In Section 5 we present our results. We find that the response of body weight to transitory changes in household income is positive for all family members, but that the elasticity decreases with age. For male adults, the elasticity is not statistically different from zero. For female children, the response of body weight to transitory income changes is markedly higher. We cannot distinguish between the responsiveness of individual health investments and the (perhaps biological) ability to transform these health inputs into body weight. Nevertheless, insofar policymakers are largely concerned with health *outcomes* rather than health *inputs*, we argue that these estimates can be quite informative for policy (a case we further develop in Section 2.2). Section 6 concludes.

2 Conceptual framework

2.1 Using rainfall as an instrument for household expenditure

The aim of this paper is estimate the response of human body weight to transitory income changes, using weather variation as an instrument for household income. Our identification strategy rests on the assumption that rainfall affects consumption outcomes, including leisure, only via the income variable and not via some omitted variable. The usage of self-reported income or observed production will not necessarily honor this assumption. As argued by Morduch (1995) – and, to some extent, verified by Kochar (1999) – an adverse shock to rural production may induce liquidity constrained households to supply more labor by decreasing leisure. A potential scenario is that households are able to completely smooth output in the event of an weather shock by adjusting their consumption of leisure. This type of "income smoothing" suggest that a realized weather shock can affect the consumption of leisure, and therefore nutritional status, without affecting observed the level of rural production.

As noted by Rosenzweig and Wolpin (2000), the credence of using weather variation as an instrument for rural income rests on how the market structure is defined, and on how expenditure and income is observed. Under some standard assumptions in this rural context, a shock to rural production will enter additively in the budget constraints if all expenditure posts, including leisure, can be observed and used to proxy the income argument in a demand equation. Only if leisure does not enter household utility, or if the labor market is completely restricted, can survey data on self-reported income or rural production be used as explanatory variables in an instrumental variable approach.

To formalize these ideas, consider first a one-member household that extracts income from farm profits and from market labor.[4] For our baseline model, we assume that the household faces binding liquidity constraints. Whether households in rural Africa are unable to save or borrow across aggregated income shocks is of course an empirical question, but the received wisdom is that rural households in the developing world are unable to achieve at least perfect intertemporal consumption smoothing (see the surveys by Townsend, 1995 and Morduch, 1995). As it turns out, rainfall has a decisive impact on expenditure in our sample, implying that households in Kagera are unable to borrow and save across transitory income shocks. If the household is completely liquidity constrained, the utility-maximization problem can be analyzed in a static, one-period setting. The household faces the following problem:

$$\max_{c,l} U = u\left[c, l, n(c,l)\right], \tag{1}$$

subject to

[4] The model can be generalized into a multi-member quite easily. Since our conceptual framework does not aim to present a testable prediction of intrahousehold behaviour – such as the income pooling hypothesis – we keep the framework simple. See Duflo and Udry (2004) and Dercon and Krishnan (2000) for extensions in that direction.

$$I = p_y F(L, r) - wL + wT = pc + wl \tag{2}$$

$$T = l + L^f + L^m \tag{3}$$

$$L = L^h + L^f, \tag{4}$$

where $\mathbf{c}$ is a vector of consumption goods, l is leisure, and n is the individual's "body weight function", determined by consumption and leisure. The household's utility thus depends on both the inputs in themselves and the way these inputs are transformed into nutritional well-being. L is total on-the-farm labor, T is the household's time endowment, L^f is the household's supply of own-farm labor, L^m is the household's supply of market labor and L^h is hired labor from outside the household; p_c, p_y and w are the prices of the consumption good, the production good and labor, respectively.

Looking at Equation 2, we see that although the realized rainfall shock r enters the production function $F(L, r)$, so does the household's labor inputs, which, in turn, enter implicitly as leisure in the utility function. Decisions on leisure will affect rural production, which in turn may affect the demand for leisure, in ways that at this point are not theoretically clear.

How shall we define income in order for a weather shock to produce a shift of the budget constraint? The correct definition of our explanatory variable depends on how we characterize the labor market structure. A first solution to the problem of identification is to assume that the labor market is "complete", and noting that the restrictions above suggest that own-supplied farm labor and hired work are perfect substitutes. In this case, Problem 1 becomes separable, and the household will maximize profits first, and then maximize utility. The budget restriction becomes

$$I = \pi^*(w, p_y, r) + wT = p_c c + wl \tag{5}$$

where π^* is maximum profits. Since leisure is typically unobservable, the income argument can not be identified using the right-hand-side of Equation 5. To the extent that rural profits (π^*) can be observed (this is not certain given that own-supplied farm labor is typically not paid for), the left-hand-side can be used to identify a shift of the budget constraint as long as wT is controlled for.

Assuming that the labor market is complete (and, consequently, that there is no unemployement) seems difficult to reconcile with the stylized facts of most developing countries. In their textbook in development microeconomics, Bardhan and Udry (1999) suggest that empirical work is best guided by more realistic assumptions. In a related paper, Duflo and Udry (2004) instead assume that leisure does not enter preferences. In this case, the problem is

$$\max_c U = u\left[c, n(c)\right]$$

subject to:

$$p_y F(L, r) - wL = p_c c, \tag{6}$$

Since there is no disutility associated with supplying labor, the problem is still separable, yielding budget constraint:

$$I = \pi^*(w, p_y, r) = p_c c$$

Under this second approach, "crude" survey measures of household expenditure will fully capture changes in household income. Again, this is based on the assumption that relative price are kept constant. Note further that under this second approach, estimates of the income elasticity of body weight based on profits should be equal to estimates based on total expenditure.

Survey data on goods expenditure is considered less noisy than self-reported rural profit. This is the case in the KHDS dataset we have available as well. In order to align our work with existing papers (most notably, the numerous estimates on the "expenditure elasticity of calories"), and to secure efficiency, we will use the traditional survey measure of total household expenditure for our baseline estimates. But it is important to note that even the absence of saving, household income measured from the expenditure side can not be used to identify a wather-induced shift of the budget constraint without some (nontrivial) assumptions about the supply of labor. The problem is that leisure is unobservable. As we will argue in the next subsection, leisure can be considered an important determinant of human calorie requirements, and we would like to compare our estimates with those that take leisure into account. For these reasons, an explanatory variable based on household profits will be used for a robustness check.[5]

[5] As it turns out, our results are to a small but non-neglible degree dependent on how we

Denoting the income argument I, the first-order conditions for utility optimization will yield the household's demand for a particular good j:

$$c_j^d = d_i(I, p_j, \mathbf{p}),$$

where $\mathbf{p}$ is a vector of the prices of other goods. The derivative of the household member's individual status with respect to household income is:

$$\frac{\partial n}{\partial I} = \sum_{j=1}^{n} \frac{\partial n}{\partial c_j^d} \frac{\partial c_j^d}{\partial I}. \tag{7}$$

2.2 The demand for nutrients

The essence of the "revisionist" critique to the analysis of malnutrition is that the costs of improving nutrition, in terms of energy intake, are easily born by even the poorest households. A first look at our sample suggests that this notion is not without support. Looking at Figures 1 and 2, and Table 4 in Appendix, about 20 % of our adult sample and 25 % of those younger than 18 years old would be considered malnourished by WHO-standards.[6] As seen, the bulk of the sample is below the standard reference value of normal weight, and not even those and the ninetieth percentile of the BMI distribution in this sample would be considered overweight. Further comparison (in Table 4 in Appendix) reveals that female children are actually somewhat better nourished than boys, although the difference is quite small.

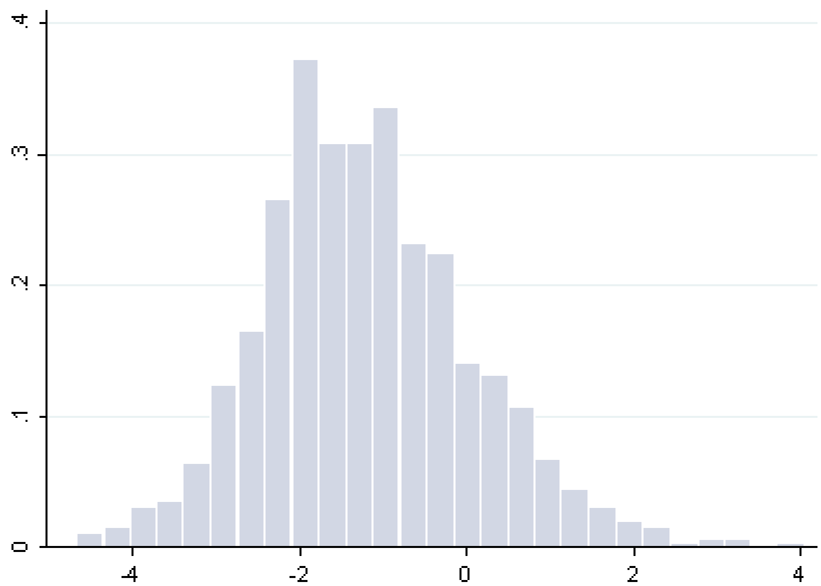
Are these figures the result of poverty? FAO (2001) estimates that the daily energy requirement needed for an "active" man of 65 kilo is about 3000 kcal; the corresponding figures are 2 500 for females and about 1 500 for children.[7] A simple (albeit rough) calculation reveals that if the household was to consume only cassava – a cheap and drought resistant root-vegetable, used as a staple crop in much of rural Africa – the cost of fulfilling the yearly energy requirement for a five-person family would be 9 125 Tnz per household member. This implies that a household at the median income level in our sample could meet the energy requirements using one sixth of the household budget; for the lowest quartile

define the explanatory income variable.

[6] For adults, a body mass index below 18.5 is a common definition of underweight; 25 is the threshold for overweight. For children (below 18 years), an analogous indicator is the normalized weight-for-age z-score. A child with a corresponding z-score below -2 is typically considered malnourished (see WHO, 1995).

[7] The FAO figures are among the most influential, but have been criticized for being to large. For our purposes, they serve well as upper bound-estimates.

of the income distribution, the corresponding figure is a little more than one fourth. This pattern is typical, and has led Behrman and Dealolikar (1989) to hypothesize that people demand luxury goods also at the very lowest levels of income.

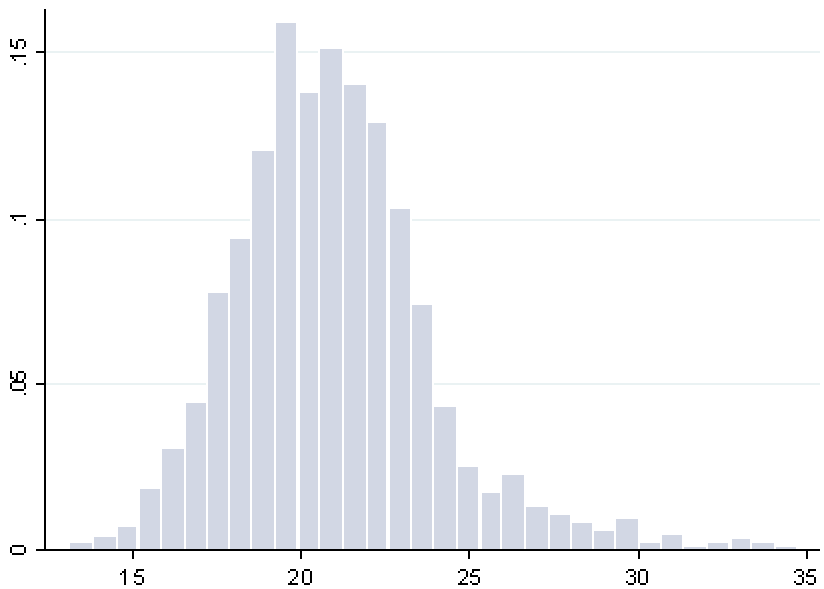


Weight-for-Age z-score distribution the KHDS sample (first wave). Age < 18.

The traditional way of testing this hypothesis has been to estimate the expenditure elasticity of calories, sometimes along with other health-improving nutrients. Data on calorie-intake have either been obtained by converting consumption data into calorie intake by using tables of conversion ratios, or by letting nutritionists observe nutrient intake directly.[8] The former approach, which relies on the assumption that all available calories are consumed by the household members, have yielded a large number of estimates, typically within the range of 0.2 to 1 (see Deaton, 1997, and the references therein). The handful

[8]Behrman and Dealolikar (1990), for example, asked the most knowledgeable person to serve the typical amount of food given to each family member on different plates, and gauged the individual calorie intake on that basis.

of papers that uses the latter approach has reported estimates that are closer to zero (Behrman and Dealolikar, 1987, 1990; Bouis and Haddad, 1992).



BMI distribution in the KHDS sample (first wave). Age > 17.

Even with this latter approach, however, the variable calorie-intake is unlikely free from measurement error. Strauss and Thomas (1999) argue that recollections of past consumption are seldom perfect, and that the respondent's inability or unwillingness to recollect can be correlated to other household characteristics. Also, by a similar reasoning, data on expenditure is likely to be plagued by measurement error as well. In the classical zero-mean case of measurement error in the independent variable, elasticity estimates will typically be "attenuated"; biased towards zero.[9]

Our reason to focus on body weight has, however, not so much to do with the fact that body weight is a more accessible statistic than nutrient-intake. Rather, what we have in mind is that the success or failure of development

[9] Of course, the rationale for estimate the *expenditure* elasticity of calories is that expenditure is considered more reliable than income in household surveys.

policy often is evaluated with a reference to health *outcomes*, not health *inputs*. The by far most utilized measure of health outcome in the developing world is a person's anthropometric status, promoted by, among others, the *World Health Organization* (WHO) and the *Food and Agriculture Organization* (FAO) of the United Nations. Anthropometric status has been recommended for econometric practice as well (Strauss and Thomas, 1998; Haddad et al, 2003). If one think of body weight as the output of a health-production function, and calories and nutrients as inputs, anthropometrics can arguably be more informative for policy in many settings – just like GDP is often (but not always) more interesting than factor endowments.

In order to illustrate the contrast between elasticity measures from the input side and from the output side, consider Equation 7. As seen, the income elasticity of body weight is determined by the sum of the income elasticity of the inputs. Since we impose no restrictions on the signs of the inputs, it is possible that some inputs are Engel goods while others are "normal" in the neoclassical sense. If, for example, leisure is a luxury good and crude calories are necessities, it is possible that the income elasticity of calories is zero or even negative, but that the expenditure elasticity of body weight is positive. Only if all calories are perfect complements can isolated estimates of the expenditure elasticity of a single input be informative for nutritional outcome. There is, however, good reason to believe that certain inputs, like leisure and calories, are substitutes. If a person increases his or her energy expenditure (i.e. physical activity) he or she will typically compensate by consuming more calories. If nutritional outcome is our primary interest, the expenditure elasticity of body weight thus has the advantage of capturing the aggregated effect of all inputs, also those that are not readily observable, such as labor supply. This is the reason why we are so anxious to allow for leisure to vary with income.

On one hand, the economics of nutrition has substantiated a concern that the millennium development goals of eradicating malnutrition might be difficult to achieve if nutrition is essentially invariant to income changes. On the other hand, that children are particularly vulnerable to drought and other shocks to regional food production throughout sub-Saharan Africa and elsewhere is verified practically every season. We therefore argue that evidence on, and magnitudes of, individual nutritional responses to exogenous shocks to household expenditure can be quite useful for outlining development policy in this area.

3 Empirical implementation

Our baseline regression equation is:

$$\ln weight_{iht} = h_{hi} + \beta_1 \ln(I_{ht}) + \mathbf{X}\boldsymbol{\beta_2} + e_{iht} \tag{8}$$

where *weight* is the body weight of individual i residing in household h at time t. The error-term e_{ih} comprises both the household-level error term and the within-household error term. There is no real reason for using the logarithm of body weight, except that elasticity-interpretations will become more straightforward. The time-constant individual controls h_{hi} are introduced non-parametrically.

In the analysis in Section 2.1, we treated prices as given. General equilibrium considerations do however suggest that relative prices will not remain unchanged if weather shocks determine rural output. If the relative prices change, decisions on consumption may be influenced by the weather shock even in the absence of a shift of the budget constraint. However, assuming that local markets are integrated implies that all households will face the same relative prices. The vector $\mathbf{X}$ contains year-district interaction dummies and seasonal dummies, which are included to control for prices for this very reason.[10]

Based on the considerations in Section 2.1, I_{ht} is defined as household expenditure (based on six-month recollections of the household's "most knowledgeable person"). For our robustness analysis, we will base our income argument on profits (we will also study the response body weight to food expenditure).[11]

The first step equation is:

$$\ln I_{ht} = \gamma_1 \ln(r_{ht-1}) + \mathbf{X}\boldsymbol{\gamma_2} + u_{ht} \tag{9}$$

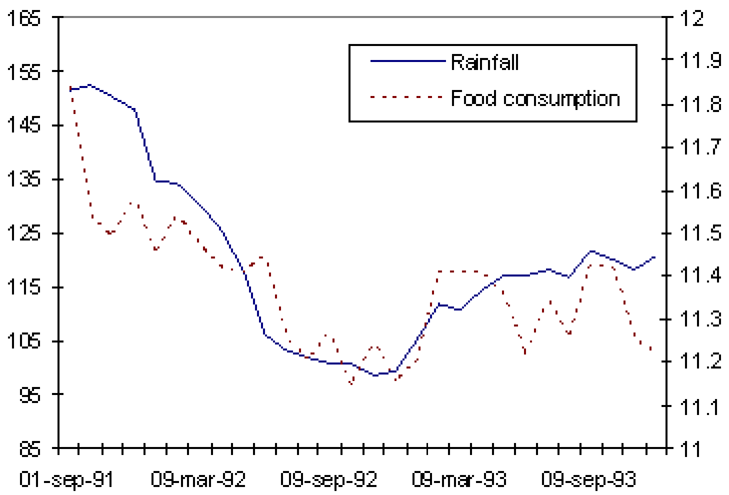
In Equation 9, $\ln(r_{t-1})$ is past rainfall, to be defined shortly. With fixed-effects, the coefficient γ_1 is to be interpreted as the percentage deviation from normal rainfall for individual i.

[10] We will test whether information on local price variation is redundant using these controls.
[11] Ackording to Equation (5), the regression equation is (assuming a linear model):

$$\begin{aligned} weight_{iht} &= h_{hi} + T_{hi} + \beta_1 I_{ht} + \beta_1 w + \mathbf{X}\boldsymbol{\beta_2} + e_{iht} = \\ &= h_{hi} + \beta_1 (I_{ht}) + \mathbf{X}\boldsymbol{\beta_2} + e_{iht} \end{aligned}$$

Notice that T_{hi} is assumed to be time-invariant, and that the wage level will be observed on the aggregate level, using time-district interaction dummies. To ease comparison, we will use a log-log variety for estimation.

Around the equator, perturbation levels are characterized by rather sharp seasonality. In the Kagera region rainfall follows a bimodal pattern, with two rain periods per year. The timing of the rain periods is however erratic, and when they come they vary in intensity. The *Kagera Health and Development Survey* datasets include monthly rainfall from 1989 collected from weather stations in the five districts of Bukoba, Karawge, Muleba, Biharamulu and Ngara. The time-series in Graph 1 represents a regional average. The years 1992 and 1993 were periods of lower rainfall levels. Comparison with longer time-series of rainfall and reports from FAO indicates that the years 1992 and 1993 were plagued by food shortages, although not as severe as the Sahelian drought in the mid-eighties or earlier dry spells. The peak in 1991 represents a "normal" year; 1992 a bad year and 1993 slightly better, but still below average. The link between (the logarithm of) food expenditure and our constructed rainfall variable is depicted in Figure 1. For some months, only a handful of households was surveyed, which can explain part of the volatility in household expenditure (since the rainfall series is a moving average, it is more smooth). The time-series of rainfall in Figure 1 represents a regional average across the five different weather stations.



Rainfall (12-month moving average) and food expenditure (last six months) in Kagera, Tanzania. Left axis is rainfall in millimeter, right axis is log of food expenditure.

Hoddinott and Kingsley (2001) and Paxson (1992) both take some care in identifying the correct sowing season and harvest seasons in order to strengthen the link between rainfall and income. Our first idea of instrumental variable was a monthly average of the last six months, hoping to extract as much variation in expenditure as possible due to the bimodal cropping regime. There are, however, several problems attached to this approach. First, since the timing of waves was not perfectly semi-annual but sometimes spanned more than nine months, the variance in household expenditure induced by rainfall might be confounded by a seasonal component. Second, cultivation techniques are somewhat dispersed, and not all households in our sample exploits the bimodal cropping regime (the traditional tree crops, like bananas, are usually harvested only once a year). This implies that the impact of a six-month instrument could be different in different periods (maybe even have the opposite sign), also for the very same household. Third, a typical observation is that consecutive weather failure is particularly severe for rural households, and longer time-series of rainfall statistics should be a stronger instrument. We chose to collapse the seasonal variation in perturbation levels, and use a twelve-month moving average as instrument.

The intertemporal variation in rainfall amounts to four points in time. The within-wave variation in rainfall has two sources. The first is the interregional variation across the five weather stations from which the rainfall data was collected. The second source of variation is cluster-based. For logistic reasons, the timing of the interviews often differed across clusters, sometimes with as much as six months. Under the assumption that the timing of the survey staff visit is uncorrelated to rainfall, after controlling for fixed effects, year and season, we can exploit this "accidental" variation within waves. For our baseline estimates, we associated the rainfall shock with the *month immediately preceding the date of the interview*. The reason not to lag the rainfall shock further and associate it with some time before the recollection period is basically that I_{ht} is defined as current income. Crop yields from tree plants (including the economically important banana tree) are determined by rainfall until the very month of harvest. If this date coincides with the recollection period, our instrument might not capture more recent consumption, especially if our assumption of binding liquidity constraints is true. This can be quite severe, since current body weight is likely to be influenced more by recent consumption than past consumption – in fact, the last week of consumption may practically determine most of the intertemporal variation in body weight. Our approach is thus to see recollected expenditure as proxy for current income, and use rainfall up to the date of the interview to correct for any measurement error the usage of such a proxy entails.

Since our instrumented variation in rainfall is based on survey-timing, our trend variables (the year and season dummies) are based on the actual date of the interview, and not on the specific wave. There were, however, some wave-specific changes to the questionnaire. In particular, in the first wave, the most knowledgeable person in the household was asked to recollect consumption for the last twelve months; in the consecutive waves, the recollection period was six months. We standardized the twelve month figure by dividing it by two. However, under the assumption that the exclusion restriction holds, this operation is inessential for our regression results, since our rainfall instrument will take care of this measurement issue. In Section 5.1 and 5.2, we explore the issue further by controlling for wave-specific questionnaire dummies. We do not find statistical support for including wave-specific dummies (they are highly correlated to the year and seasonal dummies, and we end up with a severe case of multicollinearity).

As for standard errors, the month of the interview was typically cluster-specific, implying that our instrument is measured on the cluster-level, and, consequently, that our predicted expenditure is likely to be correlated within clusters. All reported standard errors are therefore corrected for arbitrary correlation within clusters (the "cluster" is also the main stratum of selection in the KHDS dataset, see Section 4.1 and World Bank 2004). The asymptotic properties of clustered sampling have recently been subject to some interesting research (see Wooldridge 2003a; 2006; Donald and Lang, 2006), and simulation studies suggests that if the number of clusters is "large", cluster-adjusted standard errors perform well in fixed-effects analyses when an explanatory variable is a clustered variable. In our case, there were 52 clusters, which, according to this recent strand of research, is an acceptable group size. We test this hypothesis.

We now turn to the subtle issue of our exclusion restriction (i.e. that $cov(r_{ct}, e_{iht}) = 0$). If this assumption does not to hold, our IV estimates will be inconsistent. It is therefore crucial that we can maintain the assumption that rainfall affects body weight only via the expenditure channel. We have already dealt with some economic issues in Section 2.1, and, as mentioned, using observed expenditure in Equation 8 rests on the assumption that leisure is supplied inelastically, or at least that it does not enter preferences. We will relax this assumption in a robustness analysis. There are, however, some additional, non-economic, issues, that may violate the exclusion restriction regardless of how we measure income. One particular concern is that climactic factors may induce the spread of certain diseases, which, in turn, may reduce body weight. Although well-documented in epidemiological literature, this fact has been given less attention in related economic exercises (but see the discussions in Thomas and Strauss 1999 and Foster 1995). In the region of Kagera, both malaria and cholera epidemics have been triggered by rainfall in this fashion, so these concerns clearly have some merit in this context. Fortunately, the KHDS dataset is sufficiently rich for us to control for incidences of malaria and other illnesses in a stability analysis (the survey included questions of both self-reported and diagnosed health status). Under the assumption that such incidences represent random shocks in our IV-model, the baseline estimates of the elasticity should be robust to the inclusion of these variables.

Another concern is that body weight is not constant across time. Autocorrelation in our dependent variable would be no problem if our time-series of rainfall would be completely stationary, but as seen in Figure 1, there seems to be autocorrelation in rainfall as well (by construction, this is so because we use a moving average). Note first, however, that if our panel was completely balanced and every individual in the same district had the same growth rate, our year-district interaction dummies would root out general growth trends. However, human growth is not constant throughout the life cycle, and age-specific auto-correlation in body weight will be magnified if there is attrition, since household members that are observed less frequently would have had more time to "naturally" increase or reduce weight. We correct for this by including the *time since last measurement*-variable, interacted with *age at intital wave*, and the same variable squared. The variation in *time since last measurement* arises from (individual-specific) attrition and the fact that the spacing between interviewes differed – when using the unbalanced panel, this variable is individual specific. An alternative and perhaps more straightforward way to control for human growth would simply be to include *height* as an explanatory variable. However, since child growth can be retarded by malnutrition (see Hoddinott and Kinsey, 2001; Alderman et al. 2005), *height* can be viewed as a proxy for our dependent variable and therefore unsuitable as an explanatory variable. We will test whether such considerations have merit in our context, but we exclude *height* from our baseline estimates.

4 Data

4.1 Sample description

The Kagera Health and Development Survey was conducted in four rounds between 1991 and 1994 in the Kagera region, northwestern Tanzania.[12] The sample is fully longitudinal in the sense that all households that were interviewed the first period was interviewed in subsequent rounds (it was not a rotating panel, common in similar datasets in Africa). The survey staff visited each household four times between 1991 and 1994, in intervals of between six and nine months. In total, around 820 households were surveyed. 4895 individuals was observed at least twice, and the full individual sample size used for the panel analysis (N*T) was 16 640. Practically all households were engaged in subsistence farming to some degree, and as can be seen in Table 4 (in Appendix), more than three fourths produced more than 50% of their food consumption themselves.

The sample was stratified in order to capture incidences of illness, accord-

[12] A fifth round was conducted ten years afterwards, in 2004. In this paper, we will use the first four rounds.

ing to a two-step variable probability selection. The selection was based on "mortality-risk" at both the community and household level (the survey objective was to study the impact of HIV/AIDS on local economies in Africa). In the first step, clusters of households were randomly selected from predetermined PSUs corresponding to different agronomic zones. The probability that the clusters was "kept" was proportional to the level of mortality reported in the 1988 Tanzanian Census. The second stage kept individual households in a similar fashion. With stratified samples, there is always the question whether or not to use sample weights in order to infer the results to the population. Using weights comes at the price of larger standard errors, and in many settings, add little to consistency. In the model that we have derived and presented in Equation 9, the population parameter β is assumed to be identical in each population stratum. If this assumption holds, then the unweighted IV estimator will be consistent, and since 2SLS is most efficient, it is preferable. The crux is that if this assumption does not hold, then the weighted estimator will not perform any better since it, too, will be inconsistent. As pointed out by Deaton (1999), the heterogeneity lies in the population, and if we wish to estimate a behavioral model that is different for different parts of the population, neither estimator is consistent and "weighting is at best useless". The perhaps most reasonable advice is that given by DuMouchel and Duncan (1994): stick with the unweighted estimates if you cannot reject that they equal the weighted ones. Since this was the case here, we will base our results on unweighted two-step least square estimates.[13]

There was moderate attrition. About 77% of the population was available for anthropometric records in at least two periods, and only a little more than half the sample was available for all periods. As is the case in all survey-based analyses of health and mortality, it is plausible that this attrition is associated with our dependent variable. For example, insufficient calorie intake may induce illness or even death, and therefore absence. If this is the case, $\widehat{\beta}_1$ will be biased towards zero. It is fairly easy to come up with other potential correlations between absence and body weight (due to migration or work) that will distort our estimates. In Section 5.2, we will address these concerns by restricting the sample to individuals that was present at all survey round, under the null that if attrition is unimportant, the restricted and unrestricted estimates will not differ.

[13] Throughout the analysis, however, we use cluster-robust standard errors (clustered at the "cluster"-level; the main stratification unit and the level at which rainfall is observed).

Finally, there is the issue of outliers in our dependent variable. The measurement of body weight differed somewhat across cohorts and waves. Infants under the age of two were weighted using hanging Salter scales; for older cohorts, standard scales were used. In the third wave, the adult scale was replaced by a digital one, which, according to the KHDS technical report, reduced the variance of young children's weight. With individual fixed effects, our dependent variable is to be understood as the individual's body weight deviation from his or hers mean. About 90 individuals had a very high relative standard deviation in body weight across time, and about a handful had extremely high. It is reasonable to ask whether such outliers, be they the result of measurement error or not, are driving our results. To answer this question, we will compare our estimates across various sample restrictions. We will, however, use the sample "as is" for our baseline estimates, throwing away nothing but missing values.

4.2 Variable description

Our explanatory variable of interest is total household income. Our preferred estimates are base this variable on total household expenditure. The "most knowledgeable" household member was asked for total household expenditure the last six months. Since most of the crop produced was not sold but consumed, the most knowledgeable person was asked to estimate the value of own-produced consumption directly. The survey staff took some care in making this measure comparable by adjusting it to seasonal price variation, but there is scope for some mismeasurement here, if crop losses are exacerbated, for example. We show first-step results acriss household expenditure, agricultural profits and food expenditure.

In Table 4 (in appendix), we present common statistics for expenditure, household composition and anthropometrics. We also present the degree of subsistence (consumption of own-produced goods divided by total consumption) and the share of food consumption (food consumption/total consumption), and a dollar-converted measure of household expenditure. Since the dollar-shilling exchange rate fluctuated substantially in this period, conversion to dollar as of 1991 is a bit tricky. Rough calculation (using an exchange rate of 298.5 shilling/dollar; the mean exchange rate as of 1990 and 1991) suggests that median per capita expenditure in 1991 current dollar prices was about 180 $ per capita (which is similar to the IMF prices on per capita income in 1991). The budget share of food expenditure is similar to the much-cited figures in Marshall's *Principles of Economics*; about 60 percent of the total household budget is devoted to food expenditure. A little less than half the households' expenditure is produced at home, and about 75% of all food expenditure is produced at home.

Except for the year-district interaction dummies and the "time since last measurement"-variable (*time*), we will in various specifications include variables that can be thought to influence both income and body weight, as a way of checking the robustness of our results. The two indicators of illness (*malaria* and *other illness*), which can be influenced by rainfall directly and violate our exclusion restriction, are given extra attention. We also include *height* in supplementary equations as a way of capturing the autocorrelated elements in body weight. Furthermore, we will in various specifications include *schooling* ("did you attend school last week?"), *work* ("did you do any own-farm labor last week?"), *number of children in household, number of adults in household* and *resided elsewhere* ("have you resided elsewhere sometime during the last six months?"). Note that all of these variables should be excluded from our benchmark regression in order to estimate the net expenditure effect on body weight. Additional control covariates that will be used are *pregnancy*, which is known to be influenced by season, and *wave* (to control for wave-specific questionnaire and survey design issues not captured in the year and seasonal dummies).

5 Main (all-sample) results

5.1 First-step estimation

In Table 1, we present the results from the first-step regression. We divide our attention between three set of income variables: *total household expenditure* (Models 1 to 3), *agricultural profits* (Models 4 to 6) and *food expenditure* (Models 7 to 9). The usage of total expenditure is motivated by the fact that it is considered more reliable than profits, and because it eases comparison with the so-called "demand for calories"-literature. Remember again, however, for total expenditure to be equal to total income, labor supply must be inelastic to income changes. The second variable, *agricultural profits*, has higher variance, and about 20 households recorded no rural profits (note, however, that only three households stated that they did not consume any own-produced goods). As noted in Section 2.1, such a formulation allows for variable labor supply. The usage of the variable *food expenditure* for our IV estimates is based on the

Table 1: First step estimates. The impact of rainfall on household expenditure, profits and food consumption. Household sample.

	Expenditure (log)			Profits (log)			Food expenditure (log)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Rainfall	0.330***	0.411***	0.463***	0.452***	0.763***	0.672***	0.503***	0.688***	0.686***
	(0.0982)	(0.123)	(0.132)	(0.162)	(0.226)	(0.236)	(0.126)	(0.147)	(0.154)
Year-district dummies	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Seasonal dummies	No	No	Yes	No	No	Yes	No	No	Yes
Observations	2974	2974	2974	2897	2897	2897	2968	2968	2968
R^2	0.014	0.053	0.055	0.011	0.058	0.060	0.022	0.068	0.071

All models include year-district interaction terms, seasonal (quarterly) dummies, and fixed effects at the individual level. Cluster-adjusted standard errors in parenthesis. Dependent variable: log of body weight in kilos. Log (Income) equals log of household expenditure divided by current number of household members. The individual controls include total number of adults in household, total number of children in household, worked, attended school, pregant and away from household last six months.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

most restrictive assumptions, but is interesting as a reference to the discussion on whether regional food availability is the main cause of malnutrition.

As seen in Table 1, the impact of a percentage increase in rainfall implies that income, measured as total expenditure divided by the current number of household members, increases with around 0.45 percent, when controlling for year and season. The impact of rainfall on agricultural profits and food expenditure is higher. Note however, that the coefficient in the profit-specification has higher standard errors and the estimates are not statistically different from our expenditure-estimates. The impact of rainfall on food expenditure is higher, and the effect is statistically different from 0.45. We interpret this difference as evidence that the income elasticity of food is higher than one.

In Appendix, we study the first-step estimates using total expenditure in detail, and by including additional covariates. As it turns our, controlling for wave has little impact on our variable of interest, rainfall, but drives practically all other time-covariates into insignificance. Being a rather severe case of multicollinearity, we choose to exclude the wave dummies from our baseline estimates. We also control for household composition. Number of children in household and number of adults in household enter in a statistically significant way, but have no impact on the rainfall-expenditure nexus.

We have also evaluated the issue of using alternative rainfall specifications. These results are presented in the appendix. There was little support for non-linear rainfall variables. Lagging the weather variable further reduced the point estimate somewhat, implying that current expenditure and income figures are indeed given more weight when respondents are asked to recall such figures. There were some support for first-step heterogeneity (that is, that the impact of rainfall of expenditure interacts with agronomic zone and district) but by the principle of parsimony, we prefer the singe-variable instrument.

5.2 IV-estimates

In Table 2, estimates of the response of body weight to transitory income changes is presented. The reason that the first-step covariates reporten in Table 2 differs somewhat from those in Table 1 is that the IV estimates are based on the full sample of individuals (the first step results in Table 1 are based on the household sample). Further results, with results on additional covariates, can be found in the appendix.

Using the full sample and assuming that the elasticity is the same for all

Table 2: The response of body weight to transitory changes in household expenditure, using rainfall as an instrument for household income. Full sample.

	(1) OLS	(2) OLS	(3) IV	(4) IV	(5) IV	(6) IV	(7) IV	(8) IV
Expenditure (log)	0.00346	0.00171	0.0619**	0.0671***	0.0711***	0.0652***	0.0678***	0.0649**
	(0.00212)	(0.00188)	(0.0272)	(0.0241)	(0.0252)	(0.0231)	(0.0237)	(0.0292)
Height		0.00551***			0.00549***	0.00552***	0.00543***	0.00545***
		(0.00194)			(0.00192)	(0.00193)	(0.00192)	(0.00193)
Illness: Malaria		-0.00995***				-0.0102***	-0.0101***	-0.00996***
		(0.00198)				(0.00207)	(0.00202)	(0.00198)
Illness: Other		-0.00374**				-0.00451**	-0.00441**	-0.00430**
		(0.00167)				(0.00201)	(0.00201)	(0.00179)
Rainfall (*first step*)			0.456***	0.490***	0.492***	0.501***	0.493***	0.414***
			(0.123)	(0.126)	(0.126)	(0.125)	(0.129)	(0.153)
Rainfall (*reduced form*)			0.0283**	0.0329***	0.0350***	0.0327***	0.0334***	0.0269***
			(0.0112)	(0.00894)	(0.00913)	(0.00903)	(0.00897)	(0.00979)
Age-specific trend	No	Yes	No	Yes	Yes	Yes	Yes	Yes
Individual controls	No	Yes	No	No	No	No	Yes	Yes
Wave dummies	No	No	No	No	No	No	No	Yes
Observations	16638	16617	16638	16638	16617	16617	16617	16617
IV R^2	0.268	0.498	0.213	0.395	0.414	0.427	0.430	0.437
Craig Donald F			122.4	132.3	132.9	136.8	136.4	42.48

All models include year-district interaction terms, seasonal (quarterly) dummies, and fixed effects at the individual level. Cluster-adjusted standard errors in parenthesis. Dependent variable: log of body weight in kilos. The individual controls include total number of adults in household, total number of children in household, worked, attended school, pregant and away from household last six months.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

household members, our point estimate of the expenditure elasticity of body weight is around 0.065. The age-specific trend variables include "time since last survey", and this variables interacted with age and age squared at the first wave. Controlling for an age-specific growth trend adds much to precision, but does not drastically change the point estimates. Including *height* does not seem to have a large impact on our point estimates nor on efficiency. *Height* is however statistically significant, even when controlling for an age-specific linear trend.

Table 2 also reports the results obtained when including two indicators of illness. As mentioned, a concern was that rainfall could induce the spread of certain diseases – keeping expenditure constant – in which case our estimates would be inconsistent. Since many household members reported being sick but unable or unwilling to seek professional care, both of these variables are based on the "own-diagnosis"-variable in the KHDS dataset. As seen in Table 2, malaria-incidence is associated with a reduction of body weight, a statistically significant effect; for *other illness*, the effect is somewhat smaller. Our income-coefficient is only marginally changed.

It is important to note that the regressions including the illness-variables are by no means to be seen as "preferable" to the baseline estimate. The problem is that body weight is a proxy for "health" in a broad sense, and including additional health covariates will bias the elasticity downwards.[14] The reason that the illness-variables enter in a statistically significant way in our regression without affecting the point estimate (all IV-estimates of the elasticity are statistically indistinguishable from each other) implies that loosing weight is not regarded as an illness *per se* among the respondents in our sample.[15] The results conveyed in Table 2 suggests that most of the variation in reported illness is "random", and that rainfall is not associated with malaria other than via its effect on expenditure.

We also present two types of "kitchen sink"-regressions, step-wise including additional control covariates and wave dummies. The wave dummies has a dramatic impact on the F-statistic from the first step regression, which is most likely a result of the covariance between the wave dummies and the district-year and seasonal dummies. The importance of these wave controls should thus not be over-dramatized, as they only seem to decrease efficiency of our

[14] This problem is similar to the issue of adding industry or occupational dummies in a Mincer-equation if you want to estimate the net wage return of schooling.

[15] Indeed, when those who reported being ill was asked for a diagnosis, only a few reported "malnutrition".

IV-estimates, and not the coefficients. The income-effect is still statistically significant at the 5%-level. Individual coefficients from these regressions can be found in appendix, Table 8.

Finally, the strongest contrast in Table 2 is of course that between our IV estimates and the OLS estimate. The general impression one would get from regressing body weight on expenditure is that nutritional outcome is, indeed, invariant to the purchasing power of the household, since the elasticity is statistically insignificant and very close to zero. We argue that this discrepancy is most likely subject to measurement error in the main explanatory variable. As pointed out in Section 2.2 and elsewhere, expenditure figures based on recollection is likely to be associated with a non-trivial degree of error. In the classical case of zero-mean error in the explanatory variable, the attenuation bais towards zero becomes even more severe in a fixed-effects analysis. Under the assumption that our exclusion restriction holds, our IV estimates take care of this bais.

In Table 3, we present further estimates of the IV regression. All results in Table 3 are based on Equation (8), plus the age-specific trend variables. The estimates presented in Table 3 show that the estimates are robust to the definition of income, outliers and sample weight. Regressing the relationship on the "no attrition"-sample produces lower point estimates – around the magnitude of 0.04 – but the effect is still statistically significant. Again, the "no attrition" sample is likely to produce downward-biased estimates of the relationship if those who have a high elasticity are also those that are likely to be absent. Under this assumption, the "no attrition"-results can be viewed as a lower bound of the expenditure elasticity of body weight.

Although the statistical contrast is compelling, the economic significance of our results is not equally straightforward to deduce. Is an expenditure elasticity of body weight of about 0.065 high? An increase in household expenditure with ten percent will add little less than a kilo to a person that weights fifty kilo. Although log-log estimates of an elasticity is valid only for small changes in the explanatory variable, a change in household expenditure of about 50% implies that a person will increase his BMI by one index point (e.g. from 18.5 to 19.5), which is a non-trivial increase.

Table 3: Further IV results using alternative explanatory variables and sample restrictions. Full sample.

	(1) Baseline	(2) No outliers	(3) No outliers, no attrition	(4) Using Sample Weights	(5) Expl. var: Profits	(6) Expl. var: Food exp.	(7) Expl. var: Per capita exp.
Income (total expenditure unless otherwise stated)	0.0671***	0.0720***	0.0354*	0.0716**	0.0498**	0.0490***	0.0696***
	(0.0241)	(0.0232)	(0.0191)	(0.0320)	(0.0207)	(0.0165)	(0.0243)
Rainfall (*first step*)	0.490***	0.488***	0.465***	0.612***	0.639***	0.671***	0.473***
	(0.126)	(0.126)	(0.127)	(0.217)	(0.225)	(0.161)	(0.134)
Rainfall (*reduced form*)	0.0329***	0.0352***	0.0165**	0.0438***	0.0329***	0.0329***	0.0329***
	(0.00894)	(0.00841)	(0.00810)	(0.0158)	(0.00894)	(0.00894)	(0.00894)
Observations	16638	16528	11492	14613	16300	16606	16638
IV R^2	0.395	0.438	0.547	0.445	0.373	0.409	0.393
Craig Donald F	132.3	130.6	85.35	235.0	93.14	169.0	122.0

All models include year-district interaction terms, seasonal (quarterly) dummies, and fixed effects at the individual level. Cluster-adjusted standard errors in parenthesis. Dependent variable: log of body weight in kilos. To achieve meaningful comparable estimates, we exclude extreme outliers in sample weights in Model 4.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4: The response of body weight to transitory income changes. Subsample results across gender and age.

	Females					
	(1)	(2)	(3)	(4)	(5)	(6)
	All	Age>15	Age>30	Age<16	Age<11	Age<5
Total Expenditure (log)	0.0844***	0.0446*	0.0675**	0.0968**	0.126**	0.0503
	(0.0309)	(0.0261)	(0.0319)	(0.0449)	(0.0628)	(0.0772)
Rainfall (*first-step*)	0.497***	0.489***	0.477***	0.503***	0.419***	0.462***
	(0.130)	(0.142)	(0.145)	(0.129)	(0.138)	(0.169)
Rainfall (*reduced form*)	0.0419***	0.0218**	0.0322***	0.0486**	0.0528***	0.0232
	(0.0118)	(0.0105)	(0.00990)	(0.0191)	(0.0194)	(0.0359)
Observations	8651	4323	2453	4328	2944	1332
R^2	0.348	-0.085	-0.352	0.494	0.508	0.650
Craig-Donald F	70.93	32.18	16.71	38.43	17.41	9.408
	Males					
	(7)	(8)	(9)	(10)	(11)	(12)
	All	Age>15	Age>30	Age<16	Age<11	Age<5
Expenditure (log)	0.0462**	0.0213	0.0376	0.0391	0.0392	0.0239
	(0.0232)	(0.0236)	(0.0270)	(0.0277)	(0.0361)	(0.0751)
Rainfall (first-step)	0.495***	0.383***	0.452***	0.577***	0.635***	0.537***
	(0.127)	(0.137)	(0.152)	(0.138)	(0.143)	(0.165)
Rainfall (reduced form)	0.0229**	0.00822	0.0170*	0.0225	0.0250	0.0128
	(0.00941)	(0.00828)	(0.00987)	(0.0149)	(0.0223)	(0.0408)
Observations	7980	3377	1834	4603	3145	1482
R^2	0.437	0.180	-0.162	0.523	0.550	0.570
Craig-Donald F	64.34	16.45	12.92	50.30	43.35	14.38

All models include year-district interaction terms, seasonal (quarterly) dummies, and fixed effects at the individual level, and a age-specific linear trend. Cluster-adjusted standard errors in parenthesis. Dependent variable: log of body weight in kilos. "Age" is initial age at first survey round.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

5.3 Heterogeneity (subsample results)

The above estimates are valid only under the presumption that the response of human body weight to transitory changes in household expenditure is homogenous across age and gender and other subsamples. Earlier studies that assess the intrahousehold allocation of nutrients suggest that the demand for calories differ between children and adults and between females and males (see e.g. Behrman and Deolalikar, 1990). It is reasonable to ask whether our results varies in a similar fashion.

Comparing IV-estimates across subsamples are however complicated by the possibility of first-step heterogeneity (as opposed to second-step, or IV, heterogeneity). If the instrument is weaker for the subsample of households in which male adults are overrepresented, IV-estimates on the subsample of male adults will have larger standard errors (and, if the instruments are very weak, blow up a potential bias). For this reason, comparing 2SLS estimates across subsamples is a delicate task. In our case, however, this does not seem to be the case. As can be seen in Table 3, rainfall appear to have a homogenous impact on household expenditure across subsamples.

As for the IV estimates of interest, the results in Table 3 suggest that the body weight of females, and in particular female children, is more sensitive to exogenous fluctuations in income. In contrast, the body weigh response among males (Panel B in Tabel 3) is about half that of females, and there is little heterogeneity across age categories.

6 Concluding remarks

The results in this paper suggest that the nutritional status of females, and particularly female children, is subject to the household's income constraint to a non-trivial extent. Conversely, the nutritional status of adult males is largely invariant to changes in household expenditure. These findings provide important insights and implications for development policy. On one hand, the results suggest that the impact of adverse shocks are disproportionably leveled on female children. Acute emergency aid that can be thought to affect households' budget constraints indiscriminately must be outlined with this in mind. A second, more general implication of a high expenditure elasticity of body weight, short-run income-augmenting policies that manage to reach households can have a non-trivial and positive impact on the nutritional status of children. This result suggests that policies oriented towards economic growth and income are aligned with ambitions to improve the nutritional status.

References

Alderman, Harald, Hans Hoogeveen & Mariacristina Rossi. 2006. "Reducing Child Malnutrition in Tanzania: Combined Effects of Income Growth and Program Interventions." *Economics and Human Biology* 4:1–23.

Alderman, Harold, Takashi Yamano & Luc Christiaensen. 2005. "Child Growth, Shocks, and Food Aid in Rural Ethiopia." *American Journal of Agricultural Economics* 87(2):273–288.

Aromolaran, Adebayo B. 2004. "Intra-Household Redistribution of Income and Calorie Consumption in South-Western Nigeria." Working Papers 890, Economic Growth Center, Yale University.

Bank, The World. 2004. "User's Guide to the Kagera Health and Development Survey." Development Research Group.

Bardhan, Pranab & Christopher Udry. 1999. *Development Microeconomics.* Oxford University Press Inc., New York.

Beegle, Kathleen, Rajeev H Dehejia & Roberta Gatti. 2005. "Child Labour, Crop Shocks and Credit Constraints." CEPR Discussion Papers 4881, C.E.P.R. Discussion Papers.

Behrman, Jere R. & Anil B. Deolalikar. 1987. "Will Developing Country Nutrition Improve with Income? A Case Study for Rural South India." *Journal of Political Economy* 95(3):492–507.

Behrman, Jere R. & Anil B. Deolalikar. 1989. "Is Variety the Spice of Life? Implications for Calorie Intake." *Review of Economics and Statistics* 71(4):666–672.

Behrman, Jere R. & Anil B. Deolalikar. 1990. "The Intrahousehold Demand for Nutrients in Rural South India: Individual Estimates, Fixed Effects, and Permanent Income." *Journal of Human Resources* 25(4):665–696.

Bouis, Howarth E. & Lawrence J. Haddad. 1992. "Are estimates of calorie-income elasticities too high? A recalibration of the plausible range." *Journal of Development Economics* 39:333–364.

Deaton, Angus. 1997. *The analysis of household surveys: A microeconometric approach to development policy.* Baltimore, The Johns Hopkins University Press.

Dercon, Stefan & Pramila Krishnan. 2000. "In Sickness and in Health: Risk Sharing within Households in Rural Ethiopia." *Journal of Political Economy* 108(4):688–727.

Duflo, Esther & Christopher Udry. N.d. "Intrahousehold Resource Allocation in Cote d'Ivoire: Social Norms, Separate Accounts and Consumption Choices." NBER Working Papers 10498, National Bureau of Economic Research, Inc.

DuMouchel, William H. & Greg J. Duncan. 1983. "Using Sample Survey Weights in Multiple Regression Analyses of Stratified Samples." *Journal of the American Statistical Association* 78(383):535–543.

FAO. 2001. "Human Energy Requirements." Report of a joint FAO/WHO/UNU Expert Consultation, Food and Nutrition Technical Report Series, Rome.

Foster, Andrew D. 1995. "Prices, Credit Markets and Child Growth in Low-Income Rural Areas." *Economic Journal* 105(430):551–570.

Haddad, Lawrence, Harold Alderman, Simon Appleton, Lina Song & Yisehac Yohannes. 2003. "Reducing Child Malnutrition: How Far Does Income Growth Take Us?" *World Bank Economic Review* 17(1):107–131.

Hoddinott, John & Bill Kinsey. 2001. "Child growth in the time of drought." *Oxford Bulletin of Economics and Statistics* 63:409–435.

Jacoby, Hanan & Emmanuel Skoufias. 1998. "Smoothing Consumption Behaviour Using Information on Aggregate Shocks: Income Seasonality and Rainfall in Rural India." *American Journal of Agricultural Economics* 80(1):1–14.

Leibenstein, Harvey. 1957. *Economic backwardness and economic growth.* New York, Wiley.

Miguel, Edward, Shanker Satyanath & Ernest Sergenti. 2004. "Economic Shocks and Civil Conflict: An Instrumental Variables Approach." *Journal of Polictal Economy* 112(4):725–753.

Morduch, Jonathan. 1995. “Income Smoothing and Consumption Smoothing.” *Journal of Economic Perspectives* 9(3):103–114.

Paxson, Christina H. 1992. “Using Weather Variability to Estimate the Response of Savings to Transitory Income in Thailand.” *American Economic Review* 82(1):15–33.

Rose, Elaina. 1999. “Consumption Smoothing and Excess Female Mortality in Rural India.” *Review of Economics and Statistics* 81(1):41–49.

Rosenzweig, Mark R. & Kenneth I. Wolpin. 2000. “Natural ’Natural Experiments’ in Economics.” *Journal of Economic Literature* 38(4):827–874.

Rosenzweig, Mark R. & T. Paul Schultz. 1982. “Market Opportunities, Genetic Endowments, and Intrafamily Resource Distribution: Child Survival in Rural India.” *American Economic Review* 72(4):803–815.

Sen, Amartya K. 1981. *Poverty and famines: An essay on entitlement and deprivation*. Oxford, Clarendon.

Subramanian, Shankar & Angus Deaton. 1996. “The Demand for Food and Calories.” *Journal of Political Economy* 104:133–162.

Townsend, Robert M. 1995. “Consumption Insurance: An Evaluation of Risk-Bearing Systems in Low-Income Economies.” *Journal of Economic Perspectives* 9(3):83–102.

WHO. 1995. “Physical Status: The Use and Interpretation of Anthropometry.” WHO Technical Report, No.854, Geneva.

Wolpin, Kenneth I. 1982. “A New Test of the Permanent Income Hypothesis: The Impact of Weather on the Income and Consumption of Farm Households in India.” *International Economic Review* 23(4):583–94.

Wooldridge, Jeffrey M. 2003. “Cluster-Sample Methods in Applied Econometrics.” *American Economic Review* 93(2):133–138.

Wooldridge, Jeffrey M. 2006. “Cluster-Sample Methods in Applied Econometrics: An Extended Analysis.” Mimeo, Department of Economics, Michigan State University.

A Appendix: Tables

Table 5: Descriptive statistics of the KHDS sample (first wave)

	Mean	Median	First quartile	Third quartile
Household statistics				
Household expenditure (Tnz)	386 726	277 805	165 742	445 789
Per capita expenditure (Tnz)	77 635	53 771	35 236	86 517
Per capita expenditure (dollar)	260	180	118	289
Household members	5.6	5	3	7
Food exp./Total exp.	0.632	0.643	0.752	0.537
Degree of subsistence	0.446	0.463	0.318	0.595
Degree of food subsistence	0.664	0.726	0.550	0.849
Nutritional statistics				
Body mass index (Age>18)				
Females	21.46	21.06	19.17	22.97
Males	20.43	20.3	18.75	21.82
Weight-for-age z-score				
Females				
Age 0-5	-1.111	-1.28	-1.96	-0.343
Age 6-10	-1.04	-1.04	-1.72	-0.440
Age 11-18	-1.20	-1.28	-1.82	-0.666
Males				
Age 0-5	-1.32	-1.39	-2.10	-0.545
Age 6-10	-1.11	-1.15	-1.70	-0.562
Age 11-18	-1.77	-1.83	-2.39	-1.226

All expenditure variables are annual values, based on recollection. "Household members" is the number of current household members recorded on the rooster-questionnaire. Per capita expenditure is "Household expenditure" divided by "Household members". "Degree of subsistence" equals consumption of own-produced goods divided by total consumption. Food subsistence is consumption of own-produced food divided by total food consumption. Dollar exchange rate: 298.5Tnz/Dollar, based on the average exchange rate between the first quarter of 1990 and the first quarter of 1992.

Table 6: The impact of rainfall on household expenditure. Household sample.

	(1) m1	(2) m2	(3) m3	(4) m4	(5) m5
Rainfall	0.330***	0.411***	0.463***	0.464***	0.458**
	(0.0982)	(0.123)	(0.132)	(0.133)	(0.179)
District 1*Year 1		-0.441***	-0.492***	-0.505***	-0.253
		(0.123)	(0.126)	(0.136)	(0.312)
District 1*Year 2		-0.306***	-0.300***	-0.300***	-0.175
		(0.0779)	(0.0715)	(0.0690)	(0.183)
District 2*Year 1		-0.0176	-0.0542	-0.0796	0.181
		(0.0780)	(0.0836)	(0.0817)	(0.315)
District 2*Year 2		-0.0920**	-0.0858*	-0.0930**	0.0362
		(0.0443)	(0.0458)	(0.0457)	(0.165)
District 3*Year 1		0.0527	0.0110	0.00579	0.263
		(0.113)	(0.116)	(0.108)	(0.328)
District 3*Year 2		-0.144***	-0.138***	-0.137***	-0.00715
		(0.0459)	(0.0401)	(0.0416)	(0.166)
District 4*Year 1		-0.0403	-0.0984	-0.138	0.106
		(0.128)	(0.136)	(0.150)	(0.316)
District 4*Year 2		0.00152	0.00810	-0.00381	0.125
		(0.0539)	(0.0534)	(0.0605)	(0.175)
District 5*Year 1		-0.231**	-0.269**	-0.263**	-0.0112
		(0.0983)	(0.111)	(0.106)	(0.307)
District 5*Year 2		-0.0246	-0.0151	-0.00645	0.114
		(0.0573)	(0.0595)	(0.0512)	(0.161)
District 6*Year 1		-0.176**	-0.204**	-0.225***	0.0312
		(0.0734)	(0.0843)	(0.0833)	(0.295)
District 6*Year 2		-0.0183	-0.00674	-0.0136	0.108
		(0.0707)	(0.0701)	(0.0674)	(0.156)
Quarter 1			-0.0563	-0.0686*	0.0329
			(0.0386)	(0.0388)	(0.115)
Quarter 2			-0.00197	-0.00457	0.0591
			(0.0317)	(0.0310)	(0.0790)
Quarter 3			-0.0479	-0.0526	-0.0278
			(0.0337)	(0.0336)	(0.0486)
Children in hh				0.0731***	0.0734***
				(0.00917)	(0.00930)
Adults in hh				0.0729***	0.0726***
				(0.0148)	(0.0148)
Wave 1					-0.227
					(0.266)
Wave 2					-0.147
					(0.183)
Wave 3					-0.0919
					(0.0909)
Observations	2974	2974	2974	2974	2974
R^2	0.014	0.053	0.055	0.087	0.088

Cluster-adjusted standard errors in parenthesis. Fixed effects at the household level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 7: The impact of rainfall on household expenditure. Household sample. Alternative definitions of rainfall.

Rainfall	0.463***	0.761				
	(0.132)	(2.141)				
Rainfall squared		-0.0319				
		(0.232)				
Rainfall lagged 3 months			0.333***			
			(0.109)			
Rainfall lagged 6 months				0.398***		
				(0.118)		
log(rain)*Karagwe					0.0637	
					(0.149)	
log(rain)*Bukoba rur.					0.528	
					(0.330)	
log(rain)*Muleba					0.666*	
					(0.394)	
log(rain)*Biharamulu					0.996***	
					(0.352)	
log(rain)*Bukoba urb.					0.490	
					(0.319)	
Rainfall*"Tree crop"-zone						0.529*
						(0.306)
Rainfall*"Cereal"-zone						0.552**
						(0.261)
Rainfall*"Cotton"-zone						0.583***
						(0.191)
Rainfall*"Urban"-zone						0.188
						(0.203)
Observations	2974	2974	2974	2974	2974	2974
R^2	0.055	0.055	0.055	0.054	0.054	0.056

Cluster-adjusted standard errors in parenthesis. Fixed effects at the household level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 8: The response of body weight to transitory changes in expenditure, using rainfall as an instrument for household income. Full sample.

	(1)	(2)	(3)	(4)	(5)	(6)
Expenditure (log)	.0619**	.0671***	.0711***	.0652***	.0678***	.0649**
	(.0272)	(.0241)	(.0252)	(.0231)	(.0237)	(.0292)
Time		.0125***	.00911***	.00918***	.00891***	.0128***
		(.00150)	(.00226)	(.00219)	(.00219)	(.00417)
(Time)		−.675***	−.520***	−.519***	−.528***	−.527***
x(Age)		(0.0215)	(0.0585)	(0.0584)	(0.0593)	(0.0602)
(Time)		0.680***	0.531***	0.529***	0.538***	0.538***
x(Age squared)		(0.0298)	(0.0609)	(0.0608)	(0.0618)	(0.0638)
Height			.00549***	.00552***	.00543***	.00545***
			(.00192)	(.00193)	(.00192)	(.00193)
Illness: Malaria				−.0102***	−.0101***	−.00996***
				(.00207)	(.00202)	(.00198)
Illness: Other				−.00451**	−.00441**	−.00430 **
				(.00201)	(.00201)	(.00179)
Pregnant					.0455***	.0457***
					(.00487)	(.00494)
Children in hh					−.00598***	−.00577**
					(.00216)	(.00238)
Adults in hh					−.00386	−.00361
					(.00237)	(.00256)
Attended school					−.0175***	−.0174***
					(.00294)	(.00291)
Worked					−.00142	−.00131
					(.00177)	(.00174)
Away from hh					−.0122***	−.0126***
					(.00404)	(0.00402)
Wave 1						.0843
						(.0633)
Wave 2						.0513
						(.0430)
Wave 3						.0309
						(.0227)
Observations	16638	16638	16617	16617	16617	16617
R^2	0.213	0.395	0.414	0.427	0.430	0.437
Craig-Donald F	122.4	132.3	132.9	136.8	136.4	42.48

All models include year-district interaction terms, seasonal (quarterly) dummies, and fixed effects at the individual level. Cluster-adjusted standard errors in parenthesis. Dependent variable: log of body weight in kilos. Time is "time since last measurement", and age is "age at first survey round". Time*(Age) is scaled up by a factor of 1000; Time*(Age squared) by a factor of 100 000.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$