Studien und Materialien

NEW DEVELOPMENTS IN NATIONAL AND INTERNATIONAL POVERTY MEASUREMENT: PROMISE, LIMITS, AND APPLICABILITY FOR DIFFERENT DEVELOPMENT ACTORS

Bank aus Verantwortung

KfW

Author
Stephan Klasen, Simon Lange and Maria Carmela Lo Bue
University of Göttingen

on behalf of Business area KfW Development Bank

CONTENTS

List of Abbreviations

CBN	Cost-of-basic-needs method
CPI	Consumer Price Index
CSPI	Correlation Sensitive Poverty Index
DRM	Day Reconstruction Method
ECLAC	Economic Commission for Latin America and the Caribbean
EMA	Ecological Momentary assessment
ESM	Experience Sampling Method
FEI	Food-energy intake method
FGT	Foster, Greer and Thorbecke (1984)
GDP	Gross Domestic Product
HDI	Human Development Index
HPI	Human Poverty Index
ICP	International Price Comparison Project
ILO	International Labour Organization
IPL	International Poverty Lines
LSMS	Living Standards Measurement Surveys
MA	Mixed Approach
MDG	Millennium Development Goals
MIQ	Minimum Income Question
MPI	Multidimensional Poverty Index
MRS	Marginal Rates of Substitution
NAS	National Accounts Statistics
NSS	National Sample Survey
OLS	Ordinary Least Squares
PANAS	Positive and Negative Affect Schedule
PCA	Principal Component Analysis
PMT	Proxy Means Tests
PPP	Purchasing Power Parity
PPPP	Poverty-related Purchasing Power Parity
PRSP	Poverty Reduction Strategy Papers
RCS	Ravallion, Chen and Sangraula (2008)
RD	Relative Deprivation
RDV	Ravallion, Datt and Van de Walle (1991)

SMT   Simple Means Tests
SPL   Subjective Poverty Lines
SSI   Sustainable Society Index
SWB   Subjective Well-Being
SWL   Satisfaction with Life Scale
UNDP   United Nations Development Programme
VMT   Verified Means Tests
WDR   World Development Report

Summary

The last four decades have seen rapid changes in our understanding of poverty, both in terms of conceptual work as well as in terms of practical measurement. In fact, there are now so many new approaches that it is important to take stock and analyze which of the new approaches are particularly useful to be used for development cooperation, international and national debates about poverty. In this paper we review and critically appraise approaches to poverty measurement in different settings and for different purposes, evaluate the relevance of these approaches to different actors, and point out some directions for further research.

At the most general level, poverty is defined as the situation in which a minimal level of well-being is not attained. Hence, the analysis of poverty focuses on the distribution of some well-being indicator below a cut-off – the poverty line. Poverty analysis therefore requires two steps: the identification of the poor (i.e. the stipulation of some indicator of well-being and the definition of a poverty line) and the aggregation of information on shortfalls.

How exactly the concept of well-being should be defined is subject to debate. On the one hand, poverty has been defined in income terms which, under strong assumptions, can be linked to standard welfare economics. We review several methods of defining a national income poverty line, including methods that link the poverty line to nutritional inadequacy and subjective poverty lines. We argue in the paper that the cost-of-basic-needs Approach to deriving an income poverty line is most promising for most applications for national governments as well as development cooperation actors. At the same time, critical conceptual and empirical questions regarding the updating of these lines and the consistency and comparability of these lines need to be addressed.

We also review the international income poverty line ($1.25 a day), both in its absolute version (the current MDG indicator) as well as the proposed 'weakly relative' version. While we favor the 'weakly relative' version as the more relevant and promising one, we generally find that the setting and updating of the poverty line is beset with conceptual and empirical problems linked to changing purchasing power parities and uncertainties in the estimation of the line. We therefore propose that, in the medium term, it would be useful to replace this approach of generating an international poverty line with internationally coordinated national income poverty lines based on the cost-of basic-needs approach that are then aggregated at the global level to measure levels and trends of poverty in the developing world. We also suggest that for project monitoring or targeting, proxy indicators of such a national poverty line are likely to be particularly useful.

We further consider measures that are grounded in non-income notions of well-being. The leading example for these theories is Sen's *capability approach*, which focuses on individuals' freedoms to attain certain *functionings*, e.g. to lead a healthy life, to be educated, or to participate in society. These approaches naturally lead to a multi-dimensional conceptualization of well-being and poverty, even though agreement on a set of relevant capabilities (and their weights) remains controversial. Consequently, there are now a myriad of well-being measures – both aggregated and disaggregated –,

the most well-known of which are UNDP's Human Development Index, the Millennium Development Goals, as well as the recently proposed Multidimensional Poverty Index.

We discuss several approaches to multi-dimensional poverty measurement – both axiomatic and *ad hoc* – that have been suggested in the literature. In practice, the main questions are how to choose dimensions and respective cut-offs and how to decide about the weighing of different dimensions. We argue that dominance approaches is one possible way of dealing with the sensitivity of different approaches to these questions, but they are of limited relevance for policy-makers and development practicioners. Instead we argue that multi-dimensional poverty measures are clearly in need of transparent processes and public discussion to generate broadly accepted measures. Such debates and processes are most usefully taking place at the national level which will inevitably lead to different measures across countries.

Among the practical implementations of multidimensional approaches, the Multidimensional Poverty Index (MPI) has recently gained prominence. We argue that while this measure has several attractive features – it is fully decomposable, can be adjusted to country-specific needs, and also allows, in principle, for cross-country comparisons – there are a range of conceptual and empirical problems that still need to be addressed before this measure can become a useful and robust alternative for international poverty measurement. Among the conceptual issues to be addressed are ways to deal with the depth of poverty within dimensions and in the number of dimensions, while there are a range of empirical issues to do with the dimensions used, sample selection, and cut-offs that need to be carefully scrutinized.

We also review the recent and rapidly growing literature on subjective well-being and investigate to what extent they can be adapted for poverty measurement. While these approaches may serve as a useful complement to objective approaches of well-being measurement, we argue that at this stage they have not been adapted to an analysis of poverty, that doing so requires a great deal of conceptual and empirical work, and that it is unclear whether eventually these subjective approaches will lead to reliable, comparable, and consistent poverty measures that are useful for policy-makers and development cooperation.

Lastly, we summarize our results by judging the full range of poverty measures by a large set of criteria to judge their relative advantages and disadvantages. We first suggest that these advantages and disadvantages differ greatly between the measures so that it is useful to not zero in on a single approach but consider a dashboard of approaches that are suitable for different circumstances and actors. In this context, it is critical to recognize that different actors will need very different approaches, particularly also with regard to the level of disaggregation, timeliness and costs; this calls for the co-existence of comprehensive and short-cut approaches that are closely related to these comprehensive approaches. For further analytical work, we suggest, however, that refining the cost-of-basic-needs approach to national and international income poverty measurement as well as working on improvements in the MPI is particularly promising.

1 Introduction

The last four decades have seen rapid changes in our understanding of well-being and its opposite, poverty, both in terms of conceptual work as well as in terms of practical measurement. This has left us with a plethora of different concepts and indicators: is poverty best understood as insufficiency of income? Is it about access to basic services like health and education? Or is it really about basic capabilities such as being able to participate in society? What role does relative deprivation play? How reliable a guide are individuals' own perceptions of their standard of living? How important are dynamic considerations, i.e. the risk individuals face to become poor in the near future, or the long-term consequences of current modes of production and consumption on the capability of future generations to avoid poverty? Even if a consensus is established concerning the conceptual space in which well-being is to be assessed (e.g. incomes, capabilities, or primary goods), what should the actual indicator be? Lack of income might not be the only criterion by which to judge an individual's poverty status, but it could well be a useful and readily-available proxy. Put more generally, how should our understanding of poverty be operationalized in terms of indicators to assess poverty within a given population and guide policies?

These are but a few of the questions one could ask with respect to the definition and measurement of poverty. While some of these questions have been discussed for a long time (some even already in the 18th century as the discussion of the Poor Laws in England demonstrate, see Hirschman, 1991) they have gained new relevance for several reasons. First, there has been substantial progress particularly regarding conceptual work in multidimensional poverty measurement. Secondly, there is a lot more data available to be considered for poverty measures, including better income and expenditure surveys, demographic and health surveys, and information on subjective well-being, perceptions, and attitudes. Thus there are many more options now available to derive poverty measures. Lastly, poverty measurement seems to matter more than before. Many national governments, many donors, and the international community at large have made poverty reduction a key goal of government policy and international cooperation. Policy is often directly linked to particularly indicators, and sometimes involves targeted transfer to the 'poor', and thus poverty measurement matters more than before for policy.

The recent changes come after substantial shifts in the thinking about development over the past 50 years. While initially, development and per-capita economic growth were seen as largely synonymous, broader conceptions of well-being and deprivation began to take root already in the 1970s. Important steps along the way are the emphasis on basic needs in the 1970s emphasized by Streeten and others, and partly an outgrowth of work at the ILO. In the late 1980s, the Human Development Paradigm was developed by Mahbub up Haq, leading to the first Human Development Report in 1990, including also the Human Development Index (UNDP, 1990). Related to this development was the formulation of the capability approach by Amartya Sen in the 1980s which has revolutionized thinking about well-being in poverty (Sen, 1985). These two strands of the literature are important antecedents to the recent publication of the Multidimensional Poverty Index in UNDP's 2010 Human Development Report (UNDP, 2010).

While these approaches were largely focused on the measurement of well-being, the World Bank pursued a slightly different agenda. While also moving away from per-capita income as the key goal of development, it refocused its efforts to the reduction of *income poverty* across the developing world. The World Development Report 1990 is the key report signaling this change (World Bank, 1990). Using the new tool of Living Standards Measurement Surveys, income poverty measurement, analysis, and anti-poverty policy at the national and global level became a central focus of the Bank's work. The $1-a-day measure was first introduced in that WDR and became a central indicator for the MDGs 10 years later.

Poverty Reduction Strategy Papers (PRSPs) that were introduced alongside debt relief initiatives, and the formulation of the MDGs were two central tools that put (multidimensional) poverty reduction at the center of the development agenda. As the global community is now discussing how to shape a post-2015 agenda, the question of how to measure and analyze poverty is a critical question.

In this paper, we review current approaches to poverty measurement to contribute to this agenda. The paper not only seeks to review the state of the art in poverty measurement, but also emphasize the usefulness of different approaches to different users of poverty data. We start with theoretical considerations and then first discuss various unidimensional measurement approaches before moving to multidimensional approaches.

A central theme in the paper is that by now there are mature conceptual and empirical approaches to the measurement of poverty available. However, each of the measures discussed has particular advantages and disadvantages so that we will not recommend a unique measure. Instead, the choice of measure should be guided by the purpose of the exercise. In this sense our last section is of particular relevance as it highlights the strengths and weaknesses of particular approaches.

2 Theoretical Considerations

2.1 Conceptual Approaches to Poverty and Well-being

An individual is said to be poor if she does not attain a minimal level of well-being. Hence, poverty and well-being are two sides of the same coin and whatever theories we have of well-being may also be brought to bear on the analysis of poverty. With respect to different theories of well-being a few general distinctions are in order.

First, what distinguishes poverty from overall well-being and distributional analysis is its focus on the lower part of the distribution. While overall well-being analysis (as exemplified by indicators such as per-capita GDP, the HDI, or mean happiness) report on the mean (or median) of a distribution, inequality measures the spread of the distribution. Poverty analysis is related to both but focuses on the size and distribution of a population in the lower part of the distribution.

Second, a useful distinction is between *welfarist* and *non-welfarist* approaches to the measurement of well-being and poverty. In welfarist analyses, the focus is exclusively on *utilities*, the preferences of individuals themselves. Welfarist analyses usually involve

further restrictive assumptions (although some of these might be relaxed in applications), most notably rational, utility-maximizing individuals, the irrelevance of the act of choosing or the options available for welfare (called 'consequentialism', i.e. only consequences not processes matter, see Sen, 1999), and the availability of all welfare relevant commodities through competitive and complete markets (or, at least, the valuation of all welfare-relevant commodities in an 'as-if-through-the-market'-manner).

Using (most of) these further assumptions, the welfarist approach leads to measures based solely on the consumption of goods and services – including publicly provided goods and services – consumed by a household, the household's size, and its demographic composition. As Ravallion (1994, p. 8) notes, this is somewhat afar from the broader concept of utility and probably better described as 'economic welfare.'

Non-welfarist approaches are often broader in scope in the sense that they focus not exclusively on command over commodities. Sen's capability approach, which has been very influential in the discussion of poverty measurement over recent years, is a non-welfarist approach that focuses not directly on command over commodities but on the capabilities of persons 'to lead the kinds of lives they value – and have reason to value' (Sen, 1999, p. 18). Capabilities are defined as the freedom to achieve certain functionings like being well-nourished, being literature, being mobile, etc. While command over certain commodities is required for some, there are certainly others such as being free from violence or being able to breathe clean air which cannot be purchased in markets (see also Sen, 1998).[1] As there are clearly a multitude of capabilities, the capabilities approach is the main motivation behind a multidimensional conception of poverty as discussed in section 4.

2.2 Identification and Aggregation

A thorough poverty analysis usually comprises two distinct problems: (i) the identification of the poor among the total population of interest and (ii) the aggregation of information on deprivations into some overall indicator (Sen, 1976, 1979). The identification step involves the choice of a cut-off criterion of welfare (appropriately measured), the poverty line, and the comparison of individual welfare levels with respect to this criterion. Much of what follows in subsequent sections will focus on identification. However, it will be useful to briefly review the literature on the aggregation problem.

Once a cut-off level of welfare is established, one can apply different poverty measures in order to aggregate information on the shortfalls of welfare of those in poverty.[2] For simplicity, we will consider the case of an income-based poverty line (possibly adjusted for personal characteristics).

Sen (1976) proposes two basic axioms that a poverty measure should satisfy:

- MONOTONICITY AXIOM: all else equal, a reduction in income of a person below the poverty line must increase the poverty measure.

---

[1] Ravallion (1994, 1998, 2012), however, argues that the capability approach can be incorporated in a welfarist setting by letting the capabilities (or functionings) depend on commodities and postulating that utility in turn depends on commodities. Hence, capabilities can be solved out. Clearly, this is in stark contrast to the notion that some capabilities cannot be achieved through the command over commodities.

[2] For a survey see Atkinson (1987) and Ravallion (1994, pp. 44-51).

- TRANSFER AXIOM: all else equal, a transfer of income from a person below the poverty line to anyone who is richer must increase the poverty measure.

The most widely used aggregation method is the headcount ratio index, which is defined simply as the proportion of the population below the cut-off level of income. Let n be the size of the population. The headcount ratio is then defined as

$$H = \frac{q}{n},$$

where q is the number of poor. However, the headcount ratio ignores information on the distribution of the shortfalls below the poverty line and, thus, violates both the monotonicity and the transfer axiom. An alternative is the *poverty gap index*. Let z be the cut-off level of welfare and $y = (y_1, y_2, \ldots, y_q)$ the incomes of the poor in ascending order. The poverty gap is then defined as:

$$PG = \frac{1}{n} \sum_{i=1}^{q} \frac{z - y_i}{z}.$$

PG gives us the mean proportionate shortfall from the poverty level (where the non-poor have zero shortfall). If we define $I = 1 - \overline{y_p}/z$, where $\overline{y_p}$ is the mean income of the poor, PG can also be written as $H \cdot I$.

While PG satisfies monotonicity, it does not satisfy the transfer principle: all else equal, if an individual just below the poverty escapes poverty, it might well happen that the average shortfall of the poor increases. Moreover, PG does not convey any information about the distribution of incomes below the poverty cut-off.

Sen (1976) suggests a poverty measures that weighs the poor's income shortfalls according to relative position in the income distribution:

$$P_S = H\left(I + k(1 - I)G_p\right).$$

where $k = q/(q + 1)$ (converging to unity for $q \to \infty$) and G_p is the Gini index for the distribution of incomes of the poor. It is easy to see that (the above expression collapses to PG if there is no inequality among the poor.

The main shortcoming of the Sen-Index is that it does not satisfy *additivity*, i.e. aggregate poverty will not be equal to the weighted sum of poverty levels in (mutually exclusive and exhaustive) sub-groups of the population. This is an attractive property of both H and PG which allows analysts to attribute changes in overall poverty to changes in poverty in different sub-groups. A distribution-sensitive poverty measure that possesses this attribute is the *squared poverty gap index* proposed by Foster et al. (1984):

$$P_2 = \frac{1}{n} \sum_{i=1}^{q} \left(\frac{z - y_i}{z}\right)^2.$$

Note that H, PG, and P_2 possess a common structure in that they can be written as

$$P_\alpha = \frac{1}{n} \sum_{i=1}^{q} \left(\frac{z - y_i}{z} \right)^\alpha,$$

for $\alpha \geq 0$. This is the Foster-Greer-Thorbecke (FGT) class of poverty measures (Foster et al., 1984) and it can be easily checked that it reduces to H for $\alpha = 0$, to PG for $\alpha = 1$, and to P_2 for $\alpha = 2$.

From a theoretical perspective, distributional-sensitive poverty measures are clearly preferable. On the other hand, the higher-order poverty measures are generally more affected by measurement error (see Ravallion, 1994, for a discussion) and they only add information whenever inequality in welfare differs between sub-groups. Therefore, the right choice also depends on the quality of the data at hand. In any case, estimating several measures and checking the resulting profile for robustness is always good practice.

As will be briefly mentioned below, added complications arise when considering aggregation in a multidimensional poverty measure. One now has to consider aggregation across dimensions in which one may be poor and across people who may be poor in different dimensions. Poverty depth might relate to the number of dimensions one is poor in, or the depth of poverty in each dimension, which may in fact be related to the depth of poverty in different dimensions. Here a vigorous debate is currently on-going on the best ways to proceed (e.g. Rippin, 2012; Alkire and Foster, 2010; Bossert, Chakravarty, and D'Ambrosio, 2010).

2.3 Absolute vs. Relative Poverty

When setting a poverty line, a key distinction is drawn between absolute and relative poverty lines, where the former is fixed in some real value of achievements (e.g. income or education) over time and space and the latter is defined relative to, say, average achievements in a time and place. For example, the European Union uses a strictly relative concept for poverty, where the poverty line corresponds to an income of 60 percent of the median disposable income within each member country. If such a strictly relative poverty line is applied, equi-proportionate growth in average incomes will not alter the incidence of poverty,[3] whereas changes in the distribution of incomes will. Hence, the notion of a relative poverty line is closely linked to income inequality.[4]

At first glance, there is much merit in the relativist view of poverty. Customs shape what is conceived as a basic necessity within a given society (Townsend, 1979).[5] Adam Smith's famous example in *The Wealth of Nations* of what was during his times an indispensable necessity to go in public without shame (a pair of leather shoes) was

---

[3] This is true regardless of whether the poverty line is defined as a proportion of mean or median income.

[4] An even more extreme approach would be to define a fixed proportion as 'the poor', say, the bottom quintile of the income distribution. In that case, poverty will be constant over time so that the success of policies directed at poverty alleviation cannot be judged based on this measure.

[5] See also the influential work by Runciman (1966) on relative deprivation (RD).

certainly not an indispensable necessity at other points in time (or, indeed, in other societies at the same point in time).[6]

On the other hand, relying solely on a strictly relative poverty line can clearly go wrong: as pointed out by Sen (1983), a natural disaster that decreases the income of every member of society but, at the same time, attenuates inequalities would result in a *decrease* in poverty – despite the fact that everybody would be worse off.

Sen's (1983) solution is to make the decision on whether to use an absolute or relative poverty line dependent on the space in which poverty is considered. He argues that if poverty was to be defined in the space of capabilities (see section 2.1), absolute deprivation in terms of a person's capabilities should matter. In the space of commodities, however, this will in many cases relate to relative deprivation. It was not a pair of leather shoes *per se* that defined the 18th century Englishman's standard of living but the capability to go in public without shame. The escape from poverty in the form of avoiding shame might require more commodities the richer the society. The avoidance of this type of shame, on the other hand, is a requirement invariant over time and space.

3 Income-based Poverty Lines

Despite the recent tendency towards the use of multi-dimensional poverty measures such as aggregate indices like the MPI or dashboards like the MDGs (see discussion below), income poverty still figures prominently in development thinking and practice. This is not surprising: even if poverty is seen in terms of deprivation of basic capabilities and it is acknowledged that, first, not all capabilities can be achieved through the command over commodities and, second, individuals differ in terms of their ability to transform income into functionings, overall, there is little doubt that income is a major correlate of well-being.

Another advantage is the ability to link income-based conceptions of poverty to a large body of both theoretical and practical work on the measurement of economic welfare. The major theoretical contributions on welfare measures include Samuelson's (1974) work on *money-metric utility* and Blackorby and Donaldson's (1987, 1988) work on *welfare ratios* (see Ravallion, 1998, and Deaton and Zaidi, 2002, for a discussion). Both methods are welfarist in nature, i.e. they start out by postulating a utility function defined over an individual's consumption of commodities that reproduces consumer's preferences over alternative consumption bundles. The poverty line can then be defined as the expenditure necessary to achieve some reference level of utility.[7]

---

[6] The question whether relative income matters has been investigated empirically – usually by regressing self-assessed life satisfaction on own and neighbors' income (see also section 6.3). Findings suggest a negative effect in rich countries (e.g. Luttmer, 2005) and no effect – or even a mildly positive effect – in developing countries (e.g. Senik, 20054; Ravallion and Lokshin, 2010).

[7] More formally, let $u = u(q,x)$ be the individual's utility function defined over alternative consumption bundles q for an individual with characteristics x. A reference level of utility, u^r, is defined as the minimum level of utility the individual should attain in order to be counted as non-poor. The poverty line, z, is then the point on the households' expenditure function, the function returning the minimum expenditure to achieve u^r at the prevailing prices p, such that $z = e(p,x,u^r)$. Thus, the poverty line is the minimum cost at which the household can achieve the reference level of utility.

However, it should be noted that the gap between theory and practice is large as these welfarist theoretical expositions rely on a set of rather restrictive assumption that we noted above. Moreover, theoretical work is mostly based on a single set of preferences represented by an indifference map. This might be justified in some circumstances: Deaton and Zaidi (2002, p. 19) note that it might not be of great interest to allow for the idiosyncrasies of each person's tastes in welfare evaluations but this clearly depends on the idiosyncrasy at hand.[8]

Even if we were ready to assume identical tastes, two problems remain that cannot be solved without broadening the informational base upon which assessments of economic welfare are based. The first problem concerns finding the right poverty line *in utility space*, i.e. what Ravallion (1998) refers to as the *referencing problem*. The choice will clearly affect the total extent of poverty and, most likely, also the distribution across subgroups. Second, even if there is agreement on the right level of the poverty line in utility space, there is no way of knowing the parameters of the expenditure function that the above-mentioned theoretical considerations are based on when individual (or household) characteristics vary (Pollak, 1991). This is what Ravallion (1998) refers to as the *identification problem* (not to be confused with Sen's (1976) use of the term discussed in section 2.2).[9]

The methods described below in section 3.2 and 3.3 can be interpreted as attempts to solve these problems by incorporating additional information.

3.1 Practical Issues: Using Household Expenditure Surveys

Before we go on to discuss different methods of establishing poverty lines we comment on some issues that arise in practice when poverty assessments are based on household survey data. Note that some of the issues discussed in what follows also pertain to multidimensional poverty measures discussed in section 4.

3.1.1 Income or Expenditure?

An important question in practice is whether one should construct welfare measures based on income or expenditure. Welfarist theory tends to stress the importance of consumption over income as the source of economic welfare.[10] However, both variables

---

[8] For example, the fact that a particularly greedy person's income might be of little value to her compared to another person might be of little concern. On the other hand, very few people would object to the notion that a disabled person needs comparatively more income to attain a given level of welfare (see Sen, 1998).

[9] Pollak (1991) demonstrates that while the observation of demand patterns of household with different characteristics allows the identification of conditional preferences, i.e. their preferences *given characteristics*, it does not allow the identification of *unconditional preferences*. If the indirect utility function supports observed demand

$$q(p, y, x) = \frac{v_p(p, y, x)}{v_y(p, y, x)},$$

then so does every other indirect utility function $V(v(p, y, x))$ since $\frac{V_p(p, y, x)}{V_y(p, y, x)} = \frac{v_p(p, y, x)}{v_y(p, y, x)}$.

[10] The answer depends on the time-horizon over which poverty is to be assessed. The difference between income and expenditure is savings. If a long enough period is considered, total income (including

14

are found in practice, with studies set in rich world-countries (and many Latin American countries) tending to rely on income and those set in most developing countries (except Latin America) on consumption.

For several reasons, the consensus is to focus on consumption in developing country-settings. While it will always be more costly to collect data on consumption, income is often difficult to measure in settings were self-employment (including self-employment in agriculture and the informal sector) is an important source of income. For example, it is often unclear how to separate private consumption from investment and assumptions have to be made about the depreciation of assets. While this is true in both developed and developing countries, income from self-employment (inside or outside of agriculture) tends to be relatively more important in developing countries, particularly among the lower part of the income distribution. Incomes also tend to be more seasonal in developing countries, because a large portion of it is generated in agriculture. While the rural poor are often found to be constrained in their ability to smooth out fluctuations, consumption is still found to be less volatile over time (Deaton and Zaidi, 2002).

3.1.2 Expenditure as a Proxy for Welfare

Intra-household Distribution and Equivalence Scales Most available surveys on household expenditure do not contain information on the distribution of consumption within households. This constitutes an important shortcoming of poverty assessments based on household expenditure surveys: what one is ultimately interested in is individual welfare but the lack of data on intra-household distribution of consumption constrains analysts to considerations of average household welfare.[11] See Haddad and Kanbur (1990) for a discussion of the likely consequences of this assumption on the accuracy of poverty estimates, and Klasen (2007) for the consequences for gendered poverty analysis. Clearly, this points to the desirability of supplementary indicators of individual attainments that can be broken down by demographic variables in order to better understand the likely differences in resource allocations within households. Examples include data on morbidity and mortality outcomes, as well as educational attainments (e.g. Klasen, 2007).

A further point is the use of appropriate *household equivalence scales*. As we are ultimately interested in individual welfare and are, in addition, forced to accept the assumption of welfare equivalence, a rough-and-ready adjustment when comparing two households of different sizes would be to divide by the number of members. However, per capita consumption might be problematic for two reasons: first, consumption requirements differ by demographic factors such as age and gender and levels of physical activity. For example, adults will in general need more consumption expenditure

---

inheritances) must equal total consumption (including bequests), so that there is no difference between the two (Deaton and Zaidi, 2002).

[11] Consider Thorbecke's (2007) 'dictatorial bread-winner' who buys mainly tobacco and alcohol for himself: "[...] the welfare functions of such households – at least as reflected by that of a dictatorial head –yield perverse outcomes in the sense that high enough incomes to potentially escape poverty are allocated instead to yield deprivations and poverty."

in order to attain the same level of economic welfare.[12] Second, some goods exhibit *positive externalities in consumption*, i.e. households can economize on their use. This is often the case for durable consumer goods, such as cars or refrigerators. That they are used by a larger number household members does not necessarily reduce their usefulness (at least not in proportion). In consequence, "the cost of being equally well-off does not rise in proportion to the number of the people in the household" (Deaton and Zaidi, 2002).

Deaton and Zaidi (2002) review three different approaches to the calculation of equivalence scales, none of which is free from criticism:

- *The Subjective Approach* One can estimate equivalence scales from questions on perceived adequacy of consumption (e.g. Kapteyn and van Praag, 1978). This method resembles to a large degree the methods to establish subjective poverty lines to be discussed in section 3.3.

- *The Behavioral Approach* One can try to estimate to cost of children (e.g. Deaton and Muellbauer, 1986) and the degree of rivalry of goods consumed (e.g. Deaton and Paxson, 1998) from observed consumption patterns to arrive at equivalence scales.

- *The Arbitrary Approach* One can choose an arbitrary but reasonable set of equivalence scales.

The first two methods are, for conceptual and econometrics reasons, not fully convincing (Deaton and Zaidi, 2002, Deaton, 1997). Most studies to date are therefore based on arbitrary equivalence scales. In any case, it is good practice to check the robustness of results to these assumptions.

Valuation of Non-market and Durable Goods Private consumption ignores the contribution of public goods to individual welfare. Clearly, to achieve a given level of economic welfare, a household requires less total expenditure when public services like schooling and health care are free of charge. An accurate measure of welfare should therefore account for publicly provided goods, the list of which is potentially long.[13] Obviously, accounting for the contribution of publicly provided goods to well-being is more important in settings where provision varies between different sub-groups of interest. For example, provision of public goods is often cheaper in urban areas. Hence, urban households tend to have more access. Comparisons between rural and urban households will therefore depend heavily on the valuation of public goods.

In practice, it is very difficult to convincingly include such considerations into a consumption aggregate. While there are some attempts at ascertaining shadow prices for public goods, it is unclear how estimates based on these methods are related to

---

[12] This becomes evident when we think primarily in terms of food expenditure. Adults have more body mass and therefore require higher levels of caloric intake.

[13] Education and health are only but two examples. Public goods also include infrastructure like water, sanitation, and roads, as well as security provided through policing, the justice system, and national defense.

actual shadow prices (see Deaton and Zaidi, 2002, and the references cited there).[14] Deaton and Zaidi (2002) seem to think that relying on these methods will, in most cases, undermine the credibility of the welfare measure constructed, but that there might be some cases in which some adjustment is warranted even when the measurement error might be large.

Another issue is the valuation of durable goods, i.e. bicycles, TV sets, or even housing. These are purchased infrequently and often command high prices, but since it is the use of these goods rather than the purchase itself that generates welfare, they introduce measurement error in estimates of total consumption expenditure as a welfare indicator. Since use is in most cases not directly observable, a practical solution is to make reasonable assumptions about the depreciation of this good or their rental value and include these in the consumption aggregate.

Lastly, there is the issue of different prices across the country. In principle, different prices ought to be taken into account and in some poverty analyses, this is done at an impeccable level of sophistication (e.g. using price data generated alongside the living standards survey, for example in a separate community questionnaire that generates prices); in many cases, prices are not available at the sub-national level, and they are difficult and often unreliable when estimated from the surveys themselves (as the surveys often only report unit values rather than prices) (Deaton and Grosh, 2000).

On a more practical side, a key constraint to the use of household surveys is the limited comparability of the survey instruments over time which makes it very different to discern poverty trends. While some countries have been running identical surveys for decades in a row (e.g. India's National Sample Survey is a prime example of such a survey, as are a number of surveys in Latin American countries), in many countries, particularly in Africa, each survey is largely done ad hoc (usually externally funded) with different questionnaires and sampling strategies. Of particular relevance in this context is the length of the recall period for the various items in the consumption survey. As shown by Deaton and Kozel (2005), a reduction in the recall period for food in India's NSS led to a substantial spurious decline in poverty in the 1999/00. Moreover, the length of the list of items included in the consumption survey also has a sizable impact on poverty estimates. The more detailed the list, the larger typically are expenditures (Deaton and Grosh, 2000; Deaton and Zaidi, 2002).

3.2 Objective Poverty Lines

When income or expenditure is used as the relevant indicator of poverty, the question of how to set a poverty line remains. This section discusses two `objective' methods of establishing a poverty line: the food energy-intake (FEI) and the cost-of-basic-needs (CBN) methods which are widely used at the country-level.[15] In practice, both methods are commonly based on a minimum level of caloric intake, i.e. one tries to estimate the minimum level of expenditure (or income) at which certain nutritional needs are met; in this sense one tries to ground the income poverty line in some 'objective' nutritional

---

[14] What is true for public goods is also true for leisure: while most likely an important component of overall well-being, it is difficult to ascertain the shadow prices of an hour spent not working.

[15] An exposition is also given by Ravallion (1994, 1998).

requirement, thereby providing some linkage between income poverty and one key capability, hunger (see de Haen et al. 2011 for a discussion). These methods are objective in the sense that subjective assessments of own poverty status are not considered. However, they do involve a fair amount of normative judgments in determining adequate levels of consumption for a given household. Alternative methods for establishing a poverty line that explicitly incorporate subjective judgments are discussed in section 3.3.

3.2.1 The Food-energy intake Method

The FEI method proceeds by finding the consumption expenditure or income level at which food-energy intake is, in expectation, just sufficient. Once the level of caloric intake deemed adequate is determined, computing the poverty line is computationally simple. For example, assume that 2,100 calories per person are seen as just adequate. One can proceed by regressing caloric intake on consumption expenditure (both adjusted for household size). The resulting regression line will, in general, be strictly increasing as illustrated in figure 1.[16] The expenditure level z is then established as the poverty line.[17]
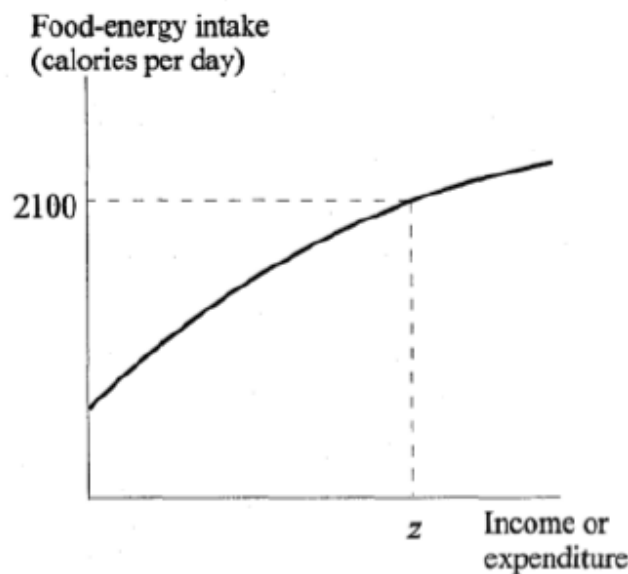
Figure 1: The food-energy intake method: in expectation, individuals above z consume more than 2,100 calories, while those below z do not. The latter are considered poor even if their actual caloric intake exceeds 2,100 calories. Source: Ravallion (1998).

Proponents of the FEI method (e.g. Osmani, 1982, Greer and Thorbecke, 1986; Paul, 1989) note two major advantages: first, it avoids the need for price data. One thus circumvents the difficult task of valuing non-market goods. All one needs is, in principle, data on total expenditure, household characteristics and total caloric intake. Second, the

---

[16] On the econometric issues involved see Bouis and Haddad (1992).

[17] Note that the FEI method is distinct from pure assessments of the extent of undernutrition: in the above example, there will be some households counted as poor despite per capita caloric intake being in excess of 2,100 calories and vice versa.

FEI poverty line has a built-in allowance for both food and non-food expenditure as long as *total expenditure* is considered in its computation.

A general problem with this approach is to choose an adequate level of food-energy intake. In practice, this is often based on an expert nutritionist's recommendations (e.g. FAO, 2001). Food energy requirements vary across adult individuals because of differences in stature and activity levels and even more variables have to be considered when it comes to determining requirements for children, adolescents, and pregnant and lactating women. Also, a possible concern here is reverse causality between the level of physical activity and poverty status. The poor may be found to be less active physically but this finding should not be taken to imply that they require less food energy. To the contrary, inactivity may be a result of insufficient intake.

It should also be noted that adequate caloric intake is not the only determinant of an adequate diet; the human body also requires nutrients like protein, iron, calcium, vitamins, etc. in certain minimum amounts (Paul, 1989). While these different requirements of the human body can at least in principle be accounted for in the CBN approach to be discussed below, there is no obvious way to do so in the FEI method.

In addition, comparisons based on the FEI poverty line are often found to be inconsistent in the sense that an increase in the command over commodities (across time or space) will not necessarily result in a corresponding decrease in poverty. The location of the regression line sketched in figure 1 depends on tastes, activity levels, relative prices, publicly-provided goods and other determinants of affluence besides consumption expenditure and there is nothing in the method to guarantee that changes in these parameters will bring about the expected change in the location of the regression line (Ravallion, 1998). This is discussed further when comparing the two approaches below.[18]

3.2.2 The Cost-of-basic-needs Method

The most common approach to setting a poverty line is the so-called cost-of-basic-needs method (CBN). The CBN method stipulates a certain bundle of basic goods which is then evaluated at relevant prices.[19] Under certain assumptions, the cost of the basic needs-consumption bundle is a utility-consistent cost-of-living index (Ravallion, 2008). In practice, however, the price of the bundle is a normative minimum income that is related as closely as possible to observed consumption patterns.

Accordingly, one usually starts out with a bundle of food items that yields a food-energy intake deemed just adequate. In a second step, the so-derived food poverty line is augmented by an allowance for basic non-food expenditure. There are typically many combinations of food items that will yield exactly the same level of food energy. A naïve approach would be to treat the problem as an exercise in linear programming, i.e. in

---

[18] Ravallion (1998) argues that because of higher relative prices for food and systematic differences in tastes and activity levels, the FEI method may be biased towards higher rates of poverty in urban areas (without the corresponding differences in actual command over commodities).

[19] This method was used by Rowntree in his seminal work on poverty in York at the end of the 19th century (see Atkinson, 1975, ch. 10). It is also the basis of the official poverty line in the US (Orshansky, 1965) and in numerous developing countries.

finding the cheapest bundle that yields the required level of food energy. However, in practice, this approach often yields a consumption bundle that is very different from the actual bundles consumed by the poor and its cost is often be found to be much too low to be interpreted as a credible food-poverty line (Deaton, 1997). A more reasonable approach is to identify the food-consumption bundle of, say, the poorest quintile of the population in each sub-group.[20] This bundle is then evaluated with respect to food energy content and scaled upward or downward so as to yield the minimum energy requirement. The cost of this bundle is then evaluated using local prices.

The second problem is to add an appropriate allowance for non-food expenditure to the food-poverty line. In principle, this could be done in much the same way as for the food-poverty line. However, for expenditure on clothes, housing, health, and education there is no obvious anchor as the adequacy of a diet in terms of nutritional requirements in the case of expenditure on food.

An often-used *ad hoc*-solution to this problem is to divide the food-poverty line by the share of food consumption in total consumption of some group of households likely to be poor in each sub-group.[21] This approach may be problematic as higher real consumption in one sub-group (e.g. urban dwellers) will usually be associated with a lower food share of the poorest 20 percent in that sub-group resulting in turn in a higher poverty line. Inconsistencies may thus arise in the sense that the same command over commodities will deem a given household poor in one sub-group but not in another. A fixed food share might be preferable if additional information is unavailable.[22] This will also be problematic, however, whenever the relative prices between food and non-food items differ across sub-groups (or over time).

In general, the question of when and how to update a poverty line using the CBN method remains unsolved. As consumption patterns change, the basket upon which the poverty line should be based on, ought to reflect that change at some point. While one should clearly not recalculate the CBN poverty line whenever a new survey appears, it is reasonable to assume that one should change the basket every 10 years or so, as is typically done with the basket for the consumer price index (which faces the same problem). In practice, however, changes in the poverty line are much more infrequent. Even in the US, they have been working for over 20 years on updating the poverty line (Citro and Michael, 1995), but it has yet to be done in practice. In India, it took 15 years to update their poverty line and the resulting line is not really based on a true update of the poverty basket (see Subramanian, 2010). China recently updated its poverty line, reflecting a different basket. But in most other countries, poverty lines have a habit of persistence, also for political reasons (Ravallion and Chen, 2009). As usually

---

[20] This can be seen as a first guess at the poverty headcount of the population studied.

[21] A similar method was used by Orshansky (1965) in which the food-poverty line is divided by the average share of expenditure on food in total expenditure over the entire population.

[22] Ravallion (1994) proposes a more elaborate procedure based on the notion of basic non-food spending in order to narrow down the plausible range of the non-food-poverty line: the idea is to use the typical non-food expenditure of those whose total spending would just suffice to purchase the basic needs-food bundle as a lower bound estimate. This is interpreted as basic non-food spending as this is expenditure that households allow to displace spending on basic food. The typical non-food expenditure of those whose spending on food just matches the cost of the basic needs-food bundle is then taken as an upper bound (see Ravallion, 1994, pp. 118-125, for details).

adjustments of the poverty basket would lead to an increase in the poverty rate, few politicians want to make such adjustments.

There is mounting empirical evidence that the two methods of deriving a poverty line (FEI versus CBN) can yield strikingly different results (e.g. Ravallion and Sen, 1996; Wodon, 1997, Tarp et al., 2002). Using Indonesian data for the 1980s and early 1990s, Ravallion and Bidani (1994) find virtually zero rank correlation between poverty profiles based on the FEI and CBN method. Wodon (1997) also shows for Bangladesh that the trend over time differs, with the former showing a decrease in poverty and the latter an increase – most likely because people switched to cheaper calories during the recession. Both studies conclude that CBN poverty lines are generally more likely to allow consistent poverty comparisons.

3.3 Subjective Poverty Lines

An alternative to the objective poverty lines discussed above are subjective poverty lines (SPLs).[23] SPLs are based on subjective views on the adequacy of income or consumption spending. This approach, originally proposed by Goedhart et al. (1977), is usually based on the *Minimum Income Question* (MIQ) (paraphrased from Kapteyn et al., 1988):

"What income level do you personally consider to be absolutely minimal? That is to say that with less you could not make ends meet."

One then fits a (sufficiently flexible) regression line to the scatter plot of the answer to the MIQ against actual income. Subjective minimum income is usually found to increase with actual income and the regression line usually has a slope between zero and unity. Hence, one can estimate the actual income at which, in expectation, respondents will claim that they dispose of just enough income to make ends meet. The actual income corresponding to this fixed point is the obvious candidate for a poverty line (see figure 2). Note that, in general, not all individuals that consider themselves poor (non-poor) will be classified as poor (non-poor), distinguishing this approach from purely subjective assessments. Rather, the method constitutes an intermediate approach between objective methods and purely subjective assessments of well-being (see section 6).

---

[23] The SPL is sometimes also referred to as the Leiden Poverty Line after its place of origin. See Goedhart et al. (1977).
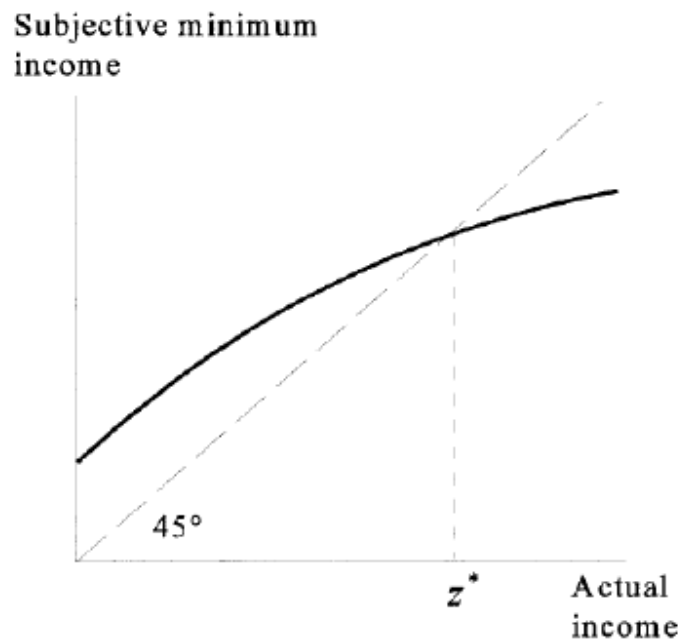
Subjective minimum
income

45°

z*   Actual
income

**Figure 2: Derivation of the subjective poverty line: individual above z*will, in expectation, deem
their incomes adequate, while those below z* do not. Source: Pradhan and Ravallion (2000).**

Several studies based on the MIQ have been conducted in industrialized countries (e.g.
Colasanto et al., 1984; Danziger et al., 1984; Kapteyn et al., 1985, 1988; Flik and van
Praag, 1991; Stanovik, 1992; Kapteyn, 1994). The MIQ is not directly applicable outside
the developed world, however, since income (or even expenditure) is not a well-defined
concept in developing countries (Deaton and Zaidi, 2002; Lokshin et al., 2006). Pradhan
and Ravallion (2000) show that SPLs can be derived using questions about the
adequacy of consumption, where answers are ranked on an ordinal scale (e.g. 'less than
adequate', 'just adequate', 'more than adequate'). Applying their method to data for
Jamaica and Nepal, they find that while the aggregate poverty headcounts accord
closely with those based on objective methods, there are significant differences in the
geographic and demographic poverty profiles. Lokshin et al. (2006) apply a slightly
modified version of the Pradhan and Ravallion-method to data from Madagascar and
also compare the resulting poverty profiles to profiles based on objective and MIQ-based
poverty lines. They find that while the former corresponds closely with profiles based on
the Pradhan and Ravallion-method, the latter results in a different (an arguably less
sensible) profile.

Some doubts about the consistency of poverty comparisons based on the MIQ remain:
Milanovic and Jovanovic (1999) study the relationship between the answer to the MIQ
and income in Russia during the depression of the 1990s. While most studies find an
elasticity of subjective minimum income with respect to actual income, the *preference
drift* parameter, of between 0.4 and 0.7 (Flik and van Praag, 1991), they estimate a
preference drift in excess of unity. In other words, the change in subjective minimum
income was larger than the change in actual income in relative terms. This caused the
poverty headcount based on the SPL to decline despite falling real incomes. While
Russia has certainly experienced an unusually deep recession during the 1990s, it can

be argued that a fully convincing method should still adequately capture changes in poverty in these circumstances.

Nevertheless, SPLs are a promising device in so far as they can be used to complement and validate objective poverty lines and identify discrepancies which can then be further investigated (Ravallion and Lokshin, 2006). However, data on perceived consumption adequacy are only available for a hand-full of developing countries and further investigations of the robustness of SPLs are warranted.

3.4 International Poverty Lines

We have so far discussed expenditure-based poverty lines to estimate poverty within a single country across time and space. With due effort devoted to the task, domestic poverty lines will be a helpful device for national policy-makers and their constituents, especially when there is public involvement in defining the basic needs-bundle. In this section we turn to international poverty lines (IPLs), i.e. income- or expenditure-based poverty lines for comparisons across all countries (or, at least, all developing countries). Our focus will be on the World Bank's IPL (referred to as $1-a-day), the official indicator to monitor progress towards the first Millennium Development Goal. After reviewing the World Bank's methodology, we turn to its critiques and discuss the potential of alternative methods.

3.4.1 The World Bank's $1-a-day IPL

The World Bank's first attempt to quantify the extent of global poverty in a comparable way dates back to the 1990 World Development Report (World Bank, 1990). Estimates of poverty for 1985 are based on a poverty line of the purchasing power equivalent of $1 per day per person in 1985 prices. The method is detailed in a background paper by Ravallion et al. (1991) (henceforth, RDV) and has remained largely the same. The poverty line was updated twice after new data on prices became available through the World Bank's International Price Comparison Project (ICP): the poverty line was shifted to $1.08 in 1993 prices in 2000 and then to $1.25 in 2008 in 2005 prices (see World Bank, 2000; Ravallion et al., 2008, respectively).

How did RDV arrive at '$1-a-day' and how is this then used to make poverty comparisons across countries and time? In all three version of the World Bank's IPL (1990, 2000, and 2008), the starting point is a sample of domestic poverty lines expressed in local currency units. In order to render these comparable, poverty lines were converted to dollar equivalents using the latest available estimates of PPPs computed from ICP data. For the 1990 exercise, the 1985 ICP data were used; for the 2000 update, the 1993 ICP data were used; and the latest estimates are based on the 2005 ICP data (Ravallion et al., 2008).

The second step is to investigate the relationship between poverty lines and per capita consumption in order to arrive at "a poverty line typical of poor countries" (Ravallion, 2010).[24] This step has changed slightly over time as discussed below, however, as we

---

[24] An earlier attempt at estimating the extent of global poverty was based solely on India's domestic povery line (see Ahluwalia et al., 1979).

will see, with large consequences in some respects.[25] In the original paper by RDV, the authors start from 33 domestic poverty lines from various sources (some official and some from academic sources) which are plotted against mean expenditure (measured by household final consumption expenditure per capita from the NAS). Both variables are measured in 1985 international dollars per month and, if necessary, countries' consumer price indices (CPIs) were used to make the data comparable. Figure 1 in their paper, reproduced here as figure 3, depicts this relationship. The main finding here is that poverty lines tend to increase in mean expenditure but less so the poorer the country. RDV subsequently focus on the ten poorest countries in their sample, which happened to have poverty lines in the neighborhood of $31 per month.
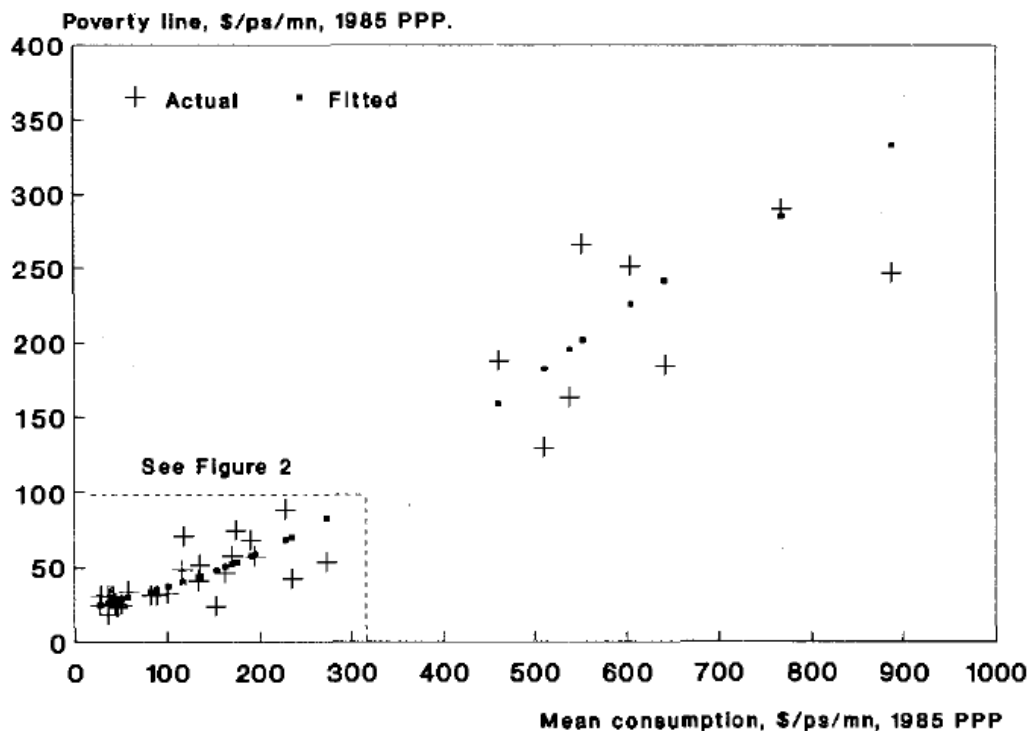
Figure 3: Scatter plot of domestic poverty liens against household final consumption expenditure per capita (from NAS) for a sample of 33 domestic poverty lines (see text for details). Source: Ravallion et al. (1991).

Their finding regarding the relationship between poverty lines and mean consumption is key to their claim that what their poverty line captures is the *absolute component* of poverty: they argue that, conceptually, poverty can be broken down into an absolute and a relative component, the former fixed over time and the latter eventually increasing in importance with mean expenditure (also see section 2.3). Hence, by focusing only on the poorest countries with poverty lines unresponsive to changes in mean consumption (countries located in the bottom left corner of figure 3, they claim to isolate the absolute component of poverty.

---

[25] For critical review, particular of this step of the World Bank's methodology, see also Klasen (2009), Srinivasan (2010), and Greb et al. (2011) among others.

Three steps are finally carried out in order to apply this poverty line to data from national household surveys: first, the IPL is converted into local currency units via PPP exchange rates for the benchmark year (1985, 1993, and 2005, respectively). Second, the poverty line is shifted backwards and forwards in time using local CPIs. Finally, the poverty headcount is calculated based on national survey data on household income or expenditure.[26]

It is important to emphasize that both the setting of the poverty line as well as the measurement of poverty using that line is based on per capita income or consumption. Thus differences in household size and composition are not taken into account. Given reasonably assumptions about needs of children and economies of scale within households, countries with larger households will have poverty rates that are biased upwards. As a result poverty rates in Africa are likely to be overestimated; conversely, declines in poverty in countries where fertility and household sizes is falling are overestimated.

Note that in each of the three rounds (1990, 2000, and 2008) poverty estimates were calculated not only for the most recent years but for all years, that is, going back to the 1980s. Therefore, there are now three sets of poverty estimates for 1985 based on different rounds of the ICP and it turns out that these imply very different patterns concerning the regional distribution of poverty and, in the case of the latest update, a different level of global poverty (Deaton, 2003; Klasen, 2009; Deaton, 2010; Reddy and Pogge, 2010). The revision from 1985 to 1993 had little effect on the total number of poor but made Sub-Saharan Africa appear to be much poorer and Latin America appear to be much less poor.

The 2005 revision not only changed our perception of the relative burden of poverty (from Sub-Saharan Africa to South and East Asia). It also increased the global poverty count by almost half a billion people. Figure 4 compares estimates of the poverty headcount using the latest two rounds. Trends in global poverty, it should be noted, remained largely unaffected by these changes. Still, as Deaton argued already after the revision in 2000, "changes of this size risk swamping real changes, and it seems impossible to make statements about changes in world poverty when the ground underneath one's feet is changing in this way" (Deaton; 2001, 2010) What causes the ground underneath our feet to change? Further below, we consider two potential culprits: (1) changes in PPP conversion factors from one round of the ICP to the next and (2) changes to the methodology of determining the level of the World Bank's IPL. However, before we do so, we comment on some general problems with the World Bank's IPL.

---

[26] There is no adjustment for household size or composition that takes into account the ability of large households to economize on resources. Therefore, poverty in countries with large (small) households might be overestimated (underestimated). While it would be challenging to define common equivalence scales that are meaningful in all countries in the developing world, this could prove to be an interesting alley for future research.
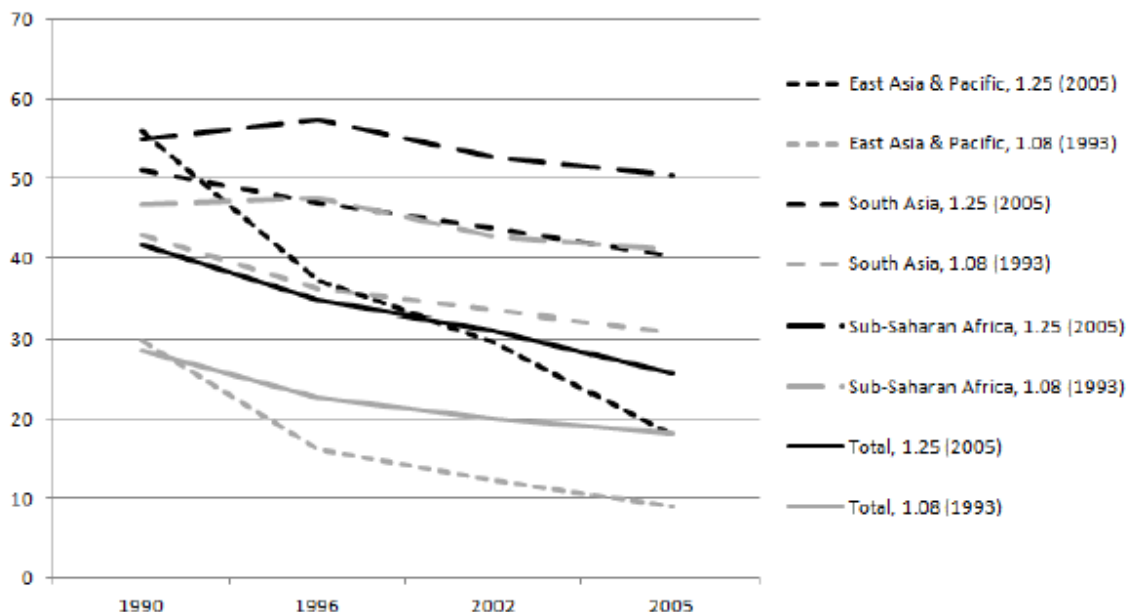
Figure 4: Share of the population below the new poverty line of $1.25 a day (black lines) and below the old poverty line of $1.08 (grey lines). The differences are, to a very small degree, also due to changes in survey data. Also note that the $1.08 figures for 2005 refer to 2004. Based on Ravallion and Chen (2007) and Chen and Ravallion (2010). See also Klasen (2009).

3.4.2 The Use of PPPs for International Poverty Comparisons

One cannot convert an international poverty line using market exchange rates. Labor is relatively cheaper in poor countries and, therefore, the prices of non-tradable goods which tend to be intensive in labor are lower relative to the prices of tradable goods. The market exchange rate between the currencies of two countries, one poor and one rich, is determined by trade. Hence, it does not adequately reflect the difference in the cost of living between the two countries. What one wants to do is to take a representative bundle of goods and price it in each country. The ratio of the cost of this bundle is a PPP exchange rate.

There are several problems with PPP estimates, some of which are inherently insolvable and others lending themselves to improvements in the future. First, PPPs are price indices and thus subject to the well-known general index number-problem: there is no best way to convert overall price levels across space or time when relative prices are changing (Diewert, 1976; Deaton, 2003; Deaton and Heston, 2010; Anand et al., 2010). These problems are familiar from within country-comparison, where, for example, the Paasche and Laspeyres indices will give different answers. However, they tend to be more pronounced in inter-country comparisons since differences in relative prices are more pronounced (Deaton and Heston, 2010).

While an in-depth analysis of these problems is beyond the scope of this problem, it is still useful to discuss one particular inconsistency concerning multilateral price indices (also noted by Reddy and Pogge, 2010): the PPP price index used by the ICP is based on the EKS-Fisher method (named after economists Eltetö, Köves, and Szulc). Its advantage is that it is *transitive* in the sense that going from country A to B via C gives

26

the same result as going from A to B directly. Transitivity, or *circularity* as it is sometimes called, is an important characteristic of a multilateral price index as it allows us to work with a PPP vector, rather than a matrix. It comes at a price, however, in that PPP comparisons between two countries depend on prices *in all other countries*.[27] Even when prices are constant in both the base country and the country of interest, the poverty line in the latter expressed in its own currency will change in response to a price change in a third country. How serious the impact of such a violation of *country irrelevance* is in practice is ultimately an empirical question. Again, the effect will likely increase if relative prices differ between the two countries being compared (Deaton and Heston, 2010).

Second, it is important to note that some aspects of the exercise are close to impossible even in theory. It should be remembered that what one is really interested in is the answer to a question like "How many pesos would it cost a landless laborer from Bihar to live as well in Mexico as she lives in India?" (Deaton, 2003) and that one is asking this type of question for all possible pairs of households. It is clear that a reasonable answer to this type of question involves several compromises. In particular, some of the goods consumed in Mexico will not be available in Bihar so that there is no price on record (or they will only be available in special shops in few larger cities).[28]

Third, the World Bank's practice is to use PPPs for household consumption, where the information on consumption patterns is extracted from the National Account Statistics (NAS). While clearly more appropriate than, say, PPPs for GDP, it might still be inappropriate as the consumption patterns of the poor might be very different from average patterns. In particular, average consumption patterns are *plutocratic* in the sense that they tend to reflect the consumption of rich households. Goods that are important to the poor will play a minor part in determining average consumption patterns.

An obvious alternative is to these plutocratic weighting scheme is to calculate PPP conversion factors based on the actual consumption patterns of the poor (e.g. Aten and Heston, 2010). The World Bank is currently working on such `poverty-related PPPs' (also called PPPPs). See http://go.worldbank.org/I79PXTG5W0 and the work cited there. The effects of re-weighting prices according to consumption patterns of the poor have also been investigated by Deaton and Dupriez (2011). They combine information on the spending patterns of the poor from household survey data – the only source of information on this issue – and combine it with prices from the 2005 ICP. They find that the difference between the so-derived PPPs and those based on NAS average consumption patterns are negligible. Indeed, the main difference seems to stem from

---

[27] If there are M countries in the system and the base country's prices index is $P_G = 1$, country c's index is defined as

$$P_F^c = \left(\prod_{i=1}^{M} P_F^{1,j} \, P_F^{1c} \right)^{\frac{1}{M}},$$

where $P_F^{jc} = (P_F^{cj})^{-1}$ is country j's Fisher price index relative to country c's. In other words, P_F^c is the geometric average over all M indirect Fisher price indices for that country.

[28] The ICP deals with this problem by aggregating prices to 110 basic headings, categories that are broadly enough defined to allow international comparison, at the regional level. While the items within each basic heading are region-specific, the basic headings are the same for all countries (Deaton, 2010).

using household consumption surveys rather than NAS data.[29] The reason for this seems to be that changes in the bundle of goods from an average bundle to a bundle relevant to the poor is not related to changes in relative prices (Deaton and Dupriez, 2011).

Finally, as mentioned, PPPs are used for different applications, but the dominant purpose thus far has been the comparison of GDP and its components, what Reddy and Pogge (2010) call "broad gauge"-PPPs. The focus of the ICP's data collection effort is therefore on comparability and this often requires as a compromise price data on products that are available in all countries. There might therefore be a bias towards urban outlets and internationally available brands. In particular, Ravallion et al. (2008) note that the 1993 ICP is generally believed to have been less rigorous in terms of assuring that the prices collected were for products of comparable quality across countries: they note that the prices collected by the 1993 ICP tended to be more typical of the items available in local markets. This has changed with the 2005 ICP round. This would have both advantages and disadvantages. On the one hand, comparability is certainly warranted. On the other hand, such a change constitutes a move away from goods that are actually relevant to the poor. Indeed, there is some empirical evidence that seems to suggest that poor people pay higher prices *even for the same product*, most likely because they buy in smaller quantities (e.g. Rao, 2000). However, it is unclear how much of a problem this is in practice.

Do changes in the methodology between the 1993 to 2005 ICP rounds explain the large increase in measured global poverty? And if yes, should these changes be seen as a welcome improvement in our understanding of poverty? The answer to the first question is that changes in the method to calculate the PPPs played a minor role (Greb et al. 2011). An increased emphasis on comparability in terms of the products included in the survey resulted in prices that are not necessarily representative for the prices the rural poor face. Hence, prices for China, which was included explicitly for the first time in the 2005 ICP, are probably too high which, in turn, increases the poverty line there. China also collected only prices in urban and peri-urban regions, which are probably also higher than average prices, not to mention the prices faced by the (largely rural) poor.

The increased coverage of the 2005 ICP round---earlier rounds would impute the data for countries not covered---necessitated some compromises as methods used earlier were no longer applicable in all countries. These compromises usually concerned what the ICP calls *comparability resistant* consumption categories such as government services and housing rentals (Deaton and Heston, 2010).[30] Deaton (2010) tries to reconcile earlier global poverty estimates with the latest figures. He concludes that a

---

[29] On the increasing divergence between NAS- and survey-based estimates of average consumption see Ravallion (2003).

[30] There is an inherent conflict in the application of PPPs to poverty comparisons: extending the coverage of international prices surveys to smaller and poorer countries is, of course, to be welcomed. On the other hand, PPP calculations need data on prices for all goods consumed. The poorer a country, the less developed its markets, the more problematic is it to accurately measure prices. This might necessitate more imputations as in the case of housing rentals.

change in the treatment of housing rentals has probably increased the number of poor by some 100 million people, roughly one fifth of the increase that one needs to explain.[31]

3.4.3 Setting the International Poverty Line

It turns out that the larger portion of the recent changes in the total number of poor can be attributed to small changes in the method applied to set the IPL. As outlined above, the first $1-a-day IPL was based on the observation that ten of the poorest countries in a very small sample had poverty lines close to $31 in international dollars (Ravallion et al., 2008, henceforth, "RCS"). Until the most recent revision, the 1990 line was updated by applying new PPP figures to the poverty lines of the same reference countries. In the latest revision, however, RCS relied on an updated dataset (comprising 74 observations on official domestic poverty lines for the 1985-2005 period) and advanced econometric methods in order to identify the threshold-consumption level after which poverty lines tend to increase. Their estimated threshold is indicated in the right panel of figure 5 through the vertical dashed line, where the mean over the poverty lines of the 15 poorest countries to the left of this threshold is about $38 per month or $1.25 per day (indicated through the horizontal dashed line).

There are several points to note about RCS's procedure. As argued by Greb et al. (2011), it is not clear whether a threshold model as estimated by RCS is at all warranted. As can be seen from figure 5, a non-linear relationship between mean household expenditure and domestic poverty lines might only be present when the logarithm of expenditure is considered (right panel). RCS fit a piece-wise linear relationship on the data in panel (a), where it is unclear whether such a relationship actually exists.[32] Note also that there is very little in economic theory that would guide the choice of whether a model in levels, a semi-logarithmic model, or a log-log-model is appropriate in this setting.

While the debate on the appropriate method to identify a set of reference countries for an IPL is not yet settled, what seems clear at this point is that no matter which method is used, estimates are surrounded by considerable uncertainty. The Bank's numbers come without confidence intervals even though there are many sources of uncertainty including sampling variance from underlying household surveys and the ICP price data, imputation methods for countries not covered by household surveys or the ICP, and the regression approach to establishing the poverty line. This was partly acknowledged, however, by RDV in their 1991 paper. In RCS, the authors state that the confidence interval for the mean of the poverty lines for the poorest 15 countries extends from $1 to

---

[31] The assumption adopted was that the quantity of housing is a fixed proportion of GDP. The expenditure on housing rental divided by the assumed quantity yields the parity for this category. Many of the (mainly African) poor countries which enter into the calculation of the level of the poverty line report low expenditure for housing in their national accounts to the point where they are no longer credible. This results in lower parities for this category relative to overall consumption parities, in particular in comparison with China and India. This inflates the global poverty count by increasing the local currency-equivalent of their poverty lines in both china and India. Dropping the category (basically, assuming that parities for housing are the same as for other consumption) increases the PPPs for these countries and, importantly, decreases the poverty line equivalents in China and India.

[32] Note that the left panel of figure 5 might be misleading because observations are staunched in the bottom-left corner of the plot.

$1.45. While other sources of uncertainty are not considered, this would already have large implications for what we can say about the extent of absolute poverty in the world.

The answer to the question why measured poverty has increased has so far been left largely unanswered. Note that the 15 countries to the left of the threshold identified by RCS are not the same countries as in earlier exercises. Notably, fast-growing and populous countries like Bangladesh, China, India, and Indonesia that have been members of this group in earlier exercises have `graduated' from the poverty line-group.[33] This is what ultimately explains the apparent increase in global poverty: from figure 5, panel (b), it can be seen that both India's and China's domestic poverty lines are lower than $1.25. Because these countries graduated from the reference group, the average poverty line in this group increased and so did the total number of poor. Thus, there seem to be more poor people in India and China today *because mean household consumption increased in these countries* – a clear violation of the basic requirement of monotonicity, that, all else equal, a decrease in poverty in any country included in the counts, should result in a decrease in global poverty as well (Deaton, 2010; see also section 2.2).
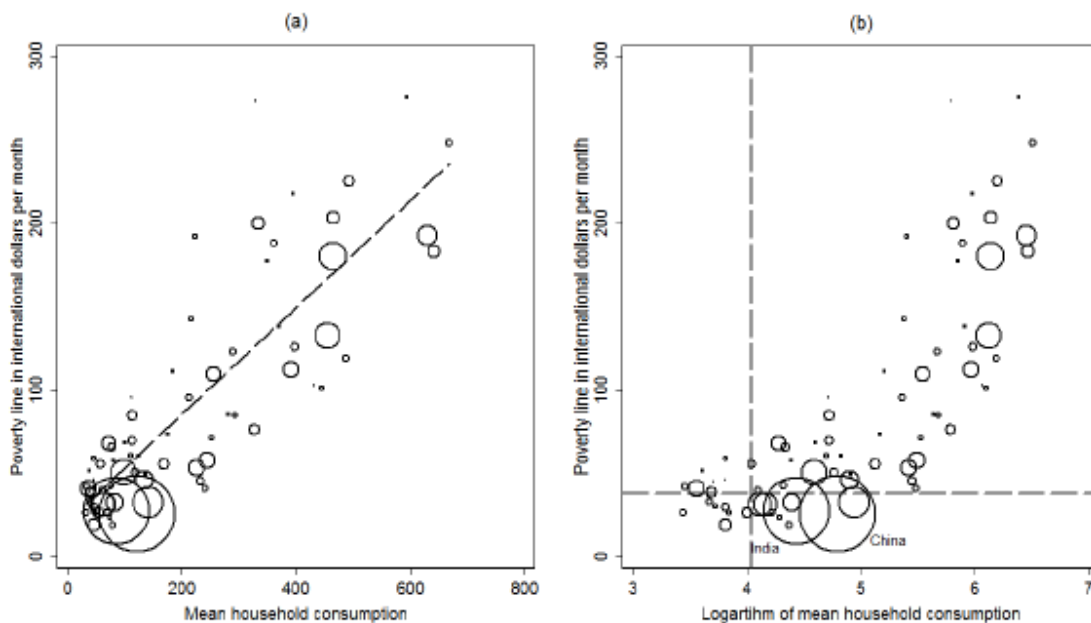
Figure 5: Scatter plots of domestic poverty lines against private consumption per capita (from NAS) (a) and the logarithm of private consumption per capita (b). The size of the markers is proportional to year 2000 population size. The vertical line in (b) indicates the threshold estimated by Ravallion et al. (2008), the horizontal line the resulting IPL at $38 per month. Data provided by Martin Ravallion.

We have noted earlier (and it is also evident from figure 4) that the change in the level of the poverty line did not have an effect on global poverty trends – a finding neatly

---

[33] Greb et al. (2011) estimate the average poverty line for the countries of the original sample (Bangladesh, China, India, Indonesia, Nepal, Pakistan, Tanzania, Thailand, Tunisia, and Zambia); this yields an IPL of $1.05 per day. Note that the apparent decline of the poverty line $1.08 to $1.05 despite inflation in the intervening years reflects the fact that the 2005 ICP finds prices levels to be relatively higher in poor count5ries than the previous round.

summarized by the title of Chen and Ravallion's 2010 paper *The Developing World is Much Poorer than We Thought, But No Less Successful in the Fight Against Poverty*. Given what we know about the relationship between growth, the poverty line, and poverty reduction (Bourguignon, 2003; Klasen and Misselhorn, 2007), this is exactly what one would expect from modest one-off increase in the level of the global poverty line.

How can the procedure to estimate the level of the IPL be changed so that changes from one round to the next are kept at bay? Deaton (2010) proposes an iterative weighting procedure where all countries in the dataset enter into the calculation of the level of the IPL with weights corresponding to their share of the global poor. While this would most likely preclude the discontinuity problems inherent in RCS's method, it is not clear why the importance of a given domestic poverty line in the calculation should be proportional to their total share in world poverty.[34] To be consistent, it would probably be preferable to rely on the same reference group of countries. However, consistency is clearly not the only criterion and there is so far no convincing procedure to sort out the problem of choosing reference countries.

3.4.4 An Alternative Approach?

Reddy and Pogge's Alternative The problems outlined above have led some observers to conclude that the Bank's approach to measuring global income poverty should be abandoned altogether (Srinivasan, 2010; Reddy and Pogge, 2010; Klasen, 2009). Reddy and Pogge's main line of criticism even goes beyond the points considered above: they argue that the World Bank's IPL is not based on some concept of human achievement, for example being adequately nourished.

Their alternative would involve a common approach to poverty measurement at the national level based on something like the CBN method described in section 3.2.2. A pre-requisite for applying this method to all countries would be the availability of comparable household surveys in developing countries. A first empirical exploration based on the version of the CBN approach used by the Government of Vietnam is given in Reddy et al. (2006), who apply this method based on the Living Standards Measurement Surveys (LSMS) for Nicaragua, Tanzania, and Vietnam. In short, they estimate the cost of a bundle of food items that yields a pre-specified level of caloric intake and is consistent with actual choices made by the poor in each of the three countries separately. They then add an allowance for non-food spending, also based on empirical observations within each country. As one would expect, the poverty lines thus established resemble more closely domestic poverty lines as most countries base poverty assessments on this method. There are slight differences though due to the use of a different caloric cut-offs and different allowances for non-food spending.

More research is warranted in order to assess the usefulness of such an approach. Note that while the World Bank's IPL is not explicitly anchored in any basic need such as being adequately nourished, most of the underlying poverty lines in the base sample are. One could thus argue that there is an implicit anchoring in basic human needs. Moreover, we have outlined several problems with the CBN approach above, which

---

[34] As one would expect, Deaton's procedure results in an IPL driven mainly by the two population-giants China and India. India's poverty line would have a weight of approximately one third (Deaton, 2010).

occur whenever relative prices vary. Such an approach would clearly violate consistency in the sense that a household would be judged poor in one country but not in another. Also see the critique of such an approach (and a defense of the World Bank's) by Ravallion (2010).

Weakly Relative Poverty We have explained above that the World Bank's traditional approach has been to try to isolate the absolute component of global poverty. However, as many developing countries particularly in Asia have recently graduated to middle income-status and keep growing at high rates while, at the same time, within country-inequality in these countries is thought by many to have increased considerably, the relevant component of poverty is deemed to become more important over time. This suggests that there might be a case for alternative IPLs that vary between countries with mean consumption expenditure.

At the heart of such an approach would be a compromise between the absolute approach to poverty measurement adopted in most developing countries and the strongly relative poverty lines found in the developed world (see section 2.3), where "strongly relative" refers to an elasticity of the poverty line with respect to mean consumption of unity. Both Atkinson and Bourguignon (2001) and Ravallion and Chen (2009) follow a hybrid approach calibrated to domestic poverty lines. The motivation behind this approached is based on the notion that the cost of social inclusion is increasing in mean consumption.

Atkinson and Bourguignon (2001) assume an absolute poverty line (i.e. invariant to mean consumption) for poor countries and an increasing component of poverty beyond some threshold-level of per capita consumption that reflects the cost of social inclusion. In their application, the relative component is strongly relative, i.e. the ratio of the relative component to mean consumption is constant, implying an elasticity of unity above some threshold. Ravallion and Chen (2009) show that an elasticity of unity above the threshold implies zero cost of social inclusion for the very poor, which can be hardly maintained in the light of findings on consumption patterns of the very poor. They combine an absolute component and a relative component that implies positive costs of social inclusion even for the poorest countries. The elasticity of the poverty line with respect to mean consumption then lies between zero and unity and is increasing in mean consumption.[35,36]

In the light of recent changes in the world distribution of incomes, these hybrid poverty lines that combine absolute and weakly relative poverty lines are an interesting approach that clearly warrants further discussion. However, two points should be noted: first, the approach explicitly violates the axiom of *sub-group anonymity*, which states that moving a person between two sub-groups should not affect the poverty measure. In a context with relative poverty lines, whether a household is deemed poor will always

---

[35] In both empirical applications, the slope of the combined poverty line above the threshold-level is usually estimated to be around one-third.

[36] The Atkinson-Bourguignon-relative poverty line is $z_{AB} = \max(z^*, kM)$, where z^* is the absolute poverty line, M, is consumption, and k (with $0 < k < 1$) a parameter to be estimated. The Ravallion-Chen-relative poverty line, which can be seen as a generalization, is $z_{RC} = \max(z^*, \alpha + kM)$, where α is an additional parameter to be estimated. For the latter, α is a positive parameter that can be interpreted as the lower bound of the cost of social inclusion. For $\alpha > 0$, z_{RC} satisfies weak relativitiy.

depends on the country in which it is located. This is reasonable only if one accepts that the poverty-level of expenditure differs between sub-groups like different societies (which will be the case if the cost of social inclusion is increasing in mean consumption). Second, there is the question of how to set such a weakly relative poverty line. While Ravallion (2010) is able to show that a current cross-section of poverty lines in developing countries is consistent with the notion of 'weakly relative poverty' in the sense that there is a group of poor countries where the poverty lines are very similar, and a group of richer countries where the poverty line rises with mean income (but the increase is less than proportionate), it is unclear that changes in poverty lines over time follow this 'weakly relative poverty' rule. In fact, as discussed above, changes in poverty lines are rather rare and do not generally have such a conceptual basis. Third, on a more general note, it is not clear that the country is the relevant comparison group for all households. While this might be appropriate for small countries, it is far from clear that an individual in one of the poorer states of India is more likely to compare herself to an average Indian or rather to a much smaller group (e.g., one's village). It is also likely that the reference group chosen by households is endogenous to the level of development. For example, increased access to mass media might change the reference group from one's direct neighbors to one's countrymen.

3.4.5 Summary of IPLs

While for poverty assessments and policy planning at the national level governments and international agencies (including the World Bank itself) will rely on domestic poverty lines, the World Bank's $1-a-day IPL has been enormously successful with audiences in the rich world. There is little doubt that this is mainly because of its apparent simplicity, it is denominated in the world's most important currency and still widely cited as '$1-a-day' rather than having been referred to with its actual level. The discussion above shows that the simplicity is illusive rather than substantial. It is much more challenging to measure income poverty at the global level as compared to measurement at the national level and analysts are frequently forced to make wide-ranging compromises.

There is some agreement on how to refine the World Bank's method. For example, poverty-related PPPs will clearly improve our understanding of the level and regional distribution of global poverty. There is much disagreement on whether an alternative that circumvents PPP-based calculations based on standardized national poverty lines is ultimately superior. It would definitely require a large effort by the international community to coordinate data collection. However, coordinating household expenditure surveys such that data can be compared across countries is a worthwhile endeavor for many reasons, the investigation of their usefulness for global poverty comparisons being only one of them.

4 Multidimensional Poverty Measures

4.1 Background

The concept and measurement of multidimensional poverty is inherent to the poverty concept itself. The most basic definition of the poor as the ones whose well-being is below a certain threshold (and therefore insufficient to meet certain requirements in

order to have a decent life) immediately begs the question of which dimensions of well-being should be considered. AS discussed above, reducing this to one dimension, e.g. income or consumption, requires a great deal of assumptions or is just a proxy (see below). Clearly, it can be useful to tackle the challenge of multidimensionality directly. The great challenge of a multidimensional approach is thus that of combining a detailed and more precise definition of the poor with appropriate and robust measures which are essential to solve the identification and aggregation problem in multiple dimensions. The conceptual and empirical problems with implementing such an approach dominate this strand of the literature.

The cornerstone of the multidimensional poverty literature can be found in the theoretical framework of capabilities and functionings elaborated by Sen and discussed in section 2.1. Two implications are worth noting at this point: first, there are several capabilities and poverty is therefore clearly a multidimensional concept. Hence, the motivation for considering different attributes in one framework. Second, one should ideally focus on capabilities, rather than functionings or the means to achieve those. Note right away that this is rarely done in practice as capabilities are much more difficult to measure. Most multidimensional poverty measures focus on functionings as these are more likely to be readily-observable in comparison to (see the discussion in Johannsen et al., 2007).

As suggested by Alkire and Foster (2007), the complexity which is inherent in the term 'poverty' should be addressed by setting up some multiple-deprivation indicator which can satisfy basic criteria as that of being understandable and easy to describe, to conform to 'common-sense' notion of poverty from a conceptual point of view and, at the same time, to be operationally viable and valid, being able to target the poor and guiding policy. Moreover, it should be a robust tool from a statistical point of view thus allowing for replicability in other settings.

In practice, the measurement of multidimensional poverty spans very diverse methodological approaches and procedures but basically relies on three core steps consisting in the specification of a list of relevant attributes, the determination of a threshold, and – in some cases – the aggregation to come up with an overall poverty index.

If the determination of cut-offs in one dimension is not a straight-forward exercise, conceptualizing, measuring and investigating poverty under a multidimensional perspective raises even more problematic issues which must be addressed in order to find appropriate and robust measures of multiple deprivations. Clearly, for some of the possible attributes of well-being (e.g. social inclusion), finding appropriate indicators and establishing thresholds involves some arbitrariness stemming from the vagueness and elusiveness of these concepts.

In addition, it is close to impossible in practice to find indicators that correspond to actual capabilities – at least on the basis of objective data – since measuring individuals' potential to assume certain functionings. Hence, most studies either focus on functionings or the means to achieve functionings (see Johannsen et al., 2007, for a review) or resort to subjective data (Anand et al., 2005). But then, if we consider a pair of leather shoes as one component of a multidimensional approach to poverty measurement rather than the ability to 'go in public without shame', we will again have the problem of context-specificity (see the discussion in section 2.3). This is summarized

succinctly by Thorbecke (2007): "The conflict between the consistency criterion and the specificity criterion that plagues the conventional income-metric approach to poverty analysis, applies equally well to a multidimensional approach to poverty analysis."

It should be noted that some of the indicators discussed below are mainly measures of well-being, i.e. average values (like the mean or the median) for the entire population, rather than poverty measures which focus on the lower part of the distribution (see section 2.1). Clearly, it would still be a useful exercise to see how these indicators looked like if applied to, say, the bottom 20% of a given population. In effect, one would then ask questions like 'How poor are the income poor in terms of life expectancy?', 'How poorly nourished are the nutrition-poor, i.e. those at the bottom of the distribution of nutritional outcomes', or 'How low is the subjective well-being of the most dissatisfied? And how does this change over time and between countries?'

4.2 'Mash-up' vs. 'Dashboards'

A preliminary crucial decision in studying a multidimensional concept of poverty would therefore be whether to collapse all information into a single number, or to keep the different dimensions of poverty separate (i.e. whether or not to do the third aggregation step mentioned above). This question is referred to in the literature as the 'mash-up'- vs. the 'dashboard'-approach (Ravallion, 2010). Of course, a single number (as, for example, the one provided by the Human Development Index) has an indisputable 'eye-catching property' (Streeten, 1994) and this communicational advantage is mainly reflected in its capacity to summarize complex problems in a simple and comprehensible way for the general public.

Another example is the Sustainable Society Index (SSI), which is based upon the definition of sustainability conceived by the Brundtland Commission, includes many different aspects of human life, such as personal development, environmental issues and extends to institutional development (Van de Kerk and Manuel, 2008). [37] It is thus much broader in scope than, say, the HDI.[38] While interesting as an indicator of well-being, it is unclear how to derive a poverty measure from it as it is based on a mix of aggregate and household-level indicators.

The Millennium Development Goals, which also include the $1-a-day IPL, would be an example of a dashboard of development indicators. They have gained broad-based support because they were time-bound, specific, and – at least superficially – easy to communicate.

---

[37] More specifically, the SSI is made of 22 indicators clustered in five categories (personal development; clean environment, well-balanced society, sustainable use of resources, sustainable world). The whole index can be computed for each country according to three different weighting systems (i.e. as unweighted average of the 22 indicators, or with equal weight for each category or also with weights on each category in order to stress the relative importance of the categories with emphasis on sustainability).Nevertheless, the whole aggregation procedure seems to be-to our view- weak and not well grounded on a solid scientific basis.

[38] However, since the SSI encompasses very diverse components – based on which Van de Kerk and Manuel claim to also capture the well-being of future generations – it also fails to be as authoritative. While the concept of sustainability, basically the measurement of well-being over several generations, is conceptually important and attractive, it involves the additional challenge of putting numbers on our uncertainty about what future generations will actually be deprived of.

Sen (1987) advocates the use of dashboards: "the passion for aggregation makes good sense in many contexts, but it can be futile or pointless in others. […] When we hear of variety, we need not invariably reach for our aggregator. […] it is difficult to see why simplicity of use should have such a priority over relevance. As it happens, the more diverse characterizations of living standard, with various components separately presented, can be used in many practical settings exercises without great difficulty" (ibid., pp. 33-34). A single number would always entail some loss of information and, moreover, in many cases it would be also is limited to the extent to which it involves a possible degree of arbitrariness in choices made during the process of aggregation. On the last point see also Duclos et al. (2006) who argue that any index resulting from an aggregative process would inevitably leave "open the possibility that two equally valid rules for aggregating across several dimensions of well-being could lead to contradictory conclusions about which groups have higher poverty."

4.3 Non-axiomatic Approaches to Aggregation

A non-axiomatic methodological approach allows for poverty comparisons which can be either based on a combination of indicators that have been previously aggregated across individuals (as the Human Poverty Index) or on data for individuals so that in a first step the welfare indicators are aggregated at the individual level and subsequently across individuals.

The Human Poverty Index (HPI) elaborated by Anand and Sen (1997) takes into account simultaneously three important dimensions in human life such as life-expectancy; education and material welfare. Formally, it can be written as:

$$HPI = (w_1 HPI_1^{\theta} + w_2 HPI_2^{\theta} + w_1 HPI_3^{\theta})^{\frac{1}{\theta}}$$

Where HPI_1 reflects the percentage of individuals whose life expectancy is less than 40 years; HPI_2 is the proportion of the adult population which is illiterate and HPI_3 results from the arithmetic means of a) the percentage of the population with no access to healthcare and b) with no access to safe water and c) the percentage of children under age five suffering from malnutrition. Moreover, the sum of the *w* weights must be equal to one whereas the parameter θ can assume values larger or equal to 1. The choice on this value basically depends on the underlying assumption on the degree of substitutability among the three components.

If the HPI undoubtedly has the advantage of being a fairly comprehensive and easy understandable measure of multidimensional poverty, there are four main weaknesses embodied in this index that should be mentioned. First, the HPI does not take into account possible correlations existing among its three components. A person who falls short of two or all three dimensions will be double counted. Second, and related to this, the index is not really a poverty measure grounded in the identification of poor individuals; a certain value of the HPI will tell us absolutely nothing about how many people are actually poor in that society, and by how much. It is an aggregate deprivation measure. Third, it was recognized that the indicators chosen were not globally relevant. Therefore two versions of the HPI with two sets of indicators were created for poor and rich countries, respectively. This generates a discontinuity which the approach of weakly relative poverty (discussed in section 3.4.4) – which could be extended to non-income

dimensions – explicitly avoids. Fourth, the way in which weights are determined entails some degree of subjectivity so that ordinal comparisons will be affected by the arbitrary values which are assigned to w or θ.

An approach to avoid arbitrary weighting rules is to use statistical techniques which let the weights be determined by the data itself. An illustration of this approach is given in Klasen (2000) who uses South African micro-data. He first assigns a score (x) in the range from 1 to 5 to each of the j attributes of individual i and then aggregates the scores for each individual:

$$\bar{x}_i = \sum_{j=1}^{j=k} w_j\, x_{i,j}$$

The weights (w_j) can then be determined in two different ways: the first one simply computes the mean of the scores by giving identical weights to all attributes ($w_j = 1/k$). Alternatively, different weights are determined through Principal Component Analysis (PCA). Two poverty measures based on the resulting vector of scores are computed and compared to poverty profiles resulting from monetary poverty lines.[39]

Similarly, Ram (1982) uses PCA in order to assign more exact weights to the different components of an index of the quality of human life proposed by Morris (1979). Callan et al. (1993) recognize the existence of a complex relationship between current income and multiple deprivations which in their study are defined as a lack of items regarded as 'necessities' by individuals. A factor analysis[40] is then applied to a set of 24 items in order to find out some sort of scale in dimensions of necessities in which 'housing and household capital deprivation' (housing quality and facilities) occupy the last seat, being preceded by 'secondary deprivations' (consumer durables and leisure activities) and 'primary deprivations' including basic items such as food and clothes. The authors then use this last dimension to construct an index of primary deprivation[41] and define the poor as those individuals featuring both low values of this index and relatively low income levels. If this type of research, exploiting the usefulness of factor analysis, provides a good insight on the relationships between different forms of deprivations and between these and income, it also raises some problematic and unsolved issues which mainly involve the definition itself of the poor. Who are the poor? Those who are experiencing both multiple deprivations and low income (as in their study)? Or those living below some poverty threshold in one or few dimensions?

---

[39] The two poverty lines correspond to the mean of the individual laying in the 20th and in the 40th percentile of $\bar{x}_i$, the vector of scores in ascending order. The same procedure is applied to the computation of monetary poverty lines.

[40] The factor analysis returns some 'new' variables (factors) from the original ones in a form of synthetic indices. These factors are obtained as linear combination of the original variables and are not correlated among each other so that ordinal comparisons of poverty levels can be made using each of these factors.

[41] Hence they exclude from their analysis the other two dimensions. Their argument in applying this arbitrary choice is that secondary and household deprivations are not "relevant to current exclusion and command over resources". In our opinion, the extent to which these two dimensions should not be considered for poverty measurement is questionable.

Other similar aggregation techniques for poverty measurement may include cluster analyses (as in Ferro-Luzzi et al., 2008), multiple correspondence analyses (Asselin and Vu Tuan Anh, 2008), structural equations models (Krishnakumar and Ballon, 2008) or the MIMIC Model applied by Abul Nagaand Bolzani (2008). This model, based on Joreskog and Goldberger (1975), interestingly takes into account the observed variables as manifestations of a latent concept as well as of other exogenous multiple causes which influence the latent factor:

$$\bar{x}_i \sum_{j=1}^{j=k} w_j\, x_{i,j} w_j w_j$$

Clearly, there are many different ways to approach the weighting issue. Statistical techniques may give an objective grounding to particular weights, but it is not clear that statistical associations should dominate an exercise that ultimately is a normative one.[42] Thus a totally different approach to the question of weights could be to make it a subject of public discussion (as advocated by Sen), polls, or surveys (e.g. Clark, 2002).

4.4 Axiomatic Approaches to Aggregation

A second important issue in aggregation, apart from the weights, is the question of complementarity or substitutability between dimensions. Can the fact that I am poor in education be compensated for by being non-poor in health? Or does my education poverty reduce my health achievement as well? Simple weighted averages of component indicators imply perfect substitutability between dimensions while more complex functional forms (e.g. geometric or harmonic means) imply partial substitutability, and yet others allow for none at all.

A partial answer to this question can be given by drawing individuals' utility functions which include the relative weights given by single attributes and also the joint welfare contributions of these attributes, hence conveying information on the direct marginal benefits of each attribute and about the interaction among these (Thorbecke, 2007).

While identification and aggregation are two separate steps in uni-dimensional indices, they are basically inseparable in multidimensional poverty as identification of the multidimensionally poor already entails some aggregation (e.g. someone is poor if the person is poor in at least 3 of 5 dimensions considered).

The related literature distinguishes between three methodological alternatives to address this question: the 'union', the 'intersection' and the 'dual cut-off' methods (see Alkire and Foster, 2010). The union approach identifies individuals as multidimensionally poor if they are deprived in at least a single dimension. Hence, different poverty attributes in this framework can be considered in some sort of sense as perfect 'complements' since a given individual cannot trade-off one attribute for another remaining at the same level of poverty. If her achievement levels fall short of the respective threshold level in at least one dimension, she would be considered poor. In other words, despite that person being non-poor in other dimensions, the complementarities among poverty dimensions are so strong that she will be identified as

---

[42] See Nguefack, Klasen, and Zucchini (2010) for a discussion of this issue in the context of the HDI.

poor. While putting a strong emphasis on achievements in single spheres of human development, this method unavoidably leads to exaggerated multidimensional poverty rates – to the point where it is simply uninformative for policy purposes.

Under the intersection approach, on the other hand, an individual is classified as poor whenever her achievement levels fall below the threshold levels of all poverty dimensions. This method implicitly considers different welfare dimensions as perfect substitutes. A person can compensate her deprivation in one sphere with being well-off in another. While this approach clearly identifies the most deprived, it can be criticized for being overly restrictive, leading to very low poverty rates. Bourguignon and Chakravarty (1997) give the example of an old beggar. If longevity and income are the two poverty dimensions, clearly he who would not be considered poor under this approach.

The social welfare approach adopted by Bourguignon and Chakravarty (2003) has offered an interesting insight on how conditions on the union and intersection can be linked to the assumed properties of the social welfare function. The authors characterize several families of multidimensional poverty indices, relying on the concept of functioning failures according to which the identification of the 'multidimensionally poor' basically implies gathering those individuals who fall below a threshold on each dimension of the individual's well-being.[43] At the same time, they draw on the concept of absolute poverty since the way they determine threshold levels is independent of the attribute distributions. Hence, by hypothesizing the existence of only two attributes of poverty, Bourguignon and Chakravarty propose a set of scalar indices aimed to combine in a single figure the information from those attributes. Their multidimensional measure of poverty is a robust and appropriate statistical tool since it can satisfy a number of desirable postulates which are generally suggested for a single poverty index[44] and – at the same time – being tailored on the possible and likely interactions between poverty attributes – it also satisfies the axiom of 'non-decreasing poverty under correlation increasing switch.'[45]

This issue of interactions among poverty attributes is central in the arguments raised by Bourguignon and Chakravarty, who seek to investigate, using a utility framework, the implications of various degrees of substitutability and complementarity between poverty dimensions. Thinking of attributes of deprivation in terms of substitutes or complements is indeed of crucial relevance not just from a theoretical point of view but also from a more practical perspective. There is a consensus that the relevant spheres which contribute to individual well-being (i.e. education, health, income…) are closely inter-linked and a strong correlation is frequently found both at the micro- and macro-level for some of them. For example, one can demonstrate the effect of nutrition on educational achievements. On the other hand, being educated raises awareness on the importance

---

[43] As they point out explicitly: "the issue of the multidimensionality of poverty arises because individuals, social observers or policy makers want to define a poverty limit on each individual attribute: income, health, education etc..." (pp. 27-28).

[44] Bourguignon and Chakravarty's measure of poverty satisfies strong focus, weak focus, symmetry, monotonicity, continuity, the principle of population, scale invariance, and subgroup-decomposability.

[45] This axiom postulates that if there is an increase in the correlation of the attributes within the population, poverty cannot decrease.

of appropriate hygiene behavior (which is of a great importance for individual health status). Hence, these linkages may suggest that the complementarity between (some of these) attributes should be strong. In a nutshell, a given person cannot easily trade-off some years of schooling for more food on her table because both the education and the health dimensions will work together and help you to function properly.

On the other hand, at least in the very short run, it could be reasonable to assume that the poverty attributes are instead so similar or close to each other that the marginal utility that a given person derives from one of those decreases when the quantity of the other increases. Hence we can think that some degree of substitutability or proximity between attributes exists.

As argued by Bourguignon and Chakravarty, the *ex ante* hypothesis of either complementarity or substitutability among dimensions is not a minor issue – very different results emerge from their study depending on which of the two assumptions is made. In the case of high substitutability, the reduction in poverty caused by a unit increase in income is more important for individuals with very low levels of education than for people who possess educational levels close to the education poverty line. Conversely, if the two attributes are supposed to be complements, the reduction in poverty appears larger for the people with higher levels of education.[46] Hence, in applying this multidimensional index, any poverty outcome would be affected by the choice on the relative weights and different degree of substitution assumed between different dimensions.

The assumption of substitutability among poverty attributes is also considered in the multidimensional well-being indicators constructed by Duclos et al. (2006). Nevertheless, the authors try to avoid any risk of arbitrariness involved in the 'aggregation curse' by means of applying a methodology grounded on the concept of stochastic dominance. In this way they offer an approach which extends the concept of a poverty line in one dimension to a poverty frontier in multiple dimensions. Their methodology has two advantages: first, it generates poverty orderings which are robust to the choice of a poverty index over a broad class of indices ('poverty-measure robustness') and, second, these orderings are also valid for the choice of any poverty frontier over broad ranges ('poverty-line robustness').

As argued by Duclos et al. (2006) the interaction and the correlation among the various dimensions of well-being included in a poverty measure would inevitably cause a divergence in the poverty outcome resulting from any univariate analysis conducted independently in different dimensions from those emerging from a multivariate analysis. Hence, their measure (based on the assumption of substitutability) incorporates the covariance between dimensions and also offers some degree of flexibility since separate poverty aversion parameters can be selected for the dimensions under scrutiny.

Their approach on multidimensional poverty measurement has the merit of taking into account the ethical arguments (as the one related to functionings and capabilities) while at the same time resting on very practical considerations. Indeed, the major here is that

---

[46] The authors apply their multidimensional indices in an analysis of the evolution of poverty in Brazil in the 1980s, experimenting with different degrees of substitutability among the two attributes income and education.

their methodology applies equally well to the union and the intersection approach and leads to poverty comparisons that are robust to aggregation procedures and to the choice of multidimensional poverty lines. Furthermore, even if their basic assumption is that of substitutability among attributes, the authors also discuss the possible implications streaming from complementarities between some attributes (i.e. nutrition status and education or education and life expectancy). Here again, there will be crucial implications for overall poverty outcomes which inevitably should be considered in the formulation and implementation of budgetary rules for the allocation of anti-poverty funds.

While conceptually elegant, it is unlikely that such stochastic dominance approaches will be able to satisfy the demands of practitioners for more precise answers to the quantitative magnitude and distribution of poverty. Just knowing whether poverty went up or down in a stochastic dominance sense might not be enough; also the problem arises that in many contexts the answers will be uncertain as stochastic dominance cannot be established. In those cases, the answer 'it depends' may not be seen as satisfactory. Clearly, there is a trade-off with the demanded precision by practitioners and the robustness of such indicators.

The two multidimensional methodologies just discussed both have the advantage of making use of the information contained in the joint distribution of achievements as it may be transpire from a matrix in which each column contains the vector of individuals' achievement levels in a single dimension, while each row records achievements levels associated with a single person[47] – an advantage, however, which is countered by the requirement of complete data from a single survey.

Another interesting contribution to multidimensional poverty measurement which – in some ways – departs from the two aforementioned studies can be found in the series of papers by Alkire and Foster (2007, 2009) who develop an alternative method of identification, the so-called 'dual cut-off method', which is the basis for the Multidimensional Poverty Index (MPI).

This method is motivated by the notion that both the union and the intersection approach can lead to misleading conclusions, being ineffective in separating the poor from the non-poor. As noted above, the union approach often leads to exaggerated poverty lines, while the intersection approach usually identifies as poor only a very small percentage of the population.[48]

As the name implies, the method proposed by Alkire and Foster is based on two cut-offs: the first identifies deprivations in each dimensions; the second corresponds to the number of poverty dimensions in which an individual has to be deprived to be poor. More specifically, the first step holds that for each dimension d, a deprivation cut-off is

---

[47]Note that, in applied work, it is in principle possible to use different data sources and ignore the link between different dimensions altogether. While this might yield useful estimates of poverty or aggregation deprivation, Alkire and Foster (2011) note that one will be "unable to identify who is multidimensionally poor - a signal benefit of a measurement methodology based on data linked by person or household."

[48] Using the intersection approach ,with ten different poverty dimensions from Indian data, Alkire and Seth (2009) find that only one tenth of one percent of the population would be poor, while under the union approach 97 percent of the population is poor.

fixed so that an individual would be considered deprived just with respect to that single dimension in which her achievement level falls short of the threshold.

Second, a vector of weights or 'deprivation values' is stipulated corresponding to the relative importance of different deprivations. The assignment of weights determines the minimum number of combinations of deprivations required to identify the poor but would also obviously alter the relative contributions of deprivations to overall poverty.

Moreover, Alkire and Foster's method takes into account the breadth of individual deprivations. This is done by introducing a vector of deprivation counts which – in the case of equal weights – simply includes the number of deprivations experienced by the individual or – in the case of different weights – is the sum of the values of the deprivations. The person is counted as poor if (and only if) her deprivation count is on or above this second poverty cut-off.

It is now easy to see how this approach is linked to the union and intersection approaches: If the second cut-off is set equal to the number of dimensions, the approach is equivalent to the intersection approach. Alternatively, if the second cut-off is less or equal to the minimum weight across all dimensions, the identification follows essentially a union approach.

Once the poor have been identified, a simple multidimensional headcount ratio can be computed as the proportion of people who are poor in at least k of all the d dimensions. Interestingly, if adding up the proportion of total deprivations each person suffers and dividing by the total number of poor persons, an average poverty gap can be obtained in order to have a closer insight on the number of deprivations experienced.

Alkire and Foster go on to define an Adjusted Headcount Ratio (M_0), which is obtained through multiplying the headcount poverty by a measure of poverty breadth as the one provided by the average poverty gap.

This measure is subgroup-decomposable, i.e. the adjusted headcount can be decomposed into the contributions of different groups of the population (e.g. ethnic groups). Also, since it adjusts for the size of the group for which it is calculated, international comparisons (across countries with different population size) can easily be done. Moreover, their measure is sensitive to the multiplicity of deprivations so that overall poverty would be affected by changes occurring in the number of people who become deprived in additional dimensions. Lastly, this measure can be broken down by dimension so that a closer inspection on the contribution of a single dimension to poverty in any specific region or by group is made possible.

4.5 The Multidimensional Poverty Index

The Multidimensional Poverty Index (MPI) developed by Alkire and Santos (2010) for the 2010 Human Development Report can be considered as a special case of Alkire and Foster's theoretical formulation outlined above. This index – which has so far been calculated for more than 100 countries –makes use of household-level data, which are then aggregated at the country-level. The MPI takes into account ten variables in three

main dimensions: health, education and living standards.[49] In analogy with the HDI, the aggregate poverty measures in these three dimensions are given identical weights (one-third each), so that a household is identified as poor if it is deprived across at least 30% of the weighted indicators.

As also argued in Ravallion (2011) and Birdsall (2011), a big advantage of the methodology proposed by Alkire and Foster is that of providing a precious insight on the way in which different poverty dimensions can be identified and shared by the same people (as it is done by considering the joint distribution of the multiple dimensions of poverty).

Nevertheless, there are some critical aspects that need to be discussed. First, even if the MPI is claimed to be a comprehensive measure, it actually leaves out many of the multiple dimensions of poverty. The set of material goods which enters into the living standards heading is rather narrow, accounting for even less dimensions of living standards than those included in the more traditional consumption-based measures.

Of course this shortcoming could eventually be justified by the fact that the trade-off between need for inclusiveness and applicability is even more demanding for the MPI because this measure requires that all the data should be obtained from the same household.

A second point concerns the degree of arbitrariness involved in the determination of the cut-off level of 30%. Rippin (2011) argues that changing upward or downward the cut-off level would certainly affect poverty rates and even country rankings. Of course, the Alkire and Foster measure has the big advantage of flexibility so that it virtually offers the chance of exploring the sensitivity of the results in terms of poverty assessment to different calibration choices. However, as Thoerbecke (2011) speculates, it is not clear whether "policy makers and analysts would benefit from such sensitivity analysis and that it would improve the quality of their decision."

A third point is related to the selection of dimensions. Ravallion (2011) argues that the distinction between their intrinsic and the instrumental value has been somehow neglected. The MPI-framework assumes that all the three dimensions have each an intrinsic value. Nevertheless, it could be easily argued that attainments in education or even in health have an instrumental importance since they can also increase individual command over commodities. Furthermore, the equal allocation of weights implies that the three dimensions have exactly the same importance. Yet again, "it is far from clear that education poverty should have the same value as health poverty" (Ravallion, 2011). Also, considering that the living standard component of the MPI simply counts the number of items which a given household lacks, there is neither any assumption of correlation among them nor between them and the health or education component.[50]

---

[49] Specifically, the full list of variables includes: malnutrition, child mortality, years of schooling, school enrolment, cooking with wood, charcoal or dung, not having a conventional toilet, lack of easy access to safe drinking water, no electricity, dirt, sand or dung flooring and not owning at least one of following: radio, TV, telephone, bike or car.

[50] Rippin (2011) points out that this assumption is not realistic: "It is rather safe to say that, for instance, proper sanitation and safe drinking water are related to health as well as education indicators."

A fourth point is the loss of information on dimension-specific shortfalls resulting from their aggregation procedure. Measures like the MPI neglect crucial information on inequalities among the poor themselves. Rippin (2010, 2011) argues that this omission has strong implications in terms of poverty assessment since, it basically implies that a transfers of items from a poor to a less poor household does not change the poverty index as long as both households remain poor according to the MPI.

Rippin (2011) proposes a Correlation Sensitive Poverty Index (CSPI) which – like the MPI – has the advantage of decomposability but also has two appealing characteristics which would vanquish some of the aforementioned weaknesses of the MPI: First, the CSPI does not require the second cut-off (of dimensions in which a household has to be deprived to be called poor). Instead of excluding households from the calculations, it weights each household according to the number of weighted items that it lacks. The absence of an arbitrary cut-off opens the door for researchers and policy-makers to deliberately choose the level of importance they want to attribute to inequality among the poor. Second, it takes into account the inequality among the poor, so that a transfer from a less poor to a poorer household would result in a reduction of inequality.

Fifth, there are a range of problems of detail associated with the version of the MPI proposed in 2010 and used by UNDP. They include the problematic treatment of childless households in the MPI. As there is, by definition, no information on child mortality and educational enrolments in these households, they are automatically considered non-deprived in these dimensions which seems problematic. Harttgen and Klasen (2011) propose an alternative approach to deal with this problem. Moreover, the cut-offs are partly based on the presence of a characteristics and partly on their absence. In particular, a household is non-deprived in educational achievements if a single household member has more than 5 years of schooling (regardless of household size); conversely a household is deprived in the dimensions of nutrition and child mortality if one person in the household is undernourished or one child has died, respectively. One may speculate that this will lead to a bias in the sense that larger households are less likely to be deprived in educational achievements and more deprived in nutrition and child mortality. One way to address this would be to consider average achievements rather than these cut-offs.

Sixth, similar to the 'weakly relative poverty' addition to international income poverty, one may ask the question whether one should also consider a relative MPM that adjusts the cut-offs in the different dimensions to average achievement levels in these dimensions.

Nevertheless, even if multidimensional indices of poverty like the MPI or the CSPI might encapsulate more information and, at the same time, have the merit of clarity and simplicity, they are not universally perceived as superior. Ravallion (2011), for example, argues that measuring and analyzing poverty under a multidimensional perspective is a too complex a task to be summed up in a single index.

The ambition for multidimensionality implied in the 'deprivation-aggregation' approaches followed by Alkire and Foster (2007, 2009), Rippin (2010), Bourguignon and Chakravarty

(2003) and Duclos et al. (2006) create a large area of tension between the coverage of the most relevant aspects and actual and possible measurability and data availability.

A radical approach to circumvent this problem is to simply rely on a set of indicators each covering a different poverty dimension, i.e. use dashboards like the MDGs, which cover multiple dimensions of human development without 'mashing-up' all the information into a single measure.

This can be combined with a more disaggregated analysis, e.g. by building individual poverty profiles which analyze the relative shortfall in each dimension (see Thorbecke, 2011). Notwithstanding the neat decomposability implied in the CSPI or in the MPI, the usefulness of aggregation in the first place is disputable. Rather, it might be more appropriate to use a conventional single-dimension income -based measure of poverty at the household level and then complement this with analyses incorporating dimensions of poverty such as schooling and health of household members which are incompletely or not at all covered through markets.

Of course, even such single measures would require some level of aggregation but the contribution made by Ravallion (2011) is related to the way in which this aggregation is made. Ravallion strongly supports the idea of grounding the analysis on the achievements or attainment space, rather than on the deprivation space. This can be done by replacing the controversial system of arbitrary weights with the use of appropriate prices for market and non-market goods, that is, shadow prices, in order to come up with a composite index for aggregate consumption to be compared to a poverty line defined in the same space.

In order to show the superiority of the attainment-aggregation approach over the deprivation-aggregation approach, Ravallion considers the concept of trade-offs between attainments that are built into any multidimensional poverty index. These trade-offs are nothing else but marginal rates of substitutions (MRS), given by the weight on one attainment relative to another. It is true, indeed, that the bulk of the literature on multidimensional poverty measurement has so far largely ignored or neglected the definition of the marginal rate of substitution among various attainments. Let us consider a simple case of two continuous attainments, x_1 and x_2 with their density functions being f_1 and f_2 respectively. In each dimensions a poverty line is defined (z_1 and z_2) and the weights on deprivation are given by w_1 and w_2 so that $w_1+w_2=1$. Also, for each of the two attainments, two prices (p_1 and p_2) are introduced.

Under the attainment-aggregation approach, the poverty measure is a function of the two poverty lines (like the usual headcount index of poverty) and the weights on the attainments are given by prices, so that the MRS will be simply equal the relative price p_1/p_2 .

If the two poverty lines have been adequately determined so that they exactly represent the point on the consumer's expenditure function, the inverse of the indirect utility function, corresponding to the poverty level of utility, any exogenous welfare change will also change poverty. In other words, attainment-aggregation measures − if calibrated

appropriately – are welfare-consistent measures that reflect the trade-offs as experienced by individuals.

Instead, for the deprivation aggregation approaches the increment to z_1 needed to compensate for less of z_2, holding the value of the aggregate poverty measure constant, would be given by:

$$MRS^D \equiv \frac{w_1 \, f_1 \, (z_1)}{w_2 \, f_2 \, (z_2)},$$

which would not necessarily coincide with the relative prices and so would not yield poverty measures that are consistent with the welfare of someone living at the poverty line. This drawback suggests that there may be a scope for using indices which rely to the greatest possible extent on shadow or market prices to aggregate across different goods and services.

A possible compromise between the approach taken by Alkire and Foster and that one proposed by Ravallion, would be to treat as separate dimensions only those for which it is impossible to get sensible estimates of relative prices (i.e. political and personal freedoms) and – as suggested by Ferreira (2011) – although the items part of the living standards dimension can clearly affect different functionings – their weights should be represented by relative prices.

Clearly, there are various reasonable lines of argumentations to reject the use of prices as long as they are deemed to be missing or unreliable. They can be on conceptual grounds (rejecting the welfarist approach inherent in this line of argument; they can also be on practical grounds, i.e. that for many welfare-relevant goods, markets are missing and prices do not exist or are heavily distorted. In response, Ravallion argues that "it is one thing to recognize that that markets and prices are missing or imperfect and quite another to ignore them in welfare and poverty measurement." The rationale behind the use of weights in lieu of prices can indeed be less strong than it is claimed to be. In particular, two questions seem to be still in abeyance. First: if market prices were available, how and why can the weights used be better than prices? Second: Do weights reasonably echo shadow prices?

But there is also an argument that the focus on finding the right weights is not only very difficult, but possibly exaggerated. As suggested by Ferreira (2011): "[...] looking at a few core, truly irreducible, dimensions, and investigating the properties of their joint distribution over the population, by means of both dominance analysis and a plurality of well thought-out indices, is likely to generate real insights about poverty, and perhaps to contribute to the design and targeting of policy actions to reduce it."

Moreover, it is worthy to note that composite indices can still have their own merits and an indisputable potential to contribute to the change in the main paradigm permeating the discourse on poverty and development. Despite the limitations inherent even in the most technically refined methodologies, composite indices like the HDI or the MPI still play a well-deserved prominent role. As highlighted by Birdsall (2011), they contribute to

more effective policies at the technical level ('technical policy rationale'), but also have a great potential to alter the discourse of priorities ('conversation changer rationale') and to communicate better in order to get more support for change ('advocacy rationale'). The challenge is therefore to retain these benefits while trying to generate multidimensional poverty measures that are technically as sounds as possible.

5 Proxy Means Indicators

The framing of poverty lines and the identification of the poor – at least according the traditional approaches (as discussed in section 2) – are strongly based on individual or household information on daily expenditure. While for pure measurement purposes, a comprehensive income-based approach to poverty measurement (discussed above, e.g. using the CBN method and a comprehensive household survey) is clearly warranted, for policy action and program evaluation often reduced approaches are required. For example, policy-makers want to target transfers to the poor and therefore need a quick way to identify the poor reasonably accurately, without relying on a full household survey with all the costs and time lags involved. Another example is that a government or a donor would like to monitor the poverty impact of a local intervention where existing household surveys provide too little coverage and a full-scale local survey is not feasible.

In these situations, one of the most common approaches used to assess poverty is known as 'means testing' which tries to ascertain the income level of a household relying on reduced information. Bitran (2000) suggests four different types: Simple Means Test (SMT), Verified Means Test (VMT), Proxy Means Tests (PMT) and Mixed Approach (MA).

Simple Means Tests have been widely used both in developed and developing countries. They are based on reported household income, size and composition. If this method is very straightforward and requires easy available administrative resources (i.e. the staff to conduct interviews and to keep record of those), the level of inaccuracy associated with it can be problematic. Moreover, strategic behavior of individuals may lead to systematic biases. In order to become eligible for a program, it is in the interest of potential beneficiaries to understate their incomes (Atkinson, 1995).

In developing countries, there are additional factors linked to the seasonality of agriculture or the instability of employment in the informal sectors which render the incomes of the poor extremely vulnerable or subject to high volatility.

Moreover, as it has been partly argued in previous sections, income is neither a perfect nor an accurate measure for individual welfare. It generally does not account for important elements such as the value of goods produced in the household, gifts, transfers or own occupied housing. In SMT, there is no attempt to value in-kind income, seasonal income or to consider individual adjustments in needs or means.

Also, apart from occasional visits to household to check household quality, reported income is not systematically verified. The incentive problem mentioned above of course

also applies to developing countries, possibly more so as verification is particularly difficult. Verifying the extent of these misstatements can be a costly and time-intensive operation especially in those contexts where reliable records do not exist.

A more sophisticated and accurate way to conduct means test are VMTs (also called 'sophisticated means tests') which adjust family income according to family size and make use of additional information such as seasonality and costs of major items like housing, university tuition, and major medical expenses.[51] Even if more accurate, these tests entail high administrative costs requiring the existence of verifiable records in the target population and the administrative capacity to process the information and update it regularly. These weaknesses may render them infeasible for poverty assessment in developing countries.

Moreover, here as well there is a risk of bias in the estimates since verification of family income may work more likely for literate respondents or for those working in the formal sector. People with very low levels of education would indeed not be able to answer properly some specific questions.

Hence, an alternative option would be that of making use of observable characteristics of the household (variables which are commonly provided in surveys) which can be thought as proxies for income or welfare. PMT or *indicator targeting* are indeed based on predictions of welfare from data on household characteristics like asset-ownership, demographic characteristics, and educational attainment and have been conceived and developed in order to identify potential beneficiaries for poverty reduction programs.

This method can be very useful in circumstances in which existing databases (i.e. Demographic and Health Surveys) cannot provide any information on income or expenditures or data on income or expenditures are not available in the form required for specific evaluation purposes. Also, PMTs can be applied in cases in which the scope (i.e. interdisciplinary research) or the (limited) budget of some intervention would not require or would just prevent the collection of detailed income or expenditure data (Johannsen, 2006).

From the methodological side, the literature on PMTs offers different possible approaches in constructing the indicators. Early theoretical contributions on PMTs include Ravallion and Chao (1989) who propose a methodology to solve the targeting problem of poverty alleviation with imperfect information on recipients' incomes. The authors develop an algorithm which makes optimal use of the (limited) available information so that the resulting targeting scheme will allocate equal transfers to all individuals within a group, but different transfers among groups with different characteristics. Their optimal targeting algorithm is applied (as an illustrative example) to

---

[51] In most cases, income, assets and expenses are recorded and verified. The level of earned and unearned income is verified using documents over the last 2-3 months such as: pay stubs, employer wage statements, benefit letters from other programs (pensions, social security, etc.). Cash and non-cash assets are checked using banking statements, life insurance policies or vehicles documentation. Concerning households' expenditures, respondents might be asked to provide most recent utility and medical bills, written statements of child care costs and real estate tax bills (see Lindert, 2005).

data from Bangladesh, Indonesia, the Philippines, and Sri Lanka in order to find the poverty-minimizing allocation. This algorithm is also applied in later studies on regional targeting (Ravallion, 1989; Datt and Ravallion, 1993; Ravallion, 1993) whose results stress how single indicators at the macro level can be of limited use.

A similar approach can be found in Glewwe (1990) who derives weights for each household variable by solving the poverty minimization problem. Nevertheless, as remarked by Grosh and Baker (1995), even if this method can be theoretically more appropriate, the results obtained from poverty minimization techniques are not very dissimilar from those based on regression analysis.

Indeed, most research on proxy means testing is based on OLS regressions. [52] These approaches aim at predicting per capita household consumption by estimating the parameters of a regression of log household consumption per person on a set of variables deemed appropriate in this exercise, e.g. geographic, demographic, and socio-economic indicators.

There will usually be some requirements concerning the explanatory variables in PMT other than having high explanatory power with respect to the welfare measure. For example, it is paramount that they be easily measurable and verifiable. Moreover, they should be incentive-compatible, i.e. they should not be amenable to manipulation by households (Atkinson, 1995, p. 26). Other considerations include their degree of variation over time (in order to allow for households to gain or lose beneficiary status) and the political feasibility[53] to construct a targeting instrument based on these variables. Practitioners will often encounter trade-offs between respecting the above criteria and targeting accuracy. In order to assess this trade-off, classifying variables into different categories along these lines and testing the prediction accuracy is good practice (e.g. Grosh and Baker, 1995).

One of the earliest contributions following this approach is the study by Glewwe and Kanaan's (1989). The authors use data from Côte d'Ivoire to demonstrate how household characteristics can be used to proxy a means test. The first step in their research is based on regression analysis to predict welfare levels based on several combinations of variables that are easy to measure. In a second step, the authors assign to each individual a transfer equal to the difference between their predicted

---

[52] It is important to note that the OLS procedure might seem inappropriate in predicting poverty. Not only are many of the explanatory variables likely to be endogenous to household welfare but the minimization of the squared errors between the true and the predicted levels of welfare is also theoretically distinct from the minimization of poverty (Grosh and Baker, 1995). Grootaert and Braithwaite (1998) show that in the specific economic context of transition economies, there might be quantitative differences between the indicator weights resulting from welfare and poverty regressions. Nevertheless, the main justification applied in the use of OLS in this specific context is built on the fact that the identification of the poor is of interest rather than causal inference. Also, OLS can be considered convenient and useful when a large number of predictor values are available. Finally, OLS produces are appealing to policy makers since the weights are easy to interpret.

[53] For example, the use of proxies indicating geographic region, ethnicity, etc. might stir political discontent (e.g. Ahmed and Bouis, 2002).

welfare level and the poverty line. Third, they use the FGT family of poverty measures[54] to compare the outcomes of various targeting schemes with untargeted and perfectly targeted transfers. As their main findings suggest, simple predictions based on OLS can considerably improve the poverty impact in comparison to untargeted transfers.

A study conceptually identical to proxying a means test is the one conducted by Haddad et al. (1994) who – by making use of household survey data for four different countries – seek to find some alternative variables that would predict food and nutrition security. Their results suggest that some readily-available variables could serve as adequate proxies for standard measures of food and nutrition security like caloric adequacy which may be more difficult to collect.

Grosh and Baker (1995) estimate OLS regressions and carry out simulations for three countries (Jamaica, Bolivia and Peru) in order to identify what kind of information can best be applied in a proxy means test and determine the level of accuracy of such tests in targeting the intended beneficiaries of a given anti-poverty program.

Per capita household consumption expenditure is chosen as the welfare measure and with a step-wise function the variables which are not statistically significant and do not increase the model's overall explanatory power are dropped so that a successively smaller set of correlates of consumption (and hence of welfare) are retained.

Their initial larger set of variables includes location variables, dwelling characteristics, family characteristics and ownership of durable goods. Four models with different possible combinations of these sets of variables are explored. Also, another model is built by using only the variables in each category which are the best predictors within their class.[55] As expected, their results suggest that while models with verifiable indicators such as location and dwelling can correctly identify only a small fraction of the individuals that should receive benefits, models with less verifiable indicators (i.e. adding family characteristics) correctly identified about half of the intended beneficiaries. Moreover, they find that despite significant undercoverage and moderate leakage, the derived targeting scheme would still have a higher impact on poverty than no targeting at all.[56]

An extension of the methodological approach proposed by Grosh and Baker can be found in the analysis devised by Johannsen (2006). The author, while relying on the

---

[54] For a more detailed discussion on FGT measures see section 2.2.

[55] From their results it appears that the variables with the highest predictive power are location of the households, household size, access to electricity, availability of toilet facilities and the ownership of a telephone.

[56] Undercoverage and leakage rates are both indicators for targeting accuracy. Undercoverage is the percentage of those whom the program is designed to cover but is not covered. This measure is based on the so called Type 1 (or false negative) error of targeting which identify the individuals whose "true" welfare level is below some cut-off threshold but whose predicted welfare is above. The leakage of a given program streams, instead, from the Type 2 ("false positive") error. It identifies the percentage of program benefits that is received by people which are truly not eligible to receive them. Of course, the cost resulting from this error can be very high since benefits are given to people who are actually not in need of them.

same conceptual assumptions, introduces some technical refinements that allow her to propose a poverty assessment tool for targeting in Peru consisting of sets of indicators characterized by high predictive power of per capita daily expenditures. More specifically, by relying on a household surveys data which nicely cover several poverty dimensions, the authors identify 142 potential predictors and makes use of the maximizing-R squared regression technique in order to identify sets of consecutively increasing numbers of indicators while maximizing the explained variance in every step. As also argued by Zeller et al. (2006), this approach proves to be well-suited to cope with a large number of both discrete and continuous variables.

With this procedure, three smaller sets of five, ten and fifteen indicators are identified and tested for their accuracy in predicting poverty by means of taking as a benchmark the national poverty line and the median poverty line (defined as the corresponding expenditure cut-off at the bottom 50% of the population below the national poverty line). Hence, the subsequent classification of households relies on a transformation of the poverty line based on a percent point function of expenditures that makes it possible to determine a poor/non poor classification which is actually based on the poverty headcount corresponding to some national or median poverty line.[57] This innovative approach implicitly takes into account the fact that the residuals in the estimated regressions might contain additional information on possible other poverty determinants so that the bias in final estimates would be less grave (see also Ravallion, 1998).

These fine methodological devices appear superior in terms of targeting accuracy.[58] Overall, a tool of indicators as the one proposed by Johannsen can have the advantage of being a less costly option to gauge poverty either in terms of the ex *ante* identification of potential beneficiaries of a given project or program or in terms of the *ex post* assessment of the impact the intervention had on the poor.

While, as discussed above, the bulk of the literature on PMT as well as that of their practical implementation has mainly relied on regression techniques, a parallel research strand bases the construction of proxy means indicators in a different methodological approach.

These procedures largely use PCA in order to derive some weighted poverty index that should be able to predict households' living conditions. Clearly, even if virtually both

---

[57] More specifically, this methodology is grounded on two alternative steps which deliver useful information for poverty assessment. First, the two empirical cumulative distribution function of the observed and the predicted expenditures are taken into account in order to get some 'percentile corrected' predicted expenditures which can be directly compared to the corresponding poverty line. Secondly, a percentile-corrected poverty line can be derived from a comparison of the percent point function of the observed expenditures to the true poverty headcount so that it would identify the value of observed expenditures that corresponds to the households laying close to the poverty headcount. The correction of the poverty line can prove to be extremely useful because – as reported by the author – "it provides the possibility of expenditure predictions using the poverty assessment tool in independent new samples without the need for information on observed expenditures, which constitutes the main purpose of an operational indicator-based poverty assessment tool."

[58] Their proposed tool of 15 indicators correctly identifies 80 to 84 percent of the poor (if the national poverty line is applied as a benchmark).

procedures are inspired by the same principle of finding an appropriate proxy for income, in practice, they are very different statistical techniques. While the economic interpretation of the welfare proxies obtained though regression analysis is direct by construction, the interpretation of the ones captured through PCA is not immediately straightforward (Castano, 2001) and the predictive capacity of the method itself is more controversial compared to regression analysis.

The rationale for using PCA to obtain some indicator is found mainly in the degree of subjectivity in weights assignment which is often inherent in the aggregation procedures entailed in the construction of an index. Compared to the regression approach, one may also argue that the PCA approach can better capture the multidimensional nature of poverty. Clearly, the regression approach is likely to be superior if the dependent variable (e.g. expenditure or income) is the only variable of interest, and if one can assume that the variable is well-measured and free of biases. If one allows for the fact that expenditure is a biased indicator of welfare (e.g. due to fluctuations, measurement error, instability of the regression model, or the multidimensionality of poverty), then a PCA approach can possibly better capture the welfare level of a household.

In this sense, PCA can indeed be a rigorous methodology since it is able to isolate and measure the different poverty components detected in the various indicators by means of assigning weights that reflect each indicator's relative contribution to the overall poverty component.

Here, as well, the information come from survey questionnaires which can cover to a large extent different poverty dimensions spanning from the ones which can be qualified as means to achieve welfare (i.e. variables which reflect the level of human, physical and social capital in the household) to those ones which qualify as achievements for satisfying present and future basic needs (e.g.. access to health services and safety nets, environmental quality, etc.) or else to psychosocial and political dimensions as they may result from subjective evaluations of poverty.

Basically, through PCA, it is possible to obtain a new variable as a linear combination of the original indicators and, thus, to derive a measure of welfare in the form of a single index summarizing all the relevant information contained in the various poverty indicators used as inputs.

In analogy with the regression approach to means testing, this method offers a low cost and flexible option to be applied across very diverse socio-economic settings in order to obtain a credible set of reliable indicators for identification of the poor. Moreover this method seems to offer a very appealing tool to policy makers whenever they aim to identify the socio-economic characteristics of the poorest section living in some district or region and to assess the extent to which these potential targets are reached by different types of interventions.

Since households can be ranked by their poverty levels, cross-area comparisons can be easily done. More interestingly, this method has the potential to track the poverty status

of a given household or individual over time thus allowing also for longitudinal comparisons and eventually offering the opportunity to make causal analysis.

On the other hand, as pointed out by Zeller et al. (2006), this method cannot provide any information on poverty as absolutely defined with respect to some national or international poverty line. It seems safe to say that it should only be applied when the resulting variable (i.e. the first principal component) is seen as an acceptable poverty measure. If (adjusted) per capita-expenditure is seen as the relevant indicator of well-being, regression techniques will in general be preferable as they are tuned to explicitly 'learn' the relationship between proxies and outcome variable.

The relative merits of the regression-based and PCA methods depend to a great extent on whether one wants to use the PCA method to simply generate a proxy for income or whether one wants to use it to generate a multidimensional poverty index. This relates back to the weights discussion we have had above.

Thus far, it has often been claimed that the proxy indicators (either through regression analysis or PCA) have often been used successfully in many countries to assess the poverty outreach of different types of interventions like micro-finance projects, health, nutrition or education programs. In particular, the method is argued to be objective and transparent (e.g. Skoufias et al., 1999; Grosh et al., 2008) and that proxies can be easily measured and verified by social workers (Coady et al., 2004a). In their meta-study of targeting schemes, Coady et al. (2004b) conclude that "[p]roxy-means testing […] show good results on average but with wide variation." These studies tend to overlook the political and social ramifications that even moderate errors in targeting might have (see Sen, 1995b, for a general discussion of the political economy of targeting, and Adato, 2000, and Adato and Roonpnaraine, 2004, for a discussion on the consequences of targeting error in Latin American programs). The optimistic view of the technical efficiency of proxy-means testing, for that matter, has recently been questioned by Kidd and Wylde (2011). Assessing the prediction accuracy of the regression-based approach with data from for different countries, they find that "PMT has high built-in errors […]". Clearly, given its popularity with practioners, more research on the effectiveness of proxy-means testing is warranted. On the other hand, a full appreciation would consider the pros and cons of alternatives, i.e. targeting mechanisms like community-based targeting.

At the same time, there might be some scope for improving on the methods used in practice. Since one is not primarily interested in obtaining causal or even unbiased estimates of welfare, it might be worthwhile experimenting with more sophisticated statistical methods that could increase prediction accuracy at the cost of allowing for some bias. However, there is clearly a limit to applying ever more sophisticated methods as the resulting instrument has to remain interpretable. Put differently, the method applied should not turn a targeting scheme into a black box.

6 Subjective Well-Being

6.1 Background

The first attempts to define and measure subjective well-being (SWB) go back to the 1950s when a large part of the psychology literature shifted its main focus of investigation from the negative emotional states to the positive emotions and subjective assessments of well-being. One of the most authoritative definitions of SWB is found in the psychology literature and holds that SWB is an attitude consisting of the two basic aspects *cognition* and *affect* (Lucas et al., 1996). The first term refers to the rational or intellectual aspects of SWB and is usually assessed with measures of satisfaction; the second refers to emotions and individual. More generally, SWB reflects people's instant evaluation of the events that occur in their lives. Hence, even if the terms happiness and SWB have often been used interchangeably, they actually describe very different concepts: SWB is a "broad category of phenomena that includes people's emotional responses, domain satisfactions, and global judgments of life satisfaction" (Diener et al., 1999, p. 277). The concept of happiness, on the other hand, is linked strictly to affect and can thus be measured from *experiential* questions, for example, about feeling happy or smiling or not being depressed in the days preceding the interview.

For a long time, economists have been reluctant to rely on subjective assessment of individual utility (e.g. Robbins, 1935). Inter-personal comparisons based on cardinality of utility were seen as meaningless. Instead, all welfare judgments were ordinal and derived by resorting to the Pareto criterion (e.g. the weak axiom of revealed preference). As a result, one can at best obtain partial orderings for different states for a given individual by observing her consumption behavior. Due to the well-known Impossibility results by Arrow and others, however, the grounding of welfare judgments on this narrow informational base has made axiomatically grounded welfare judgments impossible without adding more information or relaxing some of the stringent assumptions (see Sen, 1995a, 1999).

The emergence of research on SWB over five decades ago has challenged this view to some extent. The objectivistic assumption that utility changes should only be derived from observed choices has been questioned by the development of theories provided in the behavioral economics literature, where the main focus has been shifted from the outcome utility to the procedural utility that people gain in reaching a certain goal (see Kahneman and Thaler, 2006).

In addition, the improvements achieved in framing and conducting surveys have certainly contributed to an improvement in the measurement of well-being under this new angle. Answers to questions about personal emotions, self-esteem, and perceived status and well-being can provide the researcher with a huge amount of information on individual lives. Individuals evaluate their level of SWB under a 'relativistic' perspective, i.e. in comparison to different scenarios, and with respect to other persons or past events or even future counterfactuals. Hence, as remarked by Frey et al. (2004), measures of reported SWB can thus be thought of as a proxy for individual utility.

Given the limits of the standard objectivistic approach, a complementary view based on subjective assessment of personal utility can offer the researcher a perspective on individual welfare (and even on individuals' command over resources) which, all in all, may yield a more complete picture of reality.

In the remainder of this section, we will describe main measurement techniques of SWB (scales and methods), discuss issues related to the validity and reliability of SWB measures, and discuss the relationship between SWB and traditional economic welfare measures and its usefulness in poverty measurement.

6.2 Measurement of SWB

6.2.1 Approaches to Measurement

The field of SWB has its roots in survey research and, more specifically, in self-reports: people are asked to evaluate their lives (generally) with very simple and straightforward questions. These surveys can embody a huge and useful amount of information regarding several different dimensions which contribute to the enhancement of a people's well-being (such as income, health, family, social network or also human and political rights).

The main source of empirical evidence on life satisfaction in many countries around the world is The World Values Survey. In this survey, devised by a network of academics around the world, typical interviews may include questions like these: 'Taking all things together, would you say you are: very happy; quite happy; not very happy; not at all happy?' or 'All things considered, how satisfied are you with your life as a whole these days?'

These interviews have been conducted with individuals from more than 80 countries, representative of approximately 85% of the world population. To date, the World Values Survey has been carried on in six waves: 1981, 1990, 1995, 1999, 2005, and 2010.[59]

More recently, Gallup conducted a World Poll using questions which employ a ladder analogy. Respondents are asked to imagine a ladder in which each rung represents a successively better life. The questionnaire administered to individuals aged 15 and older from 132 countries, covers several specific dimensions of well-being but also the more general concept of life satisfaction. The advantage of the World Poll obviously lies in its universal coverage which allows for international comparisons.

Both the World Values Survey and the World Poll rely on multi-item scales in which the researcher basically uses several items to target specific parts of SWB. The first examples of this procedure is the Positive and Negative Affect Schedule (PANAS) proposed by Watson et al. (1988) and the Satisfaction with Life Scale (SWLS) devised by Diener et al. in 1985. While in the former, respondents were expected to indicate the extent to which they felt a specific emotion during a recent period, for the later they were

---

[59] For details, see http://www.worldvaluessurvey.org/.

asked to state to what extent they agree or not with a set of statements they are provided with in the questionnaire.

However, also single-item scales could – in principle – be applied. The pioneering surveys on individual well-being made use of single-item scales (Gurin, et al., 1960; Cantril, 1965; Andrews and Withey, 1976). While their advantage clearly lies in their conciseness, they also have several shortcomings. Diener (1984) argues that "with reliance on a single item, the variance due to the specific wording of the item cannot be averaged out. Because it is impossible to obtain estimates of internal consistency, usually the only estimate of reliability for these scales is temporal reliability, in which it is difficult to separate true change from measurement error." More generally, these single-item scales by definition cannot offer a broad and differentiated view of individual SWB and, consequently, result in a substantial loss of information.

Another methodological issue closely related to the framing of the questionnaire and the choice of a scale concerns the choice of the approach taken to obtain SWB reports. To date, the three most common measurement techniques employed to self-appraisals are the Experience Sampling Method (ESM), the Ecological Momentary Assessment (EMA) (Stone et al., 1999; Scollon et al., 2003) and the Day Reconstruction Method (DRM) which have been proposed recently by Kahneman et al. (2004). All of these three methods are flexible in the choice of scale adopted to assess SWB. When ESM or EMA is applied, respondents are expected to complete their reports on a regular basis while staying in their natural surroundings. Hence the two-fold advantage of such techniques lays on both place and time. The probability of obtaining distorted results due to unusual circumstances is lowered since individuals compile their reports in a familiar environment. Also, high-frequency assessments which are given 'on the spot' decrease substantially the risk of retrospective biases.

A more practical tool is the DRM which requires subjects to keep a diary in which events or episodes are recorded. Participants are asked to keep track of specific activities during different times of the day and on emotions associated with these. Hence, respondents describe each episode by indicating where and when specific episodes began and ended, what they were doing (by selecting activities from a provided list) and with whom they were interacting. In order to assess how they felt during each episode with respect to selected affective dimensions, respondents are asked to report the intensity of their feelings along nine categories on a scale from zero ('not at all') to six ('very much').

Clearly, even though these assessments are rigorous, by definition, subjective survey data are based on individuals' judgements. Hence, the risk of systematic and non-systematic errors would be relatively higher compared to other, more 'objective' measures of welfare or utility. Additionally, as argued by Frey et al. (2004), "reported subjective well being may depend on the order of questions, the wording of questions, scales applied, actual mood and the selection of information processed."

On the other hand, it is important to note that a large part of the research conducted in this field has pointed out the existence of a robust correlation between answers to SWB

questions and more objective measures of individual well-being (like physical evidence of affect such as heart rate measures or electrical activity in the brain) or supplemental subjective data on well-being, i.e. independent evaluations by friends or self-reported health. Moreover, self-appraisals on SWB have been shown to be highly correlated with major changes in life circumstances such as marriage, unemployment, divorce (see, *inter alia*, Lucas et al., 2004; Lucas, 2005; Luhmann and Eid, 2009).

Overall, individual self-assessments seem to have a meaningful interpretation and, therefore, some potential for contributing to welfare assessments. On the other hand, they should always be subjected to testing in order to determine whether the requirements of *reliability* and *validity* are fulfilled.

6.2.2 Testing SWB Data

Reliability refers to the consistency of a SWB indicator, to its ability to give the same results in repeated measurements. In the related literature, this is done by looking at the test-retest correlation of the measure of interest. It is important to note that the specific measurement scale used and the time lags between interviews are both factors which can substantially influence the results of a test.

Validity embodies several meanings. First, as argued by Larsen and Fredrickson (1999), a SWB measure should be able to capture the construct it is intended to capture (*construct validity*). Second, there should be correspondence between the indicator of interest and other indicators aimed to measure the same construct (*convergence validity*). Finally, there should be no correspondence between a given SWB indicator and other measures aimed at defining some opposing construct (*discriminant validity*).

One of the big challenges for the validity of SWB measures is their demonstrated sensitivity to minor life events (Schwarz and Strack, 1999; Bertrand and Mullainathan, 2001). External circumstances might affect the level of SWB reported.[60] However, since the nature of this type of error appears to be more random than structural, using large samples would mitigate the problem.

Another important challenge to the validity of SWB measures streams from the possible effects that differences in cultural and linguistic factors may have on SWB measures. The resulting bias will obviously affect the validity of international comparisons based on SWB measures. It is unclear how important these factors are: while Vittersø et al. (2005) show that culture can bias the results, Diener et al (1995) find that it does not explain the observed differences in SWB.

Finally, several studies have found a relationship between SWB and biological phenomena like brain activity, blood pressures, levels of salivary cortisol or neuroendocrine see, for instance, Lane et al., 1997; Levesque et al., 2003; Ryff et al., 2004; Steptoe et al. 2005; Blanchflower and Oswald, 2007). These findings strengthen the validity of the indicator and allow for better comparability across heterogeneous

---

[60] To some extent, this is even desirable if one wants to test the impact of such events; but it can be a form of measurement error if the event is unknown to the researcher.

individuals from different cultural and linguistic backgrounds. Nevertheless, research in this field is still far from establishing a consensus concerning the question of causality between SWB and these biological phenomena.

6.3 SWB and Objective Measures

If a given SWB measure is proven to be adequately valid and reliable, the next logical step would be that of modelling this measure in a function which can generally be written as:

$$W_{it} = \alpha + \beta X_{it} + \epsilon_{it},$$

where *W* is our SWB measure of interest and *X* is a vector of known variables for individual *i* at time *t*. This vector containing the correlates and the determinants of SWB can actually be very large. Indeed, the literature in this field has shown that differences in ratings can be explained by a set of very diverse variables such as demographic, institutional, and economic factors as well as psychological and contextual and situational factors.[61]

Several studies investigate the relationship between income (or, more generally, economic growth) and SWB. The seminal paper in this literature by Easterlin (1974) revealed an interesting paradox: "In all societies, more money for the individual typically means more individual happiness. However, raising the incomes of all does not increase the happiness of all. The happiness-income relation provides a classic example of the logical fallacy of composition — what is true for the individual is not true for society as a whole." In other words, his arguments suggest that there is no link between the level of economic development of a society and the overall happiness of its members. The contrasting finding that happiness increases with income at the micro-level is supported by additional empirical evidence in the literature (and quite robustly so) (e.g. Wolfers and Stevenson, 2008).

Many researchers have tried to explain this result by the hypothesis of a satiation point in happiness. Veenhoven (1991), for example, argues that more income increases happiness only until basic needs are met. If the society reaches a certain level of wealth so that people are no longer hungry, children don't die from severe diseases and absolute poverty in general has been eliminated, additional gains in income are irrelevant for happiness. This result is quite common in the literature: Deaton and Kahnemann (2010) find that emotional well-being rises with log income, but that there is no further increase beyond an annual income of $75,000. Layard (2005) finds that absolute income is only important for happiness when income is very low. In particular, after national GDP exceeds the threshold of $15.0000 per capita, the level of happiness appears to be independent from income per capita.

Obviously, the relationship between these two variables can have substantial repercussions from a policy perspective. Layard advocates aggregate SWB measures to

---

[61] Technically, as the dependent variable is commonly measured on an ordinal scale, the estimation techniques applied are based on logit or (ordered) probit regressions.

58

be the only gauges used to evaluate policy and progress.[62] If social well-being cannot be improved through economic growth beyond a certain point, governments could in principle opt for diverting more resources to the maximization of well-being, implying a greater focus on policies aimed at reducing unemployment, inequality, and other social problems.

Other findings supporting the existence of this concave relationship between income and happiness are provided by Di Tella et al. (2007). Using a micro-panel of German households and a dynamic specification, they find that income growth provides only a temporary boost to life satisfaction.

Deaton's (2008) analysis on the Gallup World Poll[63] also yields several interesting results: first, high-income countries have greater life satisfaction than low-income countries and. Second, when GDP per capita is measured in logs, there is no evidence that cross-country effects of greater income vanish as countries increase their income. Third, while all of the self-reported measures that Deaton analyses are strongly correlated with log GDP per capita, changes in self reported measures are uncorrelated with changes in log GDP per capita, consistent with Easterlin's hypothesis.[64] Fourth, in contrast with previous micro-based empirical evidence, Deaton finds a negative effect of economic growth on life expectancy *conditional on the level* of GDP per capita. Lastly, he interestingly finds that levels and changes in objective measures of health status (e.g. life expectancy and HIV prevalence) do not strongly affect life or health satisfaction.

Taken together, it seems that SWB measures can serve as a useful indication of how people experience their everyday lives. Analyzing the changes in these indicators over time would we be at least a worthwhile supplemental exercise for global monitoring of progress. As suggested by Deaton (2010), however, it will be important that the scientific community comes to a consensus on what the expectations are on the direction and degree of correlation between income and measures of SWB: "if these measures were perfectly correlated with average income, they would not be useful; we are tempting to measure something other than income, something that holds out the hope of measuring the living standards near the bottom of the distribution of the world."

---

[62] As argued by Angus Deaton (2008), Layard's view on this point is in contrast with the more accepted view based on Sen's capabilities approach which holds that different capabilities all have their own value in human life and well-being so that – in general – they cannot be aggregated into a single measure.

[63] Deaton analyzes self-reported assessments of individual well-being from the following questions: (*a*) the 'ladder'-question, which asks people to imagine a ladder whose bottom rung, 0, represents the worst possible life for them, and whose top rung, 10, the best, and to report on which rung they stand at the present time; (*b*) "Have there been times in the past twelve months when you did not have enough money to buy food that you or your family needed?" (yes/no); (*c*) "Which one of these phrases comes closest to your own feelings about your household's income these days? 'Living comfortably on present income', 'Getting by on present income', 'Finding it difficult on present income', or 'Finding it very difficult on present income'; (*d*) "Are you satisfied or dissatisfied with your standard of living, all of the things that you can buy and do?"

[64] Nevertheless, Deaton also remarks that the Easterlin hypothesis cannot be settled with these findings since his own analysis relies of a limited time span (3 years of data) and – more importantly – because both the Gallup measure and the GDP are subject to sampling and non-sampling errors.

6.4 SWB and Poverty Measurement

While SWB measures are promising indicators that can enrich our understanding of well-being, their application to the explicit measurement of poverty has been limited. In principle, one could define the poor as those whose score on happiness or life satisfaction is below a certain cut-off and then study the determinants of poverty.[65] While a way forward, there are some caveats with this approach. First, to the extent that individual responses are distorted by recent personal events, this would induce measurement errors in the analysis of poverty (particularly if these events are not recorded in the surveys). Second, the comparability over time and across space (nationally and, even more so, internationally) might be limited. Third, often only one individual within a household in asked about SWB. If that person is dissatisfied, does that make the entire household poor according to SWB? These issues have not begun to be tackled yet. Lastly, it appears that we still know too little about the long-term drivers of SWB and happiness and therefore it appears far too early to decide to orient policy towards improvements in these scores.

While keeping these limitations in mind, it should be acknowledged that subjective data can be an important source of welfare-relevant information that is usually not available from more 'objective' data. In particular, it seems worthwhile to further explore its relationship with objective measures and its usefulness in calibrating multidimensional welfare measures.

To date, there are at least three different strategies to anchor poverty measures in subjective assessments. The first one has been adopted by Ferrer-i-Carbonell and Van Praag (2005) in their comparative analyse of objective and subjective poverty measures using Russian household survey data and also by Kingdon and Knight (2006) who aim to check if their idea of subjective poverty as an 'encompassing measure' can be verified using South African surveys. Their method calibrates composite welfare indices by regressing self-related assessments on income and a set of explanatory variables so that the estimated coefficients can be used as the weights. The second approach is the one discussed in section 3.3, the computation of a poverty line in the space of per-household expenditure. Third, as outlined by Ravallion and Lokshin (2006), one can use subjective data to test objective poverty lines. This can be done by regressing subjective assessments of individual welfare on per capita household income normalized by the poverty line and a set of other variables that were used to construct the poverty line. If the objective poverty lines accord with subjective welfare, the covariates should be jointly insignificant.

As argued before, there could be variations in SWB assessments which may not be attributable to poverty status but to individual idiosyncrasies (e.g. personality traits) unrelated to well-being. In that case, individuals with the same income may provide very

---

[65] Rojas (2007), for example, relying on a large database of Mexican households, adopts a SWB approach to estimate equivalence scales which are used to make welfare comparisons for persons living under different household arrangements (i.e. in households of different sizes and with different age composition of household members). His results suggest that if scales other than the subjective one are used, there are substantial biases emerging in the assessment of Mexico's income distribution and poverty rates.

different answers when asked about their personal situation. Also, the same person may answer these questions differently at different points in time, even if her real expenditure did not change.

Hence, there is a real danger of obtaining biased and non-informative estimates if one relies solely on individual assessments. It therefore seems preferable at this point to use SWB assessments for calibration purposes only (i.e. for the derivation of poverty lines), overcoming the gap between the objective approach (as discussed in section 3) and SWB.

In short, SWB approaches have made great strides recently in advancing the measurement of aggregate well-being, and also to compare such aggregate well-being across space and time. The extension of this approach to poverty measurement is still in its infancy, and much more research and analysis is required before one can assess the merits of using SWB to define and measure poverty at the individual or household level. Our discussion suggests that many conceptual and empirical obstacles would have to be overcome first.

7 Evaluating Advantages and Disadvantages of Different Approaches

As the survey has demonstrated, there is no poverty of poverty indicators! While only 25 years ago, no data to assess poverty was available in most developing countries, researchers, policy-makers and practitioners now face an embarrassment of riches. They can choose among a wide variety of conceptual and empirical approaches to measure and track poverty at all levels of aggregation, from local to global. So a first point of note is that poverty research has come a very long way in the past two decades. The question that now arises is which of those measures to choose. Quite naturally, there is no reason to just select one of the many measures we discussed as the one and only poverty measure to use and track. In fact, the converse has more advantages. As all of these measures carry some useful information on the deprivation of people in developing countries, it is useful to consult all of them for a particularly comprehensive picture of poverty. But for many intents and purposes, it is not useful to measure poverty using the entire dashboard of approaches discussed here. Thus we will discuss in this section the relative merits and problems of the different measures we proposed which may then guide academics, policy-makers and practitioners on the choice of measures, depending on the problem at hand.

Table 1 tries to summarize this information in one table. In the columns we list different criteria by which to judge poverty measures. We consider nine criteria: whether they are conceptually grounded in some clear notion of welfare, how they deal with the issue of weights, whether they are able to consider multidimensional information, whether they are reliable in comparisons across space and time, how data intensive and costly they are, whether disaggregation is possible, whether they are suitable for global poverty monitoring, and whether they are suitable for targeting and project monitoring and evaluation.

In the rows we then consider the various approaches to measuring poverty discussed in the paper, where we do not discuss every single permutation of an approach, but focus on more general approaches. The cells than discuss to what extent the particular approach meet the criteria (or could plausible be adjusted to do so). We will go through the table by poverty concepts, beginning with various income poverty approaches.

The food energy intake method to generate an income poverty line is, as discussed above, partly grounded in utility theory but dependent on rather stringent assumptions. While it appears to be a uni-dimensional index, one can follow Ravallion's arguments made above that it really is a multidimensional index of command over commodities with prices being the implicit weights (although this requires stringent and unrealistic assumptions). While this approach does not typically consider other dimensions of poverty, it could be accommodated to do so. For example, one could not just study which income levels are required to meet a certain caloric intake, but also how much income is required to achieve a certain standard of housing, education, or health. This would, of course, then lead to multiple poverty lines but might be worth considering. The method's reliability for comparisons across space and time, as discussed above, quite limited and likely to be biased, and it is quite data-intensive (requiring a detailed household expenditure survey), costly, and involves substantial time lags. It would be hard to use this method for global poverty monitoring to the comparability problems across space and time. It can readily be disaggregated and could be used for targeting (although costs and time lags are a constraint). Its use for project monitoring and evaluation is problematic. For example, if a project succeeded in raising incomes, households might switch to more expensive calories and the project could then found to have led to a constant or even increasing poverty rate.

The closely-related income poverty measure based on the cost of basic needs (CBN) method shares a number of features with the FEI method. But its reliability over time is higher, but subject to the problem of how and when to adjust poverty baskets. Its comparability across space is also superior. In fact, one could readily imagine that this method could be used to generate internationally comparable national poverty lines using the CBN method. In fact, this is precisely what Reddy et al. (2008) did for three countries and has been proposed by them and by Klasen (2009). The challenge is the high costs involved in producing these data and the large coordination effort required (possible similar to the scale of the International Comparison of Prices Project or the UN System of National Accounts) to generate comparable national poverty lines across the world. This approach to generating a poverty line might also be used for project monitoring and evaluation. As a result, the CBN method dominates the FEI method in all dimensions except costs (as reliable price data are required for the CBN method); if these cost issues can be accommodated, the CBN method is clearly superior.

Income-based poverty measurement using a subjective poverty line has quite different advantages and disadvantages. Its grounding in welfare theory is rather tenuous, again prices are used as implicit weights and, similar to the other income-based methods could in theory be adjusted to non-income dimensions as well (by replacing the minimum income question with a minimum education, nutrition, or health question). The reliability across space and time is limited as poverty lines will change in ways that are

not easy to interpret. The data intensity is lower than with the other lines, although typically the questions needed for subjective poverty lines are simply included in standard household surveys. They could, of course, be collected using much shorter survey instruments, which would make this type of data cheaper and available more quickly. It is not suitable for global poverty monitoring and its use for targeting and project monitoring is limited. In short, subjective poverty lines are, in our view, an interesting supplementary indicator to complement other poverty measures. They particularly provide interesting insights into the local perceptions of poverty which clearly is relevant information to also consider. They can also help in an understanding of any systematic differences in poverty perceptions across space and time which might help in an assessment of the relevance of absolute versus relative poverty conceptions.

The international poverty line of $1.25 a day has, in many ways, only rather limited uses. Its particular strength is its ability to monitor poverty trends at the global level with reasonable accuracy. This is compromised, however, by the vagaries of the robustness of the various PPP rounds (the next one will publish results in 2013 and will lead to yet another revision of the numbers), and the falling relevance of such a low absolute poverty line for an increasing number of middle income countries. In many other dimensions, the line is problematic. Its grounding is weak, it cannot consider non-income dimensions, its reliability across space is limited, its aggregation method is too simplistic (not adjusted for equivalence scales). It cannot be disaggregated (without opening all the underlying household surveys many of which are not accessible), and its use for targeting and monitoring is low. While it has great intuitive appeal at the global level, it has little buy-in at the national level as most countries use their own poverty line based on the CBN or FEI method. We suggest that as soon as globally comparable income poverty lines based on the CBN method are available, there would be no further need for the current international poverty line of $1 a day.

The last of the income poverty lines is the proposed international poverty line by Ravallion (2010) based on the concept of weakly relative poverty. It of course shares most of the advantages and disadvantages of the $1.25 a day line. The one difference is that this line would automatically adjust as countries get richer. This would make this line better able to capture poverty across space, particularly in middle-income countries, and over time, as countries move up the income scale. One interesting idea would also be to extend the concept of weakly relative poverty to non-income dimensions. One could readily imagine that the poverty cut-off for education, nutrition, or housing depends on the overall education, nutrition, and housing situation in a country.

Turning now to non-income approaches, there is a plethora of approaches which we grouped into five approaches, three multidimensional poverty approaches, and two other ones. Starting with the multidimensional approaches, one way to measure poverty in a multidimensional way is to use dominance approaches. The grounding of these approaches is generally solid and the big advantage this approach has is that it is transparent about the reasonable disagreements one might have about poverty dimensions to be considered, different cut-offs for dimensions and different weights for aggregation. But while attractive for high quality poverty research, practical applicability is more limited as this approach will only provide ordinal information (i.e. can we say with

some certainty that poverty rose or fell) and often it will not be possible to say even that. Thus its practical use is limited but it can be an excellent supplementary approach to study whether claims about poverty differences and trends are really robust.

The second type of multidimensional poverty measure consists of a plethora of national approaches to develop multidimensional poverty measures. Recently, many countries have followed that route, including particularly many Latin American countries (building on the Multidimensional Poverty Index or the previous Unsatisfied Basic Needs Approach of ECLAC). At its best, these measures are adapted to local realities, choose indicators and weights that are relevant and based on some consensus, and thereby are able to capture the key aspects of poverty in that particular country context. Such indicators can therefore be useful also for targeting and monitoring. Of course, many such initiatives fall short of this ideal and can only be an ad hoc compilation of data that happens to be available using expert or statistical weights, with no clear and demonstrated relevance of the approach chosen to the country context. Thus in a sense the process of arriving at these measures is probably as important or more important than the technical sophistication of deriving poverty cut-offs and dealing with aggregation issues. This is clearly an area that is likely to expand in coming years and might also be of considerable relevance for donors to support.

The third multidimensional measure is the recently developed MPI. Its advantage is that is grounded somewhat better than many ad hoc multidimensional approaches. In principle, its great advantage is that it can be comparable across space and time and be used as an indicator for global poverty monitoring. In that sense, it is in direct 'competition' with the $1.25 a day line, but focused on multidimensional deprivation. But how well the MPI does in terms of comparisons across space and time is not clear yet and the subject of urgently needed research (which is on-going). Relative to the $ a day line it has the advantage of being based on publicly available micro data so that disaggregation is directly possible and it could, in principle, also be used for targeting and monitoring of projects. The great open question is, of course, the choice of indicators, the aggregation procedure which was discussed at length above, and the weights chosen. Clearly the MPI has substantial potential as an alternative or complement to the global income-based measures ($1.25 a day or weakly relative poverty), but it is not clear yet whether changes to the MPI are feasible, possible, and desirable. Among the open issues are the role of household size and structure in the treatment of the MPI, questions of poverty dynamics, the treatment of households with missing indicators, the rather limited set of standard of living indicators, the uneven treatment of deprivation cut-offs (e.g. an entire household is non-poor in education is *one* single member has more than 5 years of schooling, while the entire household is nutrition-poor if *one* member is undernourished), consideration of inter-dependence between dimensions (as suggested by Rippin), considering (weakly) relative deprivation cut-offs, among others.

When moving towards proxy indicators, we are now in a territory where the emphasis is on generating cheap, quick but still relevant information for practitioners and policy-makers. These methods have developed a long way and in principle such proxy indicators can be used for income-based or multi-dimensional poverty conceptions.

Their strength is their low cost and immediate suitability for the monitoring and evaluation of projects. But they are not conceptually grounded, they are highly context-specific, cannot readily be used for comparisons across space and time. In this sense they cannot replace proper poverty measurement and analysis, but can be an interesting and highly relevant supplement. In fact, whenever new poverty measures are developed, it would always be useful to immediately ask the empirical question, whether and how they can be approximated using proxy indicators in particular contexts to allow the application of these new measures via the use of such proxies.

Lastly, we consider subjective well-being measures. They are loosely based on a conception of welfare, they rely on persons properly weighing different aspects of well-being when answering these questions, comparability is somewhat limited (although may be improving), they are reasonably costly, and they are unsuitable for targeting of project monitoring. Thus at present they are not really suitable as poverty indicators. Since they have largely focused on studying the development of mean well-being (proxied by happiness or life satisfaction), they are not really poverty indicators to begin with. And this would be the key question to be addressed by further research: What are the advantages and disadvantages if one defined a poverty line as a particular score in the happiness or life satisfaction and then used the surveys to generate poverty rates, poverty depth, and study the determinants of poverty. This is largely an open question at present.

8 Research Agenda for Poverty Measurement

Table 1 has been focused on the in-principle advantages and disadvantages of these different approaches to measuring poverty using the set of criteria. One must recognize, however, that in practice some of the advantages of particular approaches are not realized, due to the actual way these measures are implemented in different contexts. Also, some of the disadvantages of particular measures could actually be tackled if those measures were suitably refined. Thus it is useful to add a Table 2 that lists the most important current problems and possible reform agendas to improve these particular measures. While there are ways to conceptually and empirically improve all of these measures, we will focus the discussion on those indicators that we have found particularly reliable and useful in Table 1. Among the income poverty measures, the one where the national poverty line is based on the CBN approach was identified as promising in Table 1, but as Table 2 suggests there are substantial open questions that need to be addressed. At the conceptual level it is important to derive a consistent and comparable procedure to update the poverty baskets across space and time. Here it is probably most promising to link the change in poverty baskets to the change in consumption baskets used to calculate the consumer price index (CPI) in each country. Since there is an established procedure based on international best practice for the adjustment of the CPI and since the poverty basket can use the same information, linking the two is most promising. In practice, this will require a much closer link between statistical offices (who usually calculate the CPI) and the authorities charged with poverty measurement. To make this approach also useful for international comparisons of poverty across space and time, it is critical to ensure that the income and consumption information of household surveys are comparable across space and

time. In the short term, one can think of empirical ways of standardizing surveys across countries and over time. Much preferred in the medium term would be an international process to standardize surveys and the calculation of poverty lines ex ante (akin to the UN System of National Accounts which did something similar for GDP calculations, see Reddy and Pogge, 2010; Klasen, 2009).

Turning to the $ a day line, one could adjust the international poverty line to account for differences in household size and structure. Similarly, improvements could be made by changing the rather ad hoc estimation procedure of calculating both the absolute as well as the weakly relative poverty line. But the dependence of this approach to the PPP adjustments as well as its rather unclear grounding will ensure that only limited improvements to this indicator are possible.

Among the non-income and multidimensional approaches, national multi-dimensional indicators hold substantial promise, but often it is not realized. The most important conceptual improvement would be to develop transparent and systematic approaches to choose indicators and weights for such measures. Here there is a clear role for the political process to lead these processes, supported by appropriate technical support. Among the empirical challenges are to reverse the currently common order that national poverty indicators are based ex post on ad hoc definitions and calculations using available data. Instead, indicators and weights should be decided upon ex ante and surveys designer ex post to generate the required information. Here one should also move beyond the commonly used very crude categorical variables.

Table 1: Judging Poverty Concepts by Criteria

	Conceptually Grounded?	Weights	Ability to consider multidimensional data	Reliability across time	Reliability across space	Data intensity/ costs/ time lags	Suitable for global poverty monitoring?	Disaggregate to sub-national levels and study determinants?	Targeting?	Project monitoring and evaluation?
Income using Food Energy Intake	Partly grounded in utility theory	Prices (implicit)	Limited (could extend to other nutrients or functionings)	Limited and biased	Limited and biased	High data intensity and costs, substantial time lags	No	Yes	Yes	No
Income using Costs of Basic Needs	Partly grounded in utility theory	Prices (implicit)	Considers food and non-food items, could possibly be extended	Good (but updating problem)	Good (but price and basket comparability problem)	Very high data intensity and costs, substantial time lags	Possible (if methods are internationally standardized)	Yes	Yes	Possible (costs and timeliness!)
Income using Subjective Poverty Line	No clear grounding	Prices (implicit)	Could possibly be extended to consider other functionings	Limited and changes in poverty line hard to interpret	Limited	Medium data intensity and costs, moderate time lags possible	No	Yes	Limited and biased	Limited
$1.25 international poverty line	Partial empirical grounding in national poverty lines	Prices (implicit)	No	Good (but PPP and updating problem)	Limited (PPP problems, relevance to MICs?)	Low (if national surveys are available), substantial time lags	Yes (esp. trends)	No	Possibly at the international level (but biases)	No
Income: Weakly Relative Poverty	Partial empirical grounding in national poverty lines and utility theory	Prices (implicit)	No (but could be used for non-income approaches as well)	Good (but PPP problem)	Good (but PPP problems)	Low (if national surveys are available), substantial time lags	Yes	No	Possibly at the international level (but biases)	No
MPM: Dominanc	Partial grounding in	No weights	Yes	Yes but only limited	Yes but only limited	High data intensity and	Limited	Limited	No	No

e Approaches	non-welfarist theories			ordinal information	ordinal information	costs, substantial time lags				
MPM: National Indices	No or limited grounding	Equal, expert or statistical weights	Yes	Yes (but changing relevance of indicators)	Limited (relevance of indicators)	High data intensity and costs, substantial time lags	Limited	Yes	Yes (but biases possible)	Yes (if adjusted to context)
Multidimensional Poverty Index	Limited grounding	Expert weights	Yes	Unclear and possibly limited	Yes (but relevance of indicators)	Low (if based on existing surveys), moderate costs, substantial time lags	Yes	Yes (when using actual micro data)	Yes (but possible biases)	Yes (but possible biases)
Proxy Indicators	No	Regression-based or PCA	Yes	Limited	No	Medium to low, costs moderate, and low time lags	No	Yes (but possible biases)	Yes	Yes
Subjective Well-Being	Very loosely	Subjective valuation of respondent	In principle	Limited	Limited	Medium to high intensity and costs, moderate time lags	No	Yes (but possible biases)l	No	No

As far as the MPI is concerned, there are a range of conceptual and empirical improvements that can be considered. At the conceptual level, the issue of how to incorporate depth of dimensions as well as the number of dimensions in an assessment is an open question. Here the work by Rippin (2011) has made a number of useful proposals. Moreover, current this measure is a purely absolute measure. While this can be justified if the dimensions were truly measuring functionings or capabilities, at least in some of the dimensions one may want to conceptualize poverty in a relative measure. For example, one may want to adjust the cut-offs for quality of housing, number of assets, and some of the other standard of living indicators to prevailing conditions in a country. But much work would need to be done to conceptualize such a relative MPI. At the empirical level, it is critical to review a number of issues, including the treatment of childless households, the determination of cut-offs, and the use of continuous rather than categorical variables. A first step should be robustness of results to changes in these issues before possibly adjusting the MPI in a second step. Lastly, the adaption of subjective well-being measures to poverty is still in its infancy. Here a number of conceptual and empirical issues need to be considered before one can assess whether this approach yields useful, consistent, and comparable results.

Table 2: (Selected) Policy Research Agenda to improve Poverty Measures

	Suggested conceptual improvements	Main empirical problems	Suggested measurement improvements
Income using Food Energy Intake	Extension to other nutrients and simultaneous consideration of many nutrients	Inconsistencies in items covered in surveys across space and time; insufficient coverage of nutrients; waste and eating-out insufficiently captured	Standardization of survey instruments; explicit consideration of eating-out and waste
Income using Costs of Basic Needs	Consistent updating procedure (similar to CPI adjustments)	Inconsistencies in income and consumption items and aggregates across space and time; inconsistencies in use of price data	Standardize existing data using common and comparable poverty lines; internationally coordinated approach for survey instruments and treatment of prices
Income using Subjective Poverty Line	Extension to other functionings (e.g. adequately nourished, healthy, etc.)		
$1.25 international poverty line	Adjust poverty line and poverty measurement for equivalence scales	Inconsistencies in income and consumption items and aggregates across space and time; inconsistencies in use of price data; insufficient coverage of poverty lines	Internationally coordinated approach for survey instruments and treatment of prices; support processes of formulating poverty lines; adjust procedure of estimating line; consider national poverty lines as benchmark;
Income: Weakly Relative Poverty	Adjust poverty line and poverty measurement for equivalence scales;	Limited scope of poverty lines considered	Expand scope of existing poverty lines; adjust procedure of estimating line
MPM: Dominance Approaches	Consider depth of poverty and number of dimensions explicitly	Often based on categorical variables; unclear robustness with respect to indicators	Broaden use of indicators
MPM: National Indices	More systematic and transparent choice of indicators and weights	Often approaches driven by available data	Define poverty measures prior to data gathering exercises
Multidimensional Poverty Index	Consider depth of poverty (within a dimension and number of dimensions); consider construction of a (weakly) relative MPI	Treatment of childless households; choice of indicators; use of categorical variables; different types of deprivation cut-offs	Testing robustness to empirical problems and suggest corrections where preferred approach is available and makes a difference.
Proxy Indicators	Need to standardize methods for more robust and comparable results	Data sets often limited in possible proxy indicators	Design and implement data sets with proxy indicators in mind.
Subjective Well-Being	Re-orient approach towards subjective poverty; clarify interpretation of trends	Unclear robustness of 'unhappiness' at a particular point in time	Examine subjective poverty in a panel data approach

	and international comparisons		

9 Preferred Poverty Indicators by Actors

Note also that the criteria used in Table 1 have different importance for different actors. Academics might particularly care about the conceptual grounding, a convincing approach to deal with weights, and the consideration of multidimensional information. National policy-makers and donors operating at the national level are likely to be particularly concerned about the reliability of comparisons, data-intensity and costs, and suitability for targeting and monitoring. Policy-makers, donors, and international organizations will be particularly concerned about the comparability of information, about the relevance for global poverty monitoring, and some about the use of multidimensional information. Table 3 therefore explicitly considers the suitability of different indicators for different actors. Often actors have needs for two types of indicators: one is to get an aggregate picture of the poverty situation in a country, and another is to assess and monitor poverty in setting where an intervention is being planned or executed. This is also considered.

International agencies are particularly interested in internationally and inter-temporally consistent and comparable poverty measures. We argue here that the absolute $1.25/day poverty indicator has increasing problems and decreasing relevance. Instead we propose that the use of internationally coordinated national poverty analysis (using the CBN method to define income poverty lines) is a more promising approach, although it requires a substantial amount of work. In addition, a (possibly revised) MPI would be an important complement to capture non-income dimensions of poverty across countries and time. Bilateral donors will be somewhat more concerned about national poverty trends and here our suggestion for the best national income poverty indicator is to use one based on the CBN method. As far as national non-income poverty is concerned, the MPI is useful but country-specific alternatives might also be very useful here as well. As these actors usually are involved in similar interventions, at the intervention level, similar indicators are required. Since most income and non-income poverty measures are costly and time-consuming to produce, short-cut and proxy indicators are typically needed.

Governments have different needs depending on whether we are talking about central governments (e.g. Planning or Finance Ministries), sectoral ministries, or regional or local governments. Central governments should be the key drivers of poverty measurement at the national level. We suggest that a income poverty line based on the CBN method is to be preferred. At the non-income level, a national MPI or a similar composite index would be best at this level of aggregation. For intervention-level indicators, similar short-cut and proxy indicators are required.

For sectoral ministries it is somewhat unclear whether poverty measures really are relevant. Instead, it would be best to monitor the sector-specific targets and outputs directly using disaggregated indicators at both the aggregate as well as the intervention level. At the same time, the use of poverty indicators can be useful for targeting purposes where again proxy indicators are likely to be required.

71

At the local level, it is often too difficult and costly to produce accurate surveys and the local and regional samples from national surveys are usually too small to be reliable. One method that can be helpful here is to use poverty mapping approaches where census information is combined with survey data to generate reliable local-level indicators (Elbers, Lanjouw, and Lanjouw 2004).

For NGOs and civil society it is even more difficult to invest in gathering a great number of indicators and surveys. Here proxy indicators will be of particular relevance.

Table 3: Preferred Indicator by Actor and Type of Analysis

	Aggregate Indicators	Intervention-Level Indicators
International Agencies (UN system, multilateral development banks, OECD)	Reformed MPI, Global assessments based on consistent national poverty lines, $ a day (weakly relative version?)	Proxy Indicators, disaggregated and aggregated multidimensional indices
Bilateral donors	National income poverty lines using CBN method, MPI or national multidimensional poverty index	Proxy Indicators, disaggregated and aggregated multidimensional indices
Central Governments	National income poverty lines using CBN method, MPI or national multidimensional poverty index	Tracking of policies using aggregate poverty indicators
Sectoral Ministries	Disaggregated sector-relevant deprivation measures;	Project-related sectoral performance indicators, poverty measures for targeting
Regional/Local Government Agencies	Regional/local income poverty using poverty mapping; local and regional MPIs	Tracking of policies using regional/local poverty indicators via poverty mapping
NGOs, Civil Society	National income poverty and MPI, use of regional/local assessments	Proxy indicators to monitor income and multidimensional poverty

10 Conclusions

In this paper we have considered the plethora of approaches to measuring and analyzing poverty that have developed over the past 30 years. Both at the conceptual as well as the empirical level, we have shown that great progress has been made so that there now are a large number of approaches to choose from. As shown in Table 1, these approaches often serve different needs and many approaches should be seen as complementary. Even so, let us first suggest some approaches that appear to offer particular advantages and identify some suggestions for further research.

In terms of national income poverty measurement, it appears that poverty lines based on the CBN method are superior to other approaches. The key issue here is to achieve greater standardization and comparability of both survey instruments (within and across countries) and methods to derive such poverty lines. Here the World Bank (particularly Ravallion and co-authors) have made great contributions along that road, but what is missing is a clear process within countries and internationally that would achieve such greater consistency and standardization. If that were to be achieved, it could well replace the current focus on the problematic international poverty line ($1.25 a day) which is in many ways much more limiting and losing relevance.

In addition, it will be important to consider the updating of poverty lines in more detail. Here the approach of weakly relative poverty lines is useful for both income and non-income poverty lines. Also, subjective poverty lines can be very helpful in seeing to what extent relative versus absolute considerations appear to shape poverty perceptions. Insights from this literature could then be used to influence the updating of CBN poverty lines as well.

While dominance approaches have little direct practical relevance to policy-makers, they are very useful to examine the robustness of multidimensional poverty measures. Thus they are an important supplement in a field where there are many uncertainties (about dimensions, weights, cut-offs, and aggregation rules).

National multidimensional poverty indices are surely an exciting way to adapt multidimensional poverty to national contexts. Here one should focus on ensuring that the process of making the difficult decisions about weights, dimensions, cut-offs, and aggregation is as transparent and inclusive as possible. The MPI is a very useful starting point for thinking about global poverty in a multi-dimensional context. At the same time, there are many questions of detail that need to be considered before the reliability and robustness can be demonstrated. Also, the question of relative poverty lines in a multidimensional space should receive more attention.

Lastly, it is critical to emphasize that further progress on this research agenda is largely hampered by data issues. Particularly in poor developing countries, household surveys are irregular, often incomparable, unreliable, and sometimes non-existing. It is a tragedy that for most of those countries, the majority of reliable surveys are driven by donors (and increasingly polling firms) rather than national statistical agencies. Thus hardly any of the improvements we advocate for poverty measurement will happen without substantial and sustained investments in the statistical capacities of developing countries. Here there is a great role also for donors, but a different form of donor engagement is necessary. Rather than funding individual idiosyncratic surveys which are part of the problem, long-term support to build up national statistical capacity is required.

A References

Abul Naga, RH. and Bolzani, E. (2008): Income, consumption and permanent income: a MIMIC approach to multidimensional poverty measurement. In: Silber, J. and Kakwani,N., eds. 2008. *Quantitative Approaches to Multidimensional Poverty Measurement.* London: Macmillan Publishers. pp. 104-117.

Adato, M. (2000): El impacto de PROGRESA sobre las relaciones socials en la comunidad. Unpublished paper, IFPRI.

Adato, M. and Roonpnaraine, T. (2004): A social analysis of the red de protección social (RPS) in Nicaragua, final report, IFPRI.

Ahluwalia, M., Carter, N. G. and. Chenery, H. B. (1979): Growth and poverty in developing countries. *Journal of Development Economics*, 6(3): 299–341.

Ahmed, A. And Bouis, H. (2002): Weighing what's practical: Proxy Means Tests for targeting food subsidies in Egypt. IFPRI Food Consumption and Nutrition Division Discussion Paper No. 132.

Alkire, S. and Santos, E. (2010): Acute multidimensional poverty: A New Index for Developing Countries. *OPHI Working Papers Series,* 38.

Alkire, S. and Foster, J. (2009): Counting and multidimensional poverty measurement. *OPHI Working Paper Series*, 32.

Alkire, S. and Seth, S. (2009): Determining BPL status: some methodological improvements. *Indian Journal of Human Development*, 2 (2): 407-24.

Alkire, S. and Foster, J. (2007): Counting and multidimensional poverty measuers. *OPHI Working Paper Series*, 7.

Anand, P., Hunter, G., and Smith, R. (2005): Capabilities and well-being: Evid3ence based on the Sen-Nussbaum approach to welfare. Social Indicators Research, 74(1): 9-55.

Anand, S., Segal, P., and Stiglitz, J. E. (2010): Introduction. In: Anand, S., Segal, P., and Stiglitz, J. E., eds. 2010. *Debates on the Measurement of Global Poverty*. Oxford University Press. pp 1–21.

Anand, S. and Sen, A. (1997): Concepts of human development and poverty: A multidimensional perspective. In: *Human Development Report 1997 Papers: Poverty and Human Development*. New York: United Nations Development Programme.

Andrews, F. M., and Withey, S. B. (1976): *Social indicators of well-being: America's perception of life quality*. New York: Plenum Press.

Asselin, L.-M., and Tuan Anh, V., (2008): Multidimensional Poverty and Multiple Correspondance Analysis. In: Kakwani, N. and Silber, J., eds. 2008. *Quantitative Approaches to Multidimensional Poverty Measurement*. London: Palgrave-Macmillan.

Aten, B. and Heston, A. (2010): Use of country purchasing power parities for international comparisons of poverty levels: Potential and limitations. In: Anand, S., Segal, P. and Stiglitz, J.E., eds. 2010. *Debates on the Measurement of Global Poverty*. Oxford University Press. pp 154–186.

Atkinson, A.B. (1995): On Targeting Social Security: Theory and Western Experience with Family Benefits. In: Van de Walle, D and Nead, K., eds. 1995. Public Spending and the Poor – Theory and Evidence. Baltimore and London: John Hopkins University Press. pp. 25-67.

Atkinson, A.B. and Bourguignon, F. (2001): Poverty and inclusion from a world perspective. In: Stiglitz, J. E., ed. 2001. *Governance, Equity, and Global Markets: The Annual Bank Conference on Development Economics - Europe*. Oxford: Oxford University Press.

Atkinson, A. B. (1987): On the measurement of poverty. *Econometrica*, 55(4): 749–764.

Atkinson, A. B. (1975): The Economics of Inequality. Oxford: Oxford University Press.

Bertrand, M. and Mullainathan, S. (2001): Do people mean what they say? Implications for subjective survey data. *American Economic Review, Papers and Proceedings*, 91 (2): 67-72.

Bitrán, R. and C. Muñoz (2000): Targeting Methodologies: Conceptual Approach and Analysis of Experiences. Partnerships for Health Reform (PHR).

Birdsall, N. (2011): Comment on multi-dimensional indices. *Journal of Economic Inequality,* 9: 489–491.

Blackorby, C. and Donaldson, D. (1988): Money metric utility: A harmless normalization? *Journal of Economic Theory*, 46: 120-129.

Blackorby, C.and Donaldson, D. (1987): Welfare ratios and distributionally sensitive cost-bene•t analysis. *Journal of Public Economics*, 34: 265–290.

Blanchflower, D. and Oswald, A. (2007): Hypertension and happiness across nations. *NBER Working Paper Series*, 12934.

Bossert, W., Chakravarty, S.R. and D'Ambrosio, C. (2009): Multidimensional poverty and material deprivation. Society for the Study of Economic Inequality (ECINEQ) Working Paper 2009-129.

Bouis, H. E. and Haddad, L. (1992): Are estimates of the calorie-income elasticities too high? A recalibratation of the plausible range. *Journal of Development Economics*, 39: 333–364.

Bourguignon, F. (2003): The growth elasticity of poverty reduction. In: Eichner; T. S. and Turnovsky, S. J., eds. 2003. *Inequality and Growth: Theory and Implications*. Cambridge, Mass: MIT Press. pp 3–26.

Bourguignon, F. and Chakravarty, S., (2003): The measurement of multidimensional poverty. *Journal of Economic Inequality,* 1: 25-49.

Callan, T., Nolan, B. and Whelan, C. (1993): Resources, deprivation and the measurement of poverty. *Journal of Social Policy*, 22(2).

Cantril, H. (1965): *The Pattern of Human Concerns*. New Brunswick, NJ: Rutgers University Press.

Castano, E. (2001): Proxy Means Test Index for Targeting Social Programs: Two Methodologies and Empirical Results. School of Economics, University of Antioquia.

Chen, S. and Ravallion, M. (2010): The developing world is poorer than we thought, but no less successful in the •ght against poverty. *The Quarterly Journal of Economics*, 125(4): 1577–1625.

Citro, C. and Michael, R. (1995): Measuring Poverty: A New Approach. Washington DC: National Academy Press.

Clark, D. A. (2002): *Visions of Development: A Study of Human Values.* Cheltenham: Edward Elgar.

Coady, D, Grosh, M and Hoddinott, J. (2004a): Targeting of transfers in developing countries. The World Bank.

Coady, D, Grosh, M and Hoddinott, J. (2004b): Targeting outcomes redux. The World Bank Research Observer, 19(1): 61-85.

Colasanto, D., Kapteyn, A., and Van der Gaag, J. (1984): Two subjective de•nitions of poverty: results from the Wisconsin Basic Needs Study. *The Journal of Human Resources*, 19: 127–138.

Danziger, S., Van der Gaag, J., Taussig, M. K. and Smolensky, E. (1984): The direct measurement of welfare levels: How much does it take to make ends meet? *Review of Economics and Statistics*, 66(3): 500–505.

Deaton, A. and Dupriez, O. (2011): Purchasing power parity exchange rates for the global poor. *American Economic Journal: Applied Economics*, 3: 137–166.

Deaton, A. (2010): Price indexes, inequality, and the measurement of world poverty. *American Economic Review*, 100(1): 5–34.

Deaton, A. and Grosh, M. (2000): Consumption. In M. Grosh and P. Glewwe (eds.): Designing Household Survey Questionnaires for Developing Countries: Lessons from 15 Years of the Living Standards Measurement Study, Vol. 1, Washington DC: The World Bank.

Deaton, A. and Heston, A. (2010): Understanding PPPs and PPP-based national accounts. *American Economic Journal: Macroeconomics*, 2(4): 1–35.

Deaton A. and Kahneman, D. (2010): High income improves evaluation of life but not emotional well-being. *Centre for Health and Well-being, PNAS*, 107(38).

Deaton, A. and Kozel, V. (2005): Data and Dogma: The Great Indian Poverty Debate. *World Bank Research Observer*, 20(2): 177-199.

Deaton, A. (2008): Income, health and well-being around the world: evidence from the Gallup World Poll. *Journal of Economic Perspectives*, 22(2): 53-72.

Deaton, A. (2003): How to monitor poverty for the Millennium Development Goals. *Journal of Human Development*, 4(3): 353–378.

Deaton, A. and Zaidi, S. (2002): Guidelines for constructing consumption aggregates for welfare analysis. *Living Standards Measurement Study Working Paper*, 135.

Deaton, A. (2001): Counting the world's poor: problems and possible solutions. *World Bank Research Observer*, 16(2): 125–147.

Deaton, A. and Paxson, C. (1998): Economies of scale, household size, and the demand for food. *Journal of Political Economy,* 106: 897–930.

Deaton, A. and Muellbauer, J. (1986): On measuring child costs: with applications to poor countries. *Journal of Political Economy*, 94(4): 720–744.

Diener, E., Suh, E., Lucas, R. and Smith, H. (1999): Subjective well- being: three decades of progress. *Psychological Bulletin,* 125 (2): 276-302.

Diener, E., Suh, E., Smith, H. and Shao, L. (1995): National differences in reported subjective well-being: why do they occur? *Social Indicators Research*, 34(1), 7-32.

Diener, E., Emmons, R., Larsen, R. and Griffin, S. (1985): The satisfaction with life scale. *Journal of Personality Assessment*, 49(1): 71-75.

Diener, E (1984): Subjective well being. *Psychological Bullettin*, 95(3): 542-575.

Di Tella, R., MacCulloch, R., and Oswald, A. (2003): The macroeconomics of happiness. *Review of Economics and Statistics*, 85(4): 809-827.

Diewert, E. (1976): Exact and superlative index numbers. *Journal of Econometrics*, 4(2): 115–145.

Duclos, J.-Y, Sahn, D. and Younger, S. (2006): Robust Multidimensional Poverty Comparisons*. Economic Journal*, 116: 943-968.

Easterlin, Richard A. (1974): Does economic growth improve the human lot? Some empirical evidence. In: David, P. and Reder, M., eds, 1974: Nations and Households in Economic Growth: Essays in Honor of Moses Abramowitz. New York: Academic Press.

FAO/WHO/UNU. (2001): Human energy requirements. Technical Report, Food and Agricultural Organization.

Ferrer-i-Carbonell, A. And Van Praag, B. (2005): Poverty in Russia. Journal of Happiness Studies 2: 147-172.

Ferreira, F., (2011): Poverty is Multidimensional. But what are we going to do about it? *Journal of Economic Inequality* 9, 493–495.

Ferro Luzzi, G., Flückiger,Y. and Weber, S. (2008): A cluster analysis of multidimensional poverty in Switzerland. In: Kakwani, N. And Silber, J., eds., 2008: Quantitative Approaches to Multidimensional Poverty Measurement. Palgrave MacMillan: 63-79.

Flik, R. J. and Van Praag, B. (1991): Subjective poverty line de•nitions. De Economist, 139(3):311–331.

Foster, J., Greer, J. and Thorbecke, E. (1984): A class of decomposable poverty measures. Econometrica, 52(3): 761–766.

Frey, B. and Stutzer, A. (2004). "Reported Subjective Well-Being: A Challenge for Economic Theory and Economic Policy,"Schmollers Jahrbuch : Journal of Applied Social Science Studies / Zeitschrift für Wirtschafts- undSozialwissenschaften, Duncker & Humblot, Berlin, vol. 124(2), pages 191-231.

Glewwe, P. (1990). "Efficient Allocation of Transfers to the Poor." Living Standards Measurement Study Working Paper No. 70. World Bank, Washington, D.C.

Glewwe, P. and Kanaan, O. (1989). "Targeting Assistance to the Poor: A Multivariate Approach Using Household Survey Data." Policy, Planning and Research Working Paper No. 225, World Bank, Washington, D.C.

Goedhart, T., Halberstadt, V., Kapteyn, A. and Van Praag, B. (1977): The poverty line: Concept and measurement. The Journal of Human Resources, 12(4):503–520.

Greb, F., Klasen, S., Hidayat Pasaribu, S. and Wiesenfarth, M. (2011): Dollar a day re-revisited. Courant Research Centre Discussion Papers No. 91, Göttingen University.

Greer, J and Thorbecke, E. (1986): A methodology for measuring food poverty applied to Kenya. Journal of Development Economics, 24:59–74.

Grootaert, C., & Braithwaite, J. (1998): Poverty correlates and indicator-based targeting in eastern Europe and the Former Soviet Union. Poverty Reduction and Economic Management Network, Environmentally and Socially Sustainable Development Network, World Bank, Washington, DC.

Grosh, M. and Baker, J. (1995): Proxy means tests for targeting social programs: Simulations and speculations. Living Standards Measurement Survey Working Paper, 118.

Gurin, G., Veroff, J., and Feld, S. (1960): Americans view their mental health. New York: Basic Books.

Haddad, L., Kennedy, E., and Sullivan, J., (1994): Choice of indicators for food security and nutrition monitoring. *Food Policy*, 19(3): 329-343.

Haddad, L. and Kanbur, R. (1990): How series is the neglect of intrahousehold inequality. *Economic Journal*, 100: 866–881.

Harttgen, K. And S. Klasen (2011). A household-level Human Development Index. *World Development*

Hirschman, A. (1991): The rhetoric of reaction. Cambridge, MA: The Belknap Press of Harvard University Press.

Johannsen, J. (2006): Operational poverty targeting in Peru - Proxy means testing with non-income indicators. UNDP International Poverty Centre Working Paper ,30.

Johannsen, J., Zeller, M. and Klasen S. (2007): The capability dilemma in operational poverty assessment. Ibero America Institute for Economic Research Discussion Paper No. 159, University of Göttingen.

Joreskog, K.G. and Goldberger, A.S. (1975): Estimation of a Model with Multiple Indicators and Multiple Causes of a Single Latent Variable. *Journal of the American Statistical Association*, 70 (351): 631-639.

Kahneman, D. and Thaler, R.H. (2006): Anomalies: utility maximization and experienced utility. *Journal of Economic Perspectives*, 20: 221-234.

Kahneman, D., Krueger, A., Schkade, D., Schwarz, N. and Stone, A. (2004): Toward national well-being accounts. *American Economic Review, Papers and Proceedings*, 94(2): 429-434.

Kapteyn, A. (1994): The measurement of household cost functions: revealed preferences versus subjective measures*. Journal of Population Economics*, 7: 333–350.

Kapteyn, A., Kooreman, P. and Willemse, R. (1988): Some methodological issues in the implementation of subjective poverty de•nitions. *The Journal of Human Resources*, 23(2): 222–242.

Kapteyn, A., Van de Goer, S. and Van de Stadt, H. (1985): The impact of changes in income and family composition on subjective measures of wellbeing. In: David, M. and Smeeding, T., eds. 1985. Horizontal Equity, Uncertainty, and Economic Well-Being. Chicago: University of Chicago Press.

Kapteyn, A. and Van Praag, B. (1978): A new approach to the construction of family equivalence scales. *European Economic Review*, 7(4): 313–335.

Kidd, S. And Wylde, E. (2011): Tageting the poorest: An assessment of the proxy means test methodology. Report prepared for AusAID/the Australian Government.

Kingdon, G. And Knight, J. (2006): Subjective well-being poverty vs. income poverty and capabilities poverty. *Journal of Development Studies*, 42(7): 1199-1224.

Klasen, S. (2009): Levels and trends in absolute poverty in the world: What we know and what we don't. In: Mack, E., Schramm, M., Klasen, S. and Pogge, T., eds. 2009. Absolute Poverty and Global Justice. Farnham: Ashgate. pp 21–36.

Klasen, S. and Misselhorn, M. (2007): Determinants of the growth semi-elasticity of poverty reduction. Courant Research Centre Discussion Papers, 176.

Klasen, S. (2000): Measuring poverty and deprivation in South Africa. *Review of Income and Wealth*, 46(1): 33-58.

Krishnakumar, J. and Ballon, P., (2008): Estimating basic capabilities: A structural equation model applied to Bolivia. *World Development*, 36(6), pages 992-1010, June.

Layard, Richard (2005). Happiness: Lessons from a New Science. London: Penguin.

Lane, R., Reiman, E., Bradley, M., Lang, P., Ahern, G., Davidson, R. and Schwartz, G. (1997): Neuroanatomical correlates of pleasant and unpleasant emotion. *Neuropsychologia*, 35 (11): 1437-1444.

Larsen, Randy J. and Barbara L. Fredrickson (1999). 'Measurement Issues in Emotion Research.' In: Kahneman, D., Diener, R. and Schwarz, N., eds.: Well-Being: The Foundations of Hedonic Psychology. New York: Russell Sage Foundation: 40-60.

Levesque, J., Joanette, Y., Mensour, B., Beaudoin, G., Leroux, J.-M., Bourgouin, P. and Beauregard, M. (2003): Neural correlates of sad feelings in healthy girls. *Neuroscience*, 121 (3): 545-551.

Lindert K (2005) Implementing Means-Tested Welfare Systems in the United States. Social Protection Discussion Paper Series No. 532. World Bank, Washington, DC.

Lucas, R.E, (2005) "Time does not heal all wounds. A longitudinal study of reaction and adaptation to divorce". Psychological science, 16: 945-950.

Lokshin, M., Umpathi, N. and Paternostro,S. (2006): Robustness of subjective welfare analysis in a poor developing country: Madagascar 2001. *Journal of Development Studies*, 42(4): 559–591.

Lucas, R.E, Clark, A.E., Georgellis, Y and Diener, E., (2004): Unemployment alters the set point for life satisfaction. *Psychological science*, 15: 8-13.

Lucas, R.E, Diener,E., and Suh, E., (1996): Discriminant validity of subjective well being measures. *Journal of Personality and Social Psychology*, 71: 616-628.

Luhmann, M. and Eid, M. (2009): Does it really fell the same. Changes in life satisfaction following repeated life events. *Journal of Personality and Social Psychology*, 97: 363-381.

Luttmer, E. (2005): Neighbors as negatives: Relative earnings and well-being. *Quarterly Journal of Economics*, 120(3): 963–1002.

Milanovic, B. and Jovanovic, B. (1999): Changes in the perception of the poverty line during the depression in Russia, 1993–96. *The World Bank Economic Review*, 13(3): 539–559.

Morris, M.D., 1979, Measuring the condition of the world's poor: The physical quality of life index (Overseas Development Council, Washington, DC).

Nguefack-Tsague, G.,Klasen, S. And Zucchini, W., 2011. "On Weighting the Components of the Human Development Index: A Statistical Justification,"Journal of Human Development and Capabilities, Taylor and Francis Journals, vol. 12(2), pages 183-202.

Orshansky, M. (1965): Counting the poor: Another look at the poverty pro•le. *Social Security Bulletin*, 28: 3-29.

Osmani, S. (1982): Economic Inequality and Group Welfare. Oxford University Press, Oxford.

Paul, S. (1989): A model of constructing the poverty line. *Journal of Development Economics*, 30: 345-361.

Pollak, R. (1991): Welfare comparisons and situation comparisons. *Journal of Econometrics*, 50(1-2): 31–48.

Pradhan, M. and Ravallion, M. (2000): Measuring poverty using qualitative perceptions of consumption adequacy. *The Review of Economics and Statistics*, 82(3): 462–471.

Ram, R. (1982). Composite indices of physical quality of life, basic needs fulfillment and income: A principal component representation. *Journal of Development Economics*, 11: 227–247.

Rao, V. (2000): Price heterogneity and 'real' inequality: A case-study of prices and poverty in rural South India. *Review of Income and Wealth*, 46(2): 201–211.

Ravallion, M. (2011): On multidimensional indices of poverty. *Journal of Economic Inequality*, 9: 235–248.

Ravallion, M. (2010): A reply to Reddy and Pogge. In: Anand, S., Segal, P. and Stiglitz, J.E., eds. Debates on the Measurement of Global Poverty, New York: Oxford University Press: 86-101.:

Ravallion, M. (2010): Mashup indices of development. Policy Research Working Paper Series No. 5432.

Ravallion, M. and Lokshin, M. (2010): Who cares about relative deprivation. *Journal of Economic Behaviour & Organization*, 73: 171–185.

Ravallion, M. and Chen, S. (2009): Weakly relative poverty. World Bank Policy Research Working Paper No. 4844.

Ravallion, M., Chen, S. and Sangraula, P. (2008): Dollar a day revisisted. World Bank Policy Research Working Paper No. 4620.

Ravallion, M. and Chen, S. (2007): Absolute poverty measures for the developing world, 1985-2004. Proceedings of the National Academy of Sciences, 104: 16757–16762.

Ravallion, M. and Lokshin, M. (2006): Testing poverty lines. *Review of Income and Wealth*, 52(3): 399–421.

Ravallion, M. (2003): Measuring aggregate welfare in developing countries: How well do national accounts and surveys agree. *The Review of Economics and Statistics*, 85(3): 645–652.

Ravallion, M. (1998): Poverty lines in theory and practice. LSMS Working Paper No. 133.

Ravallion, M. and Sen, B. (1996): When method matters: Toward a resolution of the debate about Bangladesh's poverty measures. *Economic Development and Cultural Change*, 44(4): 761–792.

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

Ravallion, M. and G. Datt. (1993): Regional disparities, targeting, and poverty alleviation in India. In: Lipton, M. and van de Gaag, J., eds. 1993. Including the Poor. Washington DC: World Bank.

Ravaillon, M. (1993): Poverty alleviation through regional targeting: A Case Study for Indonesia. In: Hoff, K., Braverman, A. and Stiglitz, J.E., eds. 1993. The Economics of Rural Organization: Theory, Practice and Policy. Oxford: Oxford University Press.

Ravallion, M., Datt, G. and Van de Walle, D. (1991): Quantifying absolute poverty in the developing world. *Review of Income and Wealth*, 37(4): 345–361.

Ravallion, M. (1989): The welfare cost of housing standards: Theory with application to Jakarta. *Journal of Urban Economics*, 26: 197-211.

Ravallion, M. and Chao, K. (1989): Targeted policies for poverty alleviation under imperfect information: algorithms and applications. *The Journal of Policy Modeling*, 11(2): 213-224.

Reddy, S. and Pogge, T. (2010): How not to count the poor. In: Anand, S., Segal, P. and Stiglitz, J.E., eds. 2010. Debates on the Measurement of Global Poverty. Oxford: Oxford University Press: 42–85.

Reddy, S., Visaria, S. and Asali, M. (2006): Inter-country comparisons of income poverty based on a capability approach. UNDP International Poverty Centre Working Paper No. 27.

Rippin, N. (2010): Poverty severity in a multidimensional framework: the issue of inequality between dimensions. Discussion Paper No. 47.

Rippin, N. (2011): A response to the weaknesses of the Multidimensional Poverty Index (MPI): the correlation Sensitive Poverty Index (CSPI). German Development Institute (DIE) briefing paper 19/2011.

Rojas, M. (2007): A subjective well-being equivalence scale for Mexico: Estimation and poverty and income-distribution implications. Oxford Development Studies, 35(3): 273-293.

Runciman, W.G. (1966): Relative Deprivation and Social Justice. London: Routledge and Kegan Paul.

Ryff, C.D., Burton H.S. and Love, G.D. (2004): Positive health: connecting well-being with biology. *Philosophical Transactions Royal Society London B*, 359: 1383-1394.

Samuelson, P. (1974): Complenentarity - An essay on the 40th anniversary of the Hicks-Allen revolution in demand theory. *Journal of Economic Literature*, 15: 24–55.

Schwarz, N. and Strack, F. (1999): Reports of subjective well-being: Judgmental processes and their methodological implications. In: Kahneman, D., Diener, E. and Schwarz, N., eds. 1999. Well-Being: The Foundations of Hedonic Psychology. New York: Russell Sage Foundation: 61-84.

Scollon, C.N, Kim-Prieto, C. and Diener, E. (2003): Experience sampling: Promises and pitfalls, strengths and weaknesses. Journal of Happiness Studies, 4(1): 5-34.

Sen, A. (1976): Poverty: An ordinal approach to measurement. *Econometrica*, 844(2): 219–231.

Sen, A. (1979): Issues in the measurement of poverty. *The Scandinavian Journal of Economics*, 81(2): 285–307.

Sen, A. (1983): Poor, relatively speaking. *Oxford Economic Papers*, 35: 153–169.

Sen, A. (1985): Commodities and capabilities. Amsterdam: North-Holland.

Sen, A. (1987): The Standard of Living. Cambridge: Cambridge University Press.

Sen, A. (1988): The concept of development. In: Chenery, H. and Srinivasan, T.N., eds. 1988. Handbook of Development Economics. Elsevier: 9-26.

Sen, A. (1995a): Rationality and social choice. *American Economic Review*, 85(1): 1-24.

Sen, A. (1995b): The political economy of targeting. In: Van de Wall, D, and Nead, K., eds., 1995. Public Spending and the Poor. Baltimore: John Hopkins University Press.

Sen, A. (1998): On Ethics and Economics. Oxford: Oxford University Press.

Sen, A. (1999): Development as Freedom. Oxford: Oxford University Press.

Sen, A. (1999): The possibility of social choice. American Economic Review, 89(3): 349-378.

Senik, C. (2004): When information dominates comparison: Learning from Russian subjective panel data. *Journal of Public Economics*, 88: 2099–2123.

Silber, J. (2011): A comment on the MPI index. *Journal of Economic Inequality*, 9: 479–481.

Skoufias, E., Davis, B. and de la Vega, S. (1999): Targeting the poor in Mexico: An evaluation of the selection of households into PROGRESA. IFPRI.

Srinivasan, T.N. (2010): Irrelevance of the $1-a-day poverty line. In: Anand, S., Segal, P. and Stiglitz, J.E., eds. 2010. Debates on the Measurement of Global Poverty. Oxford: Oxford University Press: 143–153.

Stanovnik, T. (1992): Perception of poverty and income satisfaction. An empirical analysis of slovene households. *Journal of Economic Psychology*, 13(1): 57–69.

Stevenson, B. and Wolfers, J. (2008): Economic growth and happiness: Reassessing the Easterlin Paradox. *Brookings Papers on Economic Activity*, 1–87.

Steptoe, A., Wardle, J. and Marmot, M. (2005): Positive affect and health related neuroendocrine, cardiovascular, and inflammatory processes. Proceedings of the National Academy of Sciences, 102(18): 6508-6512.

Stone, A.A., Shiffman, S.S. and Devries, M.W. (1999): Ecological momentary assessment. In: Kahneman, D., Diener, E. and Schwarz, N., eds. 1999. Well-Being: The Foundations of Hedonic Psychology. New York: Russell Sage Foundation: 26-39.

Streeten, P. (1994): Human development: Means and ends. *American Economic Review, Papers and Proceedings*, 84: 232–237.

Subramanian, S. (2010): Identifying the income-poor: Some controversies in India and elsewhere. Courant Research Centre Discussion Paper No. 46, Göttingen University.

Tarp, F., Simmler, K., Matusse, C., Heltberg, R. and Dava, G. (2002): The robustness of poverty pro•les reconsidered. *Economic Development and Cultural Change*, 51(1): 77-108.

Thorbecke, E. (2007): Multidimensional Poverty: Conceptual and Measurement Issues. In: Kakwani, N. and Silber, J., eds.: The Many Dimensions of Poverty. Palgrave-MacMillan.

Thorbecke, E. (2011): A comment on multidimensional poverty indices. *Journal of Economic Inequality*, 9: 485–487.

Townsend, P. (1979): Poverty in the United Kingdom. Harmondsworth: Penguin.

Townsend, P. (1979): The development of research on poverty. In: Social Security Research: The De•nition and Measurement of Poverty. London: Department of Health and Social Security.

Van de Kerk, G. and Manuel, A. (2008): A comprehensive index for a sustainable society: the SSI – the Sustainable Society Index. *Ecological Economics*, 66: 228-242.

Veenhoven, R. (1991). Is happiness relative? *Social Indicators Research*, 24: 1-34.

Vittersø, J., Biswas-Diener, R. and Diener, E. (2005): The divergent meanings of life satisfaction: Item response modeling of the satisfaction with life scale in Greenland and Norway. *Social Indicators Research,* 74(2): 327-348.

Watson, D., Clark, L.A. and Tellegen, A. (1988): Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, 54(6): 1063-1070.

Wodon, Q. (1997): Food energy intake and cost of basic needs: Measuring poverty in Bangladesh. *Journal of Development Studies*, 34: 66–101.

World Bank (1990): World Development Report. New York: Oxford University Press.

World Bank (2000): World Development Report. New York: Oxford University Press.

Zeller, M., Sharma, M., Henry, C., and Lapenu, C. (2006): An operational method for assessing the poverty outreach performance of development policies and projects: Results of case studies in Africa, Asia, and Latin America. *World Development*, 34(3): 446-464.