ESTIMATING CONSUMPTION-BASED POVERTY IN THE ETHIOPIA DEMOGRAPHIC AND HEALTH SURVEY[1]

Mark Schreiner[2]

Abstract

Inequalities in health are linked with poverty, but quantifying the health/poverty nexus is hampered by data constraints. In particular, the most common measure of poverty compares consumption with poverty lines, but consumption surveys often do not collect detailed health data. Conversely, the large repository of internationally-comparable Demographic and Health Surveys has detailed health data but no consumption data. This has led DHS researchers who want to control for socio-economic status use an asset index defined in terms of housing characteristics and the ownership of durable goods. While this is a valid conception of poverty, it is difficult to compare it with the more-common consumption-based measure. This paper presents a simple poverty scorecard for Ethiopia based on the poverty-mapping approach of Elbers et al. (2003). It allows researchers to estimate the likelihood that consumption is below a given poverty line using nine verifiable, inexpensive-to-collect indicators found in both Ethiopia's 2005 DHS and in the 2004/5 Household Income, Consumption, and Expenditure Survey. It turns out that the poverty scorecard and the DHS asset index do not generally rank people the same, so estimates of consumption-based poverty in the DHS should use the poverty scorecard, not the DHS asset index. The bias and precision of scorecard estimates compare well with that of other tools, suggesting that government could use it to track poverty in years between national household expenditure surveys.

Key words: Poverty measurement, asset index, poverty mapping, Africa, Ethiopia, health equity

---

[1] The final version of this article was submitted in February 07, 2013.

[2] Mark Schreiner directs Microfinance Risk Management, L.L.C., microfinance.com. He is also a Senior Scholar at the Center for Social Development in Washington University in Saint Louis.

Acknowledgements

HICE data come from Ethiopia's Central Statistical Agency; DHS data come from IFC Macro. Thanks go to Karen Foreit, Kiersten Johnson, Rachel Sanders, and Bridgette Wellington. This work was supported by MEASURE Evaluation PRH, which is funded by the U.S. Agency for International Development (USAID) through cooperative agreement associate award number GPO-A-00-09-00003-00 and is implemented by the Carolina Population Center at the University of North Carolina at Chapel Hill, in partnership with Futures Group, Management Sciences for Health, and Tulane University. The opinions expressed are those of the author and do not necessarily reflect the views of neither USAID nor the U.S. government.

1. Introduction

This paper applies a simple poverty scorecard (Table 1) to Ethiopia's 2005 Demographic and Health Survey (DHS) to estimate the likelihood that a given person has consumption below a given poverty line. This allows researchers to look at how health outcomes vary with socio-economic status as seen through a consumption lens.

Table 1: Simple scorecard for estimating consumption-based poverty in the Ethiopia DHS

Entity	Name	ID	Date (DD/MM/YY)
Client:			Joined:
Region:			Today:
Service point:			HH size:

Indicator	Value	Points
1. How many people usually live with the household?	A. Nine or more	0
	B. Eight	2
	C. Seven	6
	D. Six	9
	E. Five	14
	F. Four	21
	G. Three	27
	H. One or two	42
2. What is the highest grade the female head/spouse has completed?	A. Four or less	0
	B. No female head/spouse	0
	C. Five or six	4
	D. Seven to nine	10
	E. Ten or higher	13
3. What is the main material of the walls of the residence?	A. Stone with mud, stone with lime/cement, or cane/trunks/bamboo/reed	0
	B. Bamboo/wood, uncovered adobe, plywood, carton, no walls, or other	5
	C. Cement, bricks, cement blocks, covered adobe, or wood planks/shingles	11
4. What type of toilet facility do members of your household usually use?	A. Non-flush or none	0
	B. Flush	5
5. What type of fuel does your household mainly use for cooking?	A. Wood or straw/shrubs/grass, or animal dung	0
	B. All others	5
6. Does the household have a bed?		0
		5
7. Does the household have a radio?		0
		8
8. Does any member of this household own any land that can be used for agriculture?		0
		6
9. Does the household own any cattle, sheep or goats?	A. No	0
	B. Yes	5
Total score:		

Such a consumption-based poverty measure is useful because while the DHS surveys are the largest repository of nationally representative data on population, health, HIV, and nutrition (covering more than 75 countries, often for multiple rounds), they do not collect data on consumption. DHS researchers seeking to relate health outcomes to socio-economic status have had to rely on an asset index that in recent years comes pre-packaged with DHS data (Rutstein and Johnson, 2004).

The DHS asset index is widely used. Constructed with Principal Components Analysis (PCA), it defines socio-economic status in terms of housing characteristics, asset ownership, and agricultural employment. PCA does not explicitly model any particular conception of poverty; rather, it finds the linear combination that maximizes the explained variation among a given set of indicators. Nevertheless, the resulting DHS asset indexes seem to be related to socio-economic status, especially when this is conceived as "permanent income" or "expected long-term control over resources". The indexes are correlated in intuitive ways with outcomes such as fertility (Bollen, Glanville, and Stecklov, 2007), use of emergency obstetric care (Pitchforth et al., 2007), maternal and child mortality (Knowles et al., 2008), food security (Dekker, 2006, for Ethiopia), child health and nutrition (Sahn and Stifel, 2003), and education (Filmer and Pritchett, 2001).

Table 2 is Ethiopia's 2000 DHS asset index (Gwatkin et al., 2007). It has 23 indicators and 113-point values. The DHS index ranks people on a relative scale; a higher value of the index implies higher socio-economic status or lower poverty.[3]

The contribution of this paper is to allow DHS analysis in terms of consumption-based poverty: a person is poor if the monetized value of his/her per-capita household consumption is below a poverty line such as the Millennium Development Goals' $1.25/day at 2005 purchase power parity (PPP). Consumption-based poverty lines are commonly used by governments, the World Bank, policymakers, and others.

[3] A PCA asset index may be seen as a measure of absolute poverty as defined by its indicators and points, and as such, it can be used to measure poverty over time and across countries (Booysen et al., 2008; Sahn and Stifel, 2000). Most applications, however, treat the index as a relative measure.

Table 2: DHS asset index for Ethiopia (2000)

Question	Score if "Yes"	Score if "No"	Item score
1. In your household is/are there . . .?			
Electricity	0.16178	-0.04632	
One or more radios	0.10087	-0.03931	
One or more televisions	0.23657	-0.01390	
One or more bicycles	0.17138	-0.00198	
One or more motorcycles. Scooters	0.18746	-0.00005	
One or more cars, trucks	0.26830	-0.00248	
One or more telephones	0.24991	-0.00911	
One or more electric mitads	0.23539	-0.01491	
One or more kerosene lamps, pressure lamps	0.00624	-0.00124	
One or more beds, tables	0.04349	-0.04492	
One or more cattles, camels	-0.05021	0.05846	
One or more horses, mules, donkeys	-0.05443	0.01551	
One or more sheep, goats	-0.04996	0.02424	
2. Do any of the members of your household own . . .?			
A house	-0.03046	0.12970	
Crop land	-0.05011	0.10455	
Cash crops	-0.05072	0.01277	
3. Do the members of your household work their own or family's agricultural land?	-0.05908	0.02255	
4. What is the principal source of drinking water for your household?			
Piped water in dwelling	0.20251	-0.00039	
Piped water outside dwelling	0.20125	-0.02419	
Piped water in public faucet	0.07932	-0.01497	
Well in dwelling	-0.04956	0.00298	
Covered well	-0.02492	0.00144	
River, canal, surface water	-0.04904	0.01974	
Open spring	-0.05232	0.02165	
Covered spring	-0.02578	0.00098	
Rainwater	-0.06333	0.00012	
Other	0.01952	-0.00005	
5. What is the principal type of fuel for cooking use by your household?			
Electricity	0.23445	-0.00037	
Gas	0.26576	-0.00040	
Biogas	0.15508	-0.00020	
Kerosene	0.19738	-0.02609	
Charcoal	0.13642	-0.00459	
Wood	-0.03490	0.09703	
Dung, manure	-0.02591	0.00318	
Other	0.02881	-0.00001	
6. What is the principal type of toilet facility used by your household?			
Private flush toilet	0.25570	-0.00235	
Private latrine	0.11641	-0.03987	
VIP latrine	0.19368	-0.00174	
Bush, field as latrine	-0.04642	0.12343	
7. What is the principal material used for the floors in your household?			
Dirt, sand, dung	-0.03345	0.17974	
Cement	0.18049	-0.01260	
Wood plank	0.20963	-0.00078	
Parquet, polished wood	0.24016	-0.00250	
Bamboo, reed	0.12413	-0.00067	
Vinyl, asphalt	0.20579	-0.00944	
Tiles	0.22425	-0.00156	
Carpet	0.06956	-0.00121	
Other	0.18014	-0.00076	
8. What is the principal material used for the roof of your household?			
Corrugated iron	-0.03323	0.00290	
Cement, concrete	-0.05660	0.06001	
Wood and mud	-0.04760	0.00024	
Thatch	0.04694	-0.00004	
Bamboo, reed	-0.05413	0.00437	
Plastic sheet	0.11469	-0.05285	
Mobile roof of nomads	0.11469	-0.05285	
Other	-0.05212	0.00037	
9. How many people are there for each sleeping room in your household?	[(# people-3.89)/2.14] x -0.029		
Total household asset score (sum of individual score items):			

Source: Gwatkin et al., 2007

While the consumption-based conception of poverty is not more valid than the asset-based conception, it is more commonly used and better understood.[4] Thus, DHS research might be more relevant for policy if it compared health outcomes not only with an asset index but also with consumption-based poverty. The rub is that collecting consumption data is complex and costly (Sahn and Stifel, 2003; Deaton and Zaidi, 2002). For the example of Ethiopia's 2004/5 Household Income, Consumption, and Expenditure Survey (HICE), enumerators visited each household 16 times, asking each time about a lengthy list of consumption items. The cost of doing this explains why the DHS does not collect data on consumption.

The scorecard in this paper allows researchers to estimate consumption-based poverty in Ethiopia's 2005 DHS. The estimates come from a three-step method that follows that of poverty mapping (Elbers et al., 2003). First, potential poverty indicators are matched between a survey that collects consumption data (Ethiopia's 2004/5 HICE) and another survey that does not collect consumption data (Ethiopia's 2005 DHS). Second, a poverty scorecard is constructed based on the 2004/5 HICE, using only indicators that appear in both the 2004/5 HICE and in the 2005 DHS. Third, the scorecard is applied to the 2005 DHS to estimate consumption-based poverty.

This poverty-mapping approach rests on three difficult-to-test assumptions. The first is that scorecard indicators are well-matched across the two surveys, so that all else constant, households are as likely—for example—to report owning a radio in the 2004/5 HICE as in the 2005 DHS. Unfortunately, the strength of the match is never certain, as questions may be worded differently, offer different response options, or appear in a different context.

The second strong assumption is that the relationships between indicators and poverty are constant over time (Christiaensen et al., 2010). This is plausible for the Ethiopia data analyzed here, as the 2004/5 HICE covered July 2004 and February 2005, and the 2005 DHS covered April to August 2005. The assumption is less plausible for longer time gaps and in periods of sharper socio-economic change.

[4] Arguments in favor of the asset-based view include Carter and Barrett (2006), Schreiner and Sherraden (2006), and Sherraden (1991). In practice, the two views, though distinct, are tightly linked, as income/consumption are flows of resources received/consumed from the use of stocks of assets (wealth). The two views are low-dimensional simplifications—due to practical limits on definitions and measurement—of a higher-dimensional and more complete conception of the production of human well-being. Section 6 below discusses the correlation between health outcomes, asset indexes, and consumption-based poverty likelihoods.

The third strong assumption is that the scorecard is applied to nationally representative groups (Elbers, Lanjouw, and Leite, 2008; Tarozzi and Deaton, 2007). This holds for Ethiopia's 2005 DHS.

If all three assumptions hold, then scorecard estimates of consumption-poverty rates are unbiased, that is, their average in repeated samples matches the true rate.

The scorecard can be used to estimate the poverty likelihood of a person responding to a survey. It can also be used to estimate the poverty rate of a group of people, such as those who use public-sector health services or women whose most recent childbirth was not attended by trained personnel.

The results here suggest that the approach here has three main possible uses. First, researchers can the poverty scorecard use as a substitute for (or as a complement to) the DHS asset index when relating socio-economic status to health outcomes. Second, local government and pro-poor organizations can use the scorecard to inform their social-performance management by estimating head-count poverty rates for small regions in which the DHS has a larger sample than does a national expenditure survey. Third, national policy-makers can use the scorecard to monitor poverty in years between the less-frequent and more-costly national household expenditure surveys.[5]

The next section documents data, poverty lines, and indicator-matching. Section 3 describes scorecard construction. Section 4 defines the concept of poverty likelihood and details the estimation of consumption-based poverty rates. Section 5 compares the poverty scorecard with the pre-packaged DHS asset index in terms ranking people in Ethiopia's 2005 DHS. Section 6 places the scorecard in the context of related work. The final section is a summary.

[5] When applied to representative cross-sections at two points in time, then the scorecard measures the net change in poverty, but it does not reveal who moved up or down. When applied repeatedly to the same set of households, then the scorecard can reveal poverty dynamics, an approach with a long history in Ethiopia (Alem, 2011; Bigsten and Shimeles, 2008; Dercon, 2006; Kedir and McKay, 2005; Bigsten et al., 2003).

2. Data, poverty lines, and indicator-matching

This section briefly discusses the building blocks of the analysis. It documents the data and poverty lines used to construct and test the poverty scorecard, and it also discusses the process and results of matching indicators across surveys.

2.1 Data

The poverty scorecard is constructed from a random sample of half the people clustered in the 21,297 households in the 2004/5 HICE. The other half of the HICE is used for testing accuracy.

2004/5 HICE refers to the combination of two Ethiopian surveys, the 2004 Household Income, Consumption and Expenditure Survey (consumption data) and the February 2005 Welfare Monitoring Survey (scorecard indicators). All households in the 2005 WMS are also in the 2004 HICE, and both surveys are nationally representative and were conducted by Ethiopia's Central Statistical Agency (CSA).

2.2 Poverty lines

Ethiopia has no official, published poverty lines. Dercon (1997) uses the 1995/6 HICE and information from the CSA to derive a food line and a food-plus-non-food line in 1995/6 prices. Unfortunately, there are no regional and temporal price indexes available for converting these lines to 2004/5 prices. Other major poverty documents for Ethiopia (Woldehanna, Hoddinott, and Dercon, 2008; Ministry of Finance and Economic Development, 2006 and 2002) apparently use Dercon's lines and convert consumption to 1995/6 prices, but none of these documents provide enough information for replication. This paper uses the international "extreme" poverty line of $1.25 per person per day at 2005 purchase-power parity (Ravallion, Chen, and Sangraula, 2009), unadjusted for regional differences in cost-of-living.[6]

[6] In Ethiopia, prices vary widely by region (Gebremedhin and Whelan, 2008). Not adjusting for this overstates poverty in low-price regions and understates it elsewhere; weakening scoring's ability to reflect the real links between indicators and poverty. The World Bank normally does not adjust the $1.25/day line for intra-country price differences, so doing it here would harm inter-country comparability.

Ethiopia's $1.25/day 2005 PPP line is derived from the 2005 PPP exchange rate for "individual consumption expenditure by households" of ETB2.75 per $1.00 (World Bank, 2008) and consumer price indexes (Loening, Durevall, and Birru, 2008) of 77.43 for July 2004 and 82.88 on average for 2005, using the formula in Sillers (2006):

$$(2005\text{PPP exchange rate}) \cdot \$1.25 \cdot \left(\frac{\text{CPI}_{\text{July 2004}}}{\text{CPI}_{\text{Ave.2005}}}\right) = \left(\frac{\text{ETB2.75}}{\$1.00}\right) \cdot \$1.25 \cdot \left(\frac{77.43}{82.88}\right) = \text{ETB3.21}.$$

The scorecard is built with the $1.25/day line and person-level weights. Because policy-makers may want to use different or various poverty lines, scores from the single scorecard are also calibrated to poverty likelihoods for $1.75/day and $2.50/day.

Table 3 reports poverty rates in the 2004/5 HICE. For the $1.25/day line, 34.8 percent of Ethiopians are poor. For $1.75/day, 61.8 percent are poor, and for $2.50/day, 85.8 percent are poor. Oromiya Region has the lowest poverty rate by $1.25/day (29.3 percent), and Tigray and SNNPR have the highest (44.6 percent).

Table 3: Sample sizes, poverty lines, and poverty rates for all-Ethiopia and by construction/validation sub-samples, regions, and poverty lines

		International 2005 PPP		
	Households	$1.25/day	$1.75/day	$2.50/day
Poverty line (ETB/person/day)	N/A	3.21	4.49	6.42
Poverty rate (% people)				
All Ethiopia	21,297	34.8	61.8	85.8
Construction sub-sample	10,675	34.9	61.8	85.8
Validation sub-sample	10,622	34.8	61.8	85.8
Poverty rate (% people) by region				
Tigray	1,716	44.6	67.2	84.0
Afar	951	34.9	55.9	78.8
Amhara	3,985	37.5	66.4	89.7
Oromiya	4,622	29.3	57.9	84.8
Somali	1,152	42.1	65.3	85.7
Benshangul-Gumuz	1,064	31.9	55.7	83.4
SNNPR	3,084	44.6	67.2	84.0
Harari	635	34.9	55.9	78.8
Addis Ababa	3,417	37.5	66.4	89.7
Dire Diwa	671	35.4	56.7	76.2

Source: 2004/5 HICE. SNNPR is Southern Nations, Nationalities, and People's Region.

2.3 Matching indicators across the HICE and DHS

The validity of the estimation of consumption-based poverty in this paper rests on the assumption that indicators in the 2004/5 HICE mean the same as in the 2005 DHS. Ideally, both surveys would be fielded at the same time and their indicators would have identical wording, offer identical response options, appear in identical contexts, and elicit identical distributions of responses. Furthermore, ideal indicators would have a balanced distribution of responses (for example, half own an asset, and half do not, rather than 95 percent are non-owners) with variation linked with the poverty of people close to a given poverty line.

The Appendix documents the matching of scorecard indicators for Ethiopia. In the poverty-mapping approach, indicators are considered as "matched" only if the distributions of responses across the two surveys are not statistically different at conventional levels such as $p < 0.10$ or $p < 0.05$. Based on a Chi-squared test for the independence of the response distributions, this standard would reject eight of the nine indicators in the scorecard here, even when most of the ideal conditions appear to be met. Of course, this may be particular to these Ethiopia surveys, and it need not reflect on the usefulness or validity of the general approach.

To be able to proceed, this paper accepts as "well-matched" indicators for which the difference in response percentages across surveys does not much exceed 10 percent of the most common response.

In broad terms, the indicators are of four types. The first is household size. This is the most powerful indicator. As might be expected, larger households more likely to be poor (Lanjouw and Ravallion, 1995).

The second type of indicator is well-matched but has highly unbalanced responses and serves mostly to distinguish the least-poor five or ten percent of people from the more-poor masses. These are the education of the female head, the type of toilet facility, and the main type of cooking fuel.

The third type of matched indicator is the type of wall. Here, responses are more evenly distributed, so power is better for distinguishing among people close to the relevant poverty lines. The wording of some response options, however, differs greatly

between the surveys, and although the responses can be grouped to give similar frequencies, it is not clear how such different wordings can lead to such close matches.[7]

The fourth and final type of matched indicators concern asset ownership: beds, radios, agricultural land, and cattle, sheep, and goats. These indicators appear to fulfill many of the ideal requirements, yet reported ownership rates differ a lot across surveys.[8] Nevertheless, these indicators are included because, apart from household size and wall type, they are the only way to distinguish among the more and less poor in the non-elite masses, particularly in rural areas.

To sum up, indicators for Ethiopia do not match up as well as might be hoped. While this does not affect the construction of the scorecard from the 2004/5 HICE, it does matter for the application of the scorecard to the 2005 DHS. The weaker the match, the weaker the link between measured accuracy in the 2004/5 HICE and assumed accuracy in the 2005 DHS. Unfortunately, the extent of this inaccuracy is untestable.

3. Scorecard construction

The first step in scorecard construction is to identify potential matched indicators in the 2004/5 HICE and the 2005 DHS. About 46 potential indicators were identified in the areas of:

- Family composition (such as number of household members in an age range)
- Education (such as school attendance by children in an age range)
- Housing (such as the main type of cooking fuel)
- Ownership of durable goods (such as beds and radios)
- Ownership of agricultural assets (such as land and livestock)

Initial screening eliminated 22 potential indicators due to very weak matching, highly unbalanced response distributions, or similarity with other indicators that are simpler or more intuitive.

[7] It is possible that the wordings are more similar in local languages.

[8] See the Appendix for details. The causes of the differences is unknown, but they may hinge on differences in enumeration. For example, many households have non-functioning assets (such as broken radios, or radios with no batteries), and DHS enumerators may have handled such cases differently than HICE enumerators.

The scorecard is built using the $1.25/day 2005 PPP line and Logit regression on the construction sub-sample from the 2004/5 HICE. Indicator selection uses both judgment and statistics (forward stepwise). The first step is to use Logit to build one scorecard for each candidate indicator. Each scorecard's accuracy is taken as "c", a measure of ability to rank by poverty status (SAS Institute Inc., 2004).

One of these one-indicator scorecards is then selected based on several factors (Schreiner et al., 2004; Zeller, 2004), including improvement in accuracy, likelihood of acceptance by users (determined by simplicity, cost of collection, and "face validity" in terms of experience, theory, and common sense), sensitivity to changes in poverty, variety among indicators, applicability across geographic regions, and verifiability.

A series of two-indicator scorecards are then built, each based on the one-indicator scorecard selected from the first step, with a second candidate indicator added. The best two-indicator scorecard is then selected, again based on "c" and judgment. These steps are repeated until additional indicators do not improve power.

The final step is to transform the estimated Logit coefficients into non-negative integers such that total scores range from 0 (most likely below a poverty line) to 100 (least likely below a poverty line). This linear transformation makes the scorecard's points simple for users, and it does not affect ranks.

This algorithm is the Logit analogue to the familiar R^2-based stepwise with least-squares regression. It differs from naïve stepwise in that the criteria for selecting indicators include not only statistical accuracy but also judgment and non-statistical factors. The use of non-statistical criteria can improve robustness through time and helps ensure that indicators are simple and intuitive.

The single poverty scorecard here applies to all of Ethiopia. Tests for Indonesia (World Bank, 2012), India and Mexico (Schreiner, 2006 and 2005), Sri Lanka (Narayan and Yoshida, 2005), and Jamaica (Grosh and Baker, 1995) suggest that segmenting scorecards by urban/rural does not improve ranking accuracy much, although it may improve the accuracy of estimated poverty rates (Tarozzi and Deaton, 2007).

4. Estimates of poverty likelihoods for individuals, and estimates of poverty rates for groups

This section describes how scores are converted to poverty likelihoods, that is, the probability that an individual person has consumption below a given poverty line. It also explains how the poverty likelihoods of individuals in a group are aggregated to estimate the group's consumption-based poverty rate. The accuracy of estimates of poverty rates is measured with the validation sample of the 2004/5 HICE, providing the best-available guess of accuracy when the scorecard is applied with the 2005 DHS.

4.1 Poverty likelihoods and their calibration with scores

The sum of scorecard points for a person is called the score. As described above, scores range from 0 to 100. While higher scores indicate less likelihood of being below a poverty line, the scores themselves are only ordinal and do not have equal-interval or ratio units. For example, doubling the score does not double the likelihood of being above a poverty line.

To get equal-interval and ratio units, scores are converted to poverty likelihoods, that is, probabilities of being below a poverty line. This is done via simple look-up tables. For the example of the $1.25/day 2005 PPP line, scores of 10–14 have a poverty likelihood of 25.9 percent, scores of 15–19 have a poverty likelihood of 20.1 percent, and so on (Table 4).

The poverty likelihood associated with a score varies by poverty line. With the $1.75/day 2005 PPP line, for example, scores of 10–14 are associated with a poverty likelihood of 85.9 percent (Table 4).

Table 4: Poverty likelihoods by score and poverty line

Score	Poverty likelihood		
	$1.25/day	$1.75/day	$2.50/day
0–4	45.3	71.9	91.3
5–9	32.6	80.1	95.5
10–14	25.9	62.7	85.9
15–19	20.1	57.9	77.4
20–24	16.6	41.6	65.8
25–29	9.3	30.2	53.6
30–34	4.8	22.2	43.5
35–39	3.8	9.4	27.9
40–44	1.7	8.0	25.4
45–49	1.1	5.1	14.1
50–54	1.0	4.2	9.6
55–59	0.0	0.0	3.7
60–64	0.3	0.3	1.3
65–69	0.0	1.9	5.4
70–74	0.0	0.0	4.7
75–79	0.0	0.0	2.5
80–84	0.0	0.0	0.0
85–89	0.0	0.0	0.0
90–94	0.0	0.0	0.0
95–100	0.0	0.0	0.0

A given score is non-parametrically associated ("calibrated") with a poverty likelihood by defining the poverty likelihood as the share of households in the 2004/5 HICE construction sub-sample with the score who are below a given poverty line.

For the example of the $1.25/day 2005 PPP line (Table 5), there are 3,108 (normalized) people in the construction sub-sample with a score of 10–14, of whom 806 (normalized) are below the poverty line. The estimated poverty likelihood associated with a score of 10–14 is then 25.9 percent, because 806 ÷ 3,108 = 25.9 percent. The same method is used to calibrate scores with estimated poverty likelihoods for the other poverty lines.

Although the points in the scorecard are transformed Logit coefficients, scores are not converted to poverty likelihoods via the Logit formula of 2.718281828^{score} x $(1+2.718281828^{score})^{-1}$. This is because the Logit formula is esoteric and difficult to compute by hand. It is more intuitive to define the poverty likelihood as the share of households with a given score in the construction sample who are below a poverty line. Thus, the (transformed) Logit coefficients are used to order people by relative ranks, and the ranks are then calibrated with absolute poverty likelihoods.

Table 5: Derivation of poverty likelihoods linked with scores, example poverty line of $1.25/day 2005 PPP

Score	People below poverty line		All people at score		Poverty likelihood (%)
0–4	696	÷	1,536	=	45.3
5–9	484	÷	1,485	=	32.6
10–14	806	÷	3,108	=	25.9
15–19	956	÷	4,763	=	20.1
20–24	1,036	÷	6,250	=	16.6
25–29	712	÷	7,645	=	9.3
30–34	477	÷	9,923	=	4.8
35–39	474	÷	12,349	=	3.8
40–44	179	÷	10,595	=	1.7
45–49	117	÷	10,991	=	1.1
50–54	95	÷	9,562	=	1.0
55–59	0	÷	6,484	=	0.0
60–64	16	÷	4,840	=	0.3
65–69	0	÷	2,971	=	0.0
70–74	0	÷	2,698	=	0.0
75–79	0	÷	2,184	=	0.0
80–84	0	÷	1,736	=	0.0
85–89	0	÷	398	=	0.0
90–94	0	÷	483	=	0.0
95–100	0	÷	0	=	0.0

Number of people normalized to sum to 100,000.

4.2 Estimates of a group's poverty rate

A group's estimated poverty rate is the average of the estimated poverty likelihoods of the individuals in the group. To illustrate, suppose a program samples three people on Jan. 1, 2010 and that they have scores of 20, 30, and 40, corresponding to poverty likelihoods of 16.6, 4.8, and 1.7 percent ($1.25/day, Table 4). The group's estimated poverty rate is the people's average poverty likelihood of (0.166 + 0.048 + 0.017) ÷ 3 = 7.7 percent.[9]

4.3 Accuracy of estimates of poverty rates

As long as the relationship between indicators and poverty does not change and as long as the scorecard is applied to a representative sample of people from the same population from which it was constructed, the scorecard produces unbiased estimates of poverty rates. Unbiased means that in repeated samples, the average estimate matches the true value.

Of course, the relationship between indicators and poverty does change over time and across sub-groups within Ethiopia's population, so the scorecard will generally be biased to some unknown extent when applied—as it must be in practice—after the end of the HICE fieldwork in February 2005 or when applied to non-nationally representative groups. To the extent that indicators are mismatched, it will also be biased when applied to the 2005 DHS. Unfortunately, this bias cannot be measured, and accuracy as measured for the 2004/5 HICE validation sample is the best available approximation of accuracy for the 2005 DHS.

How accurate are scorecard estimates of poverty rates for nationally representative samples in the period when the 2004/5 HICE was in the field? Table 6 reports estimates of bias (average differences between estimated and true poverty rates) as well as precision (confidence intervals for the differences) for the scorecard applied to 1,000 bootstrap samples of size n = 16,384 from the 2004/5 HICE validation sample. For $1.25/day, the scorecard is too low by 0.8 percentage points; on average, it estimates a poverty rate of 34.0 percent for the validation sample, but the true value is 34.8 percent

[9] The group's poverty rate is not the poverty likelihood associated with the average score. Here, the average score is (20 + 30 + 40) ÷ 3 = 30, and the poverty likelihood associated with the average score is 4.8 percent. This is not the 7.7 percent found as the average of the three poverty likelihoods associated with each of the three scores.

(Table 3). For $1.75/day, bias is –1.4 percentage points, and for $2.50/day, bias is –0.3 percentage points.[10]

In terms of precision, the 90-percent confidence interval for a group's estimated poverty rate at a point in time with n = 16,384 is 0.5 percentage points or less (Table 6). This means that in 900 of 1,000 bootstraps of this size, the difference between the estimate and the true value is within ±0.5 percentage points of the average difference. In the specific case of $1.25/day and the validation sample, 90 percent of all samples of n = 16,384 produce estimates that differ from the true value in the range of –0.8 – 0.3 = –1.1 to –0.8 + 0.3 = –0.5 percentage points, as –0.8 is the average difference and ±0.3 is its 90-percent confidence interval.

As shown in Schreiner (2013), the standard error[11] of the estimated poverty rate is

$$\alpha \cdot \sqrt{\frac{p \cdot (1-p)}{n}}$$, where:

p is the proportion of sampled households below the poverty line,

n is the sample size, and
α is a factor specific to the country, scorecard, and poverty line.
α factors below 1.0 (such as 0.44 and 0.62 for the $1.25/day and $1.75/day lines, Table 6) imply that the scorecard is more precise that direct measurement, while factors above 1.0 (such as 1.04 for the $2.50/day line) imply the converse.

[10] There are differences, in spite of the estimator's unbiasedness, because the scorecard comes from a single sample. The average difference would be zero if samples were repeatedly drawn from the same population and split into sub-samples before repeating the entire scorecard-building process.
[11] This does not correct for sampling from a finite population.

Table 6: Bias, precision, and sample-size α for bootstrapped estimates of poverty rates for groups of people at a point in time for the scorecard applied to the 2004/5 HICE validation sample

	International 2005 PPP Poverty Line		
	$1.25/day	$1.75/day	$2.50/day
Estimate minus true value			
Scorecard applied to 2004/5 HICE validation sample	–0.8	–1.4	–0.3
Precision of difference			
Scorecard applied to 2004/5 HICE validation sample	0.3	0.4	0.5
α factor for standard errors			
Scorecard applied to 2004/5 HICE validation sample	0.44	0.62	1.04

Precision is measured as 90-percent confidence intervals in units of +/– percentage points.
Differences and precision estimated from 1,000 bootstraps of size n = 16,384.
α is estimated as described in Schreiner (2013).

5. Value-added by the poverty scorecard

This section asks whether the poverty scorecard and the DHS asset index produce similar rankings. If not, then the poverty scorecard may have something new and useful to offer.

Large differences in rankings are possible—at least in principle—because the poverty scorecard and the asset index define poverty differently (consumption versus assets). Also, the two tools are constructed differently; subject to usability constraints, the scorecard uses Logit to choose indicators and points to maximize the accuracy of ranking based on consumption poverty, while the asset index uses PCA to maximize the explained variance among a pre-selected set of indicators.

Nevertheless, differences might be small. After all, the two tools use many similar indicators, and the correlation between assets and consumption may be strong.

The extent of differences in rankings matters because it would be convenient if the DHS index ranked people about the same as the poverty scorecard. In that case, the DHS asset index could be calibrated to poverty likelihoods and researchers could estimate consumption-based poverty rates using the familiar asset index.

Table 7 shows the quintile correspondences for people in the 2005 Ethiopia DHS when ranked by the two tools. The sample is sorted and divided into equal-sized quintiles twice, once by the scorecard's score and once by the asset index, so each row total and each column total should be 20 percent. There are 25 cells in the matrix (5 quintiles in the rows, by 5 quintiles in the columns). Each cell contains the percentage of all people who rank in a given row quintile by the poverty scorecard and who also rank in a given column quintile by the asset index. The five cells on the diagonal show the percentage of people who fall in the same quintile by both tools (first quintile on both the scorecard and the asset index, second quintile on both the scorecard and the asset index, etc.) If the correspondence across the two rankings were perfect, all diagonal cells would be 20 percent and all off-diagonal cells would be zero. At the other extreme of no correspondence, all the cells would be 4 percent.

Table 7: Correspondences of quintile ranks for the poverty scorecard and DHS asset index applied to the 2005 Ethiopia DHS

Poverty scorecard	DHS asset index				
	1	2	3	4	5
1	6.7	4.9	3.8	2.9	1.8
2	4.6	4.3	4.3	4.0	2.8
3	3.5	4.0	4.5	4.4	3.7
4	3.1	3.7	3.9	4.9	4.4
5	2.2	3.2	3.6	3.8	7.3

The actual correspondence in Table 7 is better than random,[12] as diagonal cells always exceed 4 percent and most off-diagonal cells are less than 4 percent. Except for the first/most-poor quintile (6.7 percent) and the fifth/least-poor quintile (7.3 percent) where ranking is easiest, however, the correspondence is not much better than random. About 27.7 percent of people fall in the same quintile in both rankings (versus 20 percent if random). Large differences are common; 38 percent differ across the two rankings by two quintiles or more, and 16 percent differ by three quintiles or more. In sum, there are large differences in ranking by the two tools.

A simpler test is to replace the quintiles with a single cut-off. Table 8 shows two examples, one with a 35th-percentile cut-off (corresponding to Ethiopia's poverty rate for $1.25/day)

[12] $p < 0.01$ for a Chi-square test.

and a second example with a 85th-percentile cut-off (corresponding to Ethiopia's poverty rate for $2.50/day). Having a single cut-off increases the share of people in cells on the diagonal who are classified the same by both tools, as some cells that are off-diagonal with quintiles are now part of the (larger) diagonal cells (Friedman, 1997).

Table 8: Correspondences of ranks with cut-offs at the 35th and 85th percentiles, poverty scorecard and DHS asset index applied to the 2005 Ethiopia DHS

Cut-off at 35th percentile (poverty line of $1.25/day 2005 PPP)

		DHS asset index	
		< 35th	> = 35th
Poverty scorecard	< 35th	16.2	18.8
	> = 35th	18.8	46.2

Cut-off at 85th percentile (poverty line of $2.50/day 2005 PPP)

		DHS asset index	
		< 85th	> = 85th
Poverty scorecard	< 85th	75.2	9.8
	> = 85th	9.8	5.2

In the 2005 Ethiopia DHS, less than half of the people (46 percent = 16.2 ÷ (16.2 + 18.8)) with poverty scores below the 35th percentile also have asset scores below the 35th percentile. Thus, on the most-poor end of the scale, there are large differences between ranks between the two tools. On the least-poor end, however, agreement is greater, at 71 percent (= 46.2 ÷ (46.2+ 18.8)).

Increasing the cut-off improves accuracy for the most-poor but worsens accuracy for the least-poor. With a cut-off at the 85th percentile, there is agreement between the two

tools for 88 percent of the most-poor (75.2 ÷ (75.2 + 9.8)) but only 35 percent of the least-poor (= 5.2 ÷ (9.8 + 5.2)).

To sum up, the poverty scorecard and the DHS asset index do not generally concentrate large shares of the same people among low scores (most-poor) nor among high scores (least-poor). The two tools are not good proxies for each other. If the goal is to measure consumption-based poverty, then it is better to use an estimator tailored for that purpose than to utilize an estimator designed for asset-based poverty.

6. Estimating consumption-based poverty with the poverty-mapping approach in the DHS

This paper is not the first to build a consumption-based poverty scorecard using only indicators matched to a DHS (or DHS-like) survey. This section asks two questions of previous (non-Ethiopian) work. First, how does their accuracy and precision compare with that of the scorecard here? And second, are poverty scores more strongly linked with health outcomes than asset-index scores? Of course, the answers to these questions are related to the overall usefulness of the poverty scorecard, but they are distinct from the main point of this paper, namely, that consumption-based poverty can be estimated in the DHS, even though the DHS does not collect consumption data.

6.1 How accurate is this scorecard versus others?

This sub-section describes three cases where comparisons of bias and precision (standard errors or confidence intervals) are possible. Some papers (for example, Filmer and Pritchett, 2001; Kijima and Lanjouw, 2003) are omitted because they compare health only with true (reported) consumption (not predicted consumption) or because their scorecards use only a subset of the indicators used here.

6.1.1 Stifel and Christiaensen

Stifel and Christiaensen (2007) seek to an intuitive and inexpensive way to track changes in poverty. They build three scorecards (Nairobi, other urban, and rural) using consumption data from Kenya's 1997 Welfare Monitoring Survey and indicators matched to Kenya's DHS. The scorecards are applied to the 1993, 1998, and 2003 DHS to estimate changes in poverty in years without consumption surveys. Like most poverty scorecards—but unlike the one in this paper—Stifel and Christiaensen regress

the logarithm of per-capita household consumption against a set of indicators, many of which are similar to those in this paper.

When Stifel and Christianensen's scorecards constructed with Kenya's 1997 WMS is applied to the same data used to construct the scorecards in the first place (that is, in-sample), bias ranges from –1 to –2 percentage points. Such in-sample tests overstate the accuracy that can be expected when the estimator is applied out-of-sample to new data that was not used to construct the estimator.[13] When the poverty scorecard here is applied out-of-sample—that is, to data not used to construct the scorecard—bias ranges from –0.3 to –1.4 percentage points (Table 6). Thus, the scorecard here is not more biased than that of Stifel and Christianensen (and they understate bias).

For precision, Stifel and Christianensen report a standard error of 1.7 percentage points for an in-sample poverty-rate estimate (n = 10,639). Ignoring again the in-sample overstatement of precision, the implied α factor is about 3.5, suggesting that the scorecard here (α of 0.44 to 1.04, Table 6) is more precise; its confidence intervals are at least three times narrower.

6.1.2 Simler, Harrower, and Massingarela

Simler, Harrower, and Massingarela (2003) use poverty mapping as a simple, inexpensive way to track changes in poverty rates without complex, costly consumption surveys. They build 11 scorecards (one per province) using Mozambique's 1996/7 National Household Survey of Living Standards, using only indicators matched to Mozambique's DHS-like 2000/1 Core Welfare Indicator Survey. The scorecards predict the logarithm of consumption using indicators on education, housing, asset ownership, community averages, and GIS variables.

Based on an in-sample test with the 1996/7 consumption survey, bias is –3.9 percentage points, and the α factor is 2.3. These numbers are much larger than those for the Ethiopia scorecard here.

[13] For example, if the scorecard here is applied in sample, bias is exactly zero.

6.1.3 Azzarri et al.

Azzarri et al. (2005) construct a poverty scorecard from the 2002 Albania Living Standards Measurement Survey and then apply it to a sub-sample of households who were revisited in 2003. Thus, the indicators are perfectly matched. Like the others reviewed here, the scorecard predicts the logarithm of consumption, and indicators are selected with stepwise regression. Azzarri et al. also include some subjective indicators. In an in-sample test with the 2002 data, bias is –4.8 percentage points, much greater than for the Ethiopia scorecard here. Azzarri et al. do not report standard errors, so precision cannot be compared with that of the scorecard here.

6.2 Which tool is more closely related with health outcomes?

The poverty scorecard and the DHS asset index rank people differently, but this need not imply differences in their relationship with health (Wagstaff and Watanabe, 2003). This subsection discusses two papers that compare how health relates with poverty scores and with DHS-like asset scores. Other papers that look at health vis-à-vis true (not predicted) consumption are not discussed.

6.2.1 Sahn and Stifel

In a seminal paper covering nine countries, Sahn and Stifel (2003) look at whether child health (percentage stunted, and mean height-for-age z scores) is more closely related to ranks based on consumption from a poverty scorecard or ranks from a DHS-like asset index. As usual, they predict the logarithm of per-capita consumption.

On the one hand, Spearman correlation coefficients and correspondence indexes suggest that, "in terms of predictive capabilities, it does not matter which welfare measure is used" (Sahn and Stifel, p. 480). On the other hand, they find that—in 17 of 22 cases—the gradient with child-health outcomes between the first and fifth quintiles was greater for the asset index than for predicted consumption from a scorecard.

In the end, Sahn and Stifel fail to reject the hypothesis of no differences: "In the context of estimating models of nutrition, we find no compelling reason to believe that either reported or instrumented [predicted by a scorecard] expenditures serve as a better proxy for economic welfare than the asset index" (p. 485).

6.2.2 Filmer and Scott

Filmer and Scott (2012) compare ranks for reported consumption, scorecard-predicted consumption, and DHS-like asset indexes. Several results from their tests with 11 countries are of interest here.

First, "predicted per-capita expenditure [from a scorecard] yields the most similar household rankings to per-capita expenditure" (p. 18). This is no surprise; if matching true consumption is the goal, then poverty scorecards are better than asset indexes. Still, Filmer and Scott report that asset indexes are also highly correlated with true consumption.

Second, "despite household re-rankings, conclusions about inequalities across quintiles in education outcomes, health-care-seeking behavior, fertility, and child mortality, as well as labor-market outcomes, are not very sensitive to the particular economic-status measure used to classify households" (p. 22).[14] That is, Filmer and Scott's scorecards do better than their asset indexes at estimating consumption-based poverty and about as well as asset indexes with health outcomes.

Filmer and Scott's third point is that scorecards and asset indexes "show vastly different gradients in household composition" because scorecards do not adjust for household economies of scale (for example, one bathroom can serve five people at less than five times the cost of serving one person). Thus, the two "equivalence" results just described may not apply to the poverty scorecard here; Filmer and Scott's scorecard omits household size (and education and employment as well), but household size is by far the most powerful predictor of consumption-based poverty in Ethiopia.

The scorecard here uses more types of indicators than those of Filmer and Scott, and so their results—that asset indexes and scorecards perform about the same—may not apply here. In particular, the scorecard here uses a wider variety of types of information, so it is likely to perform better than those in Filmer and Scott. Further tests, however, are beyond the scope of this paper.

[14] This fits Wagstaff and Watanabe (2003, 19 countries), but not Lindelow (2006, one country).

7. Conclusion

The poverty scorecard provides a way to estimate consumption-based poverty for people and for groups in the 2005 Ethiopia DHS. The approach follows poverty mapping in that it constructs a scorecard based on a consumption survey (the 2004/5 HICE) using only indicators that are also in the 2005 DHS (which does not collect consumption data). Researchers can then apply the scorecard to the 2005 DHS and analyze how health outcomes vary with consumption-based poverty.

The poverty scorecard ranks people differently than the DHS asset index; asset-based poverty (a longer-term concept) is not a good proxy for consumption-based poverty (a shorter-term concept). While both conceptions of poverty are legitimate, the consumption-based definition dominates discussion among both the polity and policymakers. Thus, including consumption-based estimates may give DHS research greater policy relevance.

Like poverty mapping, poverty scoring makes three assumptions about its two data sources: that they represent the same population, that they represent the same time period, and (least tenably in the case of Ethiopia) that the indicators are well-matched. The poverty-scoring approach here improves on traditional poverty mapping in that it reports out-of-sample bias and a simple summary measure of precision.

Of course, the results here hold only for Ethiopia's 2004/5 HICE and 2005 DHS; they may or may not generalize to other countries and data sources. The approach, however, can be tested in any country with both a DHS and a national household expenditure survey.

From the point of view of policy in Ethiopia, the poverty scorecard is accurate enough in terms of bias and precision to provide useful estimates of head-count poverty rates in years when there is a DHS but no HICE.[15] More-frequent quantitative feedback on the direction of overall household well-being can help keep policy on-track from the point of view of poverty alleviation.

[15] In the same way, the approach can be used to create poverty maps, based on the DHS, in years between censuses.

References

Alem, Yonas. (2011). "Poverty Dynamics and Intra-Household Heterogeniety in Occupations: Evidence from Urban Ethiopia", efdinitiative.org/news-press/workshops-etc/workshops-archive/files-1/files-ethiopia/Yonas_Alem.pdf, retrieved 6 February 2013.

Azzarri, Carlo; Carletto, Gero; Davis, Benjamin; and Alberto Zezza. (2005). "Monitoring Poverty without Consumption Data: An Application Using the Albania Panel Survey", Eastern European Economics, Vol. 44, No. 1, pp. 59–82.

Bigsten, Arne; Kebede, Bereket; Shimeles, Abebe; and Mekonnen Taddesse. (2003). "Growth and Poverty Reduction in Ethiopia: Evidence from Household Panel Surveys", World Development, Vol. 31, No. 1, pp. 87–106.

Bigsten, Arne and Abebe Shimeles. (2008). "Poverty Transition and Persistence in Ethiopia: 1994–2004", World Development, Vol. 36, No. 9, pp. 1559–1584.

Bollen, Kenneth A.; Glanville, Jennifer L.; and Guy Stecklov. (2007). "Socio-Economic Status, Permanent Income, and Fertility: A Latent-Variable Approach", Population Studies, Vol. 61, No. 1, pp. 15–34.

Booysen, Frikkie; van der Berg, Servaas; Burger, Ronelle; von Maltitz, Michael; and Gideon du Rand. (2008). "Using an Asset Index to Assess Trends in Poverty in Seven Sub-Saharan African Countries", World Development, Vol. 36, No. 6, pp. 1113–1130.

Carter, Michael R.; and Christopher B. Barrett. (2006). "The Economics of Poverty Traps and Persistent Poverty: An Asset-Based Approach", *Journal of Development Studies*, Vol. 42, No. 2, pp. 178–199.

Christiaensen, Luc; Lanjouw, Peter; Luoto, Jill; and David Stifel. (2010). "The Reliability of Small-Area Estimation Prediction Methods to Track Poverty", wider.unu.edu/stc/repec/pdfs/wp2010/wp2010-99.pdf, retrieved 6 February 2013.

Deaton, Angus; and Salman Zaidi. (2002). "Guidelines for Constructing Consumption Aggregates for Welfare Analysis", World Bank Living Standards Measurement Survey Working Paper No. 135, Washington, D.C., http://go.worldbank.org/8YRCR9ERJ0, retrieved 6 February 2013.

Dekker, Marleen. (2006). "Estimating Wealth Effects without Expenditure Data: Evidence from Rural Ethiopia", *Ethiopian Journal of Economics*, Vol. 15, No. 1, pp. 35–54.

Dercon, Stefan. (2006). "Economic Reform, Growth, and the Poor: Evidence from Rural Ethiopia", *Journal of Development Economics*, Vol. 81, No. 1, pp. 1–24.

_____. (1997) "Poverty and Deprivation in Ethiopia", Center for the Study of African Economies, Oxford University, memo.

Elbers, Chris; Lanjouw, Peter; and Phillippe George Leite. (2008). "Brazil within Brazil: Testing the Poverty-Map Methodology in Minas Gerais", World Bank Policy Research Working Paper No. 4513, Washington, D.C., papers.ssrn.com/sol3/papers.cfm?abstract_id=1092691, retrieved 6 February 2013.

Elbers, Chris; Lanjouw, Jean O.; and Peter Lanjouw. (2003) "Micro-Level Estimation of Poverty and Inequality", *Econometrica*, Vol. 71, No. 1, pp. 355–364.

Filmer, Deon and Lant, Pritchett. (2001) "Estimating Wealth Effects without Expenditure Data—or Tears: An Application to Educational Enrolments in States of India", Demography, Vol. 38, No. 1, pp. 115–132.

Filmer, Deon and Kinnon, Scott. (2012). "Assessing Asset Indexes", Demography, Vol. 49, pp. 359–392.

Friedman, Jerome H. (1997). "On Bias, Variance, 0–1 Loss, and the Curse-of-Dimensionality", Data Mining and Knowledge Discovery, Vol. 1, pp. 55–77.

Gebremedhin, Tesfaye Alemayehu; and Stephen Whelan. (2008). "Prices and Poverty in Urban Ethiopia", *Journal of African Economies*, Vol. 17, No. 1, pp 1–33.

Grosh, Margaret; and Judy L. Baker. (1995). "Proxy Means Tests for Targeting Social Programs: Simulations and Speculation", World Bank Living Standard Measurement Survey Working Paper No. 118, Washington, D.C., go.worldbank.org/W9OWN57PD0, accessed 6 February 2013.

Gwatkin, Davidson R.; Rutstein, Shea; Johnson, Kiersten; Suliman, Eldaw; Wagstaff, Adam; and Agbessi Amouzou. (2007). "Socio-Economic Differences in Health, Nutrition, and Population: Ethiopia", Country Reports on HNP and Poverty, Washington, D.C. World Bank, go.worldbank.org/T6LCN5A340, retrieved 6 February 2013.

Kedir, Abbi M.; and Andrew McKay. (2005). "Chronic Poverty in Urban Ethiopia: Panel-Data Evidence", International Planning Studies, Vol. 10, No. 1, pp. 49–67.

Kijima, Yoko; and Peter Lanjouw. (2003). "Poverty in India during the 1990s: A Regional Perspective", World Bank Policy Research Working Paper No. 3141, ideas.repec.org/p/wbk/wbrwps/3141.html, retrieved 6 February 2013.

Knowles, James C.; Bales, Sarah; Cuong, Le Quang; Oanh, Tran Thi Mai; and Duong Huy Luong. (2008). "Health Equity in Vietnam: A Situational Analysis Focused on Maternal and Child Mortality", background paper prepared for UNICEF Consultancy on "Equity in Access to Quality Healthcare for Women and Children" (April 8-10, Ha Long City, Vietnam), jkinusa@aol.com.

Lanjouw, Peter; and Martin Ravallion. (1995). "Poverty and Household Size", *Economic Journal*, Vol. 105, pp. 1415–1434.

Lindelow, Magnus. (2006). "Sometimes More Equal Than Others: How Health Inequalities Depend on the Choice of Welfare Indicator", *Health Economics*, Vol. 15, pp. 263–279.

Loening, Josef L.; Durevall, Dick; and Yohannes Ayalew Birru. (2009). "Inflation Dynamics and Food Prices in an Agricultural Economy: The Case of Ethiopia", Working Papers in Economics No. 347, University of Gothenburg, gupea.ub.gu.se/dspace/bitstream/2077/19464/1/gupea_2077_19464_1.pdf, accessed 6 February 2013.

Ministry of Finance and Economic Development (MoFED). (2006). Ethiopia: Building on Progress, a Plan for Accelerated and Sustained Development to End Poverty (PASDEP), 2005/06–2009/10, Volume I: Main Text, Welfare Monitoring Unit, Addis Ababa, planipolis.iiep.unesco.org/upload/Ethiopia/Ethiopia_PASDEP_2005_2010.pdf , accessed 6 February 2013.

_____. (2002). "Poverty Profile of Ethiopia: Analysis based on the 1999/00 HICE and WM Survey Results", Welfare Monitoring Unit, Addis Ababa.

Narayan, Ambar; and Nobuo Yoshida. (2005). "Proxy Means Tests for Targeting Welfare Benefits in Sri Lanka", World Bank Report No. SASPR-7, Washington, D.C., documents.worldbank.org/curated/en/2005/07/6209268/proxy-means-test-targeting-welfare-benefits-sri-lanka, accessed 6 February 2013.

Pitchforth, Emma; van Teijlingen, Edwin; Graham, Wendy; and Ann Fitzmaurice. (2007). "Development of a Proxy Wealth Index for Women Utilizing Emergency Obstetric Care in Bangladesh", *Health Policy and Planning*, Vol. 22, No. 3, pp. 311–319.

Ravallion, Martin; Chen; Shaohua; and Prem Sangraula. (2009). "Dollar-a-Day Revisited", World Bank Economic Review, Vol. 23, No. 2, pp. 163–184.

Rutstein, Shea Oscar; and Kiersten Johnson. (2004). "The DHS Wealth Index", DHS Comparative Reports No. 6, Calverton, MD: ORC Macro, measuredhs.com/pubs/pdf/CR6/CR6.pdf, accessed 6 February 2013.

Sahn, David E. and David Stifel. (2003). "Exploring Alternative Measures of Welfare in the Absence of Expenditure Data", Review of Income and Wealth, Series 49, No. 4, pp. 463–489.

_____. (2000). "Poverty Comparisons over Time and across Countries in Africa", World Development, Vol. 28, No. 12, pp. 2123–2155.

SAS Institute Inc. (2004). "The LOGISTIC Procedure: Rank Correlation of Observed Responses and Predicted Probabilities", in SAS/STAT User's Guide, Version 9, Cary, NC, support.sas.com/documentation/cdl/en/statug/63033/HTML/default/statug_logistic_sect035.htm, accessed 6 February 2013.

Schreiner, Mark. (2013). "A Simple Poverty Scorecard for Peru", microfinance.com/English/Papers/Scoring_Poverty_Peru_2010_EN.pdf, accessed 6 February 2013.

_____. (2006). "Is One Simple Poverty Scorecard Enough for India?", microfinance.com/English/Papers/Scoring_Poverty_India_Segments.pdf, accessed 6 February 2013.

_____. (2005). "Un Índice de Pobreza para México", memo for Grameen Foundation, microfinance.com/Castellano/Documentos/Scoring_Pobreza_Mexico_2002.pdf, accessed 6 February 2013.

Schreiner, Mark; Matul, Michal; Pawlak, Ewa; and Sean Kline. (2004) "Poverty Scorecards: Lessons from a Microlender in Bosnia-Herzegovina", microfinance.com/English/ Papers/Scoring_Poverty_in_BiH_Short.pdf, accessed 6 February 2013.

Schreiner, Mark and Michael Sherraden. (2006). Can the Poor Save? Saving and Asset Accumulation in Individual Development Accounts, Piscataway, NJ: Transaction.

Sherraden, Michael. (1991). Assets and the Poor: A New American Welfare Policy, Armonk, NY: M.E. Sharpe.

Sillers, Don. (2006). "National and International Poverty Lines: An Overview", Washington, D.C.: United States Agency for International Development, povertytools.org/other_documents/siller-povertylines.doc, retrieved 6 February 2013.

Simler, Kenneth R.; Harrower, Sarah; and Claudio Massingarela. (2003). "Estimating Poverty Indexes from Simple Indicator Surveys", http://www.csae.ox.ac.uk/conferences/2004-GPRaHDiA/papers/3p-SimlerMassingerela-CSAE2004.pdf, retrieved 6 February 2013.

Stifel, David; and Luc Christiaensen. (2007). "Tracking Poverty over Time in the Absence of Comparable Consumption Data", World Bank Economic Review, Vol. 21, No. 2, pp. 317–341.

Tarozzi, Alessandro; and Angus Deaton. (2007). "Using Census and Survey Data to Estimate Poverty and Inequality for Small Areas", Review of Economics and Statistics, Vol. 91, No. 4, pp. 773–792.

Wagstaff, Adam; and Naoko Watanabe. (2003). "What Difference Does the Choice of SES Make in Health Inequality Measurement?", Health Economics, Vol. 12, No. 10, pp. 885–890.

Woldehanna, Tassew; Hoddinott, John; and Stefan Dercon. (2008). "Poverty and Inequality in Ethiopia: 1995/96–2004/05", ssrn.com/abstract=1259341, accessed 6 February 2013.

World Bank. (2012). Targeting Poor and Vulnerable Households in Indonesia, www-wds.worldbank.org/external/default/WDSContentServer/WDSP/IB/2012/03/01/000386194_20120301004533/Rendered/PDF/672180WP00PUBL0T00English000PUBLIC0.pdf, retrieved 6 February 2013.

_____. (2008). "International Comparison Project: Tables of Results", Washington, D.C., siteresources.worldbank.org/ICPINT/Resources/icp-final-tables.pdf, retrieved 6 February 2013.

Zeller, Manfred. (2004). "Review of Poverty Assessment Tools", Accelerated Microenterprise Advancement Project, pdf.usaid.gov/pdf_docs/PNADH120.pdf, retrieved 6 February 2013.

Appendix

Matching indicators across the 2004/5 HICE and the 2005 DHS

This appendix documents, using simple cross-tabs, how responses are grouped match indicators across the 2004/5 HICE and 2005 DHS. Only indicators in the scorecard in Table 1 are included here; documentation for other indicators is available on request.

HICE: How many people live with the household six months out of the year?
DHS: How many people usually live with the household?

	% of people with a given response		
Response	HICE	DHS	Difference
1 or 2	13.2	14.4	-1.2
3	9.9	11.5	-1.6
4	15.3	15.4	-0.1
5	16.6	16.7	-0.1
6	15.9	17.7	-1.8
7	14.0	12.5	1.5
8	9.2	8.3	0.9
9 or more	5.8	4.6	1.2

Household size is the most powerful scorecard indicator, and it seems well-matched, although in general the HICE reports larger households.

HICE: What is the highest grade the female head/spouse has completed?
DHS: What is the highest grade the female head/spouse has completed?

	% of people with a given response		
Response	HICE	DHS	Difference
Four or less, or no female head/spouse	92.5	92.7	-0.2
Five or six	2.7	2.5	0.2
Seven to nine	2.3	2.2	0.1
Ten or more	2.5	2.6	-0.1

This well-matched indicator mostly distinguishes the least-poor form the poorest ninth-tenths of Ethiopians.

HICE: What is the main construction material of the walls of the dwelling unit?
DHS: What is the main material of the walls of the residence?

HICE	Item	Group	DHS	Item	Group	Diff. Group
Mud and stone	9.6		Stone with mud	7.8		
Wood and grass	8.2		Stone with lime/ cement	1.1		
Others	2.7		Cane/ trunks/ bamboo/ reed	5.6		
			Other	3.1		
		20.4	Undefined	1.7	19.3	1.1
Wood and mud	74.9		Bamboo/ wood	78.2		
Reed and bamboo	3.4		Uncovered adobe	0.0		
No data	0.0		Plywood	0.0		
			Carton	0.0		
			No walls	0.2		
		78.3	No data	0.0	78.4	-0.1
Cement and stone	0.7		Cement	0.5		
Hollow bricks	0.5		Bricks	0.1		
Bricks	0.1		Cement blocks	0.6		
			Covered adobe	0.0		
		1.3	Wood planks/ shingles	1.1	2.3	-1.0

This indicator is well-matched in terms of the percentages across grouped responses. But it is less clear why the first group (stone with mud, stone with lime/cement, etc.) implies more poverty than the second group (bamboo/wood, uncovered adobe, etc.). Also, the responses sometimes seem different across the two surveys, even though their percentages line up nicely. Nevertheless, "wood and mud" in the HICE must match up with "bamboo/wood" in the DHS, as these account for about three-fourths of all people. Likewise, "mud and stone" in the HICE must match with "Stone and mud" or "Stone with lime/cement" in the DHS, even though that (less plausibly) implies that "wood and grass" in the HICE matches with "cane/trunks/bamboo/reed" in the DHS.

HICE: What type of toilet facility does the household use?
DHS: What kind of toilet facility do members of your household usually use?

	% of people with a given response		
Response	HICE	DHS	Difference
Non-flush or none	97.6	97.8	-0.2
Flush	2.4	2.2	0.2

This is well-matched, but, like the education of the female head/spouse, only matters for a few of the elite, failing to help differentiate socio-economic status for the masses.

HICE: What is the main source of cooking fuel?
DHS: What type of fuel does your household mainly use for cooking?

HICE	Item	Group	DHS	Item	Group	Group diff.
Mainly collected firewood	72.8		Wood or straw/ shrubs/ grass	86.8		
Mainly purchased firewood	9.5		Animal dung	7.5		
Crop residue	11.1	93.4			94.3	-0.9
Charcoal	1		Electricity	0.2		
Kerosene	2		LPG or natural gas	0.1		
Butane gas	0.4		Biogas	0		
Electricity	0.4		Kerosene	3		
Does not use cooking fuel	0.1		Charcoal	2.2		
Other	2.7		Other	0.1		
			Other	0.1		
		6.6	No data	0	5.7	0.9

Like the education of the female head/spouse and the type of toilet arrangement, the type of cooking fuel mostly serves to identify the highest end of the Ethiopian socio-economic scale. While the percentages of people in each group match well, and while it is obvious what the "high-quality" cooking fuels are, it is nevertheless odd that among low-quality fuels, the HICE response options differ greatly from those in the DHS, even though they must both be picking up virtually the same households. For example, 7.5 percent of people in the DHS live in households that mostly cook with "animal dung", but dung is not even an option in the HICE. Conversely, 11.1 percent of people in the HICE live in households that mostly cook with crop residue, but this response does not even appear in the DHS.

HICE: Does the household currently own mattresses and/or beds?
DHS: Does the household have a bed?

	% of people with a given response		
	HICE	DHS	Difference
No	44.9	40.8	4.1
Yes	55.1	59.2	-4.1

Given a rule-of-thumb that allows differences of up to 10 percent, this indicator is well-matched. Nevertheless, the differences in questionnaire wording would seem to favor

more affirmative responses in the HICE, but the data show the opposite. Still, this indicator is well-positioned to differentiate among the very poor and the merely poor.

HICE: Does the household currently own a radio?
DHS: Does the household have a radio?

	% of people with a given response		
	HICE	DHS	Difference
No	71.9	64.3	7.6
Yes	28.1	35.7	-7.6

This seemingly simple question is perhaps the least-well-matched of all the scorecard indicators. It is not clear—beyond differences in questionnaire context—why the rate of affirmative responses would be so different.

HICE: Does the household have any agricultural holdings?
DHS: Does any member of this household own any land that can be used for agriculture?

	% of people with a given response		
	HICE	DHS	Difference
No	11.6	15.7	-4.1
Yes	88.4	84.3	4.1

The differences here may well be due to differences in wording, as the DHS wording seems less likely to be interpreted as applying only to the household head or to the respondent.

HICE: Does the household currently own cattle?
Does the household currently own sheep and goats?
DHS: How many cattle does the household own?
How many goats does the household own?
How many sheep does the household own?

	% of people with a given response		
	HICE	DHS	Difference
No	24.1	30.5	-6.4
Yes	75.9	69.5	6.4

As with the previous asset indicators, the quality of the match is disappointing. Nevertheless, the indicator is used because—at least in the HICE—it helps distinguish among the poverty levels of the great mass of rural, agricultural households.

In the general, the indicators of asset ownership are kept because, without them, the scorecard would essentially identify the 10 percent or so of "elites" in Ethiopia, and estimate the poverty of the other 90 percent almost completely based on household size.