HEDG
HEALTH, ECONOMETRICS AND DATA GROUP

THE UNIVERSITY of York

WP 13/13

Income, rainfall shocks and health.
An instrumental variable approach.

David Savage and Eleonora Fichera

June 2013

york.ac.uk/res/herc/hedgwp

Income, rainfall shocks and health. An instrumental variable approach.

David Savage[1] and Eleonora Fichera[2]

Abstract

We examine whether income shocks affect a range of health outcomes and a preventative behaviour. We instrument income with rainfall measurements by matching satellite information on timing and positioning of 21 rainfall stations to longitudinal data (1991-1994) of over 4,000 individuals in 51 villages of a North Western region in Tanzania.

We find a pro-cyclical effect of income on health. A ten percent increase in income reduces by 0.2 the number of illnesses. A further finding is the positive effect on vaccinations of children under six: a ten percent increase in income implies an increase of about one vaccination, from a mean of 2.3 per child, for the four vaccinations of polio, tetanus, tuberculosis and measles. There is also some evidence of a reduction in chronic malnutrition of children under six. Our results suggest the income effect to offset the increased opportunity cost of time in this data.

JEL Classification: O12; O15; I12

Keywords:

Rainfall shocks; Income; Health; Spatial interpolation.

Word count: 5,280
Tables count: 5 (main text) + 5 (appendix)

No conflict of interest or ethical issues

---

[1] School of Environment and Development, University of Manchester. and HCL.

[2] Corresponding author: Manchester Centre for Health Economics, Institute of Population Health, University of Manchester, 4.320 Jean McFarlane Building, Oxford Road M13 9PL, Manchester (United Kingdom). Email: eleonora.fichera@manchester.ac.uk.

We thank discussant and participants of the 2013 Conference of the Centre for the Study of African Economies at Oxford University.

1. Introduction

Agricultural income in developing countries, and especially in Africa, is affected by environmental shocks such as the timing and quantity of rainfall. Exploiting local and household level shocks as an exogenous variation of income in Tanzanian individual-level panel data, we show the existence of a *pro-cyclical* effect of income on health.

The economic debate of the effect of aggregate (as opposed to idiosyncratic) shocks essentially evolves around the relative size of income and substitution effects. When income is higher, health goods are more attainable leading to potential health improvements, but opportunity cost of time is also higher leading to a potential reduction of health-promoting activities. Which of these two effects prevails is an empirical question and varies with the type of aggregate shock and the institutional setting within a country. Some studies in developed countries (see for example, Ruhm, 2003 and 2005; Dehejia and Lleras-Muney, 2004) argue in favour of a *counter-cyclical* effect of income on health. In developing countries, negative income shocks coupled with credit constraints and other market imperfections lead to incomplete consumption smoothing which, in turn, may work towards a deterioration of health through poorer nutrition and reduced consumption of health services. We argue that the income effect prevails in our data because the health-promoting activities are relatively less time-intensive and because better diet quality is available to the household.

Issues of causality and endogeneity further complicate the relationship between income and health. The attempt to infer a causal relationship between income and health is made difficult by the availability of suitable instruments. There are two potential, non-exclusive explanations for the endogeneity between income and health. Firstly, as health is both consumption and a production good, changes in income are correlated with changes in health (and vice versa). Secondly, a third, confounding, omitted variable causes changes in both income and health.

In developed countries attempts to determine a causal relation between income and health have been made either by using lottery wins or inheritances (Adda et al, 2009; Ettner, 1996; Meer et al., 2003); macroeconomic shocks such as unemployment or house price shocks (Michaud and van Soest, 2008; Apouey and Clark, 2011; Kim and Ruhm, 2012; Fichera and Gathergood, 2013) or by using public policies as natural experiments (Case, 2004; Snyder and Evans, 2006 and Frijters et al., 2005).

In many developing countries the main source of income derives from agriculture. In the United Republic of Tanzania, the country we focus on in this paper, about 38 percent of the population lives below the national basic needs poverty line (National Bureau of Statistics, Tanzania National Budget Survey, 2007). Agriculture accounts for about half of gross production and employs about 80 percent of the labour force (Thurlow and Wobst, 2003). Agriculture in Tanzania is primarily rain-fed with only two percent of arable land having irrigation infrastructure (FAO, 2009). Its main staple crops like maize are particularly susceptible to weather conditions.

Because of such highly dependence on weather events, many studies in developing countries examine the effect on health of income shocks provoked by natural disasters (see for example, Artadi, 2005; Bengtsson, 2010; Yamano et al., 2005; Miller and Urdinola, 2010; Rose, 1999; Pörtner, 2010; Datar et al., 2013; Edoka, 2013). One (conceptual) feature of these studies is the widespread focus on health outcomes, particularly for children. Only two studies have focused on health outcomes and preventative behaviours. Miller and Urdinola (2010) examine the effect of world coffee price fluctuations on infant and child mortality in Colombia. They find evidence of pro-cyclicality of mortality and of preventative care and vaccinations for children under the age of one. Such results could be explained by a reduction in the opportunity cost of time spent on childcare. Datar et al. (2011) linked three waves of the Indian National Family and Health Survey with an international database of natural disasters. They show that exposure to a natural disaster increases the likelihood of illnesses in children under the age of five by 9-18 percent and increases the likelihood of full age-appropriate immunisation by nearly 18 percent.

The second (methodological) feature entails the construction of the shock. For instance, Rose (1999) linked household level data, the Indian Additional Rural Incomes Survey, to District level rainfall data collected by the World Bank. Therefore, exogenous income variation is only identified between districts and years not between households. Bengtsson (2010), with the same Kagera data as in this paper, uses a time series of rainfall to instrument for income shocks and identify causal impacts on body weight. He finds a statistically significant increase in relative weight and a reduction in malnutrition but only on females. However, as he only exploits differences between rainfall measurements of six districts, his instrument lacks sufficient variability to measure income variation between households and between villages.

In this paper we make a number of contributions compared to the above-described literature. First, we adopt a novel construction of rainfall shocks using satellite measurements of historical rainfall data linked to individual-level longitudinal data. We use timing and positioning of 21 weather stations across 51 villages from satellite data to generate a village-specific (through spatial interpolation of distances of the stations to the centre of the village) and household-specific (through matching interview dates to historical rainfall data) rainfall measurement.

Second, we use rainfall as instrument of income and compare the effect of transitory income shocks across a wide range of health outcomes (i.e. Body Mass Index (BMI), number of self-reported illnesses; height-for-age and weight-for-height for children under the age of six). In our main specification we find no statistically significant effect of such transitory income changes on health outcomes such as BMI or weight-for-height. But we find a reduction of illnesses such as acute fever, chills, coughs, severe headaches and abdominal pain. There is also some evidence supporting a reduction of chronic malnutrition of children under the age of six.

Finally, we determine whether rainfall shocks also affect investments in preventative activities such as vaccinations for children under the age of six. The finding that vaccinations are positively affected by income changes suggests that the income effect dominates the substitution effect. Previous research has established that relaxing credit constraints in developing countries increases engagement in preventative health behaviours (see for example, Crepon et al., 2011; Mahajan et al., 2011).

The paper is structured as follows. Section 2 contains a description of the data; summary statistics are reported in section 3. The empirical strategy is explained on Section 4. Section 5 reports the results. Section 6 concludes.

2. Data description

We link two data sources to each other, namely, the Kagera Health and Development Survey (KHDS) and the Tanzania meteorological rainfall data.

2.1 Kagera Health and Development Survey (KHDS)

We use baseline data from a longitudinal Living Standards Measurement Survey (LSMS) conducted in the Kagera region of North Western Tanzania[3], the Kagera Health and Development Survey (KHDS). It is one of the few long-running surveys containing questions on individual socioeconomic characteristics, wealth and income, health behaviours and outcomes. KHDS also holds a rich set of community characteristics on health care, children education and prices in local markets.

The Kagera region is predominantly rural and lies just south of the equator bordering to the north with Uganda and to the west with Rwanda and Burundi. The

[3] The data is publicly available on the World Bank website: http://econ.worldbank.org/WBSITE/EXTERNAL/EXTDEC/EXTRESEARCH/EXTLSMS/0,,contentMDK:21991269~menuPK:4196952~pagePK:64168445~piPK:64168309~theSitePK:3358997,00.html.

population of 1.9 million people in 1991-1994 is predominantly engaged in agricultural production of banana and coffee in the north, and livestock and rain-fed annual crops, primarily cotton, maize and sorghum, in the south. The agricultural sector accounts for 45 percent of GDP. About 29 percent of all households in Kagera are below the basic needs poverty line (Kessy, 2005). In 1991, household average expenditure was US$217 per capita with a range of US$118 and US$357 within the six districts of the Kagera region.

Initially, the longitudinal survey consisted of four waves from 1991 to 1994 and then it was followed up in 2004 and 2010[4]. The first survey consisted of 915 households interviewed up to four times, from September 1991 to January 1994 (at 6-7 month intervals). Households were drawn from 51 villages ("clusters") of 16 households each in the six administrative districts of Kagera: Biharamulo, Bukoba Rural, Bukoba Urban, Karagwe, Muleba and Ngara.

The numbering of the "wave" is defined with respect to the timing of the interview across the whole sample, whereas "passage" is defined with reference to the number of interviews conducted with a specific household. Thus a replacement household during the fourth wave would be its first passage, as it would the first

[4] KHDS 1991-1994 was funded by the World Bank and Muhimbili University College of Health Sciences. Consult: http://econ.worldbank.org/WBSITE/EXTERNAL/EXTDEC/EXTRESEARCH/EXTLSMS/0,,contentMDK:21991269~menuPK:4196952~pagePK:64168445~piPK:64168309~theSitePK:3358997,00.html for a detailed description of the original data and follow-up surveys. We do not report on this paper the 2004 wave results for concerns over timing of income relative to anthropometric measures as there is over a decade gap between waves 4 and 5. But main results of this paper hold even including wave 5 and can be made available by the authors on request. Because there is no historical rainfall data linkage for the 2010 wave we have not used the latest round.

time it had been interviewed. In the first passage 840 households were interviewed. By the end of the fourth passage only 9.6 percent of the 840 had dropped out, implying a very low attrition rate for a developing country data set.

The random sample was geographically stratified according to mortality rates. A variable probability sampling selection procedure was used, based on predicted mortality: households of similar predicted mortality replaced the ones that dropped out before the fourth wave.

2.2 Tanzania Meteorological Data

We link KHDS to monthly rainfall data collected by the Tanzania Meteorological Agency[5]. Historical rainfall data are available for each month of years between 1990 and 1994. KHDS individual interview dates and the cluster where the individual is located are used to link in meteorological data. Each household member interviewed on any day of a given month is assigned the average monthly rainfall measurement of the 12 months prior to the interview date[6].

The data contains total millimetres (mm) of rain per month, total days of rain per month for 21 weather stations in the Kagera region[7]. It also reports distance of the first and second nearest weather station to the centre of the village ("cluster")

---

[5] Data is available from the Economic Development Initiatives website at http://www.edi-africa.com/research/khds/introduction.htm. Last accessed on 30th May 2013.

[6] See section 3.3 for a more detailed description of the timing of rainfall measurement.

[7] KHDS also contains monthly rainfall data averaged over the five districts (Bukoba urban and rural are combined). But for the purpose of this paper, we need to have more variation in the rainfall data which we exploit using cluster-level data, timing of rainfall linked to interview dates and distances to the nearest rainfall stations.

obtained from the Global Positioning System (GPS)[8].

Whilst the Meteorological data is quite accurate, it does contain some missing values. The first closest station to the centre of the village has a proportion of non-missing data that ranges from a minimum of 78.2 percent to a maximum of 86.6 percent depending on the month of the year. The second closest station to the centre of the village has a proportion of non-missing data ranging from 62.5 percent to 82.1 percent. We replace missing information on any one station by taking the measurement of the other closest station reaching between 88 and 94 percent of non-missing data. Remaining missing data for any particular month has been replaced with the average rainfall measurement for that weather station.

3. Summary statistics

Table 1 reports the definition of variables and the descriptive statistics for the KHDS samples across five waves.

3.1 Description of the outcome variables

We examine four health outcome variables: BMI, self-report health conditions, height-for-age and weight-for-height z-scores.

BMI is calculated as the ratio of weight (in kilograms) and the squared value of height (in metres). We take the natural logarithm of BMI. After taking out 35 observations with either very low or high BMI (under 8 or over 50), we are left with

[8] GPS is a space-based satellite navigation system that provides location and time information on or near the Earth. It was initially developed by the US Department of Defense in the early 1970s.

almost 18,000 observations across the four waves. It should be noted that only less than one percent of the sample is obese (i.e. BMI greater than 30) and, consequently, we interpret higher BMI as better health. Table 1 shows that the average logarithm of BMI is about 3. On average individuals are underweight (i.e. BMI lower than 18.5).

The household survey includes, for every individual, up to five self-reported health symptoms in the preceding four weeks: the most popular categories were acute fever, chills, coughs, severe headaches and abdominal pain. Kroeger (1983, p.465) states that “[r]esearch in industrial countries on the self-assessment of the individual’s present health status has shown that is a useful proxy measure for clinically measured health status”. 48 percent of the sample had no self-reported illnesses, 27 percent had one, 17 percent had two and about 10 percent more than two. We construct a discrete, ordinal, health variable, which indicates the number of conditions from zero to five. Table 1 reports an average of about one condition for each individual.

Anthropometric measures for children under the age of six are height-for-age and weight-for-height[9]. Height-for-age is a measure of stunting or chronic malnutrition and weight-for-height is a measure of wasting, acute or transitory malnutrition. The World Health Organisation suggests stunting to be a measure of social deprivation (WHO, 1986). Under the assumption that healthy children follow similar growth patterns across different populations, children‘s anthropometric

[9] WHO’s comparison values for children malnutrition now include children up to the age of ten. Results in this paper do not change with reference values and therefore we have chosen the five years threshold for comparability with the vaccination measure in the KHDS.

measurements are standardized according to the International Referenced Population defined by the U.S. National Centre for Health Statistics (NCHS) with the Centres for Disease Control (CDC) and the World Health Organization (WHO, 1995). Therefore anthropometric measurements are expressed as z-scores, that is, a child‘s measurements and gender is compared to those of a similar child in a reference, healthy population defined by the U.S. NCHS, who has a z-score with mean zero and standard deviation of one[10]. On average there is no evidence of chronic or transitory malnutrition as the z-scores of height-for-age and weight-for-height in Table 1 are both greater than -1. However, the standard deviation of height-for-age measurements is almost three times higher than the weight-for-height one, indicating that the effect of rainfall shocks might vary across clusters and households.

Health preventative behaviour is proxied by number of vaccinations. We construct a vaccination variable: only 7 percent of children under six (the only individuals of which the question was asked, 3,220 observations across the four waves) had no vaccinations, 57 percent had two vaccinations and 24 percent had all the four possible vaccinations against polio, tetanus, tuberculosis (TB) and measles. We construct an ordinal variable, taking values from zero, if the child has no vaccinations, to four if she has all four. Table 1 shows that on average children have been given approximately two vaccinations.

[10] A z-score between -2 and -1 indicates a child is mild stunted; between -3 and -2 indicates a child is moderate stunted; a severely stunted child has a z-score less than -3 and a z-score greater than -1 is not-stunted.

3.2 Description of the covariate variables

We consider a range of demographic variables such as household size, age in years and its squared value. The sample of individuals is relatively young with an average age of about 22 years. Table 1 reports an average number of household members of about eight.

The principal income variable we use in the regressions is the natural logarithm of real income per capita, which is calculated at the household level for each wave. Household income is defined as the sum of five components[11]: i) employment income (i.e. income received as an employee of a private individual or of an institution other than the household for remuneration in cash or in kind); ii) income from self-employment in agriculture (i.e. computed from gross revenues less costs of household level activities in farming, livestock and fishing, plus the value of home agricultural production); iii) income from rent (e.g. income from renting land, farm equipment, dwellings and rental value of owner-occupied housing); iv) transfer income from individuals and organisations; and v) other non-labour income (e.g. pension or retirement fund, insurance, interest on bank accounts, income from games, dowry and inheritance). We omit non-farm self-employment, which represents seven percent of all income, as the KDHS User Guide (World Bank, p.78) cautions that it is especially problematic, and in checking the individual data, only six out of over 3,000 data points have no income for all the other income categories.

[11] We use the most comprehensive definition of income to account for spillover effects of weather shocks onto different income sources.

To calculate per capita income, we use age- and gender-specific nutrition weightings, from a World Health Organisation reference scale (Dercon and Krishnan, 1998, p.44), for each individual within the household. For each household, we then calculate an adult-equivalent size and divide the household income by this variable, relying on the assumption that households behave as a unitary model, with income distributed across household members based on nutritional requirements. Finally, we use the national Tanzanian Consumer Price Index to deflate incomes[12].

The recall periods for questions concerning income were different across waves. In the first wave, the recall period was twelve months, whereas, for the other waves the preamble to the question was: "In the last six months or since I was last here" (World Bank, 2004, p.90). We therefore annualise the data from waves two to four by doubling it. Individuals in our sample have an average income of about US$200 per year at 2004 exchange rate[13].

Table 2 displays descriptive statistics of the outcome variables by income quintiles (from the bottom first quintile to the top fifth quintile). As income increases number of self-reported illness symptoms decrease slightly from an average of 0.96 to an average of 0.87. Number of vaccinations increases particularly between the bottom and second quintile. Comparing the bottom and top quintiles of income, it seems children become more nourished and BMI slightly increases.

---

[12] Downloaded from World Bank: http://databank.worldbank.org/ddp/home.do?Step=3&id=4. Last accessed on 30th May 2013. In 2004 the exchange rate was: 1 U.S. dollar=1,089.33 Tanzanian shillings (TZS).

[13] Note that as income can take negative values we have applied the following log transformation. Annual income has been added to a constant equal to the minimum value of income in that year.

3.3 Description of the instrumental variable

The Kagera region has a bimodal seasonal calendar. There are two main seasons: i) the "rainy season" can be split in a "long rains" season between March and May and a "short rains" season between October and December; and ii) the "dry season" between June and September. The region has been affected by severe droughts in the late 1980s, outside the period of interest to this paper. Table 1 reports that 17 percent of our observations are in the long rains season, 31 percent in the short rains season and 35 percent in the dry season.

We exploit the time and quantity of rainfall as an exogenous shock to income and analyse its effect on health outcomes and behaviours over the period of interest. We consider three dimensions of our instrumental variable, namely, its cross-sectional variation, its relevance and validity.

The cross-sectional variation in rainfall for each interview passage has two sources: timing and geography. We construct a continuous variable, which represents for each region the average monthly rainfall of the twelve months preceding the month of interview in the KHDS[14]. Timing variation occurs because the interview date, within a passage, can vary by as much as seven months. Geographical variation is generated because there are 21 weather stations and each of the 51 clusters is assigned measurements of the two closest stations. In order to determine a monthly rainfall measurement that varies between clusters we use the distances and rainfall measurements of the two closest stations to the

[14] In other words, we use each month of the 1990-1992 historical rainfall data series.

centre of the village. Spatial interpolation has been performed via Inverse Distance Weighting squared (IDW) as follows:

$$\hat{R}_{vt} = \sum_{i=1}^{N} w_i R_{it}$$

$$w_i = \frac{d_i^{-2}}{\sum_{i=1}^{N} d_i^{-2}}$$

where $\hat{R}_{vt}$ indicates the unknown rainfall measurement in millimetres (mm) at the centre of *v=1, …, 51* villages, in time *t* =1,…, 12 (months); R_{it} is the rainfall measurement (mm) of each of the *N=2* closest weather stations *i*.

The intuition of IDW is that the interpolating surface is a weighted average of the location of the weather stations and the weight assigned to each station diminishes as the distance from the centre of the village increases. This interpolation technique has been shown to be accurate in determining the actual rainfall measurement of an unknown point (see for example, Cheng and Liu, 2012; Dirks et al., 1998; Zhuang and Wang, 2003).

The relevance of the instrument can be supported by the high dependence of the population on agricultural production making them vulnerable to both the timing and the quantity of rainfall. This issue is examined more in detail in sub-section 5.1.

Our instrument is valid if rainfall affects health only through income and is not correlated with the residuals of our outcome regressions. This might not be the case if, for example, rainfall increases the incidence of malaria via an increased population of mosquitos around stagnant water. This issue will be examined in more detail in sub-section 5.2.

Table 1 shows that the logarithm of average monthly rainfall is about 4.83mm ranging between a minimum of 4mm to a maximum of 6mm.

4. Empirical strategy

Our central results are based on the two-stage least square regressions (2SLS) as follows:

$$H_{it} = \beta lnincome_{it} + \eta_i + \varepsilon_{it} \quad t = 1, 2, \ldots, T. \qquad (1)$$

where subscript *i* denotes the individual and *t* time. H_{it} is the dependent variable, which includes any of the four health outcomes and the preventative health behaviour. $lnincome_{it}$ is the independent variable of interest, the logarithm of real income per capita, β the coefficient of interest, η_i the unobserved time-invariant individual effects, and ε_{it} the error term.

There are various reasons why income might be endogenous. First, there is reverse causality as bad health might affect income. Second there is measurement error of income itself. To assert that $\hat{\beta}$ estimates the causal effect of income on health outcomes or behaviours, we use rainfall measurement for the period $t-1$ as the instrumental variable z_{it-1} in the first stage equation:

$$lnincome_{it} = \alpha z_{it-1} + \theta_i + v_{it} \qquad (2)$$

all models include a set of cluster and year interaction variables. We control for age, age squared, and household size, seasonal dummies and a linear time trend[15].

In the main models we use robust standard errors, using clustering at the "cluster" level (there are 51 clusters or villages in the sample), to control for heteroskedasticity, justified asymptotically, as the minimum number of observations is 3,200, and is typically greater than 10,000.

Following Miguel, Satyanath and Sergenti (2004), and Wooldridge (2002), we use a 2SLS linear probability model for the discrete variables, as with our continuous dependent variables, rather than an ordered model, such as logit or probit, and do not account for the categorical dependent variables to be ordered.

For each of the 2SLS regressions, we adopt a similar reporting format. A health variable is the dependent variable, representing either health outcomes or behaviours. The natural logarithm of real income per capita is the independent variable of interest. For instrumental variable, we use the monthly average rainfall measurement (mm) in the 12 months prior to the interview date.

There are two ways in which shocks can affect health outcomes and behaviours. On the one hand, there are changes in income available to the individual for health promoting activities and consumption (the income effect). On the other hand, changes in income affect the opportunity cost of time in health promoting activities.

---

[15] We do not include education or wealth (e.g. land size or number of plots) because of potential endogeneity to health and because they are highly correlated with income. Other demographic variables such as marital status, whilst important, contain a high proportion of missing data. Household size should control for family demographics. In a later robustness checks we consider within-household correlation.

At times with scarce rainfall, for instance, the income effect might work towards a deterioration of health outcomes especially if individuals are credit-constrained. The substitution effect works in the opposite direction and might result in health improvements. Which of the effects prevails in the case of health outcomes and behaviours of individuals facing income shocks is an empirical question.

5. Results

5.1 Main specifications

Table 3 shows simple linear fixed effects models of the relation between income and health outcomes and behaviours. The OLS estimates suggest a positive and statistically significant correlation between income and health, indicating an increase in BMI and weight-for-height, a reduction in number of self-reported illness symptoms and a greater uptake of vaccinations. A ten percent increase in income is associated with an increase of 0.4 in the take up of vaccinations. For a median individual with a BMI value of 17.7 a 10 percent increase in income is associated with an increase in BMI of about 0.01 points. But reverse causality and measurement error in the income/health relation may bias the results and prevent us from making any causal inference.

Tables 4 and 5 display results of the instrumental variables models. First stage regressions (i.e. equation (2)) suggest a positive and statistically significant association between rainfall and income. A one percent increase in average monthly rainfall increases income by approximately 2.7 percentage points. Kleibergen-Paap provide a Wald F-statistic for the null hypothesis of weak identification of the instrument. The critical value of 16.38 suggested by Baum et al. (2007) implies a rejection of the null.

Whilst we find no evidence of a statistically significant effect of income on BMI, we do observe a reduction in the number of self-reported illness symptoms. A ten percent increase in income decreases the number of illnesses by 0.2, less than half the effect in the OLS model.

Table 5 shows a positive but weakly statistically significant effect of income on height-for-age. The small effect can be explained by the fact that height-for-age is a typical measure of chronic malnutrition and the one-period time difference between any two waves coupled with the transitory nature of rainfall shocks might not be enough to capture the long-term effects on health. A larger effect can be noticed on the uptake of preventative health measures. A ten per cent increase in income increases the number of vaccinations of children under the age of six by about one.

In addition to the endogeneity bias, difference in the statistical significance and size of OLS and instrumental variable models can partly be explained by the transitory nature of income shocks. Rainfall shocks (other than droughts as in the case of our data) create a temporary shift in household income. Tables 4 and 5 show that transitory income shocks do not have statistically significant effects on long term measures of health outcomes such as BMI or height-for-age. But there is evidence of a reduction in the number of conditions such as cough, fever, diarrhoea and an increase in the vaccinations to children under the age of six. The latter might have long-term effects in term of life expectancy, but more data would be needed to determine whether such is the case for this population.

5.2 Robustness checks

In this sub-section we examine the robustness of our results to within-household correlation effects, issues around the validity of our instrument and attrition rates.

Tables A.1-2 in Appendix A report results of the same models of health outcomes and behaviour where the standard errors have been clustered by household. If household behave as a unitary model, food consumption and health behaviours are likely to be correlated between household members. There is no difference in the size or statistical significance of the models in Tables A.1-2 compared to Tables 5-6. However, income has a statistically significant effect at the five percent level on BMI.

One argument against the validity of our instrument could be that rainfall affects health directly through an increase in the incidence of diseases such as malaria caused by the population of mosquitoes. We test this hypothesis directly with a linear probability fixed effects model of the effect of rainfall on malaria. Table A.3 shows that rainfall has no statistically significant effect on malaria at least in this sample.

Attrition is always a concern of longitudinal household surveys, especially where health is of interest, as death or illness can be a significant determinant of non-response. As the primary objective of the KHDS survey was to “estimate the economic impact of the death of prime-age adults on surviving household members” (World Bank, 2004, p.6), it was particularly important that attrition rates be minimised: for example, if household dissolution and migration are significant coping strategies, following an adult death, then attrition might introduce material bias into the sample. By the end of the fourth passage, 9.6% of the 840 households interviewed in the first passage had dropped out (ibid).

To check for the robustness of the results to attrition bias, we re-run a sample of the regressions, using a balanced panel of households that completed all four waves, to compare whether the estimates produced are similar. Tables A.4-5 show the sign and size of the effect of income on health to be generally unchanged. We do find a larger effect of income on height-for-age. A ten percent increase in income increases the height-for-age z-score by about 0.3. This result might be determined by infant mortality or non-response across waves.

6. Conclusions

Economic studies have debated the relationship between income and health. Because health is both a production and consumption good, the relation is likely to be affected by reverse causality.

We examined whether income shocks affect a range of health outcomes and preventative behaviours by matching satellite information on timing and positioning of 21 rainfall stations to individual-longitudinal data in 51 villages of the Kagera region in Tanzania. Agriculture accounts for about half of gross production and employs about 80 percent of the labour force in Tanzania making rainfall measurements an ideal instrument for income shocks.

We find a pro-cyclical effect of income on health. A ten percent increase in income reduces by 0.2 the number of illnesses. A further finding is the positive effect on vaccinations of children under six: a 10 percent increase in income implies an increase of about one vaccination, from a mean of 2.3 per child, for the four vaccinations of polio, tetanus, tuberculosis and measles. There is also some evidence of a reduction in chronic malnutrition of children under six. These results are robust to potential attrition biases and within-household correlation in health outcomes and behaviours.

Our results suggest that the income effect offsets the increased opportunity cost of time in this data. They imply that improving the resilience of rural households in developing countries in the face of agricultural shocks, and increasing their income is likely to have significant secondary benefits in terms of improving health outcomes and investments in prevention.

References

Adda J., Banks J. and von Gaudecker H. 2009. The Impact of Income Shocks on Health: Evidence from Cohort Data. *Journal of the European Economic Association*, December 2009 7(6), pp.1361–1399

Apouey B. and Clark A. E. 2011. Winning big but feeling no better? The effect of lottery prizes on physical and mental health. Working paper, Paris School of Economics. Accessed on 16th February 2013 at http://www.econstor.eu/bitstream/10419/53295/1/646135104.pdf

Artadi E. V., 2005. Going into Labor: Earnings vs. Infant Survival in Rural Africa. Mimeograph Harvard University. Accessed on 9th June 2013 at http://www.econ.upf.edu/docs/seminars/vilartadi.pdf

Baum C., M. Schaffer and S. Stillman, 2007. Enhanced Routines for Instrumental Variables/GMM Estimation and Testing *Boston College Economics Working Paper no. 667*.

Bengtsson, N. 2010. How responsive is body weight to transitory income changes? Evidence from rural Tanzania. *Journal of Development Economics*: 92 pp. 53 – 61.

Case, A. 2004. Does money protect health status? Evidence from South African pensions. In *Perspectives in the Economics of Aging*, Wise D (ed.). University of Chicago Press: Chicago; pp. 287 – 311.

Cheng F. and Liu C., 2012. Estimation of the spatial rainfall distribution using inverse distance weighting (IDW) in the middle of Taiwan. *Paddy Water Environ.*

Crépon B., Devoto F., Duflo E. and Parienté W. 2011. Impact of microcredit in rural areas of Morocco: Evidence from a Randomized Evaluation. Working Paper. Accessed on 16th February 2013 at http://www.feedback.clearfund.org/files/DWDA%202009/IPA/Cr%C3%A9pon,%20Devoto,%20Duflo,%20Parient%C3%A9%202011.pdf

Datar A., Liu J., Linnemayr S. and Stecher C., 2013. The impact of natural disasters on child health and investments in rural India. *Social Science & Medicine,* Vol.76 pp.83-91.

Dehejia, R., and Lleras-Muney, A.L. 2004. Booms, Busts, and Babies' Health. *Quarterly Journal of Economics*; 119: 1091–1130.

Dercon, S. and Krishnan P. 1998. Changes in poverty in rural Ethiopia 1989-1995: measurement, robustness tests and decomposition, *Center for Economic Studies,* Discussion Paper Series DPS98.19, Katholieke Universiteit Leuven. Accessed on 16th February 2013 at http://www.econ.kuleuven.ac.be/eng/ew/discussionpapers/DPS98/DPS9819.pdf

Dirks KN, Hay JE, Stow CD, 1998. High resolution studies of rainfall on Norfolk Island Part II: interpolation of rainfall data. *J Hydrol* 208:187–193.

Edoka I. O., 2013. Weather shocks and nutritional status of disadvantaged children in Vietname. Health Econometrics and Data Group, The university of York, WP 13/10.

Ettner, S. L. 1996. New evidence on the relationship between income and health. *Journal of Health Economics*, 15: pp. 67 – 85.

FAO 2009 FAOSTAT Database. Food and Agriculture Organization, Rome. Available at: http://faostat.fao.org Accessed 30th May 2013.

Fichera E. and Gathergood J. 2013. House prices, home equity and health. HEDG WP 13/01 University of York, Last accessed 30th May 2013 and available at: http://www.york.ac.uk/media/economics/documents/herc/wp/13_01.pdf.

Frijters P., Haisken-DeNew J.P. and Shields M.A. 2005. The causal effect of income on health: Evidence from German reunification. *Journal of Health Economics*, 24: pp. 997 - 1017

Kessy F. 2005. Rural Income dynamics in Kagera Region, Tanzania. World Bank and Economics and Social Research Foundation (ESRF). Available at: http://www.edi-africa.com/docs/Kessy%202004,%20full%20report.pdf. Last Accessed: 30th May 2013.

Beomsoo K. and Ruhm C.J. 2012. Inheritances, health and death. *Health Economics*: 21 pp. 127 – 144.

Mahajan A., Tarozzi A., Yoong J. and Blackburn B. 2011. Bednets, Information and Malaria in Orissa. Unpublished from Duke University. Accessed at http://public.econ.duke.edu/~taroz/MahajanTarozzi2011DC.pdf on 15h February 2013.

Meer J., Miller D.J., and Rosen H. S. 2003. Exploring the health-wealth nexus. *Journal of Health Economics*: 22(5) pp. 713 – 730.

Michaud, P. and van Soest A. 2008. Health and Wealth of Elderly Couples: Causality Tests Using Dynamic Panel Data Models. *Journal of Health Economics*: 27(5) pp. 1312 – 1325.

Miguel E., Satyanath S. and Sergenti E. 2004. Economic Shocks and Civil Conflict: An Instrumental Variables Approach. *Journal of Political Economy*. 2004, vol. 112, no. 4 pp.725-53.

Miller G. and Urdinola B.P.. 2010. Cyclicality, Mortality, and the Value of Time: The Case of Coffee Price Fluctuations and Child Survival in Colombia. *Journal of Political Economy*, 2010 February; 118(1): pp.113-5.

National Bureau of Statistics, 2007. Tanzania National Household Budget Survey. Available at http://www.nbs.go.tz/tnada/index.php/catalog/2. Last accessed on 30th

May 2013.

Pörtner C. C., 2010. Natural Hazards and Child Health. University of Washington, Seattle. Available at SSRN, last accessed: 9 June 2013 at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1599432

Rose, E. 1999. Consumption Smoothing and Excess Female Mortality in Rural India. *Review of Economics and Statistics*, February 1999, 81(1) pp.41 – 49.

Ruhm C. J., 2003. Good Times Make You Sick. *Journal of Health Economics*, 22(4): 637-658.

Ruhm, C. (2005) Healthy living in hard times. *Journal of Health Economics*, 24: 341-363.

Snyder, S.E., and Evans W. 2006. The Impact of Income on Mortality: Evidence from the Social Security Notch. *Review of Economics and Statistics*: 88 pp. 482 - 495.

Thurlow J. and Wobst P. 2003. Poverty focused social accounting matrices for Tanzania. *IFPRI Discussion Paper No. 112, International Food Policy Research Institute, Washington DC.*

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

World Bank, 2004. User's guide to the Kagera health and development survey datasets. *World Bank Development Research Group*.

Yamano, T., Alderman H. and Christiaensen L. 2005. Child Growth, Shocks, and Food Aid in Rural Ethiopia. *American Journal of Agricultural Economics.* 87(2), May 2005, pp.273–288

Zhuang LW, Wang SL., 2003 Spatial interpolation methods of daily weather data in Northeast China. *Quart J Appl Meteorol* 14(5):605–616.

Table 1. Definition of variables and descriptive statistics, KHDS (1991-1994)

Variables	Variables definition	Obs.	Mean	Std. Dev.	Min.	Max.
Outcome variables:						
Ln(BMI)	Natural logarithm of Body Mass Index.	17,817	2.89	0.19	2.14	3.88
No. conditions	Self-reported health symptoms in the past 4 weeks.	17,847	0.90	1.07	0	5
Height-for-age	z-scores of height for age of children under the age of six.	3,215	0.24	3.78	-18.96	46.81
Weight-for-height	z-scores of weight for height of children under the age of six.	3,196	-0.22	1.28	-14.53	18.06
No. vaccinations	No. vaccinations against measles, tetanus, TB and polio of children under the age of six.	3,220	2.29	1.12	0	4
Covariate variables:						
Age	Age in years.	17,847	22.31	20.10	0	110
Age squared	Squared value of age in years.	17,847	901.92	1463.5	0	12100
Household size	No. of household members.	17,847	8.41	4.19	1	36
Ln(income)	Natural logarithm of real equivalised household income in TZS.	17,847	12.34	1.11	0.10	16.31
Long rains	1=long rains season March-May; 0=short rains or dry.	17,847	0.17	0.38	0	1
Short rains	1=short rains season October-December; 0=long rains or dry.	17,847	0.31	0.46	0	1
Dry season	1=dry season June-September; 0=long or short rains.	17,847	0.35	0.48	0	1
Instrumental variable:						
Ln(rainfall)	Natural logarithm of average monthly rainfall in previous 12 months (mm).	17,344	4.83	0.31	4.18	5.65

Note that statistics on covariate and instrumental variables are calculated on the no. of conditions sample.

Table 2. Outcome variables by income quintiles

Income quintiles	Bottom quintile	2nd quintile	3rd quintile	4th quintile	Top quintile
Ln(BMI)	2.88 (3,580)	2.88 (3,550)	2.89 (3,568)	2.89 (3,574)	2.91 (3,545)
No. conditions	0.96 (3,586)	0.93 (3,561)	0.86 (3,576)	0.86 (3,576)	0.87 (3,548)
Height-for-age	0.16 (628)	0.11 (650)	0.12 (656)	0.46 (678)	0.36 (603)
Weight-for-height	-0.41 (627)	-0.24 (648)	-0.11 (651)	-0.22 (673)	-0.12 (597)
No. vaccinations	2.21 (628)	2.34 (652)	2.25 (657)	2.35 (679)	2.30 (604)

Note: sample means displayed and number of observations in ().

Table 3. Basic OLS Models

	Model I: Ln(BMI)	Model II: No. conditions	Model III: Height-for-age	Model IV: Weight-for-height	Model V: No. vaccinations
Ln(income)	0.004*** (0.001)	-0.05** (0.02)	0.06 (0.07)	0.11*** (0.03)	0.38*** (0.08)
Age	0.003** (0.001)	0.03 (0.03)	-	-	0.37*** (0.08)
Age squared	-0.0001*** (0.00)	0.0001 (0.0002)	-	-	-0.07*** (0.01)
HH size	-0.001* (0.001)	0.02 (0.01)	0.07 (0.06)	-0.05** (0.02)	-0.05* (0.03)
Constant	2.83*** (0.03)	0.63 (0.51)	-0.74 (1.06)	-1.24 (0.43)	-1.70 (1.05)
N. observations	*17,817*	*17,847*	*3,215*	*3,196*	*3,220*
N. individuals	*6,130*	*6,136*	*1,322*	*1,315*	*1,324*

Note: std. errors in () clustered by villages. All models include year-cluster interaction effects, a seasonal dummy indicating the short rainy season, a month linear time trend and individual fixed effects.
****p<0.01, **p<0.05, *p<0.10.*

Table 4. 2SLS models of health outcomes

	Model I:		Model II:	
	1st stage: **Ln(Income)**	***2nd stage:*** **Ln(BMI)**	***1st stage:*** **Ln(Income)**	***2nd stage:*** **No. conditions**
Ln(rainfall)	2.77*** (0.43)	-	2.77*** (0.43)	-
Ln(income)	-	0.01 (0.004)	-	-0.17*** (0.06)
Age	0.02 (0.02)	0.004** (0.002)	0.03 (0.02)	0.03 (0.03)
Age squared	-0.001*** (0.0002)	-0.0001*** (0.00)	-0.001*** (0.0002)	0.00003 (0.0002)
HH size	0.09*** (0.02)	-0.002** (0.001)	0.09*** (0.02)	0.03* (0.01)
N. Observations	*16,077*		*16,105*	
N. individuals	*4,843*		*4,848*	
Kleibergen-Paap F-stat (H_0=weak IV)	*40.93*		*41.19*	

Note: std. errors in () clustered by villages. All models include year-cluster interaction effects, a seasonal dummy indicating the short rainy season, a month linear time trend and individual fixed effects.
****p<0.01, **p<0.05, *p<0.10.*

Table 5. 2SLS models of health outcomes and behaviours

	Model I:		**Model II:**		**Model III:**	
	1st stage: **Ln(Income)**	***2nd stage:*** **Height-for-age**	***1st stage:*** **Ln(Income)**	***2nd stage:*** **Weight-for-age**	***1st stage:*** **Ln(Income)**	***2nd stage: No. vaccinations***
Ln(rainfall)	2.66*** (0.52)	-	2.66*** (0.53)	-	2.64*** (0.52)	-
Ln(income)	-	0.33* (0.18)	-	0.02 (0.17)	-	0.86*** (0.15)
Age	-	-	-	-	0.13** (0.07)	0.30*** (0.08)
Age squared	-	-	-	-	-0.02** (0.01)	-0.06*** (0.01)
HH size	0.09*** (0.02)	0.07 (0.07)	0.09*** (0.02)	-0.04 (0.02)	0.09*** (0.02)	-0.10*** (0.03)
N. Obs.	*2,696*		*2,680*		*2,699*	
N. individuals	*899*		*899*		*899*	
Kleibergen-Paap F-stat (H_0=weak IV)	*25.66*		*25.51*		*25.54*	

Note: std. errors in () clustered by villages. All models include year-cluster interaction effects, a seasonal dummy indicating the short rainy season, a month linear time trend and individual fixed effects.
****p<0.01, **p<0.05, *p<0.10.*

Appendix A: Robustness checks

Table A.1 2SLS models of health outcomes (HH clustered std. errors)

	Model I:		**Model II:**	
	***1st stage:* Ln(Income)**	***2nd stage:* Ln(BMI)**	***1st stage:* Ln(Income)**	***2nd stage:* No. conditions**
Ln(rainfall)	2.77*** (0.23)	-	2.77*** (0.23)	-
Ln(income)	-	0.01** (0.003)	-	-0.17*** (0.05)
Age	0.02 (0.02)	0.004** (0.002)	0.03 (0.02)	0.03 (0.03)
Age squared	-0.001*** (0.0001)	-0.0001*** (0.00)	-0.001*** (0.0001)	0.00003 (0.0002)
HH size	0.09*** (0.02)	-0.002** (0.001)	0.09*** (0.02)	0.03* (0.01)
N. Observations	*16,077*		*16,105*	
N. households	*871*		*871*	
Kleibergen-Paap F-stat (H_0=weak IV)	*143.48*		*144.07*	

Note: std. errors in () clustered by households. All models include year-cluster interaction effects, a seasonal dummy indicating the short rainy season, a month linear time trend and individual fixed effects.
****p<0.01, **p<0.05, *p<0.10.*

Table A.2 2SLS models of health outcomes and behaviours (HH clustered std. errors)

	Model I:		**Model II:**		**Model III:**	
	***1st stage:* Ln(Income)**	***2nd stage:* Height-for-age**	***1st stage:* Ln(Income)**	***2nd stage:* Weight-for-age**	***1st stage:* Ln(Income)**	***2nd stage: No. vaccinations***
Ln(rainfall)	2.66*** (0.25)	-	2.66*** (0.25)	-	2.64*** (0.25)	-
Ln(income)	-	0.33 (0.23)	-	0.04 (0.02)	-	0.86*** (0.13)
Age	-	-	-	-	0.13** (0.06)	0.30*** (0.08)
Age squared	-	-	-	-	-0.02** (0.01)	-0.06*** (0.01)
HH size	0.09*** (0.03)	0.07 (0.07)	0.09*** (0.03)	-0.04 (0.02)	0.09*** (0.03)	-0.10*** (0.03)
N. Obs.	*2,696*		*2,680*		*2,699*	
N. individuals	*485*		*483*		*485*	
Kleibergen-Paap F-stat (H_0=weak IV)	*116.73*		*116.55*		*114.75*	

Note: std. errors in () clustered by households. All models include year-cluster interaction effects, a seasonal dummy indicating the short rainy season, a month linear time trend and individual fixed effects.
****p<0.01, **p<0.05, *p<0.10.*

Table A.3 Model of health outcome (Malaria)

	Malaria
Ln(rainfall)	0.08 (0.42)
Age	0.002 (0.06)
Age squared	0.0002 (0.0005)
HH size	0.05* (0.02)
N. Observations	*1,767*
N. individuals	*1,437*

Note: std. errors in () clustered by villages. All models include year-cluster interaction effects, a seasonal dummy indicating the short rainy season, a month linear time trend and individual fixed effects.
****p<0.01, **p<0.05, *p<0.10.*

Table A.4 2SLS models of health outcomes (balanced sample)

	Model I:		**Model II:**	
	1st stage: **Ln(Income)**	***2nd stage:*** **Ln(BMI)**	***1st stage:*** **Ln(Income)**	***2nd stage:*** **No. conditions**
Ln(rainfall)	2.53*** (0.53)	-	2.52*** (0.53)	-
Ln(income)	-	0.01 (0.005)	-	-0.18*** (0.07)
Age	0.02 (0.02)	0.005** (0.002)	0.02 (0.02)	0.04* (0.03)
Age squared	-0.001*** (0.0002)	-0.0001*** (0.00)	-0.001*** (0.0002)	0.0001 (0.0003)
HH size	0.09*** (0.02)	-0.002** (0.001)	0.10*** (0.02)	0.02 (0.02)
N. Observations	*11,091*		*11,105*	
N. individuals	*2,868*		*2,870*	
Kleibergen-Paap F-stat (H_0=weak IV)	*22.68*		*22.84*	

Note: std. errors in () clustered by villages. All models include year-cluster interaction effects, a seasonal dummy indicating the short rainy season, a month linear time trend and individual fixed effects.
****p<0.01, **p<0.05, *p<0.10.*

Table A.5 2SLS models of health outcomes and behaviours (balanced sample)

	Model I:		**Model II:**		**Model III:**	
	***1st stage:* Ln(Income)**	***2nd stage:* Height-for-age**	***1st stage:* Ln(Income)**	***2nd stage:* Weight-for-age**	***1st stage:* Ln(Income)**	***2nd stage: No. vaccinations***
Ln(rainfall)	2.50*** (0.60)	-	2.50*** (0.60)	-	2.50*** (0.60)	-
Ln(income)	-	0.29*** (0.11)	-	0.14 (0.25)	-	0.93*** (0.11)
Age	-	-	-	-	0.03 (0.08)	0.31*** (0.11)
Age squared	-	-	-	-	-0.005 (0.01)	-0.06*** (0.02)
HH size	0.09*** (0.03)	0.13* (0.06)	0.09*** (0.03)	-0.09** (0.04)	0.09*** (0.03)	-0.15*** (0.04)
N. Obs.	*1,691*		*1,679*		*2,699*	
N. individuals	*491*		*487*		*899*	
Kleibergen-Paap F-stat (H_0=weak IV)	*17.25*		*17.12*		*25.54*	

Note: std. errors in () clustered by villages. All models include year-cluster interaction effects, a seasonal dummy indicating the short rainy season, a month linear time trend and individual fixed effects.
****p<0.01, **p<0.05, *p<0.10.*