Name: Elaine Baker
Course: MSc Political Economy of Development
Module: Research Methods II
Lecturer: Dr Jonathan Pincus
Due date: 23rd April 2001
Word count: 6450

Contents

Introduction

Measures of the living standards, poverty or access to economic resources of representative samples of individuals or households are needed for a wide range of purposes. These include

- Observing changes in poverty over time in order to assess policies
- Observing differences in poverty in different places in order to compare and assess policies
- Identifying the characteristics of the poor in order to target policies
- Analysing the relationship between different aspects of poverty, such as income, employment, health and education
- Assessing macro relationships between poverty and other variables such as growth, distribution, trade, and quantity and types of government spending.

People meet their needs through the market by earning or receiving money (income) and buying goods and services (expenditure), through producing for their own needs, through access to state provided goods and services, through ties with family and social groups and through non-governmental organisations. An inadequate or insecure access to resources through these mechanisms is what is commonly understood to be poverty.

In most developed countries, and increasingly in developing countries, economic welfare and poverty are measured using household surveys. The largest of these are World Bank Living Standards Measurement Surveys (LSMS). These are lengthy questionnaires administered to a sample of households, a large part of which comprise questions about income and expenditure. They also include questions about remittances, income-in-kind and access to public services such as state-provided education, health, sanitation and transport infrastructure.

This paper firstly compares the assessment of household income through direct questions about income and expenditure and through questions about asset ownership. It then uses asset data from the South African Living Standards Measurement Survey (LSMS) to examine asset index issues such as weighting methods and choice of assets to include. It shows that choice of assets is the more important issue by comparing different asset indices, consumption and income data. It concludes that asset indices are an accurate and less costly proxy for income and consumption, but that large scale income and expenditure household surveys such as the South Africa LSMS should include a more suitable range of assets in order to make them complementary to more frequent, less costly surveys which use only assets as an income proxy.

Problems with direct measurement of income and expenditure

Measuring household income and expenditure directly can be inaccurate and costly due to the following problems:

Conceptual difficulties

Almost all households supply some of their own needs. It is possible for a household with greater money income and expenditure to be less well off, through having less capacity to supply its own needs (i.e. grow its own food, produce its own goods, do its own maintenance and catering work and care for children / elderly / sick people) through a lack of means of production, skills or time. Increasing specialisation means that the ratio of money income to self-provision increases, as people do less for themselves and more for the market. Measures of income/expenditure commonly attempt to include the value of production-for-own-consumption. However, it is impossible even to compile a comprehensive list of goods and services households provide for themselves. If it is decided to value only those goods and services which are also available through the market, the range of goods and services to be valued can change over time and place as markets in new goods and services emerge.

The role of time in income and expenditure also entails conceptual problems. Should the purchase of assets be added to expenditure at the time of purchase or spread over time of use? Should borrowing be added to income and repayments subtracted, or should borrowing be accounted for as a purchase of a negative asset? Different interest, depreciation and asset valuation methods can yield very different estimations. This is an especially important issue for housing in developing countries (Sahn and Stifel, 2000, p.2).

Long run income or expenditure is a more meaningful concept than short-run variations for the purposes outlined above. Measuring long run income involves either repeated surveys of the same households over a long period of time (panel surveys), or questions which cover a long period of time. Income and expenditure in the short term may not be a good indicator of economic welfare. Income tends to be especially variable, and while expenditure can be less so if people prefer smooth to variable consumption patterns and have a good knowledge of their average income, borrowing and saving constraints prevent consumption smoothing and are seldom unimportant in developing countries (Montgomery et al, 1999, p.15).

Practical problems

Measuring income is especially difficult in developing countries, as many people earn income from multiple sources such as irregular wage income, self-employment and trading (Sahn and Stifel, 2000, p.1) (Montgomery et al, 1999, p.15). In the South Africa LSMS, irregular wage work or self employment were the principal work for 21% of those classified as employed, and the survey did not include a question on combining these job types.

Expenditure is very difficult to measure as there is a trade-off between ability to recall accurately and robustness to time and seasonality (Fields, 1998). If long recall periods are used, estimates are likely to be very inaccurate. If short recall periods are used, the expenditure level is very sensitive to the season and time in which the survey is carried out. Pradhan (2000) and Scott and Amenuvegbe (1990) have shown that differences in recall period and number of commodities included in questions can create biases as well as inaccuracies.

Questions where the answers are not readily verifiable give rise to tendencies to forget and mislead (Sender and Smith, 1990, Ch.3, p.27). For example, respondents may not trust the confidentiality of the survey or its stated purposes, and so may exaggerate or understate income and expenditure, or not reveal income sources accurately.

Prices are an especially difficult issue to deal with in estimating expenditure and in estimating input and output costs for self-employment. Real markets are not perfectly competitive, do not have perfect information and do not have perfect arbitrage. In addition, real market relationships are often intertwined with political, family and other social relationships.

For these reasons prices for the same good or service vary even in the same time and place. Variations in prices across time are usually dealt with using price deflators, but these are sensitive to the weighting systems used. Prices may change at different rates for different income groups (Fields, 1998), giving rise to the need for deflators for specific classes and income groups. Variations in prices between different places are often not accounted for with regionally specific price deflators in developing countries due to the costs involved in compiling them (Sahn & Stifel, 2000, p.3) (Fields, 1998).

Converting incomes and expenditures using exchange rates in order to make international comparisons is not meaningful as the same US$ income in different countries can buy very different ranges of goods due to the many physical, logistical, regulatory and information barriers to international arbitrage. Estimating purchasing power parities in order to make international comparisons involves uniform price surveys in a range of countries.

Income from agriculture is difficult to assess, as often a range of crops is produced, some of which are consumed and some of which are sold in different markets at different times fetching different prices. Inputs of labour from household members and wage labourers, and other input costs vary over time and season (Collier et al, 1986, p.55). The myriad transactions undertaken by self-employed people and traders make it unlikely that even respondents themselves know their own costs and revenues (Deaton, 1997).

Costs of collection

As can be seen from the South Africa LSMS household questionnaire, collecting data on expenditure involves large numbers of detailed questions. It includes questions about actual or imputed rent, expenditure on utilities, food commodities (31 types), non-food commodities (13 types), occasional non-food expenditure (13 types), expenditure on education, on remittances and on transportation. The "services" implicitly provided by consumer durables are calculated and added to the total, using the dates of purchase of the durables, estimates of current value, and assumptions about depreciation (Montgomery et al, 1999, p.16).

Collecting data for income involves questions about revenue from rent and property, remittances, regular wage work, up to two casual jobs, agriculture, self-employment, and non-employment sources.

Both involve detailed data collection on prices, for expenditure and for input and output prices for self-employment and agriculture. Samples of prices of all of the above goods must be obtained from formal sources (supermarkets, department stores) and less formal sources (spazas, street vendors, "corner cafes") in each area to compile regional and time price deflators.

Collecting income and consumption data is time consuming, costly and requires a high degree of enumerator training and post-collection data processing. For these reasons, comprehensive surveys aimed at directly measuring income and consumption are carried out only occasionally in a limited range of developing countries.

Assets as a proxy for income and consumption

Advantages

Collecting information on assets, such as consumer durables, land, livestock and housing can be more accurate, more easily verifiable, and less time consuming and costly. Asset indices can be a good proxy for income and consumption levels and access to resources.

For developing countries, the number and range of assets owned, especially by the poor, is often limited. This means obtaining information about assets involves fewer and less detailed questions than obtaining information about expenditure.

Ownership of assets as an indicator does not entail recall problems and is less susceptible to errors or misleading by respondents as it can usually be immediately observed and verified (Sender and Smith, 1990, ch.3, p.27). Assets are a stock measure and so are a good indicator of long run income without entailing extensive recall.

Asset indices can be compiled without using price data through using weighting techniques outlined below. For this reason, they avoid the complexities and inaccuracies involved in controlling for prices.

Disadvantages

International comparisons of assets are more meaningful than money incomes converted using exchange rates. However asset indices do have disadvantages in this area due to the different importances of different assets for different cultures, climates, levels of infrastructure, and levels of public good provision.

Quality of assets is not easily accounted for, and a heavily depreciated asset carries the same weight as new high quality asset.

There is also the problem of the right of use versus the right of exchange. A household may have the use of an asset, but may not be able to sell it. Therefore a use right over an asset may not be a good measure of overall access to resources through sale and purchase, but still contributes to well-being, making inclusion of the asset in these cases ambiguous.

Sender and Smith (1990, Ch. 3) argue against the Chayanovian proposition that the age profile of a household has a large bearing on the assets it owns and so poor but old households could own as many assets as rich but new ones. However, time allows households to cumulatively sell assets (if poor) as well as cumulatively buy them (if rich). Sender and Smith (1990) find that the proportion of variation that can be explained by household age is very small in their area of study in Tanzania.

The availability of price data, and the level of complexity in income, expenditure or asset levels may mean different methods are appropriate in different contexts. For example in developed countries where cost constraints are less, price data is more readily available, and most people have a small number of distinct income sources, income data may be easier and less costly to collect. However in developing countries asset indices may be an important indicator of income where income and expenditure data is inaccurate or costly to collect.

Periodic surveys which collect information on both income/consumption directly and on assets, so enabling the relationship between them to be monitored in different times and places, could be continued. These could be complemented by more frequent surveys focused on different topics which use only assets as an indicator of income/consumption. However, if these frequent surveys are to draw on periodic income/consumption surveys, the sets of assets for which data is collected in both must be comparable.

Previous use of asset indices

Asset indices have been used in many studies where measuring income directly but inaccurately is seen as less suitable than using an asset index as an income proxy, or where collecting income and expenditure data would involve an overly lengthy questionnaire when the focus of the study is another household characteristic such as escape routes from poverty (Sender and Smith, 1990), education (Filmer and Pritchett, 1998), health or fertility (see studies listed in Morris et al, 1999, p.3 and App.A).

Attaching greater priority to the inclusiveness and size of a sample rather than to the length of questionnaire may also have benefits. For example, national surveys often do not capture "typical" farms, and often do not adequately include migrating groups and wage labourers. Using asset indices, for a given collection cost, it is possible to have more frequent and more widespread surveys (Collier et al, 1986 in Sender & Smith,1990 Ch.1; Sender and Smith, 1990 ch.2).

Household surveys which collect income and expenditure data directly often also include information on a number of assets. These have been used in a number of studies analysing the relationship between assets and income and expenditure data, with a view to assessing their relative merits, such as Filmer and Pritchett (1998) using LSMS from Nepal (1996), Indonesia (1994) and Pakistan (1991); Montgomery et al (1999), using LSMS in Ghana (1988), Jamaica (1989), Pakistan (1991), Peru (1994), and Tanzania (1993–94); Morris et al (1999) using data on rural households in IFPRI surveys from Mali (1998), Malawi (1998) and the LSMS in Cote d'Ivoire (1986); and Sahn and Stifel (2000) using data from LSMS in Côte d'Ivoire (1986), Ghana (1988) and Vietnam (1993).

Criteria for choice of assets to include

Assets indices which include assets which clearly differentiate between richer and poorer households, i.e., which richer households tend to have while poorer ones do not, are more likely to reflect income differences. If assets which are highly specific to particular age groups, to whether a household is rural or urban, or to the availability of supporting public infrastructure or physical environment, this must be acknowledged when the index is interpreted or used. For example, ownership of a bicycle or car can depend on the availability of public transport, of roads, and the proximity of important services, while ownership of electrical appliances can depend on the availability of an electricity supply (Knodel and Wongsith, 1991). Sahn and Stifel's observation (2000, p.11) that rural people are poorer in relation to urban people when assets rather than direct income or consumption measures are used must be interpreted carefully considering the particular assets chosen. Including producer assets specific to particular income sources may cause invalid income comparisons by income source, for example rich urban households may not own livestock or land assets. Assets that are commonly purchased through the market rather than publicly provided are a better indicator of private income. For example, whether a household has access to piped water may depend more on its location in an area with public water facilities than on its income level. If the asset index is to be used to assess the relationship between income and another variable, (e.g. health), then assets which have direct impacts on the variable (e.g. piped water supply) rather than an impact through income should be excluded.

The appropriateness of different assets will depend on the context and purposes of different studies. Some of the above criteria may present conflicts. For example in some rural areas land ownership may be a very important asset for discriminating between rich and poor households, while it would not be a good indicator to assess income differences between rural and urban areas.

The South Africa Living Standards Measurement Survey

In order to assess how asset indices can be compiled using household surveys, this paper uses data from the South Africa Living Standards Measurement Survey.

In the South Africa LSMS information on the ownership of the following consumer durable assets is collected:

- Primus cooker
- Gas stove
- Electric stove
- Kettle
- Geyser
- Fridge
- Radio
- Television
- Telephone
- Bicycle
- Motor vehicle

In other parts of the survey, there is information on:

- Number of rooms occupied by the household
- Building material of roof, floor and walls
- Type of toilet
- Source of water
- Cropped land area
- Farming assets
- Ownership of cattle, sheep, goats, pigs and poultry

The choice of which variables are most appropriate to include from this list will be discussed below. The variables used for the examples of weighting methods comprise the consumer durables only, for the sake of simplicity. However, it must be noted that of 8847 households in the South Africa LSMS, 5824 households own none of the consumer durables listed, or have no data on them.

Weighting methods

An asset index is a single value for each household indicating its degree of asset ownership. The following comprise some methods of doing this.

Asset count

The simplest form of asset index would count the number of assets each household owns. This has the advantage that a household's asset score is simple to understand and has a clear meaning. However, this does not provide discrimination between, for example, two otherwise similar households, one owning an inferior good (e.g. a Primus cooker) and the other owning a superior good (e.g. an electric stove) both of which give one "point".

Prevalence weights

Morris et al (1999, p.8) assign weights equal to the reciprocal of the proportion of the study households that owned one or more of that item, on the assumption that those assets owned by a smaller number of households are those owned by wealthy households. However, this neglects the problems of inferior goods (for example households replacing a primus cooker with an electric stove) which can be owned by a small group, but do not indicate wealth. It also neglects the possibility of less prevalent specialized assets being assigned higher weights.

Prices

Fleurent (1978 in Sender and Smith, 1990, Ch.3) suggests using cash value when new as assets weights. This would provide an estimation of the value of the assets a household owns. However, this entails all the problems in compiling price data outlined above, which the use of an asset index is seeking to avoid.

Subjective weights

Weights can be assigned subjectively based on an analysis of the pattern of accumulation of these assets. To examine this possibility, the following chart provides an indication of the pattern of accumulation of the consumer durables in the South Africa LSMS.

No. of assets owned	Number of households with this number of assets	Most common ownership pattern *Second most common ownership pattern*	No. of households with this pattern	No. of households with this pattern as % of No. of households with this number of assets
1	324	1 Radio	149	46.0%
		1 Primus Cooker	*135*	*41.7%*
2	417	1 Radio, 1 Primus Cooker	225	54.0%
		1 Radio, 1 Electric Stove	*28*	*6.7%*
3	286	1 Radio, 1 Primus Cooker, 1 Television	46	16.1%
		1 Radio, 1 Primus Cooker, 1 Bicycle	*25*	*8.7%*
4	197	1 Radio, 1 Electric Stove, 1 Television, 1 Fridge	14	7.1%
		1 Radio, 1 Electric Stove, 1 Television, 1 Electric Kettle	*12*	*6.1%*
5	184	1 Radio, 1 Electric Stove, 1 Television, 1 Fridge, 1 Electric Kettle	31	16.8%
		1 Radio, 1 Primus Cooker, 1 Television, 1 Fridge, 1 Motor Vehicle	*8*	*4.3%*
6	221	1 Radio, 1 Electric Stove, 1 Television, 1 Fridge, 1 Electric Kettle, 1 Telephone	47	21.3%
		1 Radio, 1 Electric Stove, 1 Television, 1 Fridge, 1 Electric Kettle, 1 Geyser	*36*	*16.3%*
7	271	1 Radio, 1 Electric Stove, 1 Television, 1 Fridge, 1 Electric Kettle, 1 Telephone, 1 Geyser	49	18.1%
		1 Radio, 1 Electric Stove, 1 Television, 1 Fridge, 1 Electric Kettle, 1 Geyser, 1 Motor Vehicle	*16*	*5.9%*

This table shows there seem to be typical patterns of accumulation where a household first acquires a radio and a primus cooker, then a television, then replaces the primus cooker with an electric stove and then acquires successively a fridge, electric kettle, telephone, geyser, motor vehicle, and a bicycle.

This could give rise to a weighting system such as the following, constructed so that those households owning goods which households owning more assets own receive higher scores.

Radio	10
Primus Cooker	11
Television	12
Electric Stove	13
Fridge	14
Electric kettle	15
Telephone	16
Geyser	17
Motor vehicle	18
Bicycle	19

The order of weights do not necessarily correspond to the ordering of prices that would be expected, for example, a bicycle receives a higher weight than a motor vehicle.

Principal components

The principal components technique is used in the studies by Filmer and Pritchett (1998), while Hammer (1998), and Sahn and Stifel (2000) use a variant of it called factor analysis. A more detailed description of the mathematical techniques below are provided in Appendix A.

Principal components gives higher weights to those variables which co-vary most with the other variables. For example, an asset gets a higher weight if households owning more of that asset tend to own more of other assets also.

In order to control for the units in which assets are counted, each value (the number of each asset owned by each household) is standardised, for example:

$$\frac{\text{Number of radios the household owns} - \text{mean number of radios owned}}{\text{Standard deviation of number of radios owned}}$$

The mean number of radios owned is the number of radios per household, 0.436. The standard deviation (square root of the variance) is 0.863. This standardisation means that the ownership of one radio is converted into (1 – 0.436) / 0.863 = 1.264.

This is then multiplied by the radio weight to get the radio score for a household with one radio. In the first example below, the weight for a radio is 0.128. Therefore a household with one radio gets 0.128 x 1.264 = 0.083 added to its score.

When the principal components procedure is performed for the consumer durable asset variables in the South Africa LSMS on all 8847 households, the following weights are produced:

A low weight is given to Primus cooker, as richer households often replace it with electric or gas stoves, which are given higher weights. The high weights given to geysers (the South African word for a water heater), electric kettles and televisions indicate they are assets which are important discriminators between richer and poorer households. The lower weights given to bicycles and motor vehicles indicate that these are not always owned by richer households, or may sometimes be owned by less well off ones.

Primus Cooker	0.007
Gas Stove	0.067
Bicycle	0.092
Motor Vehicle	0.126
Radio	0.128
Telephone	0.131
Electric Stove	0.131
Geyser	0.132
Electric Kettle	0.134
Television	0.135
Fridge	0.136

Each household's score is computed by adding the appropriate score for each asset according to how many the household owns. The following table illustrates the scores for owning 0, 1, 2, or 3 for each asset using the above weights.

	Motor Vehicle	Bicycle	Radio	Electric Stove	Gas Stove	Primus Cooker	Fridge	Tele vision	Geyser	Electric Kettle	Tele phone
0	-0.045	-0.025	-0.065	-0.066	-0.016	-0.002	-0.062	-0.066	-0.054	-0.064	-0.052
1	0.160	0.136	0.083	0.246	0.216	0.017	0.179	0.164	0.258	0.242	0.249
2	0.364	0.297	0.232	0.557	0.448	0.035	0.419	0.394	0.570	0.549	0.550
3	0.568	0.459	0.380	0.869	0.680	0.054	0.660	0.624	0.882	0.856	0.851

When every household is given a score using this technique, the mean score is 0 and the standard deviation of the scores is 1.

Inclusion of assets

The most important shortcoming of this list of assets is that it fails to include items which a large part of the population in South Africa own. Of 8847 households, only 3023 own any of these assets. 5824 households will therefore receive a uniform asset index score of -0.517 (the sum of the zero row) using the above range of assets.

Number of rooms

A variable corresponding to the number of rooms occupied by the household can also be included. The problems this raises in terms of numbers of members of household will be dealt with below.

Housing material, toilet facilities, water

Information on the building material of roof, floor and walls could help to discriminate among households with none of the above assets. However, including this information involves either i) giving each of the 15 material types a value or ii) including a variable for each material and giving households a 0 or 1 depending on whether it is applicable. The first method would involve deciding beforehand which materials were those of the wealthier households, i.e. giving them an order, and also deciding by how much one material was better than another, i.e. putting them on a ratio scale. Principal components technique is only valid for ordinal, and preferably ratio, variables (Bennet & Bowers, 1976, p.118). The second method would not be valid as the material dummy variables would be correlated, i.e., a "1" for asbestos roof would imply "0"s for all the other roof variables.

This information could however be included by a simple division of materials into better and worse types, using general knowledge. For example, a 1 can be given to households with roofs, walls or floors with carpet, bricks, tile, asbestos, wood, cement block or corrugated iron, and a 0 given for cardboard, prefab, wattle and daub, mixture of mud and cement, plastic, thatching, linoleum or mud. Similarly, a toilet variable could be constructed giving 0 for no toilet, 1 for a pit latrine, bucket toilet or chemical toilet and 2 for a flush toilet, where there is little ambiguity about order.

A water source variable is not included as there are twelve types of water source which cannot be unambiguously ordered, and water source may well heavily influenced by local public infrastructure as well as private income factors.

Farming assets

Asset variables for livestock assets and number of hectares of cropped land are also included initially below.

Results

Including all the variables for consumer durables, building materials, number of rooms the household occupies, livestock assets, ownership of cropped land gives the following asset weights using principal components:

Poultry	-0.017
Pigs	-0.012
Cattle	-0.009
Sheep	-0.003
Goats	0.003
Land	0.007
Primus Cooker	0.010
Roof	0.025
Walls	0.048
Gas Stove	0.053
Floor	0.065
Toilet	0.067
Bicycle	0.073
Motor Vehicle	0.101
Radio	0.104
Telephone	0.106
Geyser	0.107
Electric Stove	0.107
Electric Kettle	0.110
Television	0.110
Fridge	0.111
Number of rooms	0.112

There are negative weights for sheep, cattle, pigs and poultry, which means that the more of these a household owns, the *lower* their asset score. This is because households owning the other assets tend not to own livestock. The inclusion of livestock assets clearly worsens the asset index as an indicator of wealth across a sample of both rural and urban households.

Cropped land has a very weak correlation with the other variables (regardless of whether livestock are included) and so has a low weight – a household with no land has -0.00017 subtracted from its score while a household with 3 hectares has 0.00004 added to its score. Therefore the inclusion of land only serves the purpose of discriminating between wealthy households with a large amount of land and wealthy households with no cropped land. This is not useful as wealthy farm households may not be more wealthy than an urban household with otherwise the same assets.

Excluding land and livestock variables gives the following (very similar) set of weights.

Primus Cooker	0.010
Roof	0.025
Walls	0.047
Gas Stove	0.054
Floor	0.065
Toilet	0.067
Bicycle	0.074
Motor Vehicle	0.102
Radio	0.105
Telephone	0.106
Geyser	0.107
Electric Stove	0.108
Electric Kettle	0.110
Television	0.110

Correlations between asset indices and income/consumption measures

Giving each household a score allows a comparison of the scores households get using different methods and including different assets. The asset index scores can also be compared with the expenditure and the income data of households in the South Africa LSMS.

The following tables show the Pearson correlation coefficients, which illustrate the degree to which households' scores on one scale corresponds to their score on another scale; and the Spearman Rank correlation coefficients, which show how households' ranking or order on one scale corresponds to their ranking on another scale (see Appendix B).

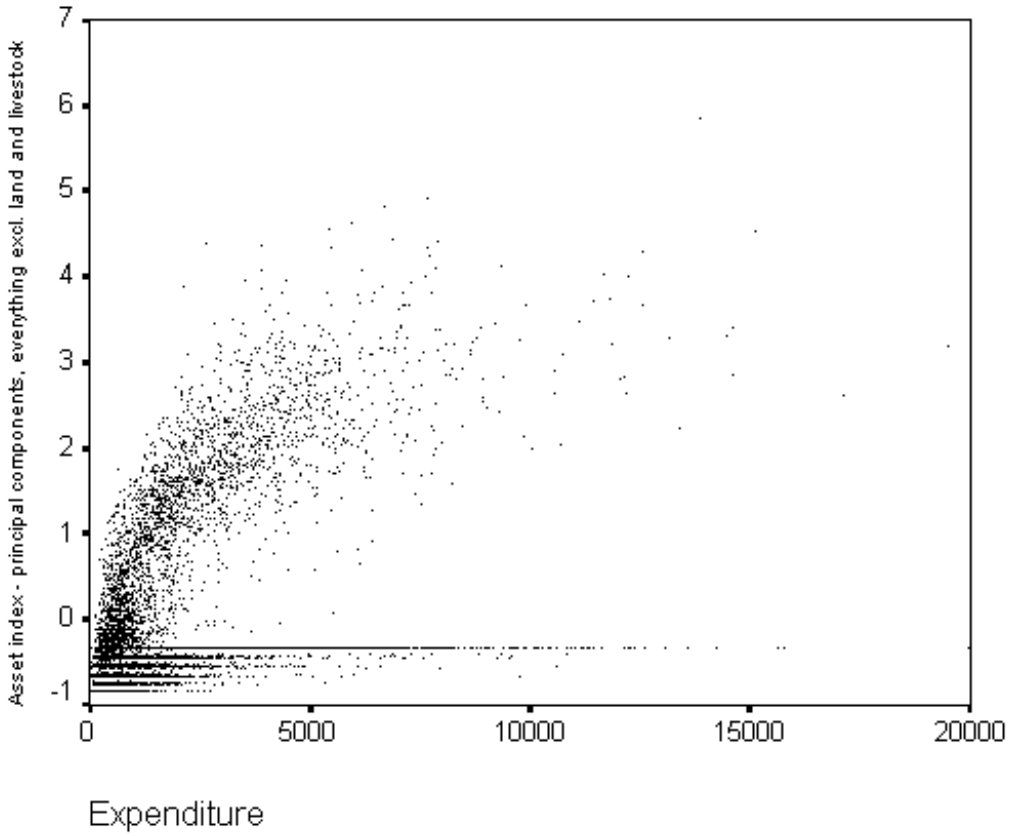
Index B assigns weights subjectively as in the second case above, with the addition of number of rooms, ten points for having each of roof, wall and floor made of a higher quality material, ten points for any toilet and twenty for a flush toilet.

Pearson correlation coefficient – all households	A Simple count of consumer durables	B Subjective weights – everything excluding land and livestock	C Principal components - consumer durables only	D Principal components - everything excluding land and livestock	E Principal components - everything	Income	Expenditure
A Simple count of consumer durables	1.000	0.980	0.993	0.983	0.982	0.317	0.395
B Subjective weights – everything excluding land and livestock	0.980	1.000	0.976	0.992	0.992	0.346	0.439
C Principal components - consumer durables only	0.993	0.976	1.000	0.988	0.987	0.322	0.403
D Principal components - everything excluding land and livestock	0.983	0.992	0.988	1.000	1.000	0.336	0.426
E Principal components – everything	0.982	0.992	0.987	1.000	1.000	0.336	0.426
Income	0.317	0.346	0.322	0.336	0.336	1.000	0.531
Expenditure	0.395	0.439	0.403	0.426	0.426	0.531	1.000

Spearman rank correlation coefficient – all households	A Simple count of consumer durables	B Subjective weights – everything excluding land and livestock	C Principal components - consumer durables only	D Principal components - everything excluding land and livestock	E Principal components - everything	Income	Expenditure
A Simple count of consumer durables	1.000	0.825	0.999	0.824	0.822	0.389	0.355
B Subjective weights – everything excluding land and livestock	0.825	1.000	0.824	0.994	0.990	0.566	0.524
C Principal components - consumer durables only	0.999	0.824	1.000	0.825	0.823	0.390	0.356
D Principal components - everything excluding land and livestock	0.824	0.994	0.825	1.000	0.995	0.563	0.521
E Principal components – everything	0.822	0.990	0.823	0.995	1.000	0.557	0.513
Income	0.389	0.566	0.390	0.563	0.557	1.000	0.767
Expenditure	0.355	0.524	0.356	0.521	0.513	0.767	1.000

This shows that the correlations between the different types of asset indices are very high – Pearson coefficients all over 0.97 and Spearman coefficients all over 0.82. Indices A and C, and indices B, D and E correlate especially highly, indicating that the assets included are more important than the weighting method.

The correlations between the asset indices and income and expenditure data are very low – Pearson coefficients between 0.31 and 0.42 and Spearman coefficients between 0.35 and 0.56. Asset indices correlate more with expenditure than income in terms of magnitude, but the ranking of households using assets corresponds more with income than expenditure. The correlation between income and expenditure is also quite low: 0.53, but rankings according to income and expenditure correspond at 0.77.



However, these correlation coefficients were calculated using all 8847 households. The above scatterplot illustrates that asset households with no consumer durable assets fall on a number of close lines (according to their scores on the other five variables), and are not spread out vertically. Some households with high expenditure may have been recorded as having no consumer durables due to inadequate data. The asset indices are not good discriminators within the lower half of the distribution due to the lack of inclusion of more basic consumer durables in the indices. The following table shows correlation coefficients when only scores for those households possessing at least one consumer durable are included:

Pearson correlation coefficient – households owning consumer durables only	A Simple count of consumer durables	B Subjective weights – everything excluding land and livestock	C Principal components - consumer durables only	D Principal components - everything excluding land and livestock	E Principal components - everything	Income	Expenditure
A Simple count of consumer durables	1.000	0.987	0.987	0.979	0.979	0.440	0.749
B Subjective weights – everything excluding land and livestock	0.987	1.000	0.982	0.985	0.984	0.443	0.755
C Principal components - consumer durables only	0.987	0.982	1.000	0.995	0.995	0.436	0.747
D Principal components - everything excluding land and livestock	0.979	0.985	0.995	1.000	1.000	0.439	0.751
E Principal components – everything	0.979	0.984	0.995	1.000	1.000	0.440	0.752
Income	0.440	0.443	0.436	0.439	0.440	1.000	0.569
Expenditure	0.749	0.755	0.747	0.751	0.752	0.569	1.000

Spearman rank correlation coefficients – households owning consumer durables only	A Simple count of consumer durables	B Subjective weights – everything excluding land and livestock	C Principal components - consumer durables only	D Principal components - everything excluding land and livestock	E Principal components - everything	Income	Expenditure
A Simple count of consumer durables	1.000	0.980	0.988	0.977	0.977	0.780	0.841
B Subjective weights – everything excluding land and livestock	0.980	1.000	0.976	0.987	0.987	0.793	0.848
C Principal components - consumer durables only	0.988	0.976	1.000	0.988	0.988	0.787	0.843
D Principal components - everything excluding land and livestock	0.977	0.987	0.988	1.000	1.000	0.798	0.855
E Principal components – everything	0.977	0.987	0.988	1.000	1.000	0.797	0.855
Income	0.780	0.793	0.787	0.798	0.797	1.000	0.836
Expenditure	0.841	0.848	0.843	0.855	0.855	0.836	1.000

This shows much higher correlation coefficients between expenditure and the asset indices, around 0.75 for scale and 0.85 for rank. This compares to asset index with expenditure correlations in the studies by Morris et al (1999), (scale correlation coefficient of 0.74 for IFPRI rural Mali data) and Sahn and Stifel (2000) (rank correlation coefficients for LSMS in Cote d'Ivoire, Ghana, Vietnam of 0.7, 0.7 and 0.5 respectively). Scale correlation between income and the asset indices improves but remains low, at around 0.44, while rank correlation increases to 0.79.

This shows that the low correlations between asset indices and income and expenditure using all 8847 households are primarily due to the inappropriate range of assets included in the questionnaires.

The high correlations in all cases between indices A and C, and indices B, D and E mean that the choice of weighting method is not as important as the choice of assets.

Recommendations

Both direct income and consumption measures and asset proxies have shortcomings, and for this reason periodic studies such as the South Africa LSMS which include both can allow them to act as checks and balances on each other. However, asset indices alone can be used for more frequent and less costly studies in developing countries. In order for both types of studies to draw on each other, they both should include an appropriate range of assets.

The South Africa LSMS is not alone among World Bank Living Standards Measurement Surveys in including only consumer durables owned by the richest households. The assets included in the surveys analysed by Filmer and Pritchett (1998) are similar to those in the South Africa LSMS, and comprise a radio, cooking appliance, television, fridge, car, toilet, bicycle, piped drinking water, number of rooms, materials of home building, clock/watch, sewing machine and lighting source.

Those analyzed by Montgomery et al (1999) are also similar, comprising a bicycle, radio, television, fridge, motorcycle, car, access to clean water, having water on premises, less than thirty minutes travel to reach water, a toilet facility, a flush toilet, non-dirt flooring and access to electricity.

Besides non-dirt flooring (78% in Ghana), some toilet facility (73% in Ghana and 96% in Tanzania) and the time to water, the largest prevalence of any of the other assets is 26% in Ghana and 44% in Tanzania. This clearly makes discrimination between the poor and the poorest dependent on very few indicators. (Montgomery et al, 1999, Table 2)

In contrast, the asset data used in smaller studies of rural Mali and Malawi collected by IFPRI (Morris, 1999, p.5) include a range of assets more likely to be owned by low income households. The Mali survey included 18 different types of agricultural implements, 18 lower cost consumer durables such as bicycles, gas lamps, tables, and chairs and 16 different kinds of kitchen equipment such as pots, cups, and calabashes.

The indicators that Sender and Smith (1990, p.7) use to identify the poorest in a rural community in Tanzania include non-ownership of basics such as a watch, light, radio, bicycle, coat, sweater, more than two pairs shoes, more than two matresses, more than three beds, more than three chairs, more than four stools, more than four rooms, more than six hectares of land, a metal roof, and non-mud walls. The assets used in Sender's Mpumalanga survey are bicycle, sewing machine, stove, primus stove, wheelbarrow, paraffin lamp, wardrobe, chair, bench, bed, watch, radio, tripod cooking port, saucepan, shoes, one pair of shoes per resident adult, post office or bank account, electricity access, nearby water tap access, house of cement or brick walls, more than two rooms, and at least one rooms per resident adult. (Sender, 2000, table 9, pp.12-13)

The household size issue

Different sizes and types of households require different levels of income to achieve the same economic welfare. In order to use income, expenditure or asset information to assess welfare, household size is often controlled for, through dividing by number of household members, or number of "equivalent adults" where a child is weighted less than an adult. If income, expenditure and assets all have the same relationship to welfare and household size, then comparisons of income, expenditure and assets should not differ from comparisons of income per person, expenditure per person and assets per person. However, some assets may be less household-size dependent than some forms of income and expenditure. For example, a household may require one and only one cooker to attain a given level of welfare, regardless of household size, whereas larger households may require higher expenditure on food to attain a given level of welfare.

For this reason, the inclusion of both household assets (eg cookers) and personal assets (eg shoes, coats) in proportions reflecting the relative importance of regular per-household expenditure (eg heating) and regular per-person expenditure (eg food) should be aimed for. The assets included in the South Africa LSMS are inadequate in this respect, as they comprise mainly household assets (only 2% of households possess more than two of any asset, regardless of household size).

Conclusion

The inclusion of a range of assets within the reach of the less poor of the poor, representing personal and household assets, and tailored to local living standards and goods availability would not be difficult in Living Standards Measurement Surveys. The range of assets included is the most important aspect of asset proxies. Questions on such assets would be quick to answer and easily verifiable in the context of a complex income and consumption questionnaire.

This would enable more frequent studies to use an index of similar assets as an accurate and less costly measure of a proxy for income/consumption, rather than an inaccurate and costly measurement of the variable itself, while understanding the relationships between the asset index and the direct measure.

Bibliography

Collier, P., Radivan, S., Wangwe, S., and Wagner, A. 1986 ‘‘Labour and poverty in rural Tanzania’’ Clarendon Press: Oxford

Deaton A.1997 ‘‘The analysis of household surveys: a microeconometric approach to development policy’’ Washington, DC: IBRD/The World Bank, 479.

Fields, G. 1998 “Income Mobility: Meaning, Measurement, and Some Evidence for the Developing World,” mimeo, Cornell University.

Filmer and Pritchett, 1998 ‘‘Estimating Wealth Effects without Expenditure Data -- or Tears: An Application to Educational Enrollments in States of India’’ World Bank Working Paper no.1994 http://econ.worldbank.org/view.php?type=5&id=707

Fleurent, P., 1978 “Farm and Market: a study of society and agriculture in Tanzania” PhD Thesis University of California

Hammer, J. 1998. “Health Outcomes across wealth groups in Brazil and India,” mimeo, DECRG, The World Bank. Washington, DC.

Knodel, J. and Wongsith, M. 1991 “Family size and children’s education in Thailand: Evidence from a national sample,” in Demography 28(1): 119–131.

Montgomery, M., Burke, K., Paredes, E. and Zaidi, S. 1999 “Measuring Living Standards with proxy variables” Research Division, The Population Council. New York, NY. http://www.popcouncil.org/publications/wp/prd/129.html

Morris, S., Carletto, C., Hoddinott, J. and Christiaensen, L. 1999 “Validity of Rapid Estimates of Household Wealth and Income for Health Surveys in Rural Africa” International Food Policy Research Institute FCND Discussion Paper No. 72 http://www.ifpri.org/divs/fcnd/dp/dp72.htm

Pradhan, M. 2000 ‘‘How many questions should be in a consumption questionnaire? Evidence from a repeated experiment in Indonesia.’’ Cornell Food and Nutrition Program Working Paper No. 112. Ithaca, NY.

Sahn, D. and Stifel, D. 2000 “Exploring Alternative Measures of Welfare in the Absence of Expenditure Data” Cornell Food and Nutrition Policy Program Working Paper No. 97. Ithaca, NY: Cornell University http://www.people.cornell.edu/pages/des16/sahn/publ.html

Scott, C and Amenuvegbe, B. 1990 ‘‘Effect of Recall Duration on Reporting of Household Expenditures: An Experimental Study in Ghana’’ Social Dimensions of Adjustment in Sub-Saharan Africa Working Paper 6, Washington DC: World Bank.

Sender, J. 2000 ‘‘Struggles to escape poverty in South Africa: Results from a purposive rural survey’’ School of Oriental and African Studies Working Paper no.107

Sender, J and Smith, S 1990 Poverty, class and gender in rural Africa : a Tanzanian case study London ; New York : Routledge

SPSS Release 10.0.5 http://www.spss.com

World Bank South Africa Living Standards Measurement Survey household questionnaire
http://www.worldbank.org/html/prdph/lsms/country/za94/docs/za94hha4.pdf

World Bank Data from South Africa Living Standards Measurement Survey

Data	**Data File**	**Household Questionnaire question number**
Housing	S4_HSV1	2.1
Water and Sanitation	S3_HSV2	2.2/2.3
Household items	M2_NFS4	4.4
Agricultural land	S4_LAND1	7.1
Livestock: cattle etc (1)	M1_AGR2	8.5.1.2
Livestock: Cattle etc (2)	S1_AGR2	8.5.1.2

Appendix A – Principal Components

To perform the principal components procedure, a variance-covariance matrix produced for the assets, which places the variance (V) of each asset on the diagonal, and the covariance (C) between each pair of assets in the appropriate row and column of the matrix. The main diagonal is then replaced with communalities (M), ie the highest covariance the asset has with any other asset. Each column is then summed (S). Each sum is divided by the square root of the sum of sums to produce factor loadings for each asset (F). These factor loadings are then put in the matrix in place of communalities and the process repeated iteratively until a desired degree of accuracy is achieved. The final factor loadings are the weights. (Bennet and Bowers, 1976, Ch.2)

$$\begin{matrix} V_1 & C_{12} & C_{13} & C_{14} \\ C_{12} & V_2 & C_{23} & C_{24} \\ C_{13} & C_{23} & V_4 & C_{34} \\ C_{14} & C_{24} & C_{34} & V_4 \end{matrix} \qquad \begin{matrix} M_1 & C_{12} & C_{13} & C_{14} \\ C_{12} & M_2 & C_{23} & C_{24} \\ C_{13} & C_{23} & M_4 & C_{34} \\ C_{14} & C_{24} & C_{34} & M_4 \end{matrix}$$

$$\begin{matrix} S_1 & S_2 & S_3 & S_4 & SS \end{matrix}$$

$$F_1 = S_1 / \sqrt{SS}$$
$$F_2 = S_2 / \sqrt{SS}$$
$$F_3 = S_3 / \sqrt{SS}$$
$$F_4 = S_4 / \sqrt{SS}$$

In the analysis performed in this paper, SPSS was used (Analyse, Data Reduction, Factor, Extraction number of factors: 1).

Appendix B – Correlation Coefficients

A variance of a variable is calculated by finding, for each household, the square of the difference between the variable and its mean, and then summing across all households and dividing by the number of households.

$$\frac{\sum (X-\bar{X})^2}{n}$$

The covariance between two variables is computed by finding, for each household, the difference between one variable and its mean, multiplied by the difference between the other variable and its mean. These are then summed across all households, and then divided by the number of households.

$$\frac{\sum (X-\bar{X})(Y-\bar{Y})}{n}$$

The Pearson correlation coefficient controls covariance for units of measurement. It is computed by dividing the covariance by the square root of the product of the two variances.

$$\frac{Covar_{xy}}{\sum (Var_x \, Var_y)}$$

The Spearman rank correlation coefficient is computed as follows. Each household is given a rank in terms of one variable and a rank in terms of the other, and the difference between the two ranks is found and squared. This is then summed across all households and adjusted according to the following formula (n is the number of households):

$$1 - \frac{6 \sum d^2}{n^2(n-1)}$$

The correlation coefficients were also calculated using SPSS (Analyse, Correlate, Bivariate, Pearson and Spearman).