On Comparability of Poverty Statistics From Different Sources and Disaggregation Levels[1]

November 2002

Introduction

The set of poverty statistics is dependent upon and as fluid as the definition of poverty, the particular circumstance, purpose or intended use, as well as the choice of the producer-user. There are myriad producers-users of poverty statistics: researcher, analyst, politician, policy maker, poverty alleviation program implementer, national statistics office, and international agency. For the aim of this note, there are advantages in focusing on the statistics to monitor the first of the eight Millennium Development Goals (MDG), which is **to eradicate extreme poverty and hunger**. It should now be well known that the MDGs were agreed upon by a great majority of countries during the 2000 UN Summit. A series of international expert consultations has led to a consensus on the statistics (or indicators or statistical indicators) to be used to monitor progress on the 18 targets under the eight goals. Those for the first goal are shown in Table 1.

Propagation of terms

There are quantitative targets for each goal. Target 1 under goal 1 is to "halve, between 1990 and 2015, the proportion of people whose income is less than one dollar a day", with the phrase in quotes exactly as stated in the Millennium Declaration. That is, **extreme poverty** is equated with per capita income below one dollar a day. Thus, this kind of poverty is income based and inclusive of both food and essential non-food needs. Three so-called *road map indicators* have been designated for monitoring progress on target 1, namely, proportion of population below $1 per day, poverty gap ratio, and share of poorest quintile in national consumption. The World Bank is the assigned compiler/custodian for these three indicators. The UNSD has set up *millenniumindicators.un.org*, an electronic web site to systematically store and update the roads map indicators together with the input statistics required to compile them; e.g. population counts, proportion and number of persons below the national poverty line (obtained from national sources).

Target two is to "halve, between 1990 and 2015, the proportion of people who suffer from hunger". While the meaning and definition of **hunger** is far from settled (e.g. see FAO ISS Proceedings, forthcoming), the road map indicators and designated sources are: prevalence of underweight children (UNICEF) and proportion of population below minimum level of dietary energy consumption (FAO). The first may

[1] Prepared by I.P. David for UN-ESCAP for presentation at the Committee on Statistics Meeting, 27-29 November 2002. The views expressed here are those of the author and do not necessarily represent the official views or opinions of ESCAP.

be considered a medium-term output indicator where the primary sources (when available) are national nutrition and anthropometrics surveys. The latter is a short-term input indicator (i.e. based on one year food supply) and narrower in the sense that it does not include basic non-food needs. Moreover, as will be discussed later, the FAO estimates are computed

Table 1. Millennium Development Goal 1: Targets and Road Map Indicators

Goal 1. Eradicate extreme poverty and hunger	**Road Map Indicators**
Target 1. Halve, between 1990 and 2015, the proportion of people whose income is less than one dollar a day.	1. Proportion of population below $1 per day 2. Poverty gap ratio (incidence x depth of poverty) 3. Share of poorest quintile in national consumption
Target 2. Halve, between 1990 and 2015, the proportion of people who suffer from hunger	4. Prevalence of underweight children (under five years of age) 5. Proportion of population below minimum level of dietary energy consumption

by the agency itself using a methodology that has no resemblance to methods used by the individual countries and other international agencies.

The methodology used by the majority of developing countries, but certainly not the only one, was described in David (2001). It involves choosing a food basket that will provide a prescribed minimum level of dietary kilocalories; estimating the cost of the food basket or the food poverty line (fpl); estimating the per capita income (or expenditure) distribution of the target population; and persons with income below fpl are called **food poor,** and their proportion is a measure of **food poverty incidence**. Some countries, e.g. Philippines, refer to the food poor as **core poor**. While some writers speak of this kind of deprivation as extreme poverty also, it should not, however, be likened to the extreme poverty referred to in MDG1, since the latter is inclusive of both food and non-food components of poverty. The countries' food poverty incidence is conceptually akin to FAO's indicator of undernourishment prevalence; it will be argued later, however, that the methodologies, and hence the results are not comparable.

A country inflates its fpl by incorporating cost estimates for basic non-food needs. The result is a total poverty line (tpl) that is used with the per capita income (or expenditure) distribution to estimate what is commonly referred to as **absolute poverty incidence**. This indicator, like the one dollar a day indicator used to monitor

MDG1-Target 1, are both inclusive of food and essential non-food deprivation; however, there are significant differences between the two in terms of methodologies and results, as will be discussed later.

The countries contribute further to the proliferation of terms. Some generate lower and upper absolute poverty incidences by defining two tpls or two fpls; e.g. Bangladesh and Vietnam. The national statistical offices in the same two countries conduct not one but two poverty surveys, resulting in two different poverty statistics series.

The Committee may wish to note and comment on the observation that all five indicators are assigned to three international agencies (WB, UNICEF, FAO) and that two, namely $1 a day poverty incidence and prevalence of dietary energy insufficiency are compiled directly by WB and FAO respectively. The poverty gap ratio, if computed using $1 a day indicator, will also be a WB product not comparable to nationally sourced indicators. Moreover, if the share of the lowest quintile in national consumption (or income) and the prevalence of underweight children are updated annually by the WB and UNICEF respectively, then there will be degrees of non-comparability with national estimates that are updated less frequently (usually once in five years).

Comparability at Sub-National Levels

Poverty alleviation in the third world and the kind being addressed in MDG 1 is absolute poverty, not relative poverty. It is desirable for sub-national estimates to be comparable for exactly the same reasons that much effort is being expended to make country estimates comparable: poverty reduction programs based on comparable sub-national estimates of the number of absolutely poor could lead to more efficient targeting and allocation of poverty alleviation resources; and simple summing or averaging leads to reasonable higher level aggregates, including national estimates. Two important and common sources of comparability loss are: the choice of initial reference poor population(s) in determining the food basket, the non-food essentials, and income (expenditure distribution); and shifts in sub-national or domain boundaries.

Reference Population

The majority of countries choose an initial segment of the population to represent the poor, e.g. households occupying the lowest 30 percent in the nominal per capita income distribution. The food basket chosen as basis for estimating food poverty and the food poverty line (fpl) as well as the upward adjustment of the latter into a total poverty line (tpl), are determined from income and expenditure survey data from this reference population. These steps are generally implemented at the sub-national level, e.g. regions or urban-rural areas, and the resulting estimates are combined consequently to form national estimates. Non-uniform implementation of the reference population cut-off to the sub-national domains can cause loss of comparability.

Suppose that a country chooses a 30 percent per capita income cut-off point that translates into $\mathbf{c}$ national currency units; i.e. $F(\mathbf{c}) = 0.30$. Suppose further that the rural and urban areas are the domains of interest. Some countries implement this by using the lowest 30 percent in the marginal per capita income distribution in each domain as reference population. That is, instead of $\mathbf{c}$, the procedure actually uses different cut-offs x for the rural domain and y for the urban domain, where $F_r(x) = 0.30$ and $F_u(y) = 0.30$ respectively. This corresponds to a 30 percent national cut-off if and only if the per capita income distributions in the two domains are the same. However, if there are proportionately more poor households in the rural areas than in the urban areas, then getting a uniform proportion will result in under-representation of the rural poor on the one hand, and over-representation of the urban poor on the other hand, in the overall sample. This will lead to non-comparability of the estimates in the sense that the cut-off point x is not observed in either of the two domains. (Similarly, biases could creep in the estimates due to discrepancies in the weights, food baskets and poverty lines between the intended and the sampled reference poor populations.) The correct procedure is to be guided by $\mathbf{c}$ in both domains, i.e. consider households that fall below $F_r(\mathbf{c})$ and $F_u(\mathbf{c})$, respectively. Another way of saying this is: determine the reference population nationally (Balisacan, 2001).

Shifting Domain Boundaries

This is best illustrated with rural-urban time series that most countries and international agencies reflect in their statistical compilations. A country's definition of an urban area is usually based on population density and/or urbanity indicators such as presence of road grid, theater, market, or health clinic. The same definition is used to reclassify villages after every population and housing census. The consequence is a shifting rural-urban map; hence statistical time series based on this map are neither continuous nor comparable. People migrate to where they think there are better economic opportunities, thereby raising population density, creating demands for more housing, health care, market access, etc. – and these processes tend to lower poverty incidence and shift the area towards the urban domain. Hence, the persistent poverty incidence gap between the urban and rural areas in most developing countries defies simple explanation. At the very least, the gap alone would not serve as a good performance indicator for programs designed to close the rural-urban poverty incidence gap.

The Philippines provides an example of the problems in interpreting rural-urban statistical time series. According to official statistics, (see Table 2, borrowed from David and Maligalig, 2002), the proportion of the urban population increased by 11 percentage points between the census years 1980 and 1990. This was brought about by an increase in the number of urban villages (barangays) from 7,700 to 10,200; given the very small probability of an urban village changing into a rural village, the increase must be accounted for by reclassifications of originally rural villages into urban villages. Conversely, although the total number of villages increased by 1,100, the number of rural villages declined by 1, 400, from 32,500 to 31,100.

Table 2. Population and Number of Barangays in 1980 and 1990, Philippines.

	1980	1990
Population (million)	48.1	60.7
Of which: Urban (million)	18.0	29.4
Urban (%)	37.5	48.5
Barangays (thousand)	40.2	41.3
Of which: Urban (thousand)	7.7	10.2
Urban (%)	19.2	24.8

Table 3. Poverty Incidences, 1985 – 1997, Philippines.

	1985	1988	1991	1994	1997
Poverty Incidence (%)	49	50	45	41	37
Urban (%)	38	34	36	28	22
Rural (%)	56	52	55	53	51

Poverty statistics are updated every three years in the Philippines (Table 3). Rural-urban dichotomies before 1990 follow the classification based on the 1980 census, while those from 1991 follow reclassifications based on the 1990 census. The results are not strictly comparable, would not represent a continuous time series, and could defy rational interpretation. Consider, in particular, the trends between 1988 and 1991 that seem arithmetically paradoxical: while the national poverty incidence appears to have declined by 5 percentage points, the urban and rural areas on the other hand both showed increased poverty incidences of 2 and 3 percentage points respectively. The usefulness of such statistics for poverty alleviation policy formulation, monitoring and evaluation are severely, perhaps unnecessarily constrained.[2] One way to rectify the situation is to continue producing parallel series for some period, one based on the previous and another on the more recent the rural-urban geographic classifications.

Users and producers of poverty statistics are often not aware of these sources or causes of non-comparability. The Committee members may wish to reflect and comment on the situations in their respective countries.

Non-Comparability of Statistics from Different Sources

In most countries the acknowledged source of official quantitative poverty measures is the national statistical system (SS). This takes on a number of arrangements. For example, the Bangladesh Bureau of Statistics (BBS) does all the required data collection and computation, and the Ministry of Planning (where the BBS belongs) vets and releases the results. In Indonesia that has a more centralized statistical system, the Central Bureau of Statistics (CBS) does practically all, including the release/publication of the results. In Malaysia, the Department of Statistics (DOS) is responsible for the surveys and production of the poverty statistics, but approval for public release is by the Economic Planning Unit (EPU) of the Prime Minister's

[2] The same comments clearly apply to other rural-urban statistical time series, as well as to those based on other sub-national domains with non-stable boundaries.

Office. In the Philippines' decentralized SS, a committee in the National Statistical Coordination Board (NSCB) recommends the official methodology for generating poverty statistics, the National Statistics Office (NSO) does the basic data collection and required calculations, and the results are publicly released by the NSCB. In Thailand's also decentralized SS, the biennial income and expenditure surveys are conducted by the NSO and passes on the cleaned basic data to the National Economic and Social Development Council (NESDB) for further processing, analysis and release of the official poverty reports.

Practically all ESCAP countries generate national and rural-urban poverty statistics; most have regional or states breakdowns; and all share a thus far unattained objective to bring the disaggregation to the next smaller administrative areas (i.e. provinces). Attempts to produce/release useful provincial estimates by some have been foiled by high sampling errors (there have not been that many documented studies on the level of non-sampling errors). The current household sample sizes at the provinces are too small to reduce the sampling errors to acceptable levels. While adopting more efficient sampling strategies is highly recommended and should provide partial relief, this should be accompanied by an increase in small area sample size (for the sampling error is of the order $n^{-1/2}$, not n^{-1}). External technical assistance could help in the former. The latter is more durably solved through an increase in the national SS budgets, which, however, is admittedly difficult given the relatively low status of the SSs in their respective government hierarchies and the financial constraints many governments presently face.

Poverty reduction (PR) program planners and implementers are the main users of poverty statistics. They demand small area statistics, often lower than province – and at more frequent intervals than normally provided by the national SS. For example, China's Office of the Leading Group for Poverty Reduction (OLGPR) has announced a plan to gradually over the next 10 years shift the PR program from one that has been county-based, to one that directly addresses PR in the townships and villages (David and Wang, 2001). From the program in which the central government allocates PR funds to designated poor counties, funds will be allocated directly to townships and villages, but with county governments being held responsible for the administration of the programs in their respective jurisdictions. This means that in addition to counties, poverty statistics will be needed for the towns and villages, first to design the programs and allocate resources accordingly, and then to monitor and evaluate the programs.

Contrast this immense information requirement to the National Statistics Bureau's output comprising statistics at the national level publicly released annually, plus breakdowns for the 592 designated poor (rural) counties made available to OLGPR and other concerned government agencies on a need to know basis. Without significantly more resources and a change in policy towards more sharing of information, it would be very difficult for NSB to provide the required small area statistics. However, program formulation and implementation cannot and will not wait for statistics to become available. This will proceed in an atmosphere of scant information – and simultaneously the agencies and local governments charged with

the poverty alleviation tasks will begin filling the information gap from other sources and quicker methods. This scenario is repeated in many other countries.

In Indonesia, the National Family Planning Board seems convinced of the need for data from every household to determine which is entitled to poverty alleviation assistance, and by how much; thus, the Board conducts (or attempts) an annual complete census to estimate the welfare level of each household. A similar complete household census is carried out annually by Vietnam's Ministry of Labor, Invalids, and Social Affairs (MOLISA). Thailand's Ministry of Interior conducts a Basic Minimum Needs Survey to compile district level poverty indicators. Cambodia and Lao PDR have conducted participatory poverty assessment surveys that are actually district/village censuses with the aim to construct so-called poverty maps using an array of indicators, including some kind of vulnerability index for every district or village.

These alternative sources of small area statistics generally use data collection methods different from those employed in the national statistical offices, viz. administrative reports, group interviews by agency staff or local officials, interviews of key informants, or self-administered questionnaires. Concepts and definitions, e.g. of sampling units such as a farm household or village, of income and of consumption, can be very different as well. In general, the indicators from these data sources are perceived to be of lesser accuracy, which when aggregated at the regional and national levels lead to numbers that are very different from the indicators from the national SS surveys. This multiplicity of non-comparable and conflicting indicators complicates the monitoring and evaluation of the poverty situation. A compromise solution would be not to carry the estimation from these alternative sources up to the domains where the national SS provides official estimates[3]. This acknowledges the fact that these sources are meant to supplement, not compete against or replace, the traditional sample surveys and censuses. Their growing popularity, particularly to project specialists in donor agencies and recipient countries alike, derives from their being cheaper, quicker, and more easily understood even by a layperson. They do not, for example, require much background on survey sampling and statistical estimation techniques; and they generally do not bother with sampling errors.

The main weakness of these methods may be the quality of the basic data obtained via self-administered questionnaires; administrative records, at times including project performance reports; and key informants. Their spreading popularity may have preceded the much needed research on the quality of the basic data generated from their non-traditional sources, of which very little seems to have been done and documented. We relate briefly here our experience on work still in progress in the Philippines, to cite some specific data quality concerns. Occasionally, the Philippines' Bureau of Agricultural Statistics (BAS) conducts a village census[4]. In this census, the

[3] There is a need, however, to test/verify coherence of small area estimates with the official ones from the national SS. This could be done by a special study that undertakes the estimation and comparison at the same small area level. The comparison should emphasize on coherence of trends and relative rankings of the small areas, rather than on the absolute levels of the estimates.

[4] A village is called a Barangay and the census is called Barangay Screening Survey (BSS). In the past, the BSS data were used mainly to update the agricultural survey sampling frames. However, the most recent BSS of 1998-2000 was used also to estimate certain quantitative parameters that were not

data collectors are the BAS field staff and the respondents in every village are "key informants", viz. the village chief, elementary school principal, chairpersons of the agriculture and fisheries committees in the village council, and traders. The choice of key informants in a village was left to the BAS field staff assigned to do the interview. The questionnaire included village level demographic items (e.g. population and number of households, of which farming or fishing), land use (e.g. crop areas, aquaculture areas), number of fruit trees by species, crop yields per hectare, and number of livestock and poultry by species. Some of this information was used to design the quarterly agricultural surveys that have about 10,000 sample villages. The households in these 10,000 villages were completely enumerated to provide household sampling frames.[5] This step is very close to a regular census operation in which all the households are enumerated completely by face-to-face interview. Thus, the totals (averages) from these sample villages can be compared with the totals (averages) from the village census (for the same variables), to assess the quality of responses one gets from key informants. The study also included comparison with the 2000 census of population and housing results (village totals of demographic characteristics) and those of the 1991 census of agriculture. The study had additional objectives, namely: to find out if a more reliable, complete and up-to-date sampling frame for the next agriculture sample census could be concocted from an amalgam of these four data sets; and, in a fit of perhaps wishful thinking, to see if BAS' village census could take the place of the next agriculture census.

Unfortunately, initial results of the study were very critical of the quality of data from the village census (key informants). Using key informants responses directly for estimation is out of the question; moreover, the results thus far put into serious question their usefulness even as sampling frame material. Differences in concepts and definitions (e.g. to a key informant a household and village could be very different from a survey household and a survey village, respectively) as well as coverage (e.g. the key informants tended to include in the village total the fruit trees on the road side that otherwise would not be covered in a household enumeration) were cited as major sources of the discrepancies that were found. Perhaps more importantly, these types of data collection tend to have overly ambitious questionnaires, thus including items that will require the key informants to hazard guesses (e.g. village inventories of livestock, areas under specific crops, and yields of the crops). As mentioned, this study is in progress; but experience from it thus far points to the need for more investigations into the quality of data from participatory poverty surveys and others that make use of key informants.

Moreover, the massive and fast growing time series databases that are utilized sub-optimally – often once only by the funding or collecting agency - represent investments not well spent. The irony is that, on the one hand the former centrally planned countries are being encouraged to adopt modern sample survey methods in favor of their so-called administrative reporting systems; on the other hand, market economies are being asked to do village censuses, participatory surveys, and other

---

covered in the regular census of agriculture, e.g. crop output and yields in the 40,000 barangays and 1,600 towns.

[5] The operation is called Household Screening Survey (HSS). The totals (averages) obtained by summing (averaging) over all the households in the sample village, has zero sampling error.

data capture methods that are not that much different from administrative reporting systems. In the process, there is a risk not only of proliferation of conflicting, non-comparable data from varied sources, but also of transplanting a "data-rich but information poor" culture.

There is much to recommend efforts of exploiting existing data collections and databases rather than starting new ones. The construction of small area poverty maps by combining population and housing census data with household income and expenditure survey data is one good example that eschews carrying out a new village census. Other sources yet to be tapped for another purpose is updated household lists in the sample villages. These could be linked with the most recent censuses through a household level econometric model, for instance to predict probable income levels, which could turn out useful in generating small area poverty measures or maps. Alternatively, village level models could be used to predict values of poverty indicators in the set of non-sample villages.
All these require more systematic and integrated planning and designing of national data collection systems; developing databases that are more user-friendly and allow matching and linking during processing/analysis; and more producer-user sharing of these databases.

Between Country Comparability

Poverty statistics may be grouped broadly into two: those in money terms, such as poverty lines; and others, such as poverty incidences as proportions of the population or actual number of poor persons. Those in the first group are inherently not comparable, as these are in various national currencies with different purchasing powers. Converting them into a common numeraire, e.g. PPP dollars is neither easy nor cheap, as experience with the four decades old ICP program shows. Moreover, if the conversion process does not correct for the differences in the countries' methodologies for calculating the poverty lines, then complete comparability is not achieved either. Those in the second group are potentially comparable in the sense that proportions are dimension-free, and one person in a country is equivalent to one person in any other country; provided, however, that the countries employed comparable methods in deriving their respective statistics. Thus, for enhancing comparability of all types of poverty measures, improving comparability of countries' respective compilation methods should have highest payoff.

Getting countries to move towards a common methodology will be difficult, but should be pursued nevertheless over a long-term horizon.

Reference Population

For the same reason given earlier with sub-national estimates, the reference poor population ideally should be determined internationally and a common cut-off point applied to all countries for the latter's estimates to retain their comparability. This is not feasible however, because it would require the national income distributions and the cut-off point to be expressed in a common currency. In practice, countries choose a cut-off point in the neighborhood of the lowest 20 – 30 percentile in their respective

per capita household income distributions. This source of non-comparability seems unavoidable; hence near perfect comparability is a pipedream.

Specification of Minimum Dietary Requirement

Most countries accept the assumption that if the energy requirement (in kilocalories) is satisfied, then the rest of the nutrient requirements are satisfied also, and vice versa, a diet deficient in kilocalories will be lacking in protein and other nutrients also. Minimum energy requirements are specified for age x sex groups comprising the national population, an average of which (e.g. using census population counts as weights) is used for simplicity as a single energy consumption cut-off. A person whose estimated kilocalorie consumption falls below the cut-off will be considered hungry, undernourished or food poor. Some countries go as far as specifying different minimum kilocalorie requirements for rural and urban dwellers, or for different classes of workers (e.g. India). Countries in the ESCAP region use cut-offs in the 2,000 - 2,100 kilocalories per capita per day range (David, 2001); hence it can be said that country practices in this respect are very nearly comparable.

Majority of the countries construct a food basket that satisfies the prescribed energy cut- off. The items in the basket are chosen from the food consumption patterns obtained from a consumption survey of the reference population. The basket composition is kept unchanged for years; it is the poverty lines that are updated based on most recent prices data. A notable exception here is the Philippines that eschews food baskets in favor of prescribed regional one-day menus. Because the menus contain much fewer ingredients than the food basket, there are fewer items to update. However, two issues have been raised regarding menus: whether the people eat according to the menus, and whether the price of the menu multiplied by 365 days will come close to the actual food expenditure the person or family incurs in one year (David and Maligalig, 2002). The latter issue in particular has not been studied empirically; however, it could be one of the main causes of the high poverty incidence estimates in the Philippines relative to those of other countries that otherwise exhibit similar levels of economic and social development; (see Table 4, first column of numbers).

Adjustment for Essential Non-Food Requirements

There are two steps here: how to define a set of essential non-food requirements and how to incorporate their cost into the food poverty line (fpl) to arrive at the total poverty line (tpl). For step 1, two procedures are followed by the majority of countries One is from the World Bank (Ravallon, 1994) and is premised on a particular definition of what constitute essential non-food basic needs: a household whose total expenditure = fpl still has to spend for items other than food, and whatever non-food goods and services the household chooses to buy can be regarded as essential. Since no households will satisfy the equality exactly, a linear regression of share of food expenditure to total expenditure (Engel's coefficient) on log (total expenditure/fpl) is run using data from the reference poor population. The intercept of the fitted equation is an estimate of the Engel's coefficient for a household with total expenditure = fpl. Other countries use a more pragmatic approach (e.g. Philippines, Lao PDR), by

computing Engel's coefficient from households with expenditures within ±10 percentage points of fpl. It is to be expected that the two approaches could lead to different results. However, the manner that the computed Engel's coefficient is applied to fpl exerts a much more significant impact on the resulting tpl, hence on comparability of the values across countries, and between values in a country's statistical series. For instance, China before 1995 used tpl = 1.40* fpl. Application of the World Bank approach described above in 1995 led to a change to tpl = 1.17*fpl. This big reduction in the adjustment factor from 40% to 17% has led to speculations that the pre-1995 estimates are not comparable to those from and including 1995, and that the latter may underestimate of the true magnitude of poverty (Park and Wang, 2000).

Conversely, the practice in Indonesia up to the early 1990s was top have a list of essential non-food items whose total estimated cost came up to 10 percent only of fpl; hence tpl = 1.10 fpl. This was one of the major reason for the very low poverty incidence reported by the country compared to, say the Philippines that at the time showed better economic and social indicators (Asra and Virola, 1993). A later application of the World Bank regression approach resulted in food expenditure to total expenditure ratios of 0.70 – 0.75 for the reference poor population (Said and Widyanti, 2001). If adopted, and the formula tpl = fpl/(Engel Coefficient) is used, this would lead to tpl = 1.4* fpl and much higher poverty incidences.

It bears noting that, since the Engel Coefficients referred to are computed from reference poor populations, it is reasonable to expect higher and more clustered values than coefficients computed from entire populations. Those currently in use tend to be clustered in the neighborhood of two-thirds to three-fourths. Moreover, these should be stable or change very slowly, as countries may tend to keep the same definitions for their reference poor populations.

Income or Expenditure?

This is another cause of non-comparability. Countries are split between per capita household income (e.g. China, Philippines) and consumption (expenditure) distributions (e.g. India, Indonesia) against which fpl and tpl are compared to determine poverty incidences and counts. Pros and cons abound: expenditure is broader and consumption smoothing through borrowings, use of savings and other transfers makes the poverty indicators more meaningful; and expenditure can be more accurately obtained. On the other hand, it has been pointed out that much fewer questions are needed to capture household income; the hard-to-get and inaccurate reputation of income is mainly from the upper income brackets; it has been argued that the income of salaried and lower income households could be obtained as accurately if not more so than expenditure; and estimating a truncated income distribution (excluding the upper percentiles) should be adequate for estimating and analyzing the extent of poverty.

What is lacking and is critically needed is empirical research comparing poverty statistics using income on the one hand, and consumption on the other hand, keeping other factors more or less constant.

Data Capture Methods

There is almost as much variation here as there are countries. Looking at household income and expenditure surveys only, the methods of data collection range from so-called "diary plus visiting" in China, where there is daily self-recording by the sample households followed by regular visits by NBS staff to check on the diaries, to a one year round of monthly interviews each time on one-twelfth of the sample (e.g. Bangladesh, Thailand), to revisiting the same sample households after six months (Philippines) so that the sum of the two visits represent a year of reference period, to annual interviews but with varying reference periods for the variables of interest. It is nearly impossible to tell which method will work best for a group of countries, or if there is a best method, even if consideration is limited to accuracy of the basic data. Practical feasibility and relative costs are important factors in the countries choice of methods. Self-administered questionnaires like diaries will not work well when illiteracy is high, for instance.

More research is needed to help national SSs make rational choices. An excellent example of such research was done in India, where it was found that shortening the recall period for food expenses from 30 days to 7 days in the 1999-2000 consumer expenditure survey resulted in a decline in the estimated poverty incidence from 26.1% to 23.3% respectively (Government of India Press Information Bureau, 2001). This is applied research built into an ongoing survey, which, with foresight and determination by the national SS leadership, can be pursued continuously with little additional resources.

Comparing information from available independent sources is another relatively inexpensive approach to doing data quality research. An example is provided by another work in progress in the Philippines that is attempting to "triangulate" basic data and statistics from the Food and Nutrition Research Institute (FNRI) and the National Statistics Office (NSO). The former is responsible for conducting quinquennial National Nutrition Surveys that consists of four components, namely, Food Consumption Survey (FCS), Anthropometrics Survey, Clinical Survey and Biochemical Survey. In the last FCS all the food items that were cooked in one day and the uneaten portions after each meal were measured by weighing scale. The 4,000 sample households were split randomly into seven groups and assigned randomly to the days of the survey week. Apart from the limitations of a one day weighing and that the observations pertain to the household and not to the individual members, the FCS basic data come very close to the gold standard in food intake or food consumption measurement. Similarly, the three other component surveys employ more objective measurement methods than NSO's Family Income and Expenditure Survey (FIES). Aside from providing checks on the quality of food consumption data obtained by interview, a combined analysis could check the coherence or consistency of food deprivation indicators coming from all these surveys – without initiating any new data collection. Of particular interest would be comparisons of the estimates of incidence of underweight children from the Anthropometrics Survey, food poverty incidences from the FCS and FIES, and undernourishment prevalence estimates from FAO. Initial results of this ongoing study, for example, raises the possibility that for

the Filipino physique, lowering the minimum energy requirement from 2000 kilocalories per capita per day to 1700-1800 could improve the concordance of the estimates of child underweight incidence from the Anthropometrics Survey on the one hand, and food poverty incidence from both FCS and FIES.

While relatively inexpensive, there are practical difficulties in conducting empirical poverty research, particularly data quality and comparability studies requiring information beyond what is published. Access to basic data is not easy; and when made available, e.g. household or village level survey data sets, these would be at varying states of computer-readiness. The difficulty is amplified when more than one data set is to be used, e.g. multiple rounds of the same survey or surveys by different agencies. In these cases, link-match-merge problems and diversities in concepts and methods used often put a curb or at times even derail the original research objectives. *The members of the Committee may wish to comment on the situation and possibilities of similar researchers in their respective agencies and countries.*

FAO Methodology for Estimating Undernourishment Prevalence

A thorough account of the FAO methodology is given by Naiken (2002) and a critique is found in David (2002). The outputs may be labeled more aptly as indicators of the proportion of the population in a country with food supplies that provide less than a prescribed minimum level of dietary energy. The methodology predates most other related indicators, including the countries methods of estimating their food poor and absolutely poor. And it does not appear to have evolved to incorporate more recent data sources and methods developed by the countries and other international agencies. In particular, the FAO methodology is based on national level estimates of annual food supply derived from production, imports, exports, change in stocks, and supply utilization summarized in food balance sheets. The latter are compiled mostly by FAO in Rome since very few developing countries, if any, compile these on an annual basis[6]. These are expressed in individual nutrient equivalents using standard conversion tables. In contrast, the countries estimate dietary energy consumption mainly from household expenditure surveys and/or from household food consumption surveys. A country-wise comparison of the estimates from these two methods/sources should be interesting, but possible only if relevant FAO and countries databases and worksheets were made available.

Furthermore, FAO's estimate of how the total energy supply is distributed to the households or population is generally not based on household survey data, but on parametric models accompanied by a slew of assumptions. On the other hand, countries estimate household per capita food or nutrient consumption distributions from the final consumers' (i.e. households) own responses from sample surveys.

[6] These are the same sources of daily per capita energy supply and per capita protein supply for each country that FAO publishes annually. The point to note is that these are estimates of supply and not intake nor consumption.

FAO's estimates are updated annually and published in the agency's flagship publication *State of Food Insecurity in the World* (SOFI).[7] These are used to monitor progress towards eradicating hunger and have been described likewise to indicate the prevalence of undernourishment. It has been noted, however, that they are based on estimates of available food supply during the year, and not on actual food intake or on consumption. The estimates for the period 1998-2000 that are carried by SOFI 2002 are shown in Table 4 for a subset of ESCAP countries. In theory, the FAO indicator is restricted to measuring food lack and in terms of dietary energy at that, so its numerical value should be below that of an indicator of absolute poverty that encompasses both food and essential non-food lack, such as the countries' indicators of poverty incidence. However, the FAO estimate for Thailand (18%) exceeds the country's official poverty incidence (13%) by a considerable margin; and the two estimates are equal for Cambodia, which could indicate overestimation by the FAO estimate. The FAO estimate for China (9%) is also higher than the official estimate of poverty incidence (5%); however, in this case the possibility has been raised that the latter value may underestimate the true poverty incidence (Park and Wang, 2000). On the other hand, Indonesia and Myanmar, which have identical estimates of absolute poverty incidence (23%), also have identical FAO estimates of undernourishment rates that appear too low (6%). In Rome where knowledge of the ESCAP countries may not be up-to-date, an undernourishment prevalence for Thailand that is almost on the same level as Vietnam, Pakistan and Sri Lanka, and which is three-fold higher than the estimates for Indonesia and Myanmar, may seem benign or acceptable; however, to someone with intimate familiarity with these countries, the indicators can look seriously out of sync with reality. This is borne by the other indicators, such prevalence of underweight children (Table 4), where the estimate for Thailand (19%) is significantly lower than those for the above-mentioned countries.

It will be interesting to hear from the Committee members whether the FAO estimates are used in their respective countries, and if so, in what specific context or applications? Have they been aware of some of the discordance between the national and FAO indicators such as those described above, and how have they dealt with them?

The World Bank's $1 a Day Indicator

Much has been written, said and made use of this indicator since it was proposed in 1990. The methodology involved converting the countries' poverty lines and mean per capita private consumption to a common currency using purchasing power parity (PPP) indexes based on consumption data in 1985 prices. From the ordinary least squares regression fit of the logarithm of the converted poverty lines on a quadratic function of the converted per capita private consumption, it was observed that a group of the poorer African and Asian countries clustered around an estimated $31/capita/month poverty line – or **$1/capita/day.** The area below $1 on a country's Lorenz curve of per capita daily income likewise adjusted to 1985 prices, was therefore used to estimate of the proportion of persons with incomes under $1/day. All these calculations were done, and are still being done, at the World Bank. They

[7] Curiously, the World Bank's MDG web site reads "data not yet available" for this indicator. It would be interesting to know the reason for this given that SOFI is now on its fifth edition.

cannot be done by any one developing country because other countries' poverty lines, income distribution data, and 1985 PPP indexes would not be accessible to it; and much revisions with the methodology must have gone into the later estimates, the details of which would be known only at the World Bank; hence the so-called 'black box' phenomenon.

The last World ICP or PPP price surveys were carried out in 1993. Thus, the $1/capita/day estimates were on 1985 base for some years after 1993; moving the estimates forward meant significant amounts of interpolation to countries that did not participate in the 1985 PPP surveys and/or did not have updated poverty lines or income distribution data, as well as extrapolations from one year to the next. The current series are now on 1993 PPP indexes. The 1993 ICP price surveys in Asia were acknowledged to suffer from rather serious data quality problems, which could partly explain why only two of the seven Southeast Asian countries listed in Table 4 have $1 a day entries. There is an ongoing initiative to launch a 2003 World ICP price survey, likewise managed at the World Bank. Even if successful, it would still mean that the $1 a day estimates will remain on 1993 base for a few more years. And until their calculations are shifted to a newer base, their quality cannot but continue to deteriorate.

In all fairness, the World Bank has made it plain that its country level estimates are not suitable for country use/analysis, but are intended as inputs to producing comparable regional and global estimates. Unfortunately, putting the country estimates in the public domain invites use, including comparisons at the country level.

Who use the WB $1 a day estimates? Do the individual countries use them? In what way? The Committee members may wish to inform the meeting of the experience in their respective countries. These are not idle questions because, as in the case of national accounts in local currency vis-à-vis internationally comparable PPP GDPs, the countries have their own-produced poverty statistics that are available also at sub-national levels. The amount of usage also should influence statistical system priorities and the allocation of statistical resources.

Estimating Distribution Functions from Survey Data

One of the three designated road map indicators for MDG 1, target 1 is the share of the poorest quintile in national consumption (or income, whichever is available or deemed more reliable). For the countries in Table 4, the indicator has values ranging from 5 percent to 10 percent. The MDG literature is not clear about how these values are to be used; i.e. the monitoring process does not state by how much or at what rate the values should increase from 1990 to 2015, or whether there is a target value for all countries to reach by 2015. It is not clear either whether the sources for the indicators will be the countries or an international agency (World Bank). This is academic if the latter compilations (in *World Development Indicators*) match those of the countries, which needs verifying. If they do not, such as if the World Bank does some reprocessing or manipulation of country data, then comparability would be an issue worth investigating.

Quintiles estimation imply estimation of the cumulative distribution function of income or consumption, which is also required in the calculation of the World Bank's one dollar a day indicator and the countries' absolute poverty indicators. FAO's indicator of undernourishment prevalence requires the estimation of the distribution of per capita dietary energy supply or consumption.

Distribution functions are estimated parametrically through models, e.g. lognormal, gamma (which is what is done in FAO). Empirical cumulative frequency distributions are non-parametric or model-free estimates. The latter are more common and almost routinely constructed during the processing/analysis of household sample surveys. In doing this, the prevailing practice may be to assign the same weight to each sample household. Intuitively, the result would be reasonable if the sample is self-weighting, i.e. the households in the target population all have equal probability of inclusion (ð) to the sample. This can be seen in the following representation.

Let $\Delta(a_i) = 1$ if $a_i \geq 0$

$= 0$ if $a_i < 0$

In a finite population U the cumulative distribution of a random variable, say x, is

$$F(t) = N^{-1} \sum_{i \in U} \Delta(t - x_i).$$

From an equal probability sample of size n,

$$\hat{F}(t) = n^{-1} \sum_{i \in n} \Delta(t - x_i)$$

is the unweighted estimate of the cumulative distribution.

However, majority of large scale sample surveys will have varying inclusion probabilities $ð_i$; i = 1, 2, …, n. In these cases, the unweighted cumulative frequency distribution from the sample will be biased. A design-unbiased estimator exists, which is a Horvitz-Thompson estimator applied to estimating cumulative distributions (Chambers and Dunstan,1986).

$$\hat{F}(t) = \sum_{i \in n} ð_i^{-1} \Delta(t - x_i) / \sum_{i \in n} ð_i^{-1}$$

In practice, the t values may be calculated for $\hat{F}$ values of interest , e.g. $\hat{F}(t) = 0.20$, …, $\hat{F}(t) = 0.80$ for quintiles, and the points connected to draw $\hat{F}$ in its entirety. Design-based estimators that utilize auxiliary information that parallel ratio-and regression-type estimators for point parameters have been developed also (Rao, Kovar and Mantel, 1990).

It will be interesting to hear from the Committee members if this design-based estimator has been used in their countries, or whether the unweighted cumulative frequency distribution is ordinarily used for all major surveys.

There is a need for empirical studies to compare the unweighted (biased) and and design-based (unbiased) estimators by calculating both from survey data and by using Monte-Carlo simulations. Such studies are needed to find out in what situations unweighted cumulative frequency distributions may or may not be recommended.

Main Conclusions and Recommendation

The WB one dollar a day indicator is inextricably linked to the International Comparison Program (ICP) which continues to have implementation problems. By WB's own caveat, the country level values are meant to build regional and global estimates and are not suitable for use in individual country analysis. By extension, their use to compare the poverty incidences between two countries should be approached with caution. There are no sub-national dollar a day estimates.

FAO's approach is commendable in eschewing money units and relying on energy units instead. If countries use the same kilocalorie threshold to distinguish the undernourished (food poor) from those that are not, then, other things being the same, the estimated proportions of undernourished are comparable and the estimated numbers of undernourished persons can simply be added up. However, FAO's continued reliance on energy supply derived from national food balance sheets instead of energy consumption estimated from household sample surveys, results in lack of coherence between the agency's estimates and those of the countries. Since there are no sub-national food balance sheets compilations in general, the FAO methodology cannot produce estimates at these levels. And the continued production of the FAO estimates in Rome does not engender the countries' collaboration or use of the indicators.

The countries will continue to produce poverty statistics that serve their individual needs. These statistics will understandably not be comparable because the needs vary, and because the managerial, technical and material resources available to the national statistical systems lead to deviations in the choice of methodologies. Nevertheless, there is a need for inter-country comparable statistics that will allow monitoring and evaluation of the poverty situation at regional and global levels. Ideally, the same statistics produced by the countries at sub-national and national levels could be improved to also serve the need at the regional and global levels.

To be acceptable (by the countries), hence durable, an international agenda to improve the comparability of poverty statistics should be woven into the existing programs of the countries. The agenda should aim at a methodology that obeys three rules: (1) it will not impose a heavy add-on data collection burden to the countries; (2) each country can compile and update the indicators from its own data, i.e. it will not be dependent on other countries' or international agencies' data or inputs; and (3) the resulting indicators are useful at sub-national, national and global levels of monitoring and analysis. Moreover, any proposed agenda should acknowledge that complete comparability is a pipedream. The aim really is proximate comparability and the agenda would be a process of iteratively getting nearer that aim.

An earlier note also commissioned by ESCAP (David, 2001) outlined a methodology that satisfies the above three rules. It is being restated here, with some refinements. At present, each country determines a threshold per capita daily kilocalorie consumption below which a person is labeled undernourished or food poor. It is suggested that the energy (i.e. daily per capita kilocalorie) consumption distribution be estimated from a food consumption survey or household expenditure survey, from which the observed proportion that consumes less than the designated threshold be used as direct estimate of undernourishment prevalence. The corresponding number of persons is the estimated number of undernourished or food poor persons in the population. The estimation can be carried separately for sub-national domains of interest. Since a calorie is a calorie and a person is a person regardless of food source or location respectively, the sub-national estimates are additive (comparable) and could be aggregated in a straightforward manner to arrive at bigger area indicators, including national level indicators. If countries adopt different energy thresholds, then their estimates would lose comparability. However, provided the empirical energy consumption distributions are available from each country, the proportions below any choice of a common threshold can be easily calculated, and the results become comparable.

To estimate absolute poverty indicators (or extreme poverty in the terminology of the MDGs), the current country practice is to estimate first the total cost of a food basket selected to provide the energy threshold; the result is a food poverty line (fpl). Although expressed in local currencies, fpls may be viewed as comparable in the sense that they represent costs needed to provide more or less **the same energy amount** in the diet. Secondly, fpl is adjusted to a total poverty line (tpl) by inflating with Engel's coefficient, i.e. dividing by (food expenditure/total expenditure) derived from a reference poor population with incomes or expenditures clustered near fpl. This second step can cause a major loss in comparability. It has been observed empirically, however, that the Engel's coefficients used in the majority of ESCAP countries have modal value around 2/3 (David and Maligalig, 2002). To circumvent the loss in comparability, it is suggested that this second step be simplified by using 2/3 uniformly; i.e. tpl = 1.5* fpl. F(tpl), where F is the estimated distribution of per capita income (or consumption), estimates absolute poverty incidence. These would be roughly comparable across countries in the sense that they are based on the same energy threshold and the same Engel's coefficient.

It is to be noted that absolute poverty measured this way is akin to extreme poverty referred to in the MDGs in the sense that both are inclusive of food and non-food basic needs.

Last but not least, it is recommended that the energy distribution and income (consumption) distribution required in the above proposed methodology be calculated using design-based estimators, as previously discussed.

References

ADB. *Key Indicators 2002*, Manila.

Asra, A. and R. Virola. ADB Consultancy Report, 1993 (unpublished).

Balisacan, A. Is what we know about the poor robust? Presented at the ADB Asia and Pacific Forum on Poverty, 5-9 February 2001, Manila.

Chambers, R. L. and R. Dunstan .(1986). Estimating distribution functions from survey data. *Biometrika,* **73,** pp.597-604.

FAO. *State of Food Insecurity in the World 2002. Rome.*

David, I. P. Discussion of "FAO Methodology for estimating the prevalence of undernourishment" by L. Naiken. In Proceedings of International Symposium on Measurement and Assessment of Food Deprivation and Under-Nutrition, 26-28 June 2002 (Forthcoming).

_____. Issues and recommendations for improving poverty statistics. Presented at the UN-ESCAP Working Group of Statistical Experts Meeting, 27-30 November 2001.

_____ and Maligalig, (2002) D. S. Issues in estimating the poverty line. Forthcoming.

_____. and Sangui Wang. (2001). Establishing a national FIVIMS in China. Presented at a National Seminar on Food Insecurity and Vulnerability Information and Mapping System, 21-22 November 2001, Beijing.

Naiken, L. FAO Methodology for estimating the prevalence of undernourishment. In Proceedings of an International Scientific Symposium on Measurement and Assessment of Food Deprivation and Under-Nutrition, 26-28 June 2002, Rome (Forthcoming).

Park, Albert and Sangui Wang (2000). China's poverty statistics. Forthcoming.

Rao, J.N.K, J. G. Kovar and H. J. Mantel. (1990). On estimating distribution functions and quantiles from survey data using auxiliary information. *Biometrika*, 77, 2, pp.365-75.

Ravallion, M. (1994). *Poverty Comparisons*. Harwood Academic Publishers.

Said, Ali and W. D. Widyanti (2001). The impact of economic crisis on poverty and inequality in Indonesia. Presented at the Philippine Institute of Development Studies – World Bank Institute Workshop on Strengthening Poverty Data Collection and Analysis, 30 April – 3 may 2001, Manila.

UNDP. *Human Development Indicators 2001*, New York.

World Bank. *World Development Indicators 2000/2001*, Washington, D.C.

Table 4. Source Poverty and Economic Indicators, Selected ESCAP Countries[a], 1998 - 2000.

Country	Poverty Incidence National Source	Undernourishment Incidence (FAO, %)	Population below $1 a day (%)	Lowest Quintile in Income or Consumption	Prevalence of underweight children (%)	GNP/Capita WB Atlas, $	GNP per Capita PPP	Memo: Population (Million)
China	5	9	18	10	10	780	4,110	1,250
Southeast Asia	27	13	…	…	…	…	…	487
Cambodia	36	36	…	8	52	260	1,280	13
Indonesia	23	6	15	8	34	580	2,440	209
Lao PDR	39	24	…	10	40	280	1,730	5
Myanmar	23	6	…	…	39	…	…	47
Philippines	39	23	…	5	28	1,020	3,810	74
Thailand	13	18	<2	6	19	1,960	5,600	62
Vietnam	37	18	…	8	39	370	1,760	77
South Asia	29	24	…	…	…	…	…	1,307
Bangladesh	50	35	29	9	56	370	1,480	135
India	26	24	44	8	53 [b]	450	2,150	993
Nepal	42	19	38	8	47	220	1,220	22
Pakistan	32	19	31	10	26 [b]	470	1,760	138
Sri Lanka	27	23	7	8	34	820	3,050	19

[a] Included in table are countries with more than 1 million population and with data available for first two columns; most data rounded to whole numbers or nearest 10 in case of per capita GNP

[b] Data differ from standard definition or refer to only part of a country.

Sources:

Poverty Incidence - As reported in ADB, *Key Indicators 2002*

Undernourishment Incidence - FAO, *State of Food Insecurity (SOFI) 2002*

Lowest Quintile Share in income or consumption, Population below $1 a day, GNP/capita (WB Atlas), GNP/capita (PPP) - As reported in World Bank, *World Development Indicators 2000/2001*

Prevalence of underweight children - As reported in UNDP, *Human Development Report 2001*