Comparable Poverty Estimates for Malawi: 1998 and 2005

This note describes the methodology used to develop comparable poverty indicators for Malawi from the two rounds of the Integrated Household Surveys. The first survey (IHS1) was administered in 1997-98. The second (IHS2) was administered in 2004-05. The issue of comparing welfare or poverty across time and surveys is not new. There is a large and expanding literature on how best to compare welfare indicators from two non-similar surveys. Changes in data collection methods, questions asked, etc., make direct comparison of poverty statistics problematic. This is the dilemma facing the analysis of the IHS data. There were some key differences in the surveys which are detailed in the Basic Information Document. As the poverty statistics rely on consumption expenditure data, one of the most critical differences in the IHS data rounds is how these data are collected. The IHS1 used a diary method to record consumption expenditure for food and frequently purchased items. The IHS2 was redesigned in part to take into account the experiences from IHS1 (both in terms of the fieldwork and subsequent analysis). A major redesign was to collect consumption expenditure information by using recall periods for all food and nonfood consumption expenditure (for example, 7 day recall for food) rather than a diary. Diaries are expensive and difficult to collect accurately especially among rural and illiterate populations. Thirty-eight percent of the 10,698 households in the IHS1 did not complete the diary (or the diary of expenditure was not consistently maintained). On the other hand, the food-recall method is, internationally, a common substitute to the diary for consumption measures to assess poverty. However, the levels of consumption between the two surveys are not directly comparable.

Given these and other revisions to the survey methodology, *__the previous poverty estimates produced from the IHS1 should not be compared to the new poverty estimates from the IHS2__*. Rather, new poverty estimates must be computed from the IHS1 following a comparable approach used with the IHS2 data.

There are numerous non-income dimensions which are directly comparable between IHS1 and IHS2, such as child malnutrition, school attendance, asset ownership, and employment activities. Nevertheless, income poverty is often taken as a primary, singular measure of progress in fighting poverty. Therefore, it is necessary to develop a poverty measure for IHS1 which can, in fact, be compared to the IHS2 poverty measure.

The approach used in developing a revised 1998 poverty estimate, one which is comparable to the IHS2 estimate, follows recently-developed statistical techniques. We develop a comparable poverty measure in IHS1 by first imputing the consumption expenditures in IHS1 (described below). The methodology originated in Elbers et. al. (2002, 2003) and has since been widely applied in different countries, in particular for poverty maps, but also for survey-to-survey imputations (as used here). For example, this method is used to estimate population prevalence of HIV in Malawi, for lack of population-based data (Ivaschenko and Montana, 2005). An example closer to the situation in this note is from India where surveys were changed between rounds. Kijima and Lanjouw (2003) use the methodology to impute poverty at the regional level in India. Finally, recent work from Uganda is also very similar to the situation in Malawi in that the surveys changed over time drawing into question the comparability of the poverty estimates (Luoto, 2005). Examples of a more simplified approach to the one used in these papers is the estimate of poverty status for households in the NSO's Core Welfare Indicator Questionnaire (CWIQ) 2002. Likewise, the IHS1 itself actually *imputed* per capita expenditure and poverty status for 38% of the 10,698 households in the survey (noted above). The imputation approach in IHS1 for these missing households is a simplified version of the approach described below. It was based on one regression model of per capita expenditure, whereas the approach here uses 4 region-specific models, 100 simulations, and estimates of otherwise omitted disturbance terms, for an improved fit.

The steps involved in computing a comparable IHS1 poverty estimate are: 1) estimate per capita expenditure for IHS1 households based on a model of per capita expenditure developed from IHS2 using a set of household characteristics measured in both surveys, and 2) estimate poverty rates for households using the imputed per capita expenditure, applying the IHS2 poverty lines. The main assumption imbedded in this approach is that the correlation between poverty and the set of household characteristics has not changed significantly over time.

To model this we first run a Generalized Least Squares (GLS) regression of the observed log per capita expenditure for household *h* as:

(1) $\ln y_h = \mathbf{x}_h \boldsymbol{\beta} + u_h,$

where $\mathbf{x}_h\boldsymbol{\beta}$ is a vector of k parameters and u_h is a disturbance term satisfying $E[u_h|x_h] = 0$. The set of household characteristics is in vector **x** (described in Table 1.1). The model in (1) is estimated separately for four regions (urban, north rural, central rural, south rural) using the IHS2 data. In a simple model of expenditure, we would impute values of per capita expenditure for IHS1 households, based on the set of **x** covariates and the estimated $\boldsymbol{\beta}$ values. However, we can improve up on this estimate as follows. Because the disturbance term for households in the target population (IHS1 in this case) are always unknown, we estimate the expected value of the indicator given the IHS1 households' observable characteristics and the model of expenditure in (1). We denote this expectation as:

(2) $\qquad \mu_v^s = E[W \mid \mathbf{X}_v^s, \boldsymbol{\xi}],$

where $\boldsymbol{\xi}$ is the vector of model parameters, including those which describe the distribution of the disturbances, and the superscript 's' indicates that the expectation is conditional on the sample of IHS1 households from district v rather than a census of households.

In constructing an estimator of μ_v^s we replace the unknown vector $\boldsymbol{\xi}$ with consistent estimators, $\hat{\boldsymbol{\xi}}$, from the IHS2 expenditure regression. This yields $\hat{\mu}_v^s$. This expectation is generally analytically intractable so simulation is used to obtain our estimator, $\tilde{\mu}_v^s$. One hundred simulated draws are performed to derive our estimator for $\hat{\mu}_v^s$ in each of the four models (urban, north rural, central rural, south rural). The four models use household weights and therefore serve as estimates for the *population* considered and not just the sample at hand. See Kijima and Lanjouw (2003) for more detail on the prediction error associated with our estimator $\tilde{\mu}_v^s$ of the expected value of per capita expenditure for a given region.

For the estimations we used the *povmap* program that builds on the methodology outlined. The *povmap* program is a program developed by the World Bank which is designed to make poverty maps based on a survey and a census. The program is still under development. As noted above, we estimate four models, which allows the estimated parameters $\boldsymbol{\beta}$ to vary across the four areas and give a better fit of the models. For the simulations, we ran 100 simulations and used non-parametric distributions for both cluster draws and household draws. The explanatory variables

used can be seen in Table 1 below, in addition to district indicator variables. The mean of the 100 simulated betas is presented in the Appendix Table.

After estimating per capita expenditure 100 times for every IHS1 household, the *povmap* program then computes the poverty indicator (poor or non-poor) for each household's simulated per capita expenditure. Thus, every household has 100 imputed values for being poor. The median value is assigned to that household for the final computation of the poverty rate in IHS1.

Table 2 shows the imputed poverty rates for IHS1 and compares them with the IHS2 poverty estimates.

Table 1: Descriptive statistics of variables to impute IHS1

	IHS1	IHS2
Demographic		
Household head is male*	0.78	0.81
Proportion of females in household	0.51	0.51
Spouse of household head is present*	0.75	0.80
Number of household members squared	37.5	39.1
Number of household members	5.57	5.71
Dependency ratio	0.45	0.48
Welfare		
Renting house*	0.12	0.07
Own water supply*	0.06	0.05
Collects firewood*	0.78	0.76
Household owns bed*	0.38	0.33
Household owns table*	0.39	0.39
Household owns chair*	0.48	0.48
Household owns refrigerator*	0.03	0.03
Household owns stove*	0.05	0.03
Household owns washing machine*	0.00	0.00
Household owns tv*	0.02	0.05
Household owns bicycle*	0.38	0.41
Household owns motorcycle*	0.01	0.01
Household owns boat*	0.01	0.01
Household owns axe*	0.60	0.68
Household owns sickle*	0.41	0.59
Household owns panga*	0.51	0.68
Household owns ox*	0.03	0.03
Education		
All kids 6-12 years old in school*	0.47	0.46
No kids in the household*	0.31	0.30
Adult females highest education: standard 5-8*	0.28	0.32
Adult females highest education: standard 8 *	0.07	0.15
Head education: standard 1-4*	0.23	0.23
Head education: standard 5-8*	0.35	0.33
Head education: higher than standard 8*	0.14	0.16
Labor market		
Household has enterprise*	0.24	0.33
Proportion of adults employed	0.56	0.55

Notes: Dummy variables for districts were also included. * indicates binary variables which are 1 if description is true for the household, else 0.

Table 2: Poverty Rates for 1998 and 2005

	1998 IHS1	2005 IHS2
Malawi	53.9	52.4
by Region		
Urban	19.6	25.4
North rural	56.0	56.3
Central rural	48.2	46.7
South Rural	67.2	64.4

References

Elbers, Chris, Jean O. Lanjouw, and Peter Lanjouw. 2003. "Micro-Level Estimation of Poverty and Inequality." *Econometrica* 71(1): 355-64.

Elbers, Chris, Jean O. Lanjouw, and Peter Lanjouw. 2002. "Micro-Level Estimates of Welfare." World Bank Policy Research Working Paper No. 2911.

Ivaschenko, Oleksiy and Livia Montana. 2005. "Using Small Area Estimation Techniques to Produce Geographic Profile of HIV/AIDS Prevalence: A Case Study of Malawi." mimeo, The World Bank.

Luoto, Jill. 2005. "A Further Analysis of the UNHS-1 and UNHS-2: An Attempt to Explain The Increase in Measured Poverty." mimeo, University of California, Berkeley.

Kijima, Yoko and Peter Lanjouw. 2003. "Poverty in India during the 1990s: A Regional Perspective." World Bank Policy Research Working Paper 3141.

Appendix Table: Mean of Simulated betas

	Urban	Rural North	Rural Central	Rural South
Demographic				
Household head is male	0.13	0.05	0.08	0.09
Proportion of females in household	-0.16	-0.12	-0.15	0.01
Spouse of household head is present	-0.12	-0.10	-0.06	-0.14
Number of household members squared	-0.28	-0.19	-0.24	-0.18
Number of household members	0.01	0.01	0.01	0.01
Dependency ratio	-0.15	-0.33	-0.15	-0.26
Welfare				
Renting house	-0.03	-0.08	0.10	0.15
Own water supply	0.20	0.19	0.09	0.19
Collects firewood	-0.29	-0.29	-0.13	-0.14
Household owns bed	0.27	0.17	0.21	0.14
Household owns table	0.06	0.01	0.04	0.07
Household owns chair	0.02	0.03	0.07	0.10
Household owns refrigerator	0.49	0.13	0.30	0.22
Household owns stove	0.31	-0.07	0.45	0.24
Household owns washing machine	0.53	0.33	0.02	0.06
Household owns tv	0.31	0.52	0.20	0.35
Household owns bicycle	0.06	0.15	0.14	0.14
Household owns motorcycle	0.05	0.31	0.28	0.21
Household owns boat	0.04	0.09	0.00	0.28
Household owns axe	0.00	-0.06	0.05	0.06
Household owns sickle	0.05	0.04	0.09	0.05
Household owns panga	0.05	0.10	0.06	0.06
Household owns ox	-0.26	0.21	0.13	0.20
Education				
All kids 6-12 years old in school	0.08	0.03	0.05	0.07
No kids in the household	0.07	-0.03	0.04	0.10
Adult females highest education: standard 5-8	0.07	0.02	0.04	0.01
Adult females highest education: standard 8	0.26	0.12	0.18	0.15
Head education: standard 1-4	0.08	-0.01	0.05	0.06
Head education: standard 5-8	0.15	0.03	0.12	0.10
Head education: higher than standard 8	0.20	0.07	0.15	0.23
Labor market				
Household has enterprise	0.07	0.13	0.10	0.08
Proportion of adults employed	0.20	-0.16	0.01	0.07
District				
Blantyre city	0.07			
Zomba city	-0.07			
Lilongwe city	0.25			
Nkhata-bay		-0.29		
Rumphi		-0.25		
Karonga		-0.25		
Chitipa		-0.26		
Ntcheu			-0.07	
Dedza			-0.05	
Salima			-0.11	

Lilongwe rural			0.14	
Mchinji			-0.16	
Kasungu			0.08	
Dowa			0.07	
Ntchisi			0.05	
Nsanje				-0.19
Chikwawa				-0.04
Blantyre rural				0.05
Zomba rural				-0.16
Thyolo				-0.16
Mulanje				-0.21
Phalombe				-0.09
Machinga				-0.25
Mangochi				-0.07
Chiradzulu				-0.05
Constant	10.38	10.98	10.50	10.15