Statistisk sentralbyrå
Statistics Norway

Challenges in predicting poverty trends using survey to survey imputation

Experiences from Malawi

Astrid Mathiassen and Bjørn K. Wold

TALL SOM FORTELLER

DISCUSSION PAPERS

900

Discussion Papers No. 900, March 2019
Statistics Norway, Research Department

Astrid Mathiassen and Bjørn K. Wold

Challenges in predicting poverty trends using survey to survey imputation.
Experiences from Malawi

Abstract:
Poverty in low-income countries is usually measured with large and infrequent household surveys. A challenge is to find methods to measure poverty more frequently. The objective of this study is to test a method for predicting poverty, based upon a statistical model utilizing consumption surveys and light annual surveys. A decade of poverty predictions and regular poverty estimates in Malawi provides us with a unique real-life experience to better understand the suitability of such approaches to monitor trends in poverty.

The analysis from Malawi suggests that a modelling approach works per se, given that information on the household's demographic composition is included in the model. The main challenge when predicting onto other surveys seems to be related to comparability between the surveys. Differences in implementation, questionnaire design and survey sample size are aspects that may contribute to incomparability of data collected between the surveys.

Keywords: Survey-to-survey imputation, poverty measurement, poverty model, household surveys, Malawi.

JEL classification: C21, C81, D12, I32

Acknowledgements: We are grateful to participants from Malawi 's National Statistical Office (NSO), IFPRI-Malawi and World Bank-Malawi, for valuable discussions in a meeting that took place in Lilongwe in March 2018 where the results from this analysis were presented. Thanks also to Talip Kilic (from World Banks LSMS-team) for discussions of the results and Mark Schreiner for reviewing the paper. We are also grateful for valuable comments from colleagues in Statistics Norway and would like to mention John Dagsvik, Julie Hass, Ellen Cathrine Kiøsterud, Vibeke Oestreich Nielsen and Terje Skjerpen. Thanks to Norad for funding this project and to NSO-Malawi for sharing the data.

Address: PO Box 8131 Dept, NO-0033 Oslo, Statistics Norway, Division for Development Cooperation. E-mail: Astrid.Mathiassen@ssb.no

Discussion Papers comprise research papers intended for international journals or books. A preprint of a Discussion Paper may be longer and more elaborate than a standard journal article, as it may include intermediate calculations and background material etc.

Abstracts with downloadable Discussion Papers
in PDF are available on the Internet:
http://www.ssb.no/en/forskning/discussion-papers
http://ideas.repec.org/s/ssb/dispap.html

ISSN 1892-753X (electronic)

Sammendrag

Utrydding av fattigdom er et hovedfokus i bærekraftsmålene. Andelen under fattigdomsgrensa er hovedindikatoren for å måle fremskritt mot dette målet. Standardmetoden for å anslå dette tallet er basert på informasjon om husholdningers detaljerte forbruk. En forbruksundersøkelse er både kostbar og tidkrevende og blir i utviklingsland ofte gjennomført kun hvert fjerde eller femte år. Offisielle fattigdomstall er derfor bare tilgjengelige med slike mellomrom og det er behov for billigere og raskere metoder for å gi en årlig oppdatering på fattigdom.

En tilnærming for mindre ressurskrevende fattigdomsrapportering går ut på å lage en modell basert på en forbruksundersøkelse som forklarer sammenhengen mellom husholdningers totalforbruk og karakteristikker (fattigdomsindikatorer). Tilnærmingen har blitt brukt til å gi årlig fattigdomsanslag i Malawi. Kort oppsummert kan metoden beskrives som følger. Forbruksundersøkelsen i Malawi i 2004-2005 (IHS2) ble brukt til å identifisere en modell for å estimere forbruket. Fattigdomsindikatorene som inngår i modellen ble samlet inn i en mindre omfattende undersøkelse, Welfare and Monitoring Survey (WMS), og ble brukt til å predikere andelen fattige. WMSer ble gjennomført årlig til den neste forbruksundersøkelsen i 2010 (IHS3). Andelen fattige basert på WMS undersøkelsene viste en gradvis nedgang fra 2005 til 2009. Fattigdommen basert på den nye forbruksundersøkelsen viste derimot ingen endring sammenliknet med nivået i 2004. Dette ledet til diskusjon både i og utenfor Malawi rundt modellberegningene og rundt de offisielle fattigdomsberegningene. På grunn av usikkerheten rundt anslagene sluttet statistikkbyrået i Malawi å bruke metoden, selv om de fortsatte å samle inn nødvendig informasjon i påfølgende WMS'er. Erfaringene og datagrunnlaget fra Malawi er unikt og brukes i denne analysen for å evaluere og justere metoden.

Analysen av Malawi i denne studien gir støtte til at en slik metode virker, per se, gitt at demografiske forklaringsvariabele er inkludert i modellen. Uten demografiske variable, som antall medlemmer i husholdet, predikere modellen systematisk for lav fattigdom. Den største utfordringen i å bruke tilnærmingen til å lage trender i fattigdomsutvikling, har derimot å gjøre med sammenlignbarhet mellom undersøkelsene. Forskjeller i implementering, spørreskjemautforming og utvalgsstørrelse er aspekter som kan bidra til problemer med sammenlikning av fattigdom over tid.

1. Introduction

Eradicating poverty was the first Millennium Development Goal (MDG) during the period 2000 – 2015 (UN, 2000) and named as the Primus inter paris among the MDGs (Kanbur, 2005).This focus has been retained by making the goal to end extreme poverty by 2030 as the first Sustainable Development Goal (SDG) by the UN General assembly in 2015 and the first of only two goals by the World Bank (2018a)[1].

It is a global consensus (UN, 2015) that in order to follow the policy goals and target of eradication and ending poverty requires the measurement of poverty headcounts. The standard approach to estimate this number is based on comprehensive survey data on households' detailed consumption. Such surveys are costly and time consuming and are, in developing countries, often undertaken only every 4th or 5th year. Cheaper and quicker methods to report on poverty on an annual basis are needed by both the national and international communities.

Survey-to-survey imputation approaches have been developed to fill this gap. The National Statistical Office (NSO) in Malawi has applied such an approach to predict annual poverty rates in Malawi, NSO (2010). In short, the Integrated Household Survey 2004[2] (IHS2) in Malawi was used to identify a model with variables (predictors) suited to predict poverty. The predictors were collected in the smaller, annual Welfare and Monitoring survey (WMS) and annual model-based poverty rates were predicted based on the IHS2 model and the predictors from 2005 till 2009.

The results suggested a gradual reduction of poverty: In 2004 the official national poverty headcount (calculated directly from the IHS2-survey) was 52 percent, whereas according to the model-based poverty estimates, poverty gradually decreased to 39 percent in 2009. This trend is consistent with an increase in real GDP per capita and an increase in production of maize, the main staple food. On the other side, official poverty numbers for 2010[3] based on a new Integrated Household Survey (IHS3), showed that poverty levels have hardly improved since 2005. This puzzle has raised the question: does the decreasing poverty

[1] The second goal of the World Bank is to promote shared prosperity by fostering the income growth of the bottom 40% for every country. Both goals require information on the total consumption across the population, an indicator requiring the same information as addressed in this document.

[2] Although the survey was undertaken over 12 months in 2004-05, we will for simplicity refer to it as 2004.

[3] Although the survey was undertaken over 12 months in 2010-11, we will for simplicity refer to it as 2010.

trend predicted by the model reflect real changes in Malawi, or was the model wrong? The present work reflects upon this question.

The prediction approach applied in Malawi builds on a poverty mapping method developed by Elbers et al. (2003), that has been moderated for survey-to-survey imputation, see for example Mathiassen (2009). This approach has been tested by predicting from one consumption survey onto other identical surveys. A series of seven households budget surveys from Uganda was used to validate the methods which showed promising results (Mathiassen, 2013). Other studies have had a similar objective; Vu and Baulch (2011) evaluate four "short cut" methods for predicting poverty[4] by using data from Vietnam; where a budget survey is used to predict onto two other budget surveys. They find that the probit method provides the most accurate prediction. The probability method tested by Vu and Baulch (2011) is similar to the approach tested in this study. Newhouse et al. (2014) found that a similar approach to the one used for Malawi fails when imputing poverty from household budget surveys into labour force surveys using data from Sri Lanka. They argue that for such a set up to produce reliable poverty estimates, a welfare tracking survey should be established. That would imply, for the Sri Lanka case, that the labour force survey included additional questions on housing and assets and that sampling design and questions used for predictors are consistent between the surveys. A welfare tracking system as recommended by Newhouse et al. (2014) is in practice what was established in Malawi.

Another related method is the Scorocs (TM) Simple Poverty Scorecard® poverty-assessment tool. It collects 10 verifiable indicators to estimate poverty likelihood using a model based on a budget survey (Schreiner, 2014). It has been developed and is used for programming purposes in several countries.[5,6]

Other studies have tried to understand the puzzle in Malawi where there is a stagnant (official) poverty level between the IHS2 and IHS3 surveys despite other economic indicators suggesting improvements in this period. A number of methodological issues in setting the poverty threshold and estimating poverty in Malawi are considered in a recent work by Pauw

---

[4] Poverty probability method, ordinary least squares, principal component and quantile regression.

[5] See www.simplepovertyscorecard.com for the list of countries.

6 Schreiner (2014) measures the accuracy for the scorecard between two surveys with compatible definitions of consumption and poverty lines for 19 countries. In general, he finds "… accuracy to be less than I hoped and often less than would appear useful (for example, signs are wrong, or errors exceed 5 percentage points)" (personal communication, September 5, 2018).

et al. (2016). Contrary to the official estimates showing almost no changes in poverty between 2004 and 2010, Pauw et al. (2016) estimate that poverty declined by 8.4 percentage points. This study also documents improvements in a number of other non-monetary welfare indicators consistent with a decline in poverty level. The survey experiment documented in Kilic and Sohnesen (2019) is another attempt to understand the poverty puzzle in Malawi. The experiment was undertaken with the collection of IHS3 data, aiming at understanding how context affects answers to the same question. Their experiment shows that questionnaire design has consequences for the underlying predictors and could move the poverty level predicted by a similar model used for the poverty trend in Malawi, with 3 to 7 percent. Thus, suggesting that the downward trend predicted by the WMS surveys were, at least partly, due to differences in the survey instruments.

Because of the uncertainty around the model-based predictions based on WMS2005-WMS2009 the Malawi NSO stopped calculating such numbers from the following WMSes, although they continued collecting the information necessary for doing so. After WMS2009 three additional surveys are available, and it is possible to calculate poverty trends including official poverty numbers and model-based predictions for the period from 2004 to 2014. This survey material, including a total of six WMSes and three IHSes will be used to validate the model-based predictions.

There are two main ways the present study approaches the validation. The first is to test results within the same context, i.e. predicting within or onto another IHS survey. The second is to test results when predicting in another context, i.e. predicting onto WMS surveys.

Three approaches are applied for the tests within the same context: Firstly, predicting within IHS sample and comparing to the known (actual) poverty level in the other half sample is a direct test of how well the models work, everything else being equal. Secondly, predicting from one IHS-survey to the other is a test of the models' stability over time. Even if the models work well at the same point of time, the relationship between the variables in the models may change. The test onto another IHS-survey will provide us with indications, but not solid proofs, as there are comparability issues even between the same type of survey. Thirdly, the analysis discusses different types of predictors best suited to predict poverty, again this is done by comparing predictions within the sample. The question is: Do some predictors bias the predicted poverty level – compared to the actual poverty level?

Even if a model is suited to predict within the sample or onto another identical survey, there may be other challenges when predicting onto another type of survey, i.e. another context. This is discussed by comparing WMS trends predicted by the IHS2 and IHS3 models. Will models developed from different surveys provide different prediction trends? Such analysis will help us to understand whether the models are stable over time, or whether the relationship between predictors and household welfare changes over time. The WMS surveys do not cover a full year and the season for the survey period was not properly accounted for previously. Rather, the model-based approach only predicted for the season covered in the WMS. This paper develops a way to include seasonality in the model. Further, it discusses the effect of differences in questionnaire design. The implication of the findings in Kilic and Sohnesen (2019) is that the same information collected in IHS and WMS may differ – not due to real changes, but due to the context the questions were framed in. Although the questions to capture the poverty predictors are the same, the WMS questionnaire is much shorter than the IHS, and the questions are not followed up with additional probing. For example, regarding food consumption, the households in both WMS and IHS were asked a yes/no question to whether they consumed the specific food, while in IHS there were additional questions regarding how much they consumed. Thus, both responders fatigue, as well as the elaboration on questions can affect the answers, see Lavrakas (2008) for a review around this theme. The analysis in the present study discusses this element by comparing model-based poverty trends with and without predictors that are expected to be most affected by the context in the questionnaire.

Although hard to quantify, we also discuss whether the trends may have been affected by differences in implementation of the surveys. If, for example, training of enumerators and the organisation of data collection differ this will have a bearing on the results. Such differences may not be easy to measure but the following factors may affect the results: size of surveys; type of survey, and donor support.

The understanding of Malawi's experience is important for policy makers and statisticians in Malawi. It is also important as the poverty scorecard method has been developed and is used for programming purposes in Malawi, Schreiner (2015). A better understanding of the Malawian case is also valuable for the international community, as poverty models are increasingly applied and are potentially useful for annual reporting on SDGs.

The next section gives some background about Malawi. Section 3 describes the data and Section 4 explains the methodology used. The results are presented in Section 5 and discussed in Section 6. Section 7 provides some concluding remarks.

2. Background/context

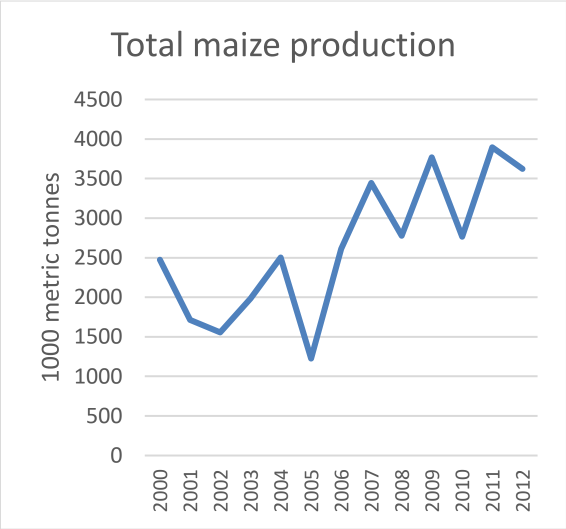
Malawi is a developing country in Sub-Saharan Africa, with the majority of its 18 million people living in rural areas (NSO, 2016). About 80 percent of the population is engaged in agriculture which is Malawi's main economic sector generating about 30 percent of the gross domestic product, GDP (NSO, 2016). The main agricultural strategy in Malawi has, for many years, been to produce tobacco for export and to produce maize to ensure food security for the rural and urban population. Maize is cultivated across the country and the value of the production is twice that of tobacco and accounts for about 25 percent of the agricultural economy (NSO, 2016). High dependency on agriculture, and on one crop in particular, makes Malawi vulnerable to climatic variability and there are droughts or floods or both almost every year (Government of Malawi, 2015). In 2004, the Government of Malawi introduced a smallholder-targeted fertilizer subsidy program (FISP) with the purpose of improving food security and welfare. Malawian smallholders were to be provided with sufficient fertilizer and seeds to satisfy the maize consumption needs of an average-sized family (Pauwet al., 2016). In practice about half of all farmers, irrespective of landholding size, benefitted from this program in 2009 (Kilic et al., 2013).

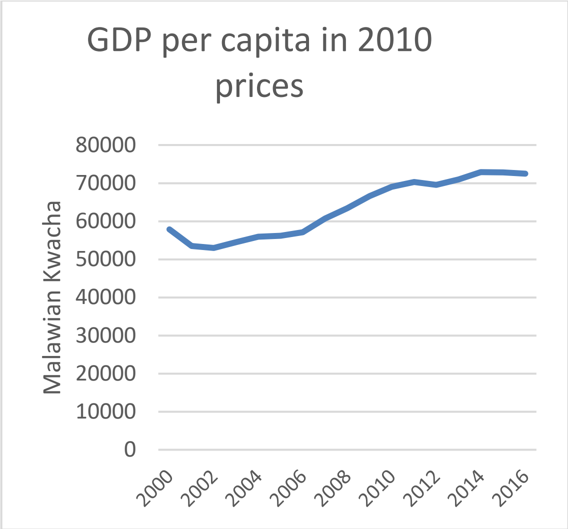
A number of studies have argued that the program has had a positive impact on yield and food security (Chirwa and Dorward, 2013; Carr, 2014; Pauw et al., 2014 and Haug and Wold, 2017). Arndt et al. (2014) estimate that the direct effect of each dollar spent on FISP generates 1.65 US dollars in direct welfare benefits, and that the indirect effects would increase the benefits with another 70 percent. Haug and Wold (2017) argue that the FISP has proven to be the cheapest approach to ensure food security over the years, as the FISP program yielded a surplus for all farmers in years with good climate conditions and even created a buffer for seasons with drought.

Since 2005/06 (when the FISP was introduced), relatively favorable weather conditions, combined with the input subsidies seem to have led to rapidly increasing maize yields.

As shown in Figure 1, in the same period, the GDP per capita steadily increased. From 2004 to 2010 maize production increased by 11 percent and GDP per capita by 23 percent.

Figure 1. Maize production and GDP per capita

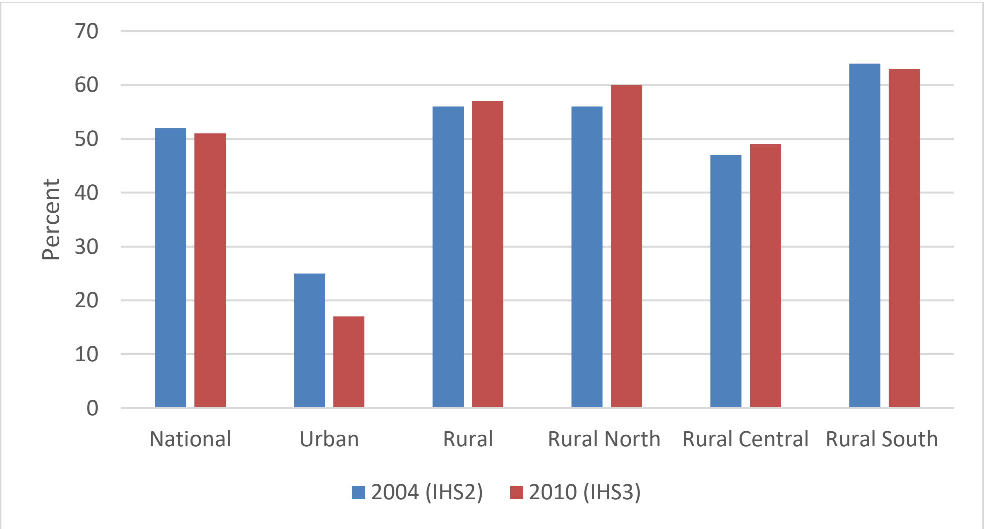




Source: Respectively World Bank (2018b) and MOAIWD (1997–2015).

Figure 2 shows that the official poverty level calculated from IHS2 and IHS3, however, did not reflect the agricultural and economic improvements. Only in urban areas was there a significant reduction in poverty level between 2004 and 2010 – where poverty dropped by about 8 percentage points.

Figure 2. Official (actual) poverty in Malawi



Source: NSO (2011).

3. Data

About the dataset

The data sets used in the analysis consist of two Integrated Household Surveys (IHS), one Integrated Household Panel Survey (IHPS) and six Welfare and Monitoring Surveys (WMS), See **Table *1*** for an overview of these surveys. The IHS2 (2004) and IHS3 (2010) are large surveys covering respectively, 11,280 and 12,271 households. The IHPS was the smallest survey covering 4,000 households. The questionnaires for the IHS2, IHS3 and IHPS are almost identical, with only minor changes. They contain detailed information about consumption and expenditures, and can be used to calculate total consumption for the households, and therefore poverty. The WMS surveys were conducted annually from 2005 to 2009, and again in 2011 and 2014. These are lighter surveys that do not provide information on consumption expenditure, but aim to track welfare in a number of areas, such as education, health, employment and asset ownership. The questionnaires for the WMSs remained largely unchanged from 2005 till 2009. There were some changes in 2011 and 2014 compared to the previous WMS questionnaires, in particular with respect to the placement of modules. In addition, a large module on peace and governance was added to the WMS2014. In 2014, data collection was for the first time done electronically by using CAPI[7] technique. The survey sample varied a great deal, from 5,234 (2005) to 29,389 (2007) households. There was some obvious quality flaws with WMS2011, and it had to be dropped from the further analysis: About 20 percent of the households did not report any information regarding food consumption.

Table 1. About the surveys

Name and year	IHS2 2004/5	WMS 2005	WMS 2006	WMS 2007	WMS 2008	WMS 2009	IHS3 2010/11	IHPS 2013	WMS 2014
Number of households	11280	5234	5287	29389	17857	20673	12271	4000	14198
Type of survey	IHS	WMS	WMS	WMS+ NACAL	WMS	WMS	IHS	IHS-panel survey	WMS
Institutions involved	NSO/WB	NSO/SSB	NSO/SSB	NSO/SSB	NSO/SSB	NSO/SSB	NSO/WB	NSO/WB	NSO/SSB
Seasons	All	Q3	Q3	Q3	Q3, Q4	Q3, Q4	All	Q3, Q4	Q1, Q4
Direct measure of poverty	yes						yes	yes	

Note: Q denotes quarter of the year, i.e. Q1=1st quarter (January-March).

[7] Computer Assisted Personal Interviewing

Table 1 also shows the type of survey. WMS2007 was attached to an agricultural census (NACAL). An extra sample had to be drawn for the WMS to include landless households. In the end this double sampling approach made it necessary to recalculate the household weights for 2007. The IHPS was a panel survey where 3,246 households from the IHS3-survey were revisited in the 2013. Individuals, rather than households were followed and if one individual moved into another household also that household would be sampled in IHPS.

Not all surveys covered the whole year, and the fourth row in Table 1 shows the quarters (also referred to as seasons) covered in each survey.[8] The WMS was initially designed only to cover the months from July-September, season 3. However, for 2008 and 2009 the fieldwork also spanned into season 4, and in 2014 the survey in fact covered parts of 2013 (season 4) and parts of 2014 (season 1). Note also that the IHPS in 2013 only covered two seasons.

In addition to the NSO, the World Bank (WB) and Statistics Norway (SSB) were involved in implementing the surveys as presented in the 4th row in Table 1. WB was supporting the IHS surveys with respect to questionnaire design, sampling, fieldwork and preparation of data. SSB was giving support to all WMS's but to various degrees. The Norwegians involvement in the WMS-surveys was strong in 2005, 2006 and 2007 and included support in questionnaire design, sampling, fieldwork and preparation of data. The work was supported by a long-term advisor from Statistics Norway. For the following WMSs the technical support from Statistics Norway was limited to an advisory role. However, in 2014 Statistics Norway supported the redesign of the questionnaire into an electronic format and the pilots in using tablets.

Only for the integrated household surveys covering a full consumption expenditure module, is it possible to directly measure poverty, see the 6th row in Table 1.

There were some changes in the way the consumption aggregate was calculated IHS-surveys, which may cause issues in the comparability of consumption and poverty between the IHSes. The IHS3 developed new and improved conversion factors for transformation of all non-standard units into kilograms, at the same time they kept the conversion factors used in IHS2 ("old" conversion factors). In the end the "old" set of conversion factors were kept

[8] Season1 (Q1) covers January to March, Season 2 (Q2) April-June, Season 3 (Q3) July-September and Season 4 (Q4) covers October-December.

for the IHS3 analysis with one exception: The factors for "pails" of normal and refined maize flour were replaced by new factors estimated from a supplementary survey conducted in markets in all districts in the country during February and March 2011. According to IHS3-survey report (NSO, 2012) page 228: "The reasons for this revision were that the previous factors were not considered to be accurate enough and that a significantly larger proportion of households in the IHS3 (compared to the IHS2), reported the consumption of maize flour in pails".

In IHPS (2013), the new conversion factors were used for all foods. In addition, a set of new price indices to adjust nominal consumption for cost of living differences was estimated. These two changes imply that the consumption and poverty status of the panel households are not comparable to the poverty estimated in IHS2 and IHS3. Thus, we do not compare the predicted poverty numbers for IHPS to the actual poverty calculated from this survey.

Even with no changes in methodology, the poverty line used with the different surveys may affect comparability. The standard approach to set the poverty line (the Cost of Basic Needs) has some elements of relativity in it, being anchored in the consumption pattern of the poor as observed from the survey (Ravallion, 1998). Consequently, if the consumption pattern changes, so will the poverty line. In Malawi the poverty line was estimated based on the IHS2 survey and updated in IHS3 and IHPS to account for changes in prices.[9] As time passes it can be argued that this poverty line is no longer relevant as it may no longer reflect the consumption pattern of the poorer part of the population. This was indeed one of the points raised by Pauw et al. (2016). Another point they considered in their recalculation of poverty in Malawi, was a revised set of conversion factors to convert food consumption into kilograms. The revised set of conversion factors was developed by Verduzco-Gallo et al. (2014) and was applied to both IHS2 and IHS3. Pauw et al. (2016) do not separate out the effects on poverty of the various aspects of change in methodology and all in all they estimated a decrease in poverty level from IHS2 till the IHS3 at more than eight percentage points at national level.

While the methodology used for calculating total consumption in the household is important for the analysis on whether the survey-to-survey imputation works, the level of the poverty line will not affect the estimation.

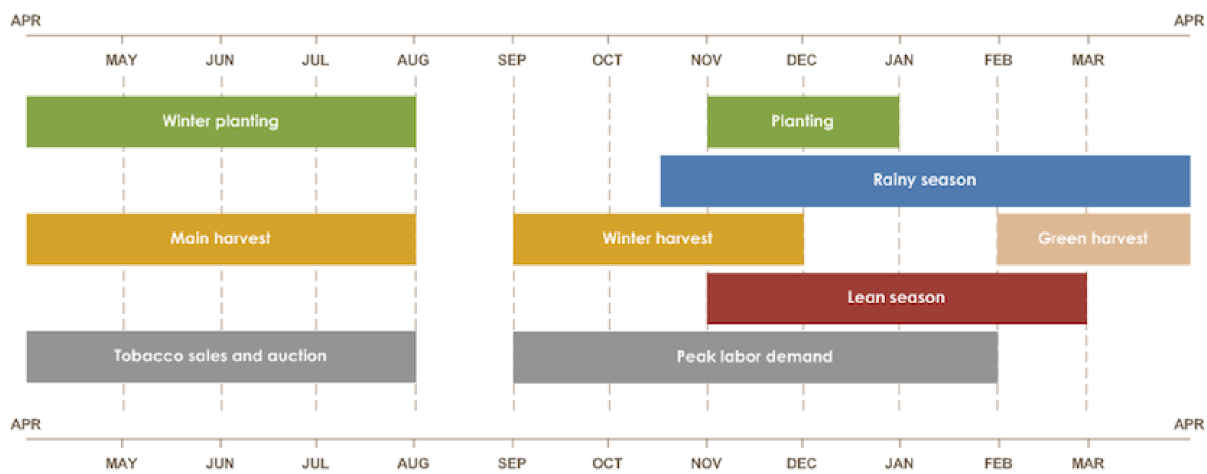
[9] See World Bank (2018c) for details on how the poverty line in Malawi was calculated.

Seasonality

Figure 3 shows the seasonal calendar for Malawi. The majority of households in Malawi are rural smallholders with a consumption pattern following the seasonal variation. The main planting season starts with the rainfall in the fourth quarter. Wild plants and green maize become available in the first quarter, while the main harvest starts in the second quarter. Crop produce are in abundance and hence cheap in the end of the second and in the third quarter. Hence one may expect an increased volume consumption. Stores are running out in the fourth quarter giving high prices. This is also the start of the hunger months period which lasts into the first quarter. One would expect the producers to have money from the seasonal sale in third quarter and therefore able to buy food in the fourth quarter. But during the first quarter both food stocks and money may be short, hence this is called the main hunger period. Even the population in urban areas would usually grow some maize. But here the volume may be smaller and rather consumed in the third quarter. For the non-farming urban households, food will be cheaper during the main harvest period (season 2).

Figure 3. Seasonal Calendar – typical year

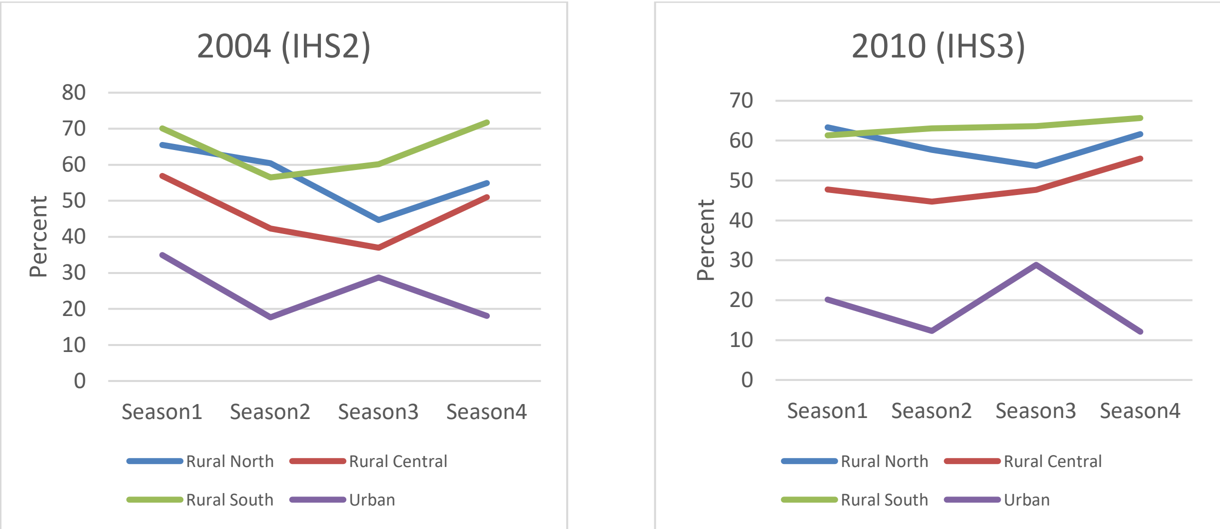
2nd quarter (season 2) | 3rd quarter (season 3) | 4th quarter (season 4) | 1st quarter (season 1)



Source: Fewsnet http://www.fews.net/southern-africa/malawi/seasonal-calendar/december-2013.

Figure 4 shows the poverty headcount across seasons and regions calculated using IHS2 and IHS3. The variation follows the expected seasonal pattern in rural areas, with high poverty levels in the lean season (1st and 4th quarter) as expected. Seasonality in poverty in rural areas was less pronounced in IHS3 than in IHS2. Poverty is also high in urban areas during the 1st quarter due to the lean season when prices are high. In urban areas poverty is also relatively high in the third quarter while turning lower again in the fourth quarter. This may reflect high prices and lower consumption in the third quarter. It is however more difficult to interpret why the urban poverty level is lower in the fourth quarter.

Figure 4. Actual poverty in Malawi, IHS2 and IHS3 by season



Source: based on authors' own calculations.

4. Methodology

The approach to predict poverty builds on the method outlined in detail in Mathiassen (2009). In short, a full household budget survey is used to estimate models for consumption per capita with a set of explanatory variables/predictors. The explanatory variables in Malawi included in the models can be divided into the following groups: core demographic variables; characteristics of head of household; education; housing characteristics; assets ownership; food consumption (yes/no of specific food items); non-food consumption (yes/no of specific non-food items) and two indicators regarding possessions of head which we refer to as subjective welfare predictors.[10] In addition, controls for districts and seasons are included. The variables were selected among a large set of relevant candidates in IHS2 by using a stepwise approach. Information on the selected predictors was collected in a WMS survey – replicating exactly the same phrasing of the questions as in the IHS survey. Together with the estimated

[10] As they are taken from the section named «subjective assessment of well-being» in the IHS-questionnaires.

parameters these predictors are used to predict consumption per capita. A probit function is used to calculate the probability that a household is "poor" given the predicted consumption and the poverty line.

The approach is extended to account for seasonality in the following way. The approach is to first estimate the model:

$$(1)\ Y_{is} = \alpha + \beta X_i + \delta Z_{is} + \sum_{S=1}^{3} \gamma_S D_{iS} + e_{iS}$$

where Y_{is} denotes total consumption per capita in household i in season s, *X* is a vector of predictors that does not vary across seasons, *Z* is a vector of predictors that varies across individuals and seasons, and *D* denotes a dummy to capture unexplained seasonal variation across the year. For example, D_1 is 1 if season=1 and 0 else. $\alpha, \beta, \delta, \gamma$ are parameters in the model and *e* is a i.i.d. error term, with known cumulative distribution function ϕ, zero mean and σ variance. Assume further that e_i are uncorrelated with X_i, and Z_i.

To solve the problem that the WMS does not cover all seasons, we predict the *average* consumption per capita over the year, $\bar{Y}$, while assuming that the relative seasonal variation in the *Z* variables are the same in the IHS and WMS years. Let $\tilde{Z}_{iS}$ denote the value of the Z variable in season S in the WMS survey. We only observe $\tilde{Z}_{iS}$ for S = 1 (example). For the sake of predicting $\tilde{Z}_{iS}$ for the other seasons we assume that

$$(2)\ \tilde{Z}_{iS} = \tilde{Z}_{i1} \frac{\bar{\bar{Z}}}{\bar{Z}_1} + \tau_{iS} \tilde{Z}_{i1}$$

where $\bar{\bar{Z}} = \frac{1}{n}\frac{1}{4}\sum_i \sum_S Z_{iS}$, $\bar{Z}_1 = \frac{1}{n}\sum_i Z_{i1}$ and τ is an i.i.d. error term with zero mean and constant variance, uncorrelated with Z_i.

Using the parameters estimated from equation (1) we can predict the average consumption for household *i* as:

$$(3)\ \hat{\bar{Y}}_i = \frac{1}{4}\sum_{S=1}^{4}\left(\hat{\alpha} + \hat{\beta} X_i + \sum_{S=1}^{3} \hat{\gamma}_S D_{iS} + \hat{\delta}\tilde{Z}_{is}\right) = \hat{\alpha} + \hat{\beta} X_i + \frac{1}{4}\sum_{S=1}^{3} \hat{\gamma}_S + \hat{\delta}\tilde{Z}_{i1} \frac{\bar{\bar{Z}}}{\bar{Z}_1}$$

The probability of being poor can then be written as:

$$(4)\ P_i = \Phi\left(\frac{\hat{\bar{Y}}_i - povline}{\hat{\sigma}/\sqrt{4}}\right)$$

Formulas for calculation of the standard deviation can be found in Mathiassen (2009).

5. Results

Results in this chapter is based on models estimated from IHS2 and IHS3. Separate models were applied for urban areas and for each of the three rural regions (North, Central and South). The R-squared for the models including predictors from all the groups listed in the section above, range from 57 to 84 percent, in the following also referred to as “full” models. See Table A6 – Table A13 in the Appendix for results of full model estimate based on the entire sample.

Test of the method, predicting for IHS-samples

To discuss how the method described in the previous section works we compare actual poverty to predicted poverty figures. To ensure that the contexts we are comparing are the same, each IHS sample was randomly divided into two equal subsamples. Model parameters were estimated from one subsample and were used to predict poverty for the other subsample of the same survey, and vice versa. In this way we can compare poverty predicted to actual poverty for the same households. The results shown in Table 2 are based on the average of these two predictions. It shows the actual poverty, poverty predicted using a full model without seasonal adjustments and poverty predicted using a full model with seasonal adjustments. The latter is included as a test for how the suggested seasonal adjustment works, although not needed in this case as the IHSes cover the whole year. Table 2 shows that poverty is closely predicted, with and without seasonal adjustments, and none of the figures differ significant from the other.

Table 2. Actual poverty and poverty predictions within IHS2 and IHS3 sample. Standard deviations /errors in parenthesis

		(1)	(2)	(3)	t-values		
		Mean actual poverty	Mean predicted poverty	Mean predicted seasonal adjusted poverty	Difference between: (1) and (2)	Difference between: (1) and (3)	Difference between: (2) and (3)
IHS2	Urban	27 (2.9)	27 (3.6)	25 (3.4)	0.2	-0.1	0.3
	Rural North	57 (3.1)	55 (4.7)	59 (4.1)	-0.2	0.4	-0.5
	Rural Central	47 (1.8)	47 (3.0)	48 (2.3)	0.2	0.4	-0.1
	Rural South	64 (1.7)	64 (2.8)	67 (2.4)	-0.1	0.9	-0.8
IHS3	Urban	17 (3.0)	20 (3.0)	19 (3.3)	0.7	0.4	0.3
	Rural North	60 (2.8)	60 (3.7)	62 (3.2)	0.1	0.5	-0.4
	Rural Central	49 (1.9)	50 (2.8)	53 (2.4)	0.5	1.3	-0.6
	Rural South	63 (1.5)	64 (2.5)	67 (2.0)	0.2	1.3	-0.8

Source: based on authors’ own calculations.

To get a better understanding whether some types of variables are more important to include in the model than others, Column (1) – (8) in Table 3 show the differences between the actual poverty and the predicted poverty when excluding one or more groups of explanatory variables. Excluding only demographic variables (column (6)) has a large impact on predicted poverty in rural areas, systematically predicting lower poverty compared to the actual level. The bias is even larger when excluding them from a model without consumption variables (column (7)) and without assets and housing (column (8)). Excluding other variables does not have a systematic impact on the results and causes only smaller changes in overall poverty (see column (2)- (5)).

Table 3. Percentage points differences betwee actual and predicted poverty when dropping some explanatory variables

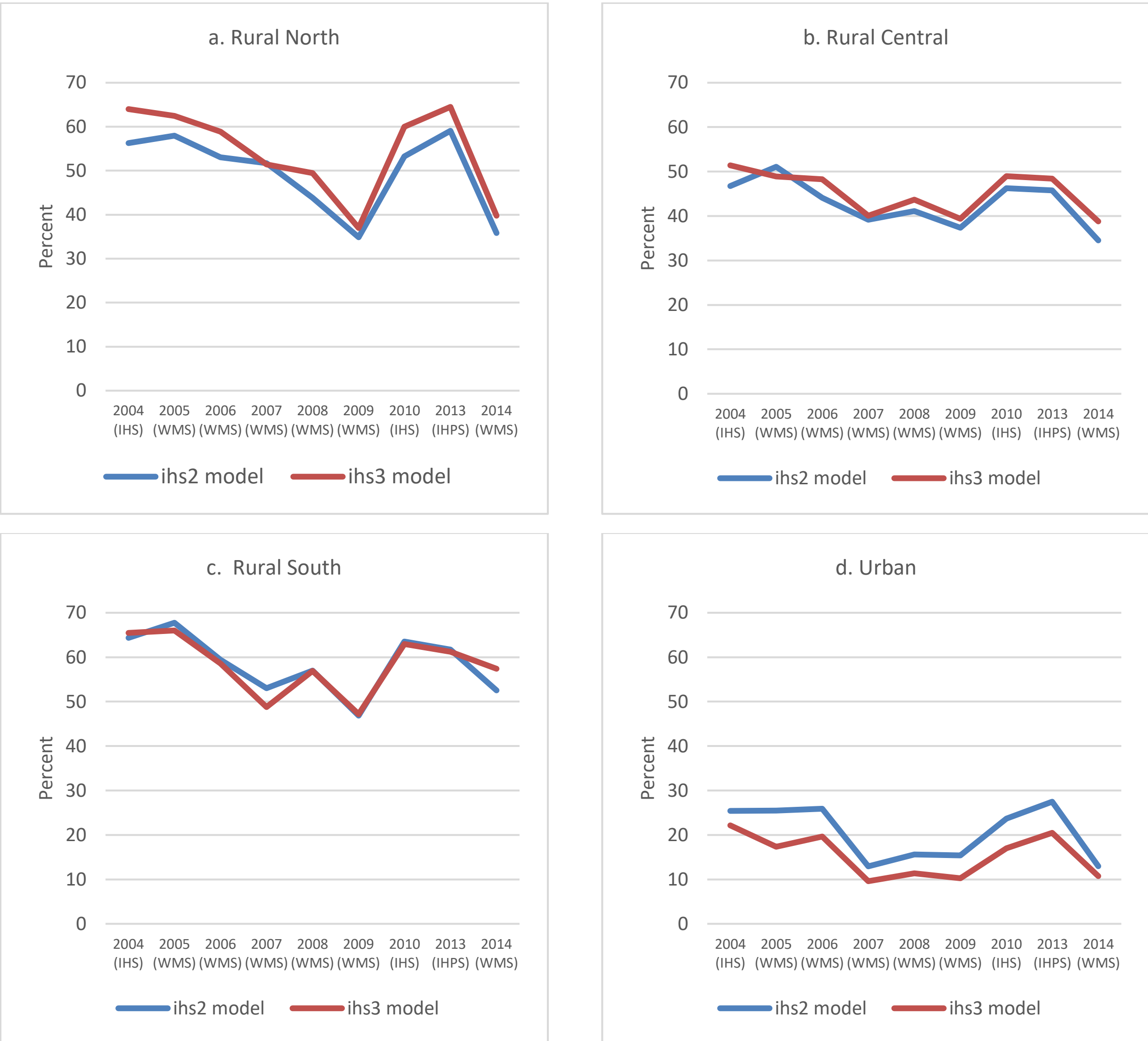
	(1) All var	(2) WO assets, housing	(3) WO educ	(4) WO cons	(5) WO cons, welf, assets, housing	(6) WO demo	(7) WO cons, welf, demo	(8) WO demo, assets, housing
IHS2								
Urban	-1	-1	-1	-1	-1	3	3	6
Rural North	1	1	1	2	1	10	10	13
Rural Central	-1	0	0	1	0	7	8	10
Rural South	0	0	0	1	1	7	9	13
IHS3								
Urban	-3	-3	-3	-2	-1	-1	1	2
Rural North	0	0	0	0	0	9	9	11
Rural Central	-2	-1	-1	0	1	4	5	6
Rural South	-1	-1	0	0	0	5	7	9

Source: based on authors' own calculations.
Note: We employ the following abbrivations: WO=without, educ=education, cons=consumption, welf=subjective welfare, demo=demographic, var=variables.

Predicting poverty trends

Figure 6 (a, b, c, and d) shows the predicted trends for the four areas, using models based upon IHS2 and IHS3. In the figures the predictions for IHS2, using the IHS2 model, are in fact not the numbers predicted by a model, but the actual poverty level calculated directly from the survey. And the same is the case for IHS3, using the IHS3 model. As all seasons were covered in IHS2 and IHS3 it is not necessary to adjust the predictions onto these surveys for seasonality. The predictions for WMS2005-WMS2009 using the IHS2 models differ from the published estimates because of the new adjustment for seasonality and because of some variables originally included have been taken out of the model: Two expenditure variables (expenditure for sugar and cooking oil) were taken out as the CPI used was questioned, cell phone was taken out as it is not considered a stable poverty predictor and whether household paid for public transport was removed as the instruction on how to ask the question had changed in the subsequent surveys. However, the trend still shows decreasing poverty level from 2005 to 2009, although not as much as the published poverty trend in NSO Malawi (2010) (shown in Appendix, Table A1). The tables with the predictions and standard errors for the predictions are found in Table A2 and Table A3 in the Appendix.

Figure 5 Prediction trends using full models adjusted for seasonality



Source: based on authors' own calculations

In Rural North, poverty predicted using the WMS surveys, shows a gradual declining poverty trend with the lowest level of poverty obtained for 2009 and 2014. The poverty levels predicted for 2010 (IHS3) and 2013 (IHPS) do not fit within this trend with relative high poverty levels. Also for Rural Central, the predicted poverty using WMS-surveys shows a decline in poverty over the period, although less pronounced. Neither here do poverty levels for 2010 (IHS3) and 2013 (IHPS) fit within the trend. In Rural South there has been a general decrease in poverty according to the WMS predictions, again the IHS3 and IHPS estimates are out of line with the others. Finally, for urban areas the WMS predictions suggest a sharp decline in poverty from 2006 to 2007 whereas afterwards the WMS predicted poverty levels have remained stable. Again, poverty levels predicted using IHS3 and IHPS are higher than for the WMS.

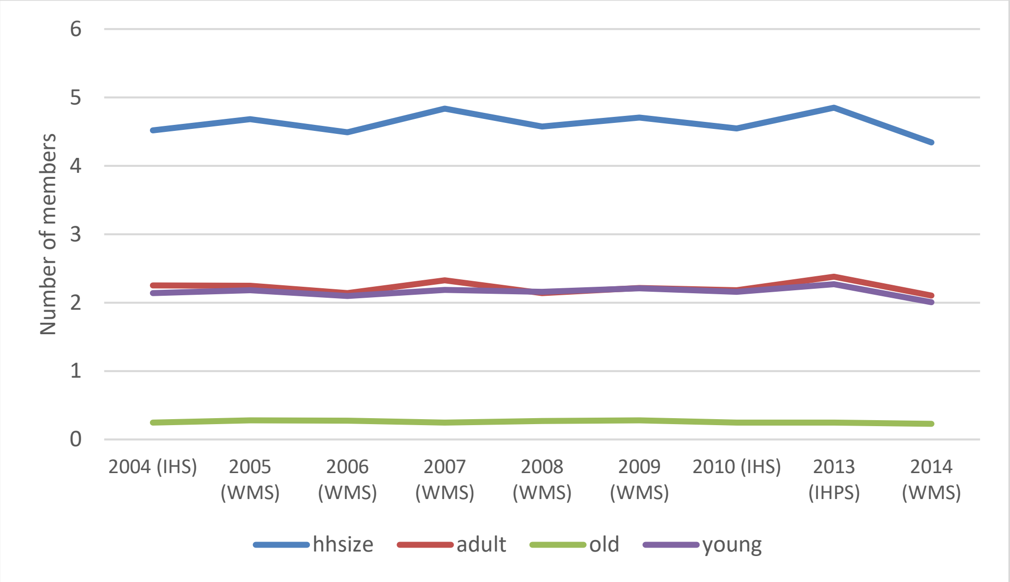
For all regions the two IHS models predict the same changes/trend in poverty – and only small differences in the predicted level. This is illustrated by the t-values of the difference in the prediction between the two models, see Table A4 in the Appendix. While comparing the predicted poverty level to the actual in the 8 cases when predicting onto IHS2 or IHS3 we find that the differences are not significant at the 5 percent level, except for the rural North when using IHS3-model to predict for IHS2. When comparing the two predictions, respectively, based on the IHS2 and IHS3 model for the WMSes only two out of 56 cases differ significantly.

Trends in poverty predictors

This section discusses the variables in the model– the drivers behind the poverty predictions. We refer to them as poverty predictors as they can be self-standing signals of changes in poverty. The aim is to see whether they signal consistent trends in poverty changes and to identify whether there are patterns suggesting that the effect of some poverty predictors is dependent on survey design.

Figure 7 shows the average household size, as well as the average number of members in three age groups; below 15 (young), between 15 and 60 (adult) and above 60 (old) years old. These are variables that we would expect to not fluctuate. The relatively high household sizes in 2007 (WMS) and 2013 (IHPS) significantly differ from the other years. On the other hand, household size in 2009 (WMS) is significant lower than the other years, see Table A5 in the Appendix. A closer inspection of Table A5 shows that all adjacent surveys provide significantly different figures for household size, and there is no trend in any directions over time. This is not systematically assigned to type of survey; however the IHS2 and IHS3 numbers are not significantly different from each other.

Figure 6. Average number of household member, adults, old and young persons in households

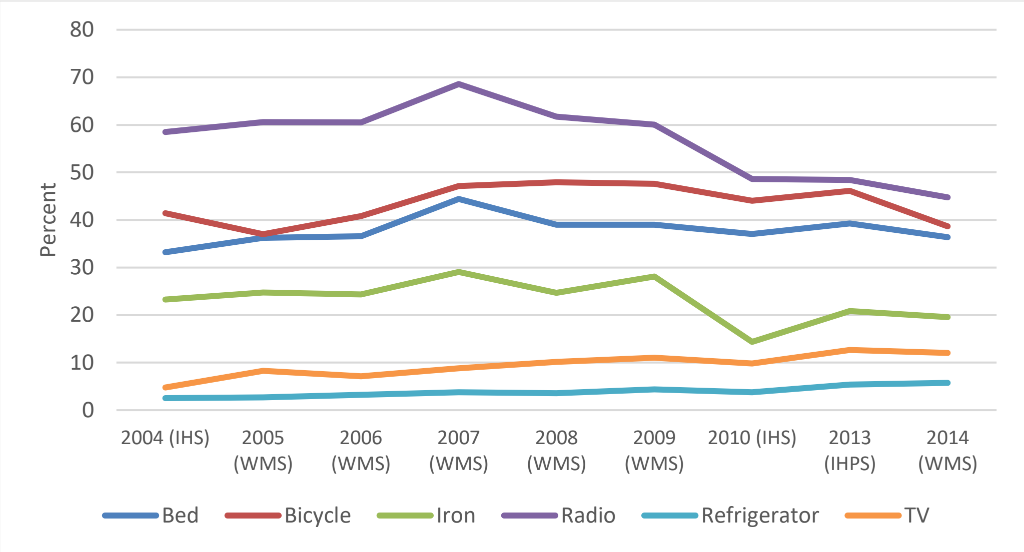


Source: based on authors' own calculations.

Figure 8 shows the percentage of the population with ownership of various assets. There seems to be a downward slope in the ownership of radios – but not a smooth trend. 2007 has a high peak value and 2014 the lowest value. This decrease may be associated with a high and steady increase in mobile phone and tv ownership over the period (other means of information and music). The rate of ownership of mobile phones increased from less than 5 in 2004 to almost 55 percent in 2014. Ownership of refrigerators and tv are slightly increasing. Ownership of bicycles varies much; between 38 and 59 percent in the period. The overall trend in ownership of bed is stable while iron ownership has been decreasing since 2007.

Only assets which are likely to be owned by wealthier households (tv and refrigerator) are steadily increasing. Ownership of less expensive assets, in general, is lower, or about the same, comparing the beginning and the end of the period.

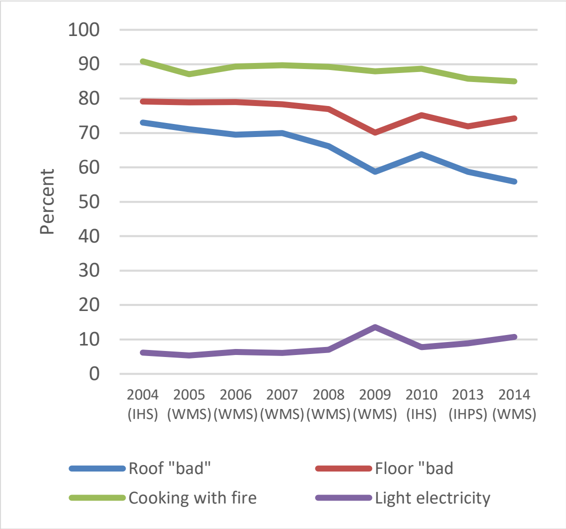
Figure 7. Percentage of households that own various assets

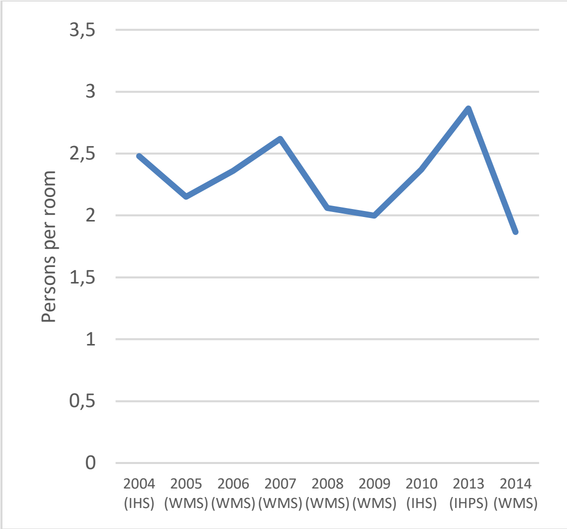


Source: based on authors' own calculations.

As shown in Figure 9 below there is a slight increase in the percentage of population using electricity for lightening over the period. Quality of floor and roof seems to steadily improve, as the percentage with poor quality of these housing conditions decreases over the period. There is only a small decrease in the percentage whose main source of cooking fuel is fire-wood. 2009 seems to be at odds with the other surveys with respect to quality of floor, roof and electricity: it is not plausible with such high annual fluctuation in these variables as the WMS2009 shows. Persons per room in households vary much and not systematically with time or survey type in the period. There were some differences in how this question was asked which may affect the outcome.

Figure 8 Housing condition variables

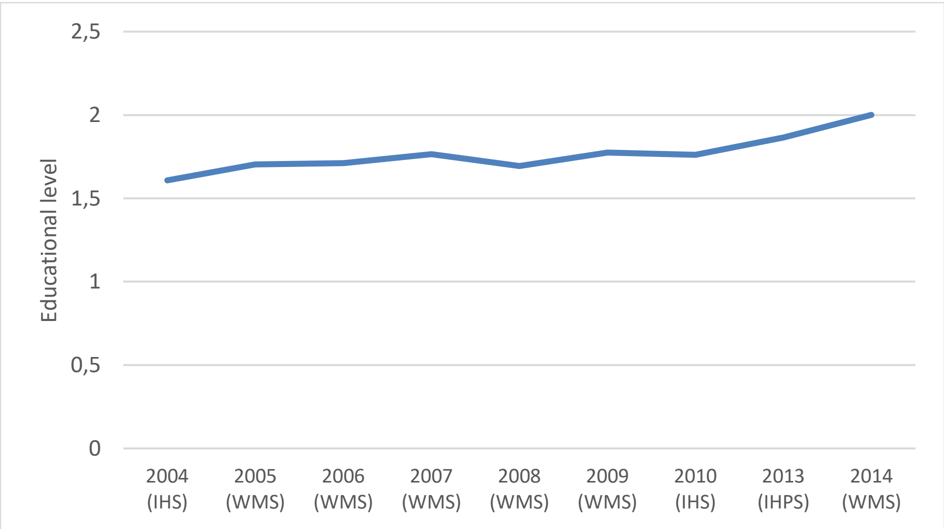




Source: based on authors' own calculations

Educational qualifications in 7 categories[11] among all household members above 5 years, are reported in the survey. The average maximum household qualification is shown in Figure 10, zero denotes that no education certificate was achieved among the household members and 6 denotes that at least one person in the household holds a post graduate degree. There is an increasing trend towards higher education in households.

Figure 9. Average maximum education level among household members

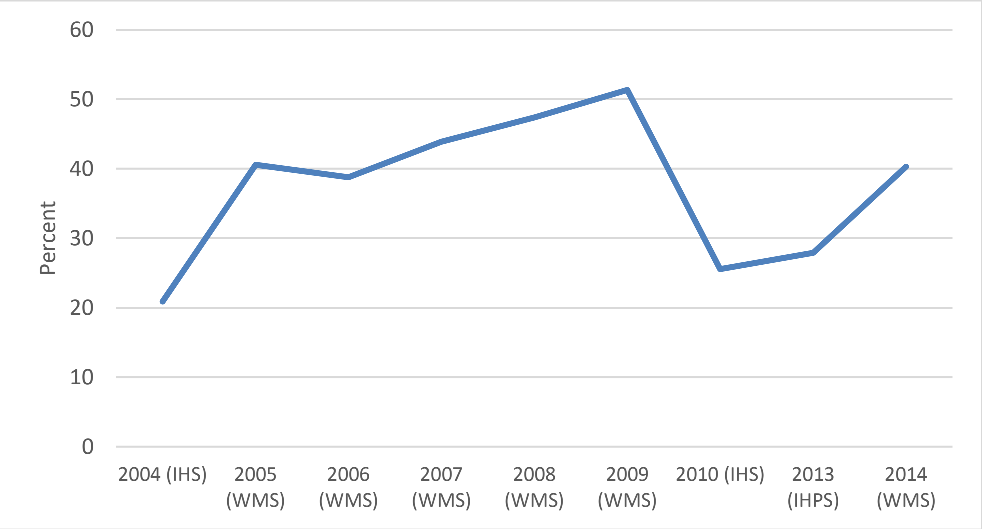


Source: based on authors' own calculations.

[11] (0) None; (1) Primary School Leaving Certificate; (2) Junior Certificate Examination; (3) Malawi School Certificate Examination; (4) Non-University Diploma; (5) University Diploma Degree; (6) Post graduate Degree.

As shown in Figure 11, there are large differences in the percentage of households purchasing toothpaste in the period. It seems to be a systematic difference between the two survey types, with a much higher percentage reporting purchase of toothpaste in the WMS surveys. Part of this can be explained by seasonality – but far from all. The purchase of toothpaste in the IHS-surveys varies with 10 percentage points across the four seasons. It is no obvious reason for purchase of toothpaste to vary so much.

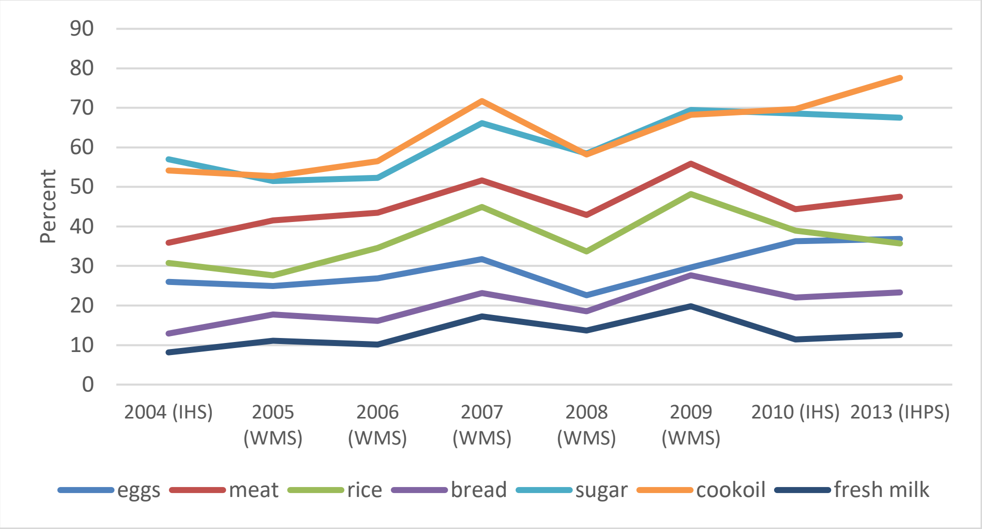
Figure 10. Percentage of households buying toothpaste



Source: based on authors' own calculations.

Figure 12 shows the percentage of households who consumed various food items in the last 7 days before the interview. These numbers will be affected by seasonality and only season 3, covered in all but one survey, is shown. WMS2014 did not include this season and thud does not occur in the figure. Generally, over the period there seems to be a tendency to increased food consumption. However, there are ups and downs with particularly WMS2007 and WMS2009 reporting seemingly high consumption.

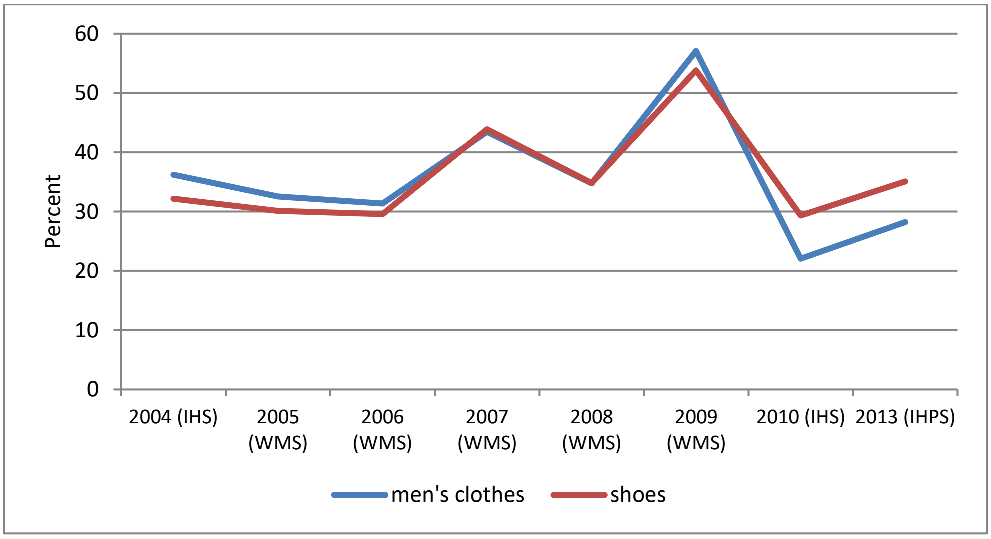
Figure 11. Percentage of households consuming food items in season 3



Source: based on authors' own calculations.

Figure 13 shows that there is a large variation in purchase of men's clothing and shoes (to both gender) in the last three months. These variables are the sum over respectively five different types of mens clothing and four types of shoes. WMS2009 level is much higher than the others, and WMS2007 is also high. There is no overall trend throughout the period, and the two poverty predictors closely track each other.

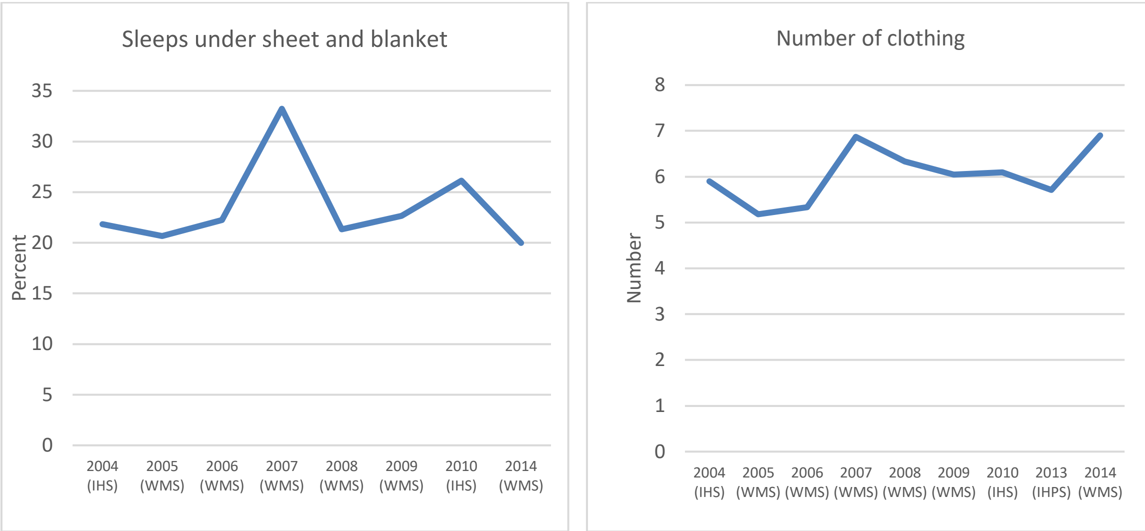
Figure 12. Consumption of men's clothes and shoes Season 3



Source: based on authors' own calculations.

Figure 14 shows that there is no clear trend to whether household head sleeps with blanket *and* sheet in the cold season over time. This prevalence is however, much higher in 2007 compared to the other years. Similarly, there is no systematic pattern with respect to type of survey, in the number of clothes the household head owns. This variable varies much, and the values are particularly high in 2007 and 2014. A hypothesis is that, since these questions are not standard, as are the other predictors included in the models, the enumerators may not have been trained to ensure a consistent field approach.

Figure 13. Welfare predictors concerning head of household



Source: based on authors' own calculations.

Some of the variables that are included in the models are also available from the Census which took place in 2008. Including these variables can give us an additional validation on whether the WMS or IHS provide systematically different estimates. Table 4 shows the development in these variables including the neighbor surveys. Cooking with open fire is constant over the three years period the table includes. Electricity for light is unexpectedly high in WMS2009. The share that owns a bicycle is a little higher in the WMSs and radio ownership is much lower in IHS3 than in the other sources. Thus, with respect to these indicators, there is no systematic pattern to be observed.

Table 4. Housing and asset variables compared to Census. Percent

	2008 (WMS)	2009 (Census)	2009 (WMS)	2010 (IHS3)
Cooking with fire	89	88	88	89
Light from electricity	7	7	14	8
No toilet	7	12	9	8
Bicycle	48	45	48	44
Radio	62	64	60	49

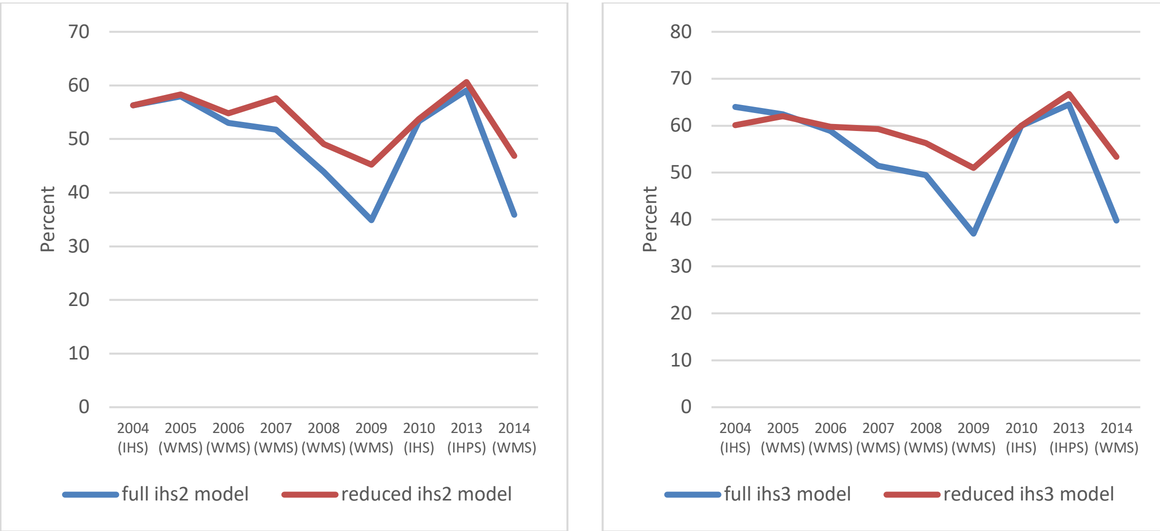
Source: based on authors' own calculations.

Poverty trends with reduced set of poverty predictors

The previous analysis on trends in the poverty predictors, showed that some are more "troublesome", i.e. they show an unlikely variation across the surveys. In particular, number of rooms in household; non-food consumption variables; and the two variables concerning subjective welfare. Also the binary food consumption variables tend to be systematic lower in the IHS-surveys compared to WMS. This is in accordance with the findings in Kilic and Sohnesen (2019), suggesting that particularly food and non-food consumption as well as the welfare variables were affected by the questionnaire context. We are left with demographic variables; assets; housing; education and geographic controls. We will refer to the models with the fewer explanatory variables as the "reduced models".

Figure 15 to Figure 18 present the poverty trends using the using the reduced model in the same graph as the full model to easily visualize the effect of excluding the mentioned variables from the model. With the reduced variable model, poverty declines in Rural North up to 2009 – but not as much as when all variables are included in the model, as seen in Figure 15. The reduced variable models predict higher poverty for the WMS'es after 2006 than the full model. Poverty predicted for IHS2, IHS3 and IHPS is nearly unchanged.

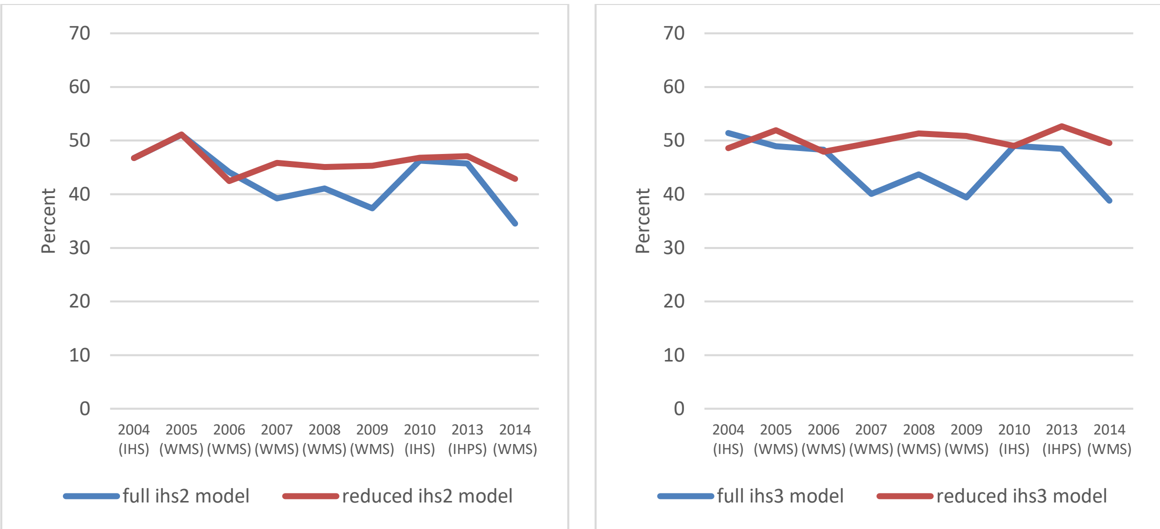
Figure 14. Poverty trends, full models and reduced models. Rural North



Source: based on authors' own calculations.

There is less change in poverty in Rural Central over the survey period with the reduced variable model. Poverty predicted with IHS2 as well as with IHS3 reduced models are higher for WMS2007-WMS2014 compared to the respective full models, as well as higher for IHPS using the IHS3 model, as seen in Figure 16.

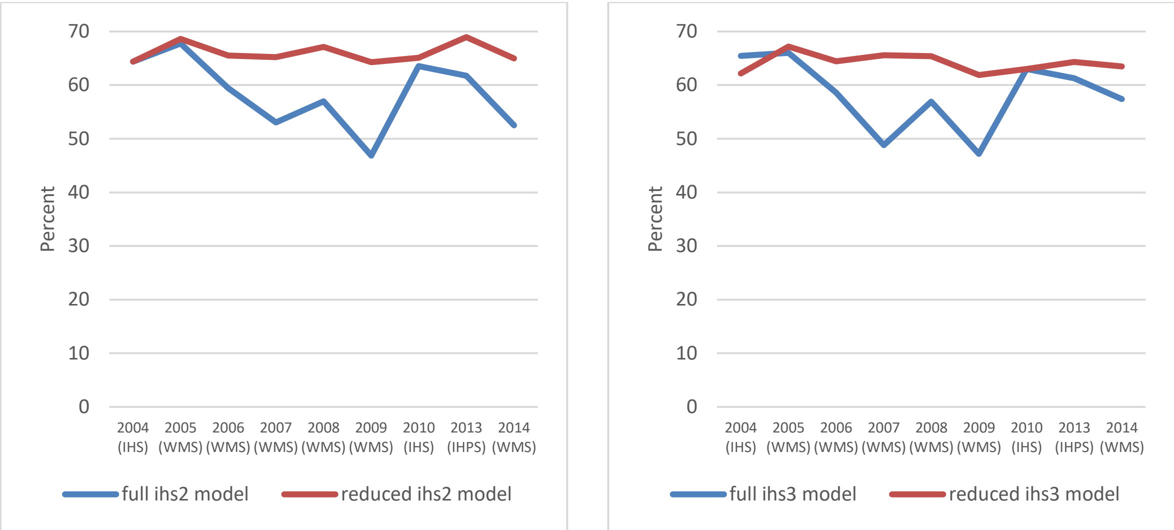
Figure 15. Poverty trends, full models and reduced models. Rural Central



Source: based on authors' own calculations.

Also for Rural South the reduced variable model reveals less changes in poverty in the period. Compared to the full models, poverty predicted by the reduced variable model is higher for most of the WMS surveys, and also for IHPS using IHS2 model (Figure 17). For 2005 there is no impact of reducing the number of predictors in the model, whereas for 2006 the impact is relatively small.

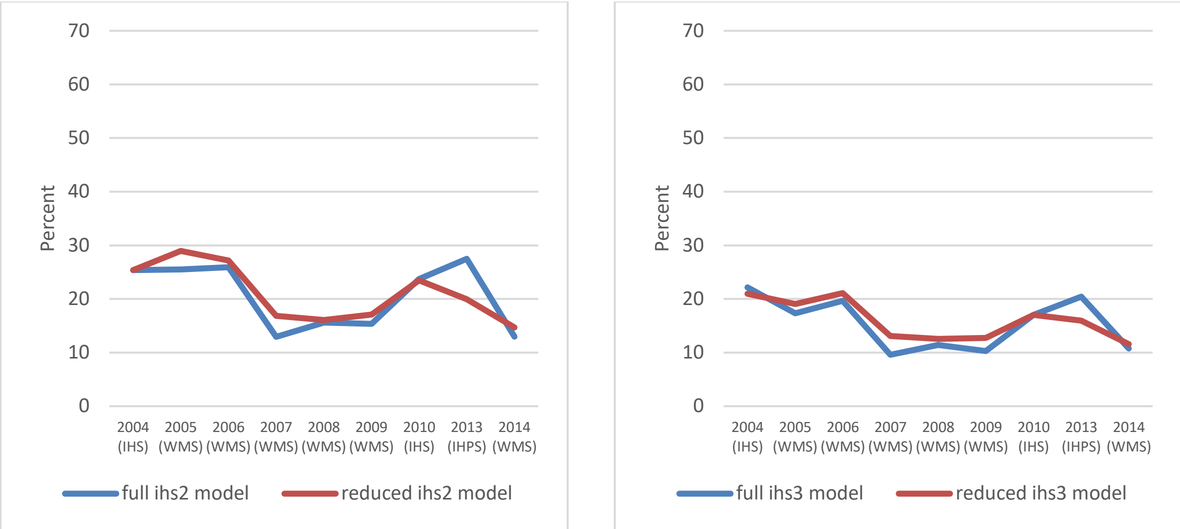
Figure 16. Poverty trends, full models and reduced models. Rural South



Source: based on authors' own calculations.

As seen from Figure 18, for the urban area, comparing the predictions from the two types of models shows a more mixed picture. The reduced variable model predicts lower poverty for IHPS. For the other surveys poverty predicted is consistent with the findings in the other areas; taking out the consumption variables and poverty predictors from the model produces a higher estimate for most of the WMS-surveys. The predictions onto IHS2 and IHS3 are the same using the reduced and the full models.

Figure 17. Poverty trends, full models and reduced models. Urban



Source: based on authors' own calculations from the surveys.

Overall, from the graphs above, we can see that excluding these variables from the model smoothens out the trend as the poverty predicted for the WMS increases, but hardly affect the poverty levels predicted from the IHSes.

6. Discussion

There is no clear answer to what was the "true" trend in the poverty headcount in Malawi. To ensure a consistent annual poverty trend we need strictly comparable surveys every year. There should be no change in methodology considering all aspects from the survey instruments, how to undertake the interview and analyzing the data. Unfortunately, this is not the case in Malawi.

However, as self-standing signals of poverty, the analysis in the present study shows small but steady improvements over the period in the share of the population with good roofs and floors, access to electricity and a small reduction in the share cooking with firewood. Ownership to assets that typically only the wealthier have, i.e., refrigerator and TV, is also increasing. There is no clear conclusion on other assets, such as bicycle, beds and iron. In terms of food consumption, the overall patterns also suggest better conditions. Education level is increasing as well. However, these improvements may well have taken place only among the non-poor, in which case, it would not affect the welfare among the poor.

The present study has had as a main objective to assess whether a survey to survey imputation technique can be used to predict poverty changes over time with different survey tools. The validation of the model has been done to the extent that is possible, i.e. predicting within the same survey. By randomly dividing the sample and predicting onto the other half we can be sure that the methodology is strictly comparable, i.e that the way data was collected and analyzed is the same. The analysis within sample shows that the method provides a good fit. The new addition to the methodological approach to correct for the fact that the WMS surveys only cover a short period and not a whole year as in the IHS surveys, also works well when tested within sample. Predicting onto an "identical" survey, from IHS2 to IHS3 and vice versa, mostly works well, except one out of eight cases where the difference in the predicted versus actual poverty level is significant at 5 percent level. This is a test of whether the model is stable over time; as time passes one may expect the relation between the predictors and the consumption aggregate to change. For example, households may shift from using radio as mobile phones rapidly entering the market provides them with a means to listen to music and news. Such changes may distort the predictions. However, as new conversion factors were partly used (for maize, the main staple) in IHS3, the two poverty levels calculated directly from the surveys are not strictly comparable. Thus, we cannot make a firm conclusion based on predicting onto the other IHS-survey. The official (actual) poverty estimates for IHPS, as published in World Bank (2018c) are not comparable to the official poverty estimates from the previous IHS surveys as it uses an entirely new set of conversion factors for all food items. Therefore, the model-based poverty predicted for IHPS will neither be comparable to the actual poverty level in IHPS, and we cannot say whether the models "works" in this case. Another test of stability of the model over time is provided by comparing the two trends predicted by the two IHS-models for the WMS and the IHPS- surveys. The trends provided by the two models are the same within each area in the 10 years period we look at. This is the same that was found for a study in Uganda (Mathiassen, 2012).

When correcting for the element of seasonality the resulting decline in poverty is less than the official poverty numbers that were published using the WMSs up to 2009. This is as expected as the seasons covered in the WMS2005-WMS2009 are expected to be relatively good times during the year. However, there is still a clear downward trend in poverty predicted for the WMS surveys, including the WMS2014 which was analyzed for the first time in the present study. There is a clear break in the trend with higher predicted poverty for IHS3 and IHPS compared to the predictions for the WMS surveys.

Part of the reason for this break in trend may be assigned to differences in the way the IHS versus the WMS questionnaire were designed. Some explanatory variables seem to be particularly influenced by the survey tool, i.e. consumption and subjective welfare predictors. There is a relatively large impact on the model-based predictions for the WMS surveys from 2007 and onwards, when excluding these variables from the model. At the same time there is hardly any impact on the model-based predictions for the IHS surveys of excluding the same variables. This supports the hypothesis that the respondents report "overly" positive on such variables in the light survey tool compared to the IHS everything else equal, confirming the finding in Kilic and Sohnesen (2019). For example; in the three IHSes, households were asked about the household's consumption of more than hundred food items. First, they were asked whether they consumed the specific item. If yes, the question is followed up with questions on amount, price and source of consumption. On the other side, in the WMS-surveys, the households are only asked whether they consumed eleven specified food items. The question on whether they consumed the food item or not is phrased in the same way as in the IHS-surveys, there are however, no follow-up questions.

Excluding the variables that are most likely to be systematically affected by survey design results in a much flatter development in poverty, and in rural Central and South confirming the unchanged poverty level as given by the IHS-surveys. However, it may be argued that a model without such variables will be less able to pick up fluctuations in poverty as the remaining variables are more associated with long run development.

An analysis of the trends in the predictors sometimes reveals variations that are not plausible. Some of this variation can be assigned to the questionnaire design as discussed above. This pattern is visible in the higher consumption of toothpaste in all WMS surveys and clothes and shoes from the WMS2007 and onwards. Also, two other variables, from the so-called "subjective welfare predictor" section in the IHS questionnaire, showed a large and unsystematic variation. With respect to food consumption, two surveys WMS2007 and WMS2009, deviate from the trend with relatively high prevalence of the population reporting consumption of the food items. Still there is an overall trend showing a more diverse food consumption pattern in the period.

Other variations cannot be attributed to questionnaire design. In WMS2009 there is an unexpected high proportion of households that report using electricity as the main source of light. The figure is out of the trend and significantly different from the other surveys. This is not a variable that fluctuates from year to year as it requires large investments. The same year two other poverty predictors, which are not expected to change fast, i.e. good quality of roof and floor, indicate improvements in this year that are out of trend with the adjacent years. The systematic improvements in these poverty predictors, which are not expected to fluctuate, suggest that there are implementation differences here.

Household composition, such as number of household members in different age groups and households size are key indicators to include in the model. Calculation of these indicators requires a detailed roster section and clear definition of who is to be counted as a household member. In principle, household member was defined in the same way in the IHS and WMS-surveys, but the year to year significant change in the number of people in the household, suggests that the implementation differs. The roster module requires solid training and testing of the enumerators to ensure consistency.

Household composition stood out in the analysis as a critical explanatory variable in the rural models, which makes it even more important to get this variable right. Omitting this group of variables provides a systematic lower prediction of poverty in rural areas. No other variable group has such a systematic impact on the poverty numbers. Why is it so important to include household size for predicting correctly? One explanation is that there is not much heterogeneity among households, in particular in rural areas. In that case the quality of house, consumption pattern, assets and education are not sufficient to distinguish between households. Rather it is the point of the life cycle, the number of household members to which to divide the asset and consumption between, that is most important for the poverty level. Another explanation can be that household composition creates a bias in the prediction because it is more difficult to collect consumption information correctly for larger households. For example, it is likely that it is difficult to capture all food consumed in households with many adults that frequently eat outside the household.

The effect of differences in questionnaire design in WMS versus IHS has been discussed but there are also other features with the surveys influencing the non-sampling errors and therefore the results. Additional factors that we have some information on and that vary between the surveys are: institutions supporting the implementation of the surveys, size of the field operation and data collection tools. The IHS-surveys were followed by the same team in the WB, and the questionnaire remained the same – however, the last round in 2013 included a much smaller sample and a follow up of a panel from IHS3. There seem, however, to be more inconsistencies with respect to how the WMS-surveys were implemented. A main apparent difference between the WMS-surveys is the sample size. Whereas, the two first WMSes were relatively small, around 5,000 households, the other WMS surveys were much larger, varying between 14,200 and almost 30,000 households. It is interesting to note that the smaller surveys, WMS2005 and WMS2006 are hardly affected by omitting the consumption variables from the model in contrast to the larger WMS surveys. This suggests that there may be additional differences which are associated with the size of the field operation. A large field operation requires a larger number of team leaders and enumerators and therefore a larger requirement for training. The training would, in such cases, be done in two steps, first training the trainers at central level, followed by the newly trained trainers teaching the enumerators at a lower level. The result may be larger heterogeneity across the interviewers, with respect to how the questions are asked and how errors are identified. According to UNSD (2008) non-sampling errors may increase by sample size.

WMS2007, the largest survey, was attached to an agricultural Census (the NACAL). For this survey, most of the enumerators were agriculture supervisors receiving some additional training to serve as enumerators for the WMS module. This change of training and enumerators may lead to a difference compared to the other WMS surveys. The NACAL included several visits and ten modules, where the WMS was the sixth. This affects the interview burden and may therefore have led to respondent fatigue and affected the answers. From 2008 to 2014, the sample sizes remained large, and Statistics Norway had only an advisory role. The questionnaire changed in 2011 and in 2014. Further, the WMS2014 data were collected by electronic means. This allows for identifying errors during the time of the interview, which can be corrected at the spot. Although we are not able to quantify the effect of the various factors we can conclude that there have been changes that are likely to have had an effect on the final output from the surveys.

7. Concluding remarks

This paper has shown that several factors point towards too low poverty levels predicted by the model in Malawi. It has illustrated that there are substantial challenges in providing comparable time series in core welfare variables even between similar surveys, and thus the need to design a good and stable system of collecting and analyzing data. It is fundamental to focus on quality in all phases of a survey, and to stick to the same methodology, instructions and practical implementation. It can, however be difficult to maintain the same questionnaire when the NSO is dependent on financial support from stakeholders that want to have their say on the questionnaire. If a survey program is to provide annual poverty trends the ideal would be to collect annual consumption data in a systematic way. However, if less costly means based on survey to survey imputation is used it is important to ensure that the poverty predictor module is included in the same way and placed in the same order both in an IHS and in a WMS. If this is not possible, a second-best option would be to provide time series in the light surveys that signal the poverty trend, as well as poverty distribution across space – recognizing that poverty between the different types of surveys are difficult to compare.

References

Arndt, C., Pauw, K. and Thurlow, J. (2014): "The Economy Wide Impacts and Risks of Malawi's Farm Input Subsidy Programme" *WIDER Working Paper* 2014/099, Helsinki: World Institute for Development Economics Research of the United Nations University (UNU-WIDER).

Beck, U., Pauw, K. Mussa, R. (2015): "Methods matter. The sensitivity of Malawian poverty estimates to definitions, data, and assumptions" *WIDER Working Paper* 2015/126, Helsinki: World Institute for Development Economics Research of the United Nations University (UNU-WIDER).

Carr, S. (2014): "The Challenge of Africa's Nitrogen Drought: Some Indications from the Malawian Experience" *MaSSP Policy Note* 19, Washington DC: International Food Policy Research Institute.

Chirwa, E. and Dorward, A. (2013): "*Agricultural Input Subsidies. The Recent Malawi Experience*" Oxford University Press.

CIA (2018): "the World Fact Book" https://www.cia.gov/library/publications/the-world-factbook/geos/mi.html Retrieved April 2018.

Elbers, C., J. O. Lanjouw, and P. Lanjouw, (2003): "Micro-Level Estimation of Poverty and Inequality," *Econometrica*, 71, 355–64.

Government of Malawi (2015): "Malawi 2015 flood Post Disaster Needs Assessment Report, Malawi": Ministry of Disaster Management Affairs.

Haug, R. and Wold, B. (2017): "Social Protection or Humanitarian Assistance: Contested Input Subsidies and Climat Adaptation in Malawi" *IDS Bulletin* Vol. 48 (4), 93-111.

Kanbur, R. (2005): "Growth, Inequality and Poverty: Some Hard Questions" *Journal of International Affairs*, Vol. 58, 223-232.

Kilic, T., Whitney, E. and Winters, P. (2013): "Decentralized Beneficiary Targeting in Large scale Development Programs. Insights from the Malawi Farm Input Subsidy Program" *Policy Research Working Paper* 6713. World Bank, Washington.

Kilic, T. and Sohnesen, T.P (2019): "Same Question but Different Answer: Experimental Evidence on Questionnarie Design's Impact on Poverty Measured by Proxies" *Review of Income and Wealth* 65 (1), 144-165.

Lavrakas, P. J. (2008): "*Encyclopedia of survey research methods*" Thousand Oaks, CA: Sage Publications, Inc. doi: 10.4135/9781412963947P.J.

Mathiassen, A. (2009): "A model based approach for predicting poverty without expenditure data" *Journal of Economic Inequality* 7 (2), 117-135.

Mathiassen, A. (2013): "Testing predictive performance of poverty models: Empirical evidence from Uganda " *Review of Income and Wealth*, 59 (1), 91-112.

MOAIWD (1997–2015): Annual National Crop Estimates, spreadsheet available from MOAIWD, Malawi: Ministry of Agriculture, Irrigation and Water Development.

NSO (2010): "The Welfare Monitoring Survey 2009 Main Report" http://catalog.IHSn.org/index.php/catalog/4580/download/58291. Retrieved April 2018

NSO (2012): "Integrated Household Survey 2010-2011. Household socio-economic characteristics report" Republic of Malawi.

NSO (2016): "Statistical Yearbook 2016" Republic of Malawi.

Newhouse, D. Shivakumara, S., Takamatsu, S. and Yoshida, N. (2014): "How Survey-to-Survey Imputation Can Fail" *Policy Research Working Paper* 6961, World Bank Group.

Pauw, K., U. Beck, and R. Mussa (2016). 'Did Rapid Smallholder-Led Agricultural Growth Fail to Reduce Rural Poverty?'. In C. Arndt, A. McKay, and F. Tarp (eds.), *Growth and Poverty in Sub-Saharan Africa*, pp. 89-112. Oxford: Oxford University Press.

Ravallion, M. (1998): "Poverty Lines in Theory and Practice" *Living Standard Measurement Study Working Paper* 133, Washington DC, World Bank.

Schreiner, M. (2014): "The Process of Poverty Scoring Analysis" http://www.simplepovertyscorecard.com/Process_Poverty_Scoring_Analysis.pdf Retrieved September 2018.

Schreiner, M. (2015): "Simple Poverty Scorecard. Poverty Assessment tool. Malawi" http://www.simplepovertyscorecard.com/MWI_2010_ENG.pdf Retrieved April 2018.

Schreiner, M. (2018): "Simple Poverty Scorecard® Tool. Kenya" http://www.simplepovertyscorecard.com/KEN_2015_ENG.pdf Retrieved September 2018.

UNSD (2008): "Designing Household Survey Samples: Practical Guidelines" *Studies in Methods*, Series F No. 98, Department of Economic and Social Affairs, Statistics Division United Nations; New York.
https://unstats.un.org/unsd/demographic/sources/surveys/Handbook23June05.pdf Retrieved March 2019

UN (2000): United Nations Millennium Declaration, Resolution 55/2 adopted by the UN General Assembly 18 September 2000 http://www.un.org/millennium/declaration/ares552e.pdf Retrieved September 2018.

UN (2015): Transforming our world: the 2030 Agenda for Sustainable Development Resolution 70/1 adopted by the UN General Assembly 21 October 2015 http://www.un.org/ga/search/view_doc.asp?symbol=A/RES/70/1&Lang=E

Verduzco-Gallo, I. Ecker, O. and Pauw, K. (2014): "Changes in food and nutrition security in Malawi: Analysis of recent survey evidence." *MaSSP Working Paper* 6. Washington, D.C.: International Food Policy Research Institute (IFPRI).

Vu, L. and Baulch, B. (2011): "Assessing Alternative Poverty Proxy Methods in Rural Vietnam" *Oxford Development Studies*, Vol. 39 (3), 341-362.

World Bank (2018a): “What we do” http://www.worldbank.org/en/about/what-we-do
Retriever October 2018.

World Bank (2018b): “Country Statistics” https://data.worldbank.org/country Retrieved April 2018.

World Bank (2018c): “Methodology for Poverty Analysis in Malawi in 2010-2013”
http://siteresources.worldbank.org/INTLSMS/Resources/3358986-1233781970982/5800988-1271185595871/6964312-1404828635943/Methodology_for_Poverty_Analysis_in_Malawi_2010_2013_for_Dissemination.pdf
Retrieved September 2018.

Appendix

Table A1.Published poverty numbers. Percent

	IHS2 2004	WMS 2005	WMS 2006	WMS 2007	WMS 2008	WMS 2009
Malawi	52	50	45	40	40	39
Urban	25	24	25	11	13	14
Rural	..	53	47	44	44	43
Rural Northern	56	51	46	46	35	31
Rural Central	47	46	40	36	40	41
Rural Southern	64	60	55	51	51	51

Source: NSO (2010).

Table A2. Poverty predictions

	2004 (IHS)	2005 (WMS)	2006 (WMS)	2007 (WMS)	2008 (WMS)	2009 (WMS)	2010 (IHS)	2013 (IHPS)	2014 (WMS)
Full model (all variables included)									
IHS2 model									
Urban	25	25	26	13	16	15	24	27	13
Rural North	56	58	53	52	44	35	53	59	36
Rural Central	47	51	44	39	41	37	46	46	35
Rural South	64	68	59	53	57	47	64	62	53
IHS3 model									
Urban	22	17	20	10	11	10	17	20	11
Rural North	64	62	59	51	49	37	60	65	40
Rural Central	51	49	48	40	44	39	49	48	39
Rural South	65	66	59	49	57	47	63	61	57
Reduced model (excluded consumption and "subjective" poverty variables)									
IHS2 model									
Urban	25	29	27	17	16	17	23	20	15
Rural North	56	58	55	58	49	45	54	61	47
Rural Central	47	51	42	46	46	45	47	47	43
Rural South	64	69	65	65	67	64	65	69	65
IHS3 model									
Urban	21	19	21	13	13	13	17	16	12
Rural North	60	62	60	59	56	51	60	67	53
Rural Central	51	52	48	50	51	51	49	53	50
Rural South	62	67	64	66	65	62	63	64	64

Source: based on authors' own calculations.

Note: The results for IHS2 using IHS2-model and for IHS3 using IHS3-model, are the actual poverty and not model based predicted as the other figures.

Table A3. Standard errors

	2004 (IHS)	2005 (WMS)	2006 (WMS)	2007 (WMS)	2008 (WMS)	2009 (WMS)	2010 (IHS)	2013 (IHPS)	2014 (WMS)
Full model (all variables included)									
IHS2 model									
Urban	2.8	2.9	4.6	2.1	2.5	2.2	2.8	3.2	1.8
Rural North	2.7	3.4	3.2	3.7	3.1	2.6	3.4	3.3	3.2
Rural Central	1.6	2.3	2.4	2.1	1.9	1.9	2.4	2.8	1.9
Rural South	1.5	2.3	2.4	2.2	2.0	1.9	2.0	2.9	1.9
IHS3 model									
Urban	3.0	2.5	4.2	1.8	2.3	2.0	2.5	2.8	1.5
Rural North	2.7	2.9	2.8	3.3	2.6	2.7	2.3	2.9	2.6
Rural Central	2.2	2.2	2.4	2.1	2.0	2.0	1.6	2.7	1.9
Rural South	1.8	2.2	2.2	2.2	1.9	1.8	1.3	3.6	1.8
Reduced model (excluded consumption and "subjective" poverty variables)									
IHS2 model									
Urban	2.8	2.6	2.2	2.5	2.3	2.4	2.4	3.0	2.0
Rural North	2.7	3.6	3.6	4.1	3.6	3.6	3.9	3.8	3.7
Rural Central	1.6	2.5	2.5	2.5	2.1	2.4	2.6	2.7	2.2
Rural South	1.5	2.4	2.6	2.3	2.2	2.3	2.4	2.4	2.1
IHS3 model									
Urban	3.1	2.6	4.3	2.2	2.3	2.1	2.5	3.2	1.7
Rural North	3.4	2.8	2.8	3.3	3.4	2.5	2.3	3.7	3.2
Rural Central	2.4	2.2	3.7	2.6	2.3	2.4	1.6	3.0	1.2
Rural South	1.7	2.3	2.5	2.2	2.2	2.1	1.3	3.0	1.9

Source: based on authors' own calculations.
Note: The standard errors for the IHS2 and IHS2 estimates using respectively IHS2- and IHS3- model are the actual standard errors.

Table A4. t-values. Differences in predictions between IHS2 and IHS3 models. Full models

	2004 (IHS)	2005 (WMS)	2006 (WMS)	2007 (WMS)	2008 (WMS)	2009 (WMS)	2010 (IHS)	2013 (IHPS)	2014 (WMS)
Urban	0.01	0.02	0.01	0.01	0.01	0.02	0.02	0.02	0.01
Rural North	-0.02	-0.01	-0.02	0.00	-0.01	0.00	-0.02	-0.01	-0.01
Rural Central	-0.02	-0.01	-0.02	-0.02	-0.01	-0.01	-0.01	-0.01	-0.01
Rural South	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	-0.02

Source: based on authors' own calculations.

Table A5. Lower and upper limits for 95% confidence interval for household size

	2004 (IHS)	2005 (WMS)	2006 (WMS)	2007 (WMS)	2008 (WMS)	2009 (WMS)	2010 (IHS)	2013 (IHPS)	2014 (WMS)
Lower limit for Confidence Interval	4.48	4.62	4.43	4.86	4.55	4.68	4.51	4.78	4.31
Upper limit for Confidence Interval	4.56	4.74	4.55	4.92	4.61	4.73	4.59	4.92	4.38

Source: based on authors' own calculations.
Note: in the following tables controls for season and district are not shown.

Table A6. IHS2-full model. Rural North

	Parameter estimate	Standard error	t-value
Intercept	10.19	0.059	171.4
dependency ratio	-0.18	0.054	-3.3
if grass or plastic roof	-0.12	0.031	-4.0
no of clothes for head	0.02	0.003	6.5
if hh bought shoes last 3 months	0.08	0.025	3.2
if hh owns a bed	0.12	0.026	4.7
number of radios hh own	0.25	0.050	4.9
if hh consumed rice last 7 days	0.13	0.030	4.4
if hh consumed fresh milk last 7 days	0.11	0.031	3.6
if hh consumed meat last 7 days	0.21	0.025	8.5
if hh consumed cookingoil last 7 days	0.12	0.025	4.5
max education among members, squared	0.01	0.003	3.8
age of head, squared	0.00	0.000	1.3
log of hhsize	-0.59	0.023	-25.7
Adjusted R-squared	0.56		
Number of observations	1433		

Source: based on authors' own calculations.

Table A7. IHS2-full model. Rural Central

	Parameter estimate	Standard error	t-value
Intercept	10.43	0.032	321.0
dependency ratio	-0.12	0.026	-4.9
members per room'	-0.02	0.004	-3.8
if sand or smoothed mud floor	-0.16	0.021	-7.8
max education among members	0.05	0.007	6.2
if hh bought shoes last 3 months	0.10	0.013	7.3
if hh bought mens clothes last 3 months	0.09	0.013	7.2
if hh owns an iron	0.08	0.018	4.5
if hh owns a radio	0.07	0.013	5.5
if head sleeps under blankets and sheets in cold season	0.06	0.019	3.2
if hh owns a bed	0.12	0.018	6.4
if hh consumed eggs last 7 days	0.09	0.015	5.8
if hh consumed meat last 7 days	0.20	0.013	15.3
if hh consumed rice last 7 days	0.13	0.017	7.9
if hh consumed fresh milk last 7 days	0.09	0.026	3.5
if hh consumed bread last 7 days	0.12	0.022	5.5
if hh consumed cookingoil last 7 days	0.13	0.013	9.9
if hh bought toothpaste last month	0.11	0.016	6.5
log of hhsize	-0.58	0.013	-44.5
Adjusted R-squared	0.64		
Number of observations	3822		

Source: based on authors' own calculations.

Table A8. IHS2-full model. Rural South

	Parameter estimate	Standard error	t-value
Intercept	9.83	0.130	75.5
dependency ratio	-0.17	0.024	-7.2
hhsize	0.03	0.007	4.5
members per room	-0.08	0.013	-6.5
if grass or plastic roof	-0.13	0.014	-9.0
if hh owns a radio	0.05	0.011	4.4
if hh bought mens clothes last 3 months	0.12	0.013	9.6
if hh owns an iron	0.12	0.016	7.5
no. of clothes for head	0.01	0.001	7.7
if hh bought shoes last 3 months	0.09	0.014	6.4
if hh owns a bed	0.08	0.015	5.6
if hh consumed eggs last 7 days	0.10	0.014	7.2
if hh consumed meat last 7 days	0.22	0.013	16.8
if hh consumed rice last 7 days	0.14	0.014	10.4
if hh consumed bread last 7 days	0.22	0.023	9.5
if hh consumed cookingoil last 7 days	0.15	0.012	11.8
if hh consumed sugar last 7 days	0.11	0.012	9.4
if hh bought toothpaste last month	0.12	0.016	7.2
members per room, squared	0.01	0.002	4.5
age of head, squared	0.00	0.000	-2.3
log of hhsize	-0.70	0.027	-26.2
log of age of head	0.14	0.040	3.6
Adjusted R-squared	0.70		
Number of observations	4546		

Source: based on authors' own calculations.

Table A9. IHS2-full model. Urban

	Parameter estimate	Standard error	t-value
intercept	10.63	0.049	218.5
dependency ratio	-0.20	0.044	-4.6
members per room	-0.17	0.027	-6.3
if grass or plastic roof	-0.13	0.023	-5.7
max education among members	-0.07	0.021	-3.2
if hh owns an iron	0.09	0.022	4.2
if hh bought mens clothes last 3 months	0.07	0.019	3.8
no of clothes for head	0.02	0.002	7.1
if hh bought shoes last 3 months	0.06	0.020	3.1
if head sleeps under blankets and sheets in cold season	0.13	0.021	5.9
if hh consumed eggs last 7 days	0.11	0.020	5.6
if hh consumed meat last 7 days	0.26	0.021	12.6
if hh consumed rice last 7 days	0.08	0.020	4.2
if hh consumed bread last 7 days	0.13	0.021	6.3
if hh consumed fresh milk last 7 days	0.18	0.023	7.9
if hh bought toothpaste last month	0.13	0.021	6.3
if hh owns a refrigerator	0.55	0.036	15.2
members per room, squared	0.02	0.004	4.2
max education among members, squared	0.02	0.003	6.9
log of hhsize	-0.56	0.024	-23.7
Adjusted R-squared	0.84		
Number of observations	1431		

Source: based on authors' own calculations.

Table A10. IHS3-full model. Rural North

	Parameter estimate	Standard error	t-value
Intercept	8.25	0.051	160.3
dependency ratio	-0.19	0.045	-4.1
if grass or plastic roof	-0.16	0.023	-7.0
number of radios hh own	0.33	0.046	7.2
if hh owns a bed	0.05	0.022	2.3
no. of clothes for head	0.02	0.003	5.6
if hh bought shoes last 3 months	0.15	0.023	6.5
if hh consumed rice last 7 days	0.20	0.022	9.3
if hh consumed fresh milk last 7 days	0.20	0.028	7.3
if hh consumed rice last 7 days	0.33	0.021	16.0
if hh consumed cookingoil last 7 days	0.18	0.022	8.1
max education among members, squared	0.01	0.002	5.8
age of head, squared	0.00	0.000	1.7
log of hhsize	-0.53	0.019	-28.4
Adjusted R-squared	0.61		
Number of observations	1755		

Source: based on authors' own calculations.

Table A11. IHS3-full model. Rural Central

	Parameter estimate	Standard error	t-value
Intercept	8.64	0.038	229.9
dependency ratio	-0.18	0.029	-6.1
members per room	-0.04	0.006	-7.4
if sand or smoothed mud floor	-0.14	0.023	-5.9
if hh owns a radio	0.11	0.014	7.6
if hh bought mens clothes last 3 months	0.14	0.018	7.8
if hh bought shoes last 3 months	0.12	0.016	7.3
if hh owns an iron	0.18	0.025	6.9
max education among members	0.04	0.007	6.0
if head sleeps under blankets and sheets in cold season	-0.03	0.019	-1.6
if hh owns a bed	0.10	0.019	5.3
if hh consumed eggs last 7 days	0.17	0.016	10.6
if hh consumed meat last 7 days	0.32	0.015	22.1
if hh consumed rice last 7 days	0.20	0.018	11.3
if hh consumed bread last 7 days	0.16	0.024	6.7
if hh consumed fresh milk last 7 days	0.15	0.026	5.7
if hh consumed cookingoil last 7 days	0.15	0.015	10.0
if hh bought toothpaste last month	0.13	0.019	7.0
log of hhsize	-0.52	0.015	-34.5
Adjusted R-squared	0.68		
Number of observations	3484		

Source: based on authors' own calculations.

Table A12. IHS3-full model. Rural South

	Parameter estimate	Standard error	t-value
Intercept	7.77	0.133	58.4
dependency ratio	-0.15	0.025	-5.9
hhsize	0.00	0.007	-0.5
members per room	-0.09	0.014	-6.2
if grass or plastic roof	-0.13	0.014	-9.7
if hh owns a radio	0.10	0.012	8.2
if hh bought mens clothes last 3 months	0.11	0.015	6.9
if hh owns an iron	0.17	0.021	8.3
max education among members	0.04	0.006	6.0
no. of clothes for head	0.01	0.002	7.9
if hh bought shoes last 3 months	0.15	0.015	10.0
if hh owns a bed	0.10	0.015	6.7
if hh consumed eggs last 7 days	0.15	0.014	10.5
if hh consumed meat last 7 days	0.23	0.013	17.9
if hh consumed rice last 7 days	0.15	0.014	10.9
if hh consumed bread last 7 days	0.17	0.021	7.8
if hh consumed fresh milk last 7 days	0.10	0.027	3.7
if hh consumed cookingoil last 7 days	0.14	0.013	10.9
if hh consumed sugar last 7 days	0.17	0.013	13.2
if hh bought toothpaste last month	0.08	0.016	5.2
members per room, squared	0.01	0.002	4.6
age of head, squared	0	0.000	-2.9
log of hhsize	-0.60	0.029	-20.8
log of age of head	0.20	0.041	4.9
Adjusted R-squared	0.73		
Number of observations	4790		

Source: based on authors' own calculations.

Table A13. IHS3-full model. Urban

	Parameter estimate	Standard error	t-value
Intercept	9.10	0.048	188.8
dependency ratio	-0.22	0.043	-5.2
members per room	-0.19	0.024	-7.7
if sand or smoothed mud floor	-0.21	0.022	-9.4
max education among members	-0.12	0.021	-5.8
if hh bought mens clothes last 3 months	0.08	0.020	4.3
no. of clothes for head	0.00	0.001	2.8
if hh bought shoes last 3 months	0.08	0.019	4.1
if head sleeps under blankets and sheets in cold season	0.11	0.020	5.7
if hh owns an iron	0.15	0.022	6.9
if hh consumed eggs last 7 days	0.15	0.020	7.3
if hh consumed meat last 7 days	0.21	0.021	9.9
if hh consumed rice last 7 days	0.13	0.021	6.1
if hh consumed bread last 7 days	0.16	0.022	7.2
if hh consumed fresh milk last 7 days	0.14	0.021	6.5
if hh bought toothpaste last month	0.11	0.020	5.6
if hh owns a refrigerator	0.43	0.029	14.8
members per room, squared	0.02	0.003	5.1
max education among members, squared	0.03	0.003	10.1
log of hhsize	-0.49	0.023	-21.5
Adjusted R-squared	0.78		
Number of observations	2232		

Source: based on authors' own calculations.