Policy Research Working Paper 10800

Multidimensional Well-Being Measurement Practices

A Review Focused on Improving Global Multidimensional Poverty Indicators

Benoit Decerf

WORLD BANK GROUP
Development Research Group &
Poverty and Equity Global Practice
June 2024

Abstract

Multidimensional well-being indicators have the potential to reduce the "bias" associated to monetary indicators. However, they face stringent data constraints. This paper studies the construction of indicators that strike a balance between (i) reliability in approximating conceptually sound well-being comparisons and (ii) simplicity of application and communication. The recommendations focus on global multidimensional poverty measures. The paper identifies three potential sources of improvements: "wasting" less data, better filtering the data, and further developing multidimensional analysis. Less information would be "wasted" by avoiding needlessly dichotomizing all the variables, using the available mortality data, and combining variables from separate surveys. To filter the data better, "equal weights" could be replaced by weights selected from external information on preferences. When the data permit, the unit of analysis should be switched from household level to individual level. Finally, multidimensional indicators should be used to help move beyond a suboptimal "dimension-by-dimension" approach to policy making.

This paper is a product of the Development Research Group, Development Economics and the Poverty and Equity Global Practice. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://www.worldbank.org/prwp. The author may be contacted at bdecerf@worldbank.org.

Multidimensional Well-Being Measurement Practices: A Review Focused on Improving Global Multidimensional Poverty Indicators*

Benoit Decerf
DECPI, World Bank

JEL: I32.
Keywords: Multidimensional Well-Being, Poverty Measurement, Data constraints.

*Acknowledgments : I am grateful to Paola Ballon, Mery Ferrando, Jed Friedman, Balint Menyhert and Berk Ozler for helpful suggestions. I thank all the participants at two internal seminars of the World Bank. This paper was prepared as part of the Research Frontiers in Poverty Measurement series organized by the Global Solutions Group Data for Policy of the Poverty and Equity Global Practice and DEC, led by Utz Pape and Jed Friedman. The findings, interpretations, and conclusions expressed in this paper are entirely those of the author and should not be attributed in any manner to the World Bank, to its affiliated organizations, or to members of its Board of Executive Directors or the countries they represent. The World Bank does not guarantee the accuracy of the data included in this paper and accepts no responsibility for any consequence of their use.

1 Introduction

Well-being indicators are important tools for evidence-based policy making. They allow targeting the worst-off, allocating efforts towards regions with the lowest well-being and monitoring progress. These tasks are most often done with monetary indicators. However, monetary indicators typically yield biased well-being comparisons because they do not account for the multidimensional nature of well-being. Multidimensional well-being indicators, which also account for non-monetary outcomes, can potentially reduce this bias.

In practice, the construction of multidimensional indicators faces stringent data constraints. First and foremost, there are no well-defined prices for non-monetary outcomes or "achievements". As a result, the practitioner typically has limited or no information on the preference that is relevant to rank achievement vectors. Second, the joint distribution between key dimensions of well-being is often missing. Hence, the practitioner may observe the partial distributions, but does not know whether the same individuals cumulate low achievements in several dimensions. These two constraints prevent practitioners from constructing conceptually-ideal multidimensional indicators.

These data constraints have led to two opposite extreme reactions. The first reaction consists in ditching a preference-based notion of well-being as well as non-paternalism. Many pragmatic theories for multidimensional indicators deal with the absence of information on the relevant preference by setting aside preference theory. This is for instance the case of the theory grounding the multidimensional poverty measures most used in practice.[1] The second extreme reaction, which is perhaps triggered by the first extreme reaction, is to disqualify multidimensional indicators. According to this view, a dashboard of indicators is sufficient for policy making and a summary multidimensional indicator brings no added value. Admittedly, the conceptual foundations for pragmatic multidimensional indicators seem less solid than those of monetary indicators. However, this does not imply that they necessarily yield less reliable well-being comparisons. Arguably, neither of these extreme reactions is fully satisfactory.

This review aims at identifying opportunities to further improve multidimensional well-being measurement practices. The perspective taken is pragmatic although it strives to remain conceptually sound. It is conceptually sound in the sense that it is informed by a general and classic preference-based notion of well-being. It is pragmatic because it starts from current practices, takes data constraints as given and searches for improvements while keeping an eye on simplicity of implementation and communication. Under this nuanced perspective, the objective is to construct indicators that most reliably yield well-being comparisons in line with the theory. The best indicator thus minimizes the

[1] Both the functional form and the parameters values of many "pragmatic" multidimensional indicators are selected without much regard for whether their well-being comparisons somewhat align with individual preferences.

incorrect comparisons coming from data constraints. Loosely speaking, this indicator yields the largest "signal-to-noise ratio" when filtering available data.

The review focuses on global multidimensional poverty measures. First, multidimensional poverty measures are arguably the multidimensional indicators most used for policy making, perhaps because of the simplicity of the Alkire-Foster (AF) methodology. Second, global applications face the most stringent data constraints. However, most of the improvement opportunities we identify still make sense at the national or regional level. The conceptual framework proposed can also guide well-being measurement practices beyond poverty.

We identify three sources of potential improvements: "wasting" less of the available data, better filtering the data used, and better harnessing the added-value of multidimensional poverty measures. The first two sources relate to the construction of the indicator while the third source relates to the use of the indicator.

The main source of potential improvements relates to "wasting" less data. Ironically, in spite of stringent data constraints, large amounts of relevant data are simply ignored. To a large extent, this problem originates from conceptual rigidities constraining the indices used to aggregate across dimensions. We identify three ways of reducing the current "waste" of data.

First, the AF methodology to identify the multidimensionally poor is based on dichotomous deprivation statuses, i.e., either being deprived or non-deprived in a given variable. However, not all variables are dichotomous. For instance, this is neither the case of consumption, which is a cardinal variable, nor the case of some health or education outcomes, which are captured by categorical variables. Dichotomozing such variables leads to a loss of information. We identify in the literature two conceptually sound solutions to reduce this loss of information and show how these solutions can be implemented even in the absence of information on the relevant preference.

Second, in spite of its large intrinsic value, longevity is ignored by all global multidimensional poverty measures. Mortality data are often available but these data are not used. One difficulty is that the joint distribution between mortality and other achievements is almost always missing. However, this does not mean that indicators that ignore mortality yield better cross-regions comparisons, quite the contrary. There exists several conceptually sound indices that integrate mortality into poverty measures while assuming that all individuals within a given region face the same mortality risks. These indices are "hybrid" in the sense that they account for the joint distribution between some but not all variables considered.[2]

[2]For instance, assume that the joint distribution between two variables x_1 and x_2 (say income and education) is known, but the joint distribution between variable x_3 (mortality) and the other two variables is unknown. Consider an index that aggregates x_1, x_2 and x_3 while taking into account the joint distribution between x_1 and x_2. This index is "hybrid" in the sense that it accounts for the joint distribution between some but not all variables considered.

Third, the joint distribution between consumption and health outcomes is often missing because monetary and non-monetary outcomes are typically collected through separate surveys. The classical solution used by global multidimensional poverty measures is to ignore the data on one of these two key aspects of well-being. We identify three alternative methods to construct a hybrid index accounting for data coming from separate surveys. The first two methods consist in assuming a value for the missing joint distribution, while the last method assumes equal non-monetary outcomes at the local level. The first method assumes the overlap between the monetary and the non-monetary poverty head-count ratios. We show in one context that this method yields better well-being comparisons than the classical solution regardless of the exact value assumed for the (unobserved) overlap. Shockingly, "making up" overlap data can improve well-being comparisons. The reason is that it allows increasing the amount of data used. The second method consists in imputing consumption into the non-monetary survey. We argue that the relevant benchmark to evaluate the relevance of such imputation approach is *not* whether the joint distribution assumed is sufficiently correct but rather whether it improves well-being comparisons. The third method consists in constructing a "total consumption" variable that aggregates monetary consumption with monetary valuations of non-monetary outcomes, which are considered equally distributed at local level. We propose one way of constructing the "total consumption" variable that recognizes that each core dimension of well-being is essential.

The second source of potential improvements relates to better filtering the data used. We identify two ways in which this could be done. The first is to move beyond the default procedure that assigns equal weights to all dimensions. This procedure has the advantage of being applicable without information on the relevant preference, but runs the risk of making implausible trade-offs. Some information on the relevant preference can be obtained from subjective well-being data or from dedicated preference elicitation surveys. We succinctly present a procedure to select the weights from these data sources.[3] Importantly, this procedure accounts for the fact that the weights of multidimensional poverty measures are *not* marginal weights of substitution. This fact implies that seemingly natural procedures are not fit for purpose.[4] The second is to switch the unit of analysis from the household level to the individual level as soon as *some* variables are collected at individual level. We discuss the advantages of this switch as well as its associated misinterpretation risk.

The third source of potential improvements relates to the way in which the indicator is used. Merely reporting that monetary poor individuals tend to cumulate non-monetary deprivations gives the impression that the exercise has no added value. Multidimensional

[3]To be sure, it is possible to use *fixed* weights even when their value is selected from information on preferences.

[4]Multidimensional poverty measures will make incorrect well-being comparisons when the values for their weights are selected to correspond to marginal rates of substitution.

indicators, when properly used, have the potential to improve the identification of the worst-offs and to improve policy making. By means of a toy example, we show that dimension-by-dimension policy making is suboptimal and that multidimensional indicators can yield valuable insights for policy makers. It would be worth reviewing the most relevant types of analysis to perform with multidimensional poverty measures. This could help standardize such analysis.

The remainder is organized as follows. Section 2 presents our pragmatic approach to designing indicators that best approximate a conceptually-sound notion of well-being. Section 3 details the opportunities we identify to further improve global multidimensional poverty measures. Section 4 summarizes the opportunities identified and briefly discusses their relative relevance. Section 5 lists research papers to study these opportunities in practice.

2 Well-being measurement in practice

2.1 General notion of well-being

We consider throughout a general notion of well-being.

Let $x_i = (x_{ij})_{j \in J}$ denote the achievement vector of individual i. For each well-being relevant dimension $j \in J$, this vector specifies individual i's outcome x_{ij}. This outcome could capture the quantity she consumes of good j, as in the monetary case, or it could measure some nutritional intake or yet some level of rights secured, capability or functioning achieved.

Any meaningful notion of individual well-being yields a ranking on achievement vectors. This ranking follows from the individual's preference under a non-paternalistic view. Without loss of generality,[5] this ranking can be represented by a utility function u_i. In this paper, we stick to ordinal comparisons of achievement vectors and only use function u_i as a convenient tool for representing these comparisons. Assuming the same ranking for all individuals,[6] individual well-being can be compared across individuals and across time with

$$u(x_i). \qquad \text{(Individual well-being)}$$

A standard notion of social well-being can be defined as the average value of some

[5]The notion of individual well-being considered is *ordinal* because it only depends on the underlying ranking. In particular, it is much more general than a welfaristic notion. It also include notions following from minimal rights or capabilities (Sen, 1999; Nussbaum, 2009).

[6]This is a strong assumption because individuals are likely to hold heterogeneous preferences. However, this assumption is almost always made in practice given the lack of information on preferences. For a recent application that does not make this assumption, see Boarini et al. (2022).

monotonic transformation of individual well-being (Piacquadio, 2017).[7]

$$\frac{1}{N}\sum_i f(u(x_i)), \qquad \text{(Social well-being)}$$

where N is the number of individuals. Let U denote *social well-being* when function f is an increasing transformation. For instance, we get average individual well-being when function f is defined as $f(u(x_i)) = u(x_i)$. Let P denote *poverty* when function f is a decreasing transformation based on a reference "low well-being" threshold $\underline{u}$. For instance, we get the poverty head-count ratio when function f is defined as $f(u(x_i)) = 1$ when $u(x_i) < \underline{u}$ and 0 otherwise.

2.2 Why aggregate across dimensions?

Ravallion (2011) observes that aggregating across dimensions is not necessary for many policy purposes. However, this does not imply that aggregating is never helpful for policy making, as is sometimes heard. Indeed, aggregating can be useful for at least three reasons, all related to evidenced-based policy making.

First, prioritarian policy makers target their efforts towards individuals with lowest well-being. Identifying those individuals requires making inter-personal well-being comparisons, typically

$$u(x_i) > u(x_{i'}) \qquad (i' \text{ is worse off than } i)$$

$$\underline{u} > u(x_{i'}) \qquad (i' \text{ has low well-being.})$$

Observe that these comparisons depend on the whole achievement vector. For instance, in Figure 1, i is better off than i' even though $x_{i1} < x_{i'1}$.

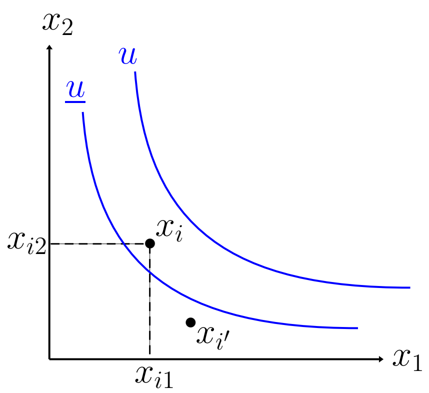

Figure 1: Inter-personal well-being comparisons require aggregation.

[7] Even though the individual well-being index $u(x_i)$ is ordinal, its monotonic transformation $f(u(x_i))$ is cardinal, i.e., differences in the level of $f(u(x_i))$ are assumed meaningful.

Second, prioritarian policy makers allocate their efforts and budgets towards the regions with lowest social well-being. Identifying those regions requires making cross-regions well-being comparisons, typically

$$U\left((x_i)_{i\in R}\right) > U\left((x_i)_{i\in R'}\right) \qquad \text{(Region } R \text{ is better off than region } R'\text{)}$$
$$P\left((x_i)_{i\in R}\right) < P\left((x_i)_{i\in R'}\right) \qquad \text{(Region } R \text{ is less poor than region } R'\text{)}$$

Third, progress is monitored by following the trend of indicators U or P. Forming an objective idea on whether well-being has improved over time also requires integrating across dimensions. Indeed, a dashboard of dimension-specific indicators is silent on the direction of progress when several of its indicators follow opposite trends.

2.3 Limits of the monetary approach

Well-being measurement is relatively simple in theory but quite complex in practice. One key reason is that, in practice, neither u nor x_i are directly observed. Instead, the practitioner constructs well-being indicators based on imperfect data.

Well-being measurement has long been dominated by the monetary approach. This approach focuses on "market" dimensions (food and non-food goods and services) for which well-defined market prices are available. Formally, the set of well-being dimensions, denoted by J, is restricted to J^{market}. The data available is $d_i = (p_i, q_i)$, where p_i is the price vector and q_i is the bundle, which summarizes the quantities of goods and services consumed. Bundle q_i thus captures the relevant outcome $(\hat{x}_i)$[8] and the price vector p_i provides information on the relevant preference u. Indeed, when assuming rational choice, p_i provides the marginal rate of substitution of u at $\hat{x}_i$. The welfare aggregate of i is defined as

$$\hat{u}(d_i) = \frac{\sum_{j\in J^{market}} p_{ij} q_{ij}}{\text{Price index}} \tag{1}$$

where an appropriate price index is used for deflation purposes (Diewert, 1998; Mancini and Vecchi, 2022).

The monetary approach has good conceptual foundations. Under a series of strong assumptions, the welfare aggregate $\hat{u}(d_i)$ is ordinally equivalent to individual well-being $u(x_i)$, which means that

$$\hat{u}(d_i) > \hat{u}(d_{i'}) \qquad \Leftrightarrow \qquad u(x_i) > u(x_{i'}).$$

More precisely, we call $\hat{u}(d_i)$ money-metric welfare when $J = J^{market}$, individuals behave

[8] The "hat" in our notation $\hat{x}_i$ is meant to emphasize that data only provide noisy information on the relevant outcomes x_i.

rationally on markets, preference u is homothetic, etc. See Ravallion (2016) for a broad view of the monetary approach.

In practice, the monetary approach suffers from several limitations. For our purposes, we mention one of its key limitations.[9] Well-being is widely recognized as a multidimensional phenomenon (Stiglitz et al., 2009). This multidimensionality implies that the monetary approach yields biased well-being comparisons.

We illustrate these biased comparisons in Figure 2. In the figure, the x-axis captures the monetary welfare aggregate while the y-axis captures some non-monetary achievement, say health, i.e., $(x_{i1}, x_{i2}) = (\hat{u}(d_i), health_i)$. The monetary approach classifies individuals as poor or not based on a poverty line, which is a threshold monetary amount z (dashed red). Individual i is monetary poor while individual i' is not. However, the "true" classification of low well-being should rather be based on a well-being threshold $\underline{u}$ (blue). Individual i does not have low well-being, while individual i' has low well-being. The monetary approach makes an "inclusion error" on i and an "exclusion error" on i'. Its well-being comparison is flawed as $\hat{u}(d_i) < \hat{u}(d_{i'})$ although $u(x_i) > u(x_{i'})$.

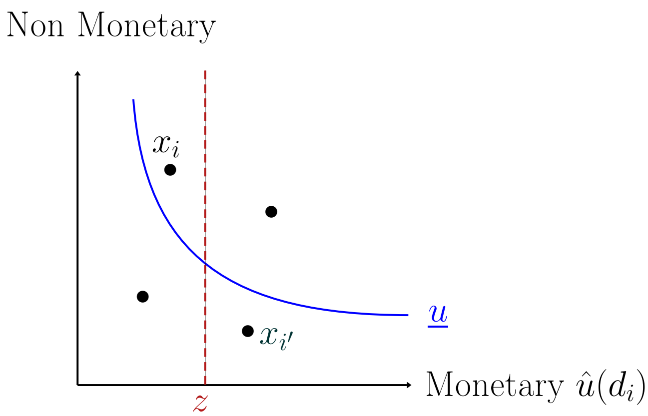

Figure 2: The monetary approach yields biased well-being comparisons.

2.4 Data constraints on non-monetary dimensions

One obvious way to improve on the monetary approach is to construct indicators that also include information on non-monetary outcomes, such that $J = J^{market} \cup J^{non-market}$. One key difficulty in practice is that data constraints on $J^{market} \cup J^{non-market}$ are much more stringent than on J^{market}. Three types of data-constraints are particularly problematic: noisy data on non-monetary outcomes, missing joint distributions and lack of data on preference u.

[9] Another limitation is for instance that the monetary approach often ignores intra-household inequalities. The reason is that its data are mostly available at the household level, not at the individual level. Over the last two decades, a series of contributions started exploring the distribution of consumption within households (Browning and Chiappori, 1998; Cherchye et al., 2015) and found substantial inequality (Lise and Seitz, 2011; Dunbar et al., 2013).

First, the data available on non-monetary outcomes x_{ij} can be noisier than the data available on monetary outcomes. When selecting a variable $\hat{x}_{ij}$ capturing a non-monetary dimension j, there often is a trade-off between at least three conflicting goals. We present these three goals when assuming that health is the dimension j for which variables must be selected.

1. $\hat{x}_{ij}$ approximates x_{ij} with limited noise.

   Self-reported health surveys lack the objectivity of medical check-ups and thus noisily capture the health status of respondents. Many variables like vaccines administred or access to clinics relate to health *inputs* and thus noisily capture health *outcomes*.

2. $\hat{x}_{ij}$ offers a good coverage of dimension j.

   Having received a vaccine or having delivered a baby in a clinic only capture very limited aspects of a person's health. Indeed, they both ignore disabilities, nutritional status, chronic diseases, etc. Arguably, self-reported surveys like EQ5D provide a larger coverage of a person's health (Szende et al., 2014).

3. $\hat{x}_{ij}$ is available or cheaply collected.

   Medical check-ups could provide a broad coverage of a person's health with limited noise, but their cost could be prohibitive.

In most cases, the variable $\hat{x}_{ij}$ that provides the best trade-off between these three conflicting goals is not cardinal, but rather ordinal or even categorical. As a result, the practitioner must rely on indicators that can accommodate ordinal data.

Second, the joint distribution between outcomes in different dimensions j and j' is often missing. The inter-personal comparisons presented in Figure 2 require not only partial distributions $(x_{i1})_{i\in R}$ and $(x_{i2})_{i\in R}$ but also the joint distribution. Without the joint distribution, an individual's achievement vector (x_{i1}, x_{i2}) is not known and we thus ignore whether an individual cumulates low achievements in several dimensions. One problem is that non-monetary outcomes are often collected through different surveys than monetary outcomes. At global level, monetary outcomes are collected through LSMS-type surveys while non-monetary outcomes are collected through DHS or MICS surveys.[10] As a result, the joint distribution between monetary outcomes (consumption or income) and non-monetary outcomes is often missing. A similar problem arises for the mortality dimension. The joint distribution between mortality risks and other well-

[10]At country level, this constraint may not be present, which allows some countries like Mexico to construct multidimensional poverty indicators capturing both monetary outcomes and non-monetary outcomes (Bank et al., 2021).

being relevant outcomes like consumption is almost systematically missing.[11] As a result, expected longevity cannot be estimated by consumption categories.

Last but not least, the data contains very limited information on the preference u over non-monetary outcomes. The problem is that there are no well-defined prices on non-market dimensions. More generally, the non-monetary outcomes cannot reveal preference u when individuals cannot (or do not) optimally select these outcomes. As a result, it is harder to estimate a preference $\hat{u}$ over $J^{non-market}$. It is therefore difficult for practitioners to select appropriate weights between variables $\hat{x}_{ij}$ and $\hat{x}_{ij'}$. This contributes to the popularity of ad-hoc methods for the selection of these weights, like the "equal weights" procedure.

2.5 Multidimensional indices under data constraints

The stringent data constraints described just above help categorize the types of multidimensional indices used in practice to aggregate well-being across dimensions. This is in particular the case for the latter two constraints, namely the missing joint distributions and the scarce information on preference u. We illustrate this in Table 1, in which the two rows differ in the availability of joint distributions and the two columns differ in the availability of data on preference u.

Table 1: Multidimensional indices classified by data constraints.

		Data on preference u	
		None (Mash-up "weights")	Available (Calibrated "weights")
Data on joint distribution	None (Synthetic index)	HDI	Jones and Klenow (2016), Boarini et al. (2016),
	Available (Objective index)	MPM global MPI	Decancq et al. (2019) Boarini et al. (2022)

Consider first the rows in Table 1. Following Fleurbaey (2009), we call "synthetic" the indices that are constructed from partial distributions only. Synthetic indices aggregate within dimension first and across dimensions second. In other words, they first construct dimension-specific statistics over the population and then aggregate these statistics across dimensions. This is for instance the case of the Human Development Index (HDI) (Anand and Sen, 1994). The HDI captures three dimensions, namely incomes, longevity and education. The partial distributions for the former two dimensions are summarized as mean income and life expectancy. These statistics are then normalized and aggregated

[11] Death is a rare event that is hard to properly capture in surveys, which partly explains why this joint distribution is missing.

using equal weights.

We call "objective" the indices that are constructed using the joint distributions of outcomes. Objective indices first aggregate across dimensions at the individual level to generate an individual index $\hat{u}(d_i)$. Then, the individual indices are aggregated across individuals, typically by averaging them. This is for instance the case of the two main global multidimensional poverty indices: the global MPI of UNDP-OPHI and the World Bank's MPM, both defined in Appendix 6.1. These indices generate their individual indices $\hat{u}(d_i)$ following the Alkire-Foster aggregation methodology (Alkire and Foster, 2011a). This methodology makes a weighted sum of dichotomous deprivation statuses in a number of dimension-specific variables (see Section 3.2).

From a conceptual perspective, objective indices are more appealing than synthetic indices because the former account for well-being inequalities across individuals (Fleurbaey, 2009). Synthetic indices ignore that some individuals cumulate low achievements in several dimensions. Moreover, synthetic indices do not generate an individual index $\hat{u}(d_i)$ and can therefore not be used for inter-personal well-being comparisons. However, synthetic indices can be computed even when the joint distribution is missing.

Consider then the columns in Table 1. In absence of data on preference u, the typical practice consists in using ad-hoc methods to set weights, like the popular "equal-weights" procedure.[12] The risk when resorting to this ad-hoc procedure is to select weights that are not plausible, which casts doubt on the added-value of the multi-dimensional indicator (Ravallion, 2012). Going beyond ad-hoc methods may help increase the confidence in well-being comparisons obtained from multidimensional indicators. This requires selecting the values of weights from some empirical source, like dedicated preference-elicitation surveys (willingness-to-pay, stated preferences, etc.), subjective well-being data or external sources in the literature. For instance, the dimension-specific statistics feeding synthetic indices are sometimes aggregated using a parametric utility function. Its parameters can be calibrated using empirical values taken from specialized literatures. For the trade-off between income and longevity, one can draw on empirical estimates for the elasticity of inter-temporal substitution and the value of a statistical life (see section 3.3).

[12] In the case of multidimensional poverty measures, the equal-weights procedure is more nuanced than what its name may suggest. The reason is that, in practice, each dimension is captured by several dimension-specific variables. The equal-weights practice attributes equal weights between dimensions and also attributes equal weights among the variables capturing the same dimension. Hence, the weight received by each variable is given by $\frac{1}{n_{dim}} * \frac{1}{n_{var}}$ where n_{dim} denotes the number of dimensions and n_{var} denotes the number of variables capturing the dimension. For instance, the global MPI has three dimensions: education, health and living standards. Health is captured by two variables (nutrition and child mortality) that each receives a weight of $\frac{1}{6} = \frac{1}{3} * \frac{1}{2}$. Living standards is captured by six variables that each receives a weight of $\frac{1}{18}$.

2.6 Minimizing incorrect well-being comparisons

The previous sections introduced different approaches to well-being comparisons. Data constraints limit the choices available to the practitioner. However, beyond data constraints, the practitioner must still exercise judgment when selecting the most appropriate well-being indicator. The reason is that different approaches come with different trade-offs. In this section, we describe a conceptual framework that help navigating the trade-offs presented by alternative indicators. Among these trade-offs, we have

- The monetary approach ignores outcomes on $J^{non-market}$ but aggregates outcomes on J^{market} using individual indices $\hat{u}(d_i)$ that are related to individual preferences.
- Multidimensional poverty measures account for some outcomes on $J^{non-market}$, but they ignore other key outcomes on $J^{non-market}$ like longevity, for which the joint distribution is missing.
- Synthetic indices can potentially account for a large number of outcomes on $J^{non-market}$, but they mostly ignore unequal outcomes across individuals and do not allow inter-personal well-being comparisons.

We argue that to be policy-relevant, these trade-offs should be navigated with the aim of constructing an indicator whose incorrect well-being comparisons are minimized. Loosely speaking, this indicator should have maximal "signal-to-noise ratio". Well-being is a latent variable, which is never directly observed. It is only indirectly observed through its determinants like income, longevity, health, etc. Some imperfect data on these determinants are available. A well-being indicator "filters" some of the available data with the objective to make comparisons that are, as much as possible, in line with the latent well-being comparisons. From section 2.2, two types of well-being comparisons are policy relevant. First, a well-being indicator ($\hat{u}$) should make inter-personal comparisons that are reliably in line with the "true" comparisons (with u). That is, such indicator should as much as possible yield

$$\begin{array}{rcll} \hat{u}(d_i) > \hat{u}(d_{i'}) & \text{when} & u(x_i) > u(x_{i'}) & \text{(Objective 1)} \\ \underline{u} > \hat{u}(d_{i'}) & \text{when} & \underline{u} > u(x_{i'}). & \end{array}$$

Second, a well-being indicator ($\hat{U}$ or $\hat{P}$) should make cross-regions comparisons that are reliably in line with the "true" comparisons (with U or P). That is, such indicator should as much as possible yield

$$\begin{array}{rcll} \hat{U}(d_R) > \hat{U}(d_{R'}) & \text{when} & U((x_i)_{i\in R}) > U((x_i)_{i\in R'}) & \text{(Objective 2)} \\ \hat{P}(d_R) < \hat{P}(d_{R'}) & \text{when} & P((x_i)_{i\in R}) > P((x_i)_{i\in R'}). & \end{array}$$

Crucially, given the imperfect data on dimension-specific outcomes and the data constraints on joint distributions and preference u, a well-being indicator will sometimes make "errors". For instance, yield $\hat{u}(d_i) < \hat{u}(d_{i'})$ when in fact $u(x_i) > u(x_{i'})$. We argue that the practitioner must select the data and the multidimensional index so as to minimize these errors. Importantly, minimizing these errors is not merely of conceptual importance, but it also improves policy making. Indeed, better well-being comparisons directly yield better targeting of policies, be it across individuals (Objective 1) or across regions (Objective 2).

Maximizing an indicator's signal-to-noise ratio requires judgment because some trade-offs may appear. For instance, adding more data is not always a good idea. The additional data might *reduce* the signal-to-noise ratio. This can happen if the additional data is too noisy or if the multidimensional index does not filter properly the additional information. To understand the intuition for this, assume for instance that $u(x_1, x_2) = \alpha x_1 + (1-\alpha)x_2$, where α represents the "true" weight of dimension 1.[13] Assume the data available is $d = (\hat{x}_1, \hat{x}_2)$ where the dimension-specific variables are normally distributed $\hat{x}_1 \sim N(\mu_1, \sigma_1)$ and $\hat{x}_2 \sim N(\mu_2, \sigma_2)$ with $\mu_1 = x_1$ and $\mu_2 = x_2$. Should we use indicator $\hat{u}(d) = \hat{x}_1$ or rather $\hat{u}'(d) = \hat{\alpha}\hat{x}_1 + (1-\hat{\alpha})\hat{x}_2$? Indicator $\hat{u}'$ uses more information, but more data can potentially reduce the signal-to-noise ratio when

- Noisy data.

  Indicator $\hat{u}$ might be preferable to $\hat{u}'$ if the noise σ_2 of variable $\hat{x}_2$ is much larger than the noise σ_1 of variable $\hat{x}_1$.

- Implausible weight.

  Indicator $\hat{u}$ might be preferable to $\hat{u}'$ if dimension 1 is key to well-being ($\alpha >> 1/2$) but its assumed weight is implausibly low ($\alpha >> \hat{\alpha}$).

Observe that these two considerations interract and may be in conflict.

- Noisy data on key well-being dimension

  Indicator $\hat{u}'$ might be preferable to $\hat{u}$ even when the noise σ_2 is larger than the noise σ_1, at least if dimension 2 is sufficiently important ($\alpha << 1/2$).

The correlation with unobserved variables may also be relevant. For instance, it might be worth adding a variables that has little relevance to well-being if this variable is strongly correlated to another unobserved variable that is key to well-being.[14] Conversely, $\hat{u}$

[13]The same points could be made for cross-country comparisons. We take a linear expression for simplicity.

[14]Assume for instance that $u(x_1, x_2, x_3) = \alpha x_1 + \beta x_2 + (1-\alpha-\beta)x_3$ with $\alpha = 1/3$, and β is close to zero. Assume again that $d = (\hat{x}_1, \hat{x}_2)$, meaning there is no variable capturing dimension 3. If $\hat{x}_2$ is strongly correlated to x_3, indicator $\hat{u}'$ might be preferable to $\hat{u}$. Given the importance of dimension 3 ($1/3 < 1-\alpha-\beta$), it might be good to select a high weight to variable $\hat{x}_2$, namely $\hat{\alpha} \leq 1/2$. Such

might be preferable to $\hat{u}'$ if the two variables $\hat{x}_1$ and $\hat{x}_2$ are highly correlated. Indeed, the additional complexity of building indicator $\hat{u}'$ might not be worth its additional benefits.

Clearly, this framework aimed at minimizing well-being comparison errors can be applied to well-being measurement at large. It can also guide the improvement of monetary indicators, which also face data constraints and thus yield erroneous comparisons in practice.

2.7 Improving on current indicators

It seems unlikely that the monetary approach yields indicators whose well-being comparison errors are minimized. Even though income or consumption are important determinants of well-being, several non-monetary dimensions like health and longevity are arguably at least as important.[15] Furthermore, longevity has been regularly shown to substantially affect cross-country well-being comparisons.[16] The fact that multidimensional poverty measures ignore longevity was placed under crude light during the Covid-19 crisis.[17] Also, consumption and income indicators capture in a noisy way the control that people exert over economic resources.[18] Therefore, the challenge is to build a multidimensional index making fewer well-being comparisons errors than monetary indicators.

Consider first the case when the goal is to identify low well-being across individuals *within* a given region, say for the purpose of targeting social spendings. For this case, the indicator should make reliable inter-personal comparisons (Objective 1). Thus, synthetic indicators are of no use and one must consider objective indicators. As a result, one must use a set of variables for which the joint distribution is available. At global level, given the current data constraints discussed in Section 2.4, this means to either ignore the monetary dimension or ignore the health dimension. The global MPI made the former choice while the World Bank's MPM made the latter choice.[19] Currently, a key challenge

reasoning may provide a rationale for attributing a quite high weight to education in the global MPI, even though the intrinsic value of education to well-being is disputed (Ravallion, 2011), at least if education is strongly correlated with consumption.

[15]Subjective well-being data suggests that good health and longevity are key determinant of well-being (Adler et al., 2017; Diener et al., 2018). The causal relationships behind the statistical association between subjective well-being, health and mortality are still an active area of research (Steptoe et al., 2015).

[16]A string of applied papers suggest that adding longevity significantly affects cross-country well-being comparisons (Becker et al., 2005; Jones and Klenow, 2016; Boarini et al., 2016).

[17]See Chapter 3 of the Poverty and Shared Prosperity Report (Bank, 2022) where the multidimensional impact of Covid-19 is analyzed through lenses different from the World Bank's MPM.

[18]For instance, prices are imperfectly captured through unit values, rents implicitly perceived by homeowners are imperfectly captured through hedonic regressions, recall periods used in household consumption surveys affects welfare aggregates, etc.

[19]At this stage, it is an open question which of these two choices should yield the larger "signal-to-noise ratio" given the data available. On the one hand, the global MPI include a few variables capturing "standard of living", which may serve as a noisy proxy for consumption. On the other hand, the variables capturing "health" in the global MPI provide a limited coverage of that dimension, as discussed in Section 2.4. It would be interesting to compare the correlation at household level between subjective well-being

is to increase the signal-to-noise ratio that these indicators get from the data they use. As we discuss below, this can be done by improving on their ad-hoc weights or by increasing the information they extract from the variables they use.

Consider then the case where the goal is to identify which region has low well-being (Objective 2). Both synthetic indicators and objective indicators could be relevant for cross-region comparisons. Synthetic indicators allow accounting for some non-monetary variables for which the joint distribution is missing. Objective indicators allow accounting for the joint-distribution for the subset of variables for which this information is available. As we discuss below, hybrid indicators that are part-synthetic and part-objective would allow harnessing both types of data. These hybrid indicators have the potential to simultaneously account for consumption, health and longevity.

3 Opportunities for improvement

We present alternative opportunities that we believe could help further improve global well-being measurement practices. We start by explaining the criteria used to select these opportunities.

3.1 Criteria for selecting opportunities

Improving the "signal-to-noise ratio" of mainstream global well-being indicators can be done in many different ways. A very large literature explores different routes, notably through the use of new data sources such as mobile phone, remote sensing or biometric data (Steele et al., 2017; Chi et al., 2022; Blanchflower and Bryson, 2022), aggregation methods to account for heterogeneous preferences (Fleurbaey and Maniquet, 2011; Piacquadio, 2017; Decancq et al., 2019; Boarini et al., 2022) or yet alternative preference elicitation methods (Benjamin et al., 2014b; Decancq and Nys, 2021). In this review, we only consider a small subset of opportunities, which we selected based on a few criteria. In a nutshell, these criteria relate to (1) the communicability of the indicator, (2) the simplicity of its construction method and (3) the potential for short term application, i.e., data availability and maturity.[20] We shortly detail these criteria, which admittedly all include elements of subjectivity.

First, we focus on indicators whose value can easily be communicated to policy makers and the public. This excludes indicators whose units are not intuitive to the lay person. We focus mostly on poverty measures because we believe they are the most successful distribution-sensitive well-being indicators.[21] Arguably, everyone can make an intuitive

indicators and these two poverty measures.

[20]The first two criteria are echoed in the expert report on poverty measurement (Atkinson, 2016).

[21]The two main social well-being indicators discussed beyond academic circles are mean income and poverty head-count ratios (Kraay et al., 2023). The former does not account for inequalities across

sense of the poverty head-count ratio. This is much less the case for social well-being indicators based on complex mathematical formulas, like the welfare indices proposed by Atkinson (1970).

Second, we focus on construction methods that are relatively simple to implement and explain. A successful method typically yields a good balance between conceptual validity and simplicity of implementation.[22] Also, the trade-offs that the indicator makes across dimensions should not appear as "black box" to policy makers. Aggregating the outcomes in different dimensions through utility function representing individual preferences may be conceptually satisfactory. However, it seems challenging to transparently explain the implicit trade-offs under which a given achievement vector is deemed "equivalent" to some monetary value in a reference situation (Boarini et al., 2022). With such aggregation, the policy maker has no direct way of checking whether the implicit trade-offs make sense to her.[23] For this reason, we mostly focus on aggregation methods based on weights.

Third, we focus on opportunities with the potential to be applied in the near future. We exclude avenues whose implementation would require data coming from new recurrent surveys. The data feeding the well-being indicator must come from available surveys. Note that this does not necessarily exclude opportunities that requires a one-off dedicated survey, say to fix the values of weights. We also exclude new data sources like bio-indicators or cheaper data sources like remote sensing because we believe more work is necessary to firmly establish their relationship with determinants of well-being.

3.2 Identification not based on dichotomous variables

In this section, we consider two solutions to waste less data from variables that are already taken into account.

Both the global MPI and the World Bank's MPM are based on the Alkire-Foster (AF) methodology to identify the multidimensionally poor (Alkire and Foster, 2011a).[24] In a nutshell, the AF methodology starts from a list of dichotomous deprivation variables $s_{ij} \in \{0, 1\}$ where 1 means that i is deprived in variable j and 0 means i is not deprived. The AF methodology identifies individual i as (multidimensionally) poor when the weighted

individuals. The latter gives full priority to individuals at the bottom of the well-being distribution. From a normative perspective, ignoring entirely the progress taking place above the poverty line is not satisfactory. However, from a pragmatic perspective, a poverty line allows constructing an easy-to-communicate distribution-sensitive social well-being indicator: the poverty head-count ratio.

[22]For instance, in the monetary approach, the deflation methods based on demand-system à la Deaton and Muellbauer (1980) are rarely used in practice. Instead, simpler deflation methods based on price indices are routinely used in spite of being less conceptually satisfactory. Arguably, one key reason is the difficulty and the experience required to successfully implement the former methods.

[23]Relatedly, many parameters entering the definition of standard social well-being indicators have no intuitive meaning attached to their values, even for practitioners. This is arguably the case of coefficients of risk aversion or elasticities of inter-temporal substitution.

[24]See Appendix 6.1 for a succinct definition of both of these global multidimensional poverty measures.

sum of her deprivation statuses is too large, i.e.,

$$\sum_{j \in J} w_j s_{ij} \geq k, \tag{2}$$

where w_j is the weight given to deprivation status in variable j and k is the identification threshold. In practice, the weights are often selected using the "equal-weights" procedure. That is, all variables are partitioned into a few number of dimensions and equal weights are given across dimensions and within each dimension (see footnote 12). This methodology has many advantages, including its simplicity. However, because the AF methodology requires dichotomous variable, it yields a rather crude identification of the multidimensionally poor. Moving away from dichotomous variables may thus reduce misidentification errors.

This limitation of the AF methodology can be illustrated graphically using a simple classification of the poor. Decerf (2023b) observes that there are two types of multidimensionally poor individuals, at least under two mild assumptions on the ranking u of achievement vectors. The first assumption is that dimensions are minimally substitutable (Ravallion, 2011). That is, an individual is always willing to decrease her achievement in one dimension provided her achievements in the other dimensions are sufficiently increased.[25] The second assumption is that this substitutability is limited because each dimension is essential (Nussbaum, 2009; Alkire and Foster, 2011b). That is, there exists an extreme achievement threshold e_j below which an individual is poor regardless of her achievements in the other dimensions.[26] From there, some individuals are poor because at least one of their achievements is extremely low ($x_{ij} < e_j$ for some j), while others are poor because they cumulate moderately low achievements in several dimensions. We call the former "extremors" and the latter "cumulators". In Figure 3, i is an extremor and i' is a cumulator.

Because it is based on dichotomous variables, the AF methodology struggles to simultaneously identify both types of poor individuals. Consider the case with two dimensions only, for which the AF methodology boils down to either the union approach or the intersection approach.[27] Let z_j denote the dimension-specific cutoff used to define deprivation status in j. As explained in Decerf (2023b), the pracitioner can select small cutoffs z_j and use the union approach, in which case she correctly identifies extremors but makes exclusion errors on cumulators. Alternatively, she can select intermediate cutoffs z_j and use the intersection approach, in which case she correctly identifies cumulators but makes exclusion errors on extremors. This limitation comes from the use of dichotomous

[25]Graphically, the iso-wellbeing curve $\underline{u}$ defining the poverty threshold does not have horizontal or vertical segments.

[26]Graphically, the iso-wellbeing curve $\underline{u}$ admit asymptotes in e_j.

[27]See Figure 7 in section 3.6 for a graphical illustration of the union and intersection approaches.

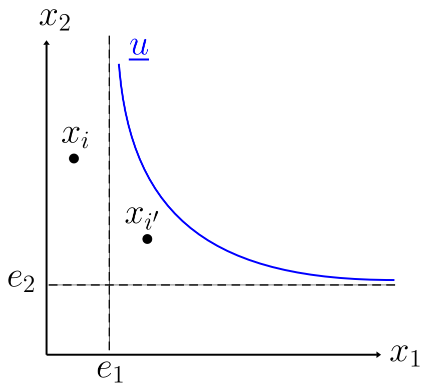

Figure 3: Two types of multidimensionally poor individuals: "extremor" i and "cumulator" i'.

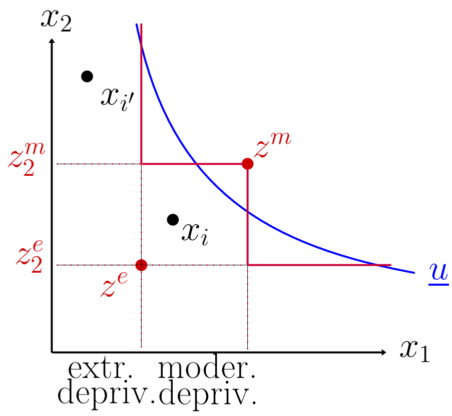

Figure 4: Trichotomous identification of the poor with "extreme" and "moderate" deprivations

variables, which cannot convey information on the depth of deprivation in dimension j. With a unique dimension-specific cutoff, a dichotomous variable either reveals extreme deprivation or the fact that the individual is moderately-or-extremely deprived.

Two recent contributions propose ways to identify the poor with more information. Importantly, they can be implemented with limited or no information on preference u. Indeed, the challenge is not merely to waste less data using conceptually sound solutions. It is also important to make sure that the solutions can be implemented under the data constraint faced by the practitioner.

First, a simple solution consists in using two dimension-specific cutoffs, one extreme cutoff z_j^e and one moderate cutoff z_j^m (Decerf, 2023b). This generates three mutually exclusive deprivation statuses: extreme deprivation ($x_{ij} < z_j^e$), moderate deprivation ($z_j^e \leq x_{ij} < z_j^m$), or non-deprivation ($z_j^m \leq x_{ij}$). The extremors can be identified using the union approach from extreme deprivations and the cumulators can be identified using some version of the intersection approach from moderate deprivations.[28] This is illustrated in Figure 4 for the case with two dimensions. The identification contour, which shows the frontier of achievement vectors identified as poor, is shown in red. Observe that this solution can be applied when variables are cardinal, but also when variables are categorical. An illustration of one way to implement this solution is proposed in Decerf and Fonton (2023).[29] One could think of other simple ways of implementing this

[28]Note that this solution is conceptually different from the identification of "destitute" individuals proposed by Alkire et al. (2014). Although both this solution and the destitute approach rely on two cutoffs in each variable, they do different things. This solution suggests using two cutoffs in order to change the set of individuals identified as multidimensionally poor. In contrast, destitutes are a subset of the individuals identified as multidimensionally poor (when using a single cutoff per dimension).

[29]For instance, for the cardinal monetary variable, the extreme cutoff could be taken to be the extreme International Poverty Line (Jolliffe et al., 2022) and the moderate cutoff could be taken to be a higher poverty line such as a weakly relative poverty line (Ravallion and Chen, 2011). For the categorical crime variable, which tracks whether households have been victim of a crime, the practitioner can partition crime types between those leading to extreme consequences, like murder, rape or abduction, and those

solution.[30] One could think of other solutions for reducing the waste of data due to the dichotomization of cardinal variables, but some of these alternative solutions may be harder to implement under the data constraints on preference u.[31]

Second, for a cardinal variable j' like income, another solution based on a single cutoff $z_{j'}$ has been proposed by Pattanaik and Xu (2018), henceforth PX. Instead of using the AF deprivation status s_{ij}, this solution suggests computing the deprivation variable $d_{ij'}$ defined as

$$d_{ij'} = \frac{z_{j'} - x_{ij'}}{z_{j'}} \qquad \text{when } x_{ij'} < z_{j'} \qquad \text{otherwise } d_{ij'} = 0. \tag{3}$$

For cardinal variables, Decerf (2023b) shows that this solution leads to fewer identification errors than using deprivation status. The intuition is easy to grasp in Figure 5, where both dimensions are captured through cardinal variables. In a nutshell, the identification contour associated to the PX methodology (in red) is a better approximation of indifference curve $\underline{u}$ than either the intersection or the union approach. We suggest here a simple way to implement this solution with limited information on u. Importantly, this proposal works also when some of the variables are cardinal while some are dichotomous. Indeed, it suffices to identify an individual i as poor when

$$\sum_{j \in J^{dichot.}} w_j s_{ij} + \sum_{j' \in J^{card.}} w_{j'} d_{ij'} \geq k, \tag{4}$$

where $w_{j'}$ is the weight given to the deprivation in the variable j'. Here is a natural way to select cutoff $z_{j'}$ and weight $w_{j'}$, which is illustrated in Figure 5 where income is dimension 1.[32] Consider a low poverty line $z^e_{j'}$ below which individuals can safely be identified as multidimensionally poor, e.g., the extreme International Poverty Line. Cutoff $z_{j'}$ can be set at a higher poverty line below which individuals are moderately monetary deprived, but are not necessarily multidimensionally poor, e.g., the International Poverty Line for LMICs or UMICs, or the national poverty line. A natural value for the weight is $w'_j = k * \frac{z_{j'}}{z_{j'} - z^e_{j'}}$. Under these choices for $z_{j'}$ and $w_{j'}$, an individual i' is identified to be

leading to "moderate" consequences, like theft.

[30] Another possibility is to use the "equal-weight" procedure on moderate deprivations and attribute a weight $w^e_j = k$ on extreme deprivations. This is in fact the solution implicitly used by the World Bank's MPM.

[31] When possible, using three dimension-specific cutoffs rather than two would further reduce the waste of data. However, the selection of the weights given to the four types of deprivation may become harder for the practitioner.

[32] To be sure, if dimension 1 is captured by a cardinal variable and dimension 2 by a categorical variable, then the identification contour associated to Eq. (4) will not have the same shape as that depicted in Figure 5. If the variable capturing dimension 2 is trichotomous, then the identification contour has the shape depicted in Figure 4. In that case, the PX identification contour associated with Eq. (4) can better approximate indifference curve $\underline{u}$ than the identification contour associated with Eq. (2). Eq. (4) can also yields an improvement if there are at least two dichotomous variables on top of the cardinal variable.

multidimensionally poor when her income is below the low poverty line $z^e_{j'}$, regardless of her achievements in the other dimensions. In contrast, when the income of an individual i is above the low poverty line $z^e_{j'}$ but below the cutoff poverty line $z_{j'}$, i is identified as multidimensionally poor only if she also suffers from (enough) deprivation in non-monetary dimensions, as is the case in Figure 5.

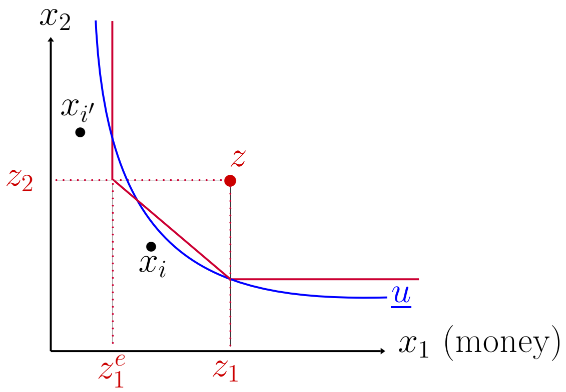

Figure 5: Deprivation-based identification makes fewer mistakes.

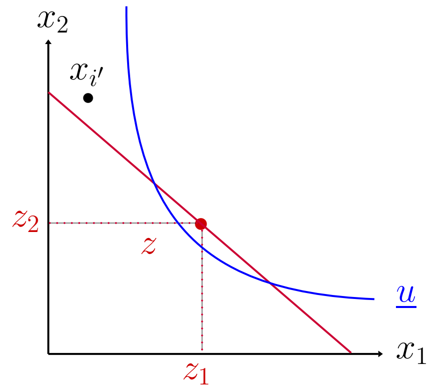

Figure 6: Achievement-based identification makes exclusion errors on extremors

Observe that Eq. (4) uses information on the depth of deprivation (in the cardinal variable) in order to improve the *identification* of the multidimensionally poor. This approach, which deviates from the "counting" approach, could also be used when measuring the *intensity* of multidimensional poverty, although here we only considered its *incidence*.

Finally, we warn against one key limitation of another solution to increase the information content considered in the identification methodology. A seemingly natural solution is to take a weighted sum of achievements, namely identifying the multidimensionally poor with $\sum_{j \in J^{card}} w_j x_{ij} \geq k$. This solution was proposed in Ravallion (2011). This solution would also seem to be a natural way of identifying the multidimensionally poor from multidimensional well-being indices proposed by Benjamin et al. (2014b) or from the total consumption approach (see Section 3.4.3). The limitation of this approach is that it assumes an infinite substitutability across dimensions and thus does not acknowledge the essentiality of well-being dimensions. As a result, it systematically makes exclusion errors on extremors. The intuition is easy to grasp in Figure 6, where extremor i' is (mistakenly) not identified as poor because her achievement vector is above the linear identification contour associated to such an achievement-based identification methodology (in red).

It would be interesting to study empirically the differences that these two solutions make when compared with a more traditional AF identification. For example, on could look at how well-being comparisons by the World Bank's MPM would be affected by these solutions. In the remainder of this section, we shortly discuss how to quantify the impact of these solutions.

Let $\hat{u}$ and $\hat{P}$ denote the preferred indicators, respectively for individual well-being and social well-being. Let $\hat{\hat{u}}$ and $\hat{\hat{P}}$ denote the benchmark indicators, respectively for individual well-being and social well-being. In this case, the benchmark indicators are those obtained from the AF methodology while the preferred indicators are those obtained from one of the two solutions presented above, or a mixture of the two. The individual well-being indicator is whether i is identified as multidimensionally poor and the social well-being indicator is the multidimensional poverty measure, such as the head-count ratio.

The evaluation should quantify the deviations between inter-personal comparisons (Objective 1) or between cross-regions comparisons (Objective 2), as defined in Section 2.6. In the case of interpersonal comparisons, we thus check how reliably we have $\underline{u} > \hat{\hat{u}}(d_i)$ when $\underline{u} > \hat{u}(d_i)$. In the case of cross-regions comparisons, we thus check how reliably we have $\hat{\hat{P}}(d_R) > \hat{\hat{P}}(d_{R'})$ when $\hat{P}(d_R) > \hat{P}(d_{R'})$.

There are different ways in which these quantifications can be done. One way is to compute the frequency of "reversals", i.e., the fraction of all pairwise comparisons for which the two indicators yield opposite comparisons. One drawback of this approach is that the frequency of reversals depends on how similar are the units being compared. Indeed, you mechanically find more frequent reversals when contrasting any two countries in Sub-Saharan Africa than when contrasting any two countries in the world. For this reason, it is desirable to check robustness by also using another method that avoids this drawback when quantifying the difference between two indicators. We present here one such other method, which interprets the two indicators as-if they have cardinal meaning.[33] Intuitively, the idea relies on the distribution of a counterfactual budget across the regions being compared, where the distribution is made proportionally to the poverty in each region. One can then compare how differently the two indicators would distribute the counterfactual budget across regions, say by computing the fraction of the total budget that would need to be re-allocated across regions when moving from one indicator to the other. These two types of quantifications are for instance used in Decerf and Fonton (2023). There exist other quantification methods, e.g., comparing the profile of individuals in the bottom of the well-being distribution, as in Decancq and Neumann (2014).

3.3 Hybrid indices that integrate longevity

In this section and the next, we consider solutions to add more variables, which are not yet taken into account.

The case for integrating mortality into well-being measurement seems strong. From a conceptual perspective, mortality determines longevity, which arguably has high intrinsic

[33]Such cardinal interpretation is meaningful for the poverty head-count ratio.

value to well-being.[34] "Being alive" is the most fundamental functioning. Also, although mortality is correlated with income, empirical analyses show that integrating mortality substantially affect cross-country well-being comparisons (Becker et al., 2005; Jones and Klenow, 2016). Another pragmatic reason is that mortality is highly correlated with key dimensions of $J^{non-market}$ that are ignored or poorly covered by current indicators, like health and security.

A synthetic aggregation of mortality into the social well-being indicator can make sense. First, from a pragmatic perspective, the joint distribution between mortality and income is almost always missing (Section 2.4). One solution to this data constraint is to use a synthetic aggregation, which assumes that all individuals within a given society face the same mortality risks. Second, from a conceptual perspective, longevity is a peculiar dimension that may require a specific aggregation into poverty indicators. Mortality determines the *quantity* of life. In contrast, other types of deprivations reduce the *quality* of life. A person can cumulate deprivations in several dimensions, but she cannot be dead and simultaneously be deprived in other dimensions.

There exist simple hybrid indices that integrate mortality with (multidimensional) poverty measures.[35] Some of these indices attribute (negative) intrinsic value to all deaths (Riumallo-Herl et al., 2018; Baland et al., 2023), while other indices only attribute (negative) intrinsic value to deaths taking place below some age threshold ($\hat{a}$) defining premature mortality (Baland et al., 2021, 2023). These indices are all expressed in years of human life, e.g., years spent out of poverty, years spent in poverty or years of life prematurely lost. All these indices are based on the same normative parameter, which tunes the weight given to one year of life spent in poverty (vs one year of life lost). We illustrate this weight by presenting the Poverty-Adjusted Life-Expectancy (PALE) indicator (Baland et al., 2023; Bank, 2022). This indicator sums years of life spent out of poverty with years of life spent in poverty and weighs down the latter. Formally, this synthetic indicator takes the perspective of a newborn who assumes she will be confronted throughout her life with the poverty and mortality observed in the year in which social well-being is assessed, namely

$$PALE = LE * (1 - \theta * H),$$

where LE is life-expectancy at birth, H is the (multidimensional) head-count ratio and the normative parameter $\theta \in [0, 1]$. When one year spent in poverty is considered as bad as one year of life lost ($\theta = 1$), $PALE$ boils down to the Poverty-Free Life-Expectancy indicator proposed by Riumallo-Herl et al. (2018). When one year spent in poverty is

[34] Longevity is associated to higher subjective well-being (Diener et al., 2018).

[35] The literature on the mortality paradox (Kanbur and Mukherjee, 2007; Lefebvre et al., 2018) is not relevant here because it is only interested in the *instrumental* impact that mortality has on the poverty measure (Decerf, 2023a). It disregards the *intrinsic* impact that mortality has on well-being.

considered negligible compared to one year of life lost ($\theta = 0$), $PALE$ boils down to LE.

Observe that $PALE$ is a hybrid index. The multidimensional head-count ratio H is an "objective" index, which relies on the joint distribution between dimensions capturing quality of life. Yet, longevity is aggregated into $PALE$ in a synthetic way, assuming that all individuals face the same mortality risks summarized in LE. Arguably, $PALE$ has a larger signal-to-noise ratio than H when performing cross-country comparisons (Objective 2). At least this should be the case if LE captures mortality risks with limited noise and if longevity captures an important dimension of well-being.

Baland et al. (2021) show empirically that monetary poverty measures that ignore mortality lead to substantially biased well-being comparisons. They find that mortality is not dwarfed by poverty even when assuming a conservatively low age threshold ($\hat{a} =$ 50 years) and the largest plausible weight to monetary poverty ($\theta = 1$).[36] Furthermore, the correlation between monetary poverty and mortality is far from perfect. Depending on the value attributed to the weight θ, they find that between 8 and 27% of country trends are reversed over five-years periods when integrating mortality.

This points to one key open question: what is the range of plausible values for parameter θ? Narrowing down this range within the $[0, 1]$ interval would allow sharpening the conclusions obtained when integrating mortality.[37] There are at least two routes one can take to investigate plausible values for parameter θ.[38] One could try to survey people's attitudes on these two parameters.[39] Another route would be to calibrate values for θ based on a parametric expression for a Bernoulli utility function, as illustrated in the Appendices of Baland et al. (2023) and applied for a large set of countries in Decerf et al. (2024). This approach could be further refined by calibrating the parametric expression for the Bernoulli utility from the Value of Statistical Life literature (Kniesner and Viscusi, 2019), as for instance done in Becker et al. (2005); Jones and Klenow (2016); Boarini et al. (2016).

However, besides the selection of its parameter's value, the $PALE$ index is ready-to-use. The same holds for other indices based on a premature mortality threshold.[40] The mortality data required are often available at country level and sometimes at subnational level. Accounting for longevity could very well be the lowest hanging fruit for practitioners

[36] Decerf and Fonton (2023) also find that mortality is sizable compared to multidimensional poverty when contrasting well-being across Nigerian states in 2019.

[37] Baland et al. (2023) study the theoretical conditions and document the empirical frequency with which poverty comparisons are reversed for all values of $\theta \in [0, 1]$.

[38] The value for θ depends on the poverty standard considered. A more austere poverty line implies a larger weight to poverty and thus a larger value for θ.

[39] For instance, one could ask to non-poor individuals how many years of their remaining life they would be willing to spend in poverty to extend their longevity by one year.

[40] It could still be interesting to investigate how to extend these indices for contexts where some information is available on the joint distribution between mortality and income, e.g., information on selective mortality, i.e., the differential mortality risks faced by the poor and the non-poor. Another question is how to extend these indices for gap-sensitive poverty indices.

interested in improving the signal-to-noise ratio of their multidimensional poverty measure when aiming at Objective 2.

3.4 Combining data from separate surveys

One of the main data constraint at global level is the missing distribution between incomes and key non-monetary outcomes such as health (Section 2.4). A priori, this constraint is very consequential because both health and income are arguably among the most important dimensions of well-being.[41]

The classical approaches to deal with this data constraint "waste" substantial amount of available data. One popular solution consists in constructing an objective well-being indicator based on a single survey. If income data comes from one survey and health data comes from another survey, this solution wastes all the data in the ignored survey. This is the solution used by global multidimensional poverty measures.[42] An alternative solution is to construct a synthetic indicator that ignores inequalities in health and inequalities in incomes. This is the solution used by the HDI, at least when interpreting mortality data as a signal on a population's health. This solution wastes the data on income inequalities that is available in the income survey.

In this section, we outline three alternative methods that waste less data by combining information from separate surveys using a hybrid index. The first two methods rely on making assumptions on the missing joint distribution. These methods seem relevant when cross-regional well-being comparisons are expected to depend mainly on the partial distributions, with the joint distribution playing a limited role. The third method assumes equal non-monetary outcomes on small areas to compute household's "total" consumption. This method seems relevant when non-monetary inequalities across areas are expected to be larger than non-monetary inequalities within areas.

3.4.1 Assuming the overlap between two types of poverty

We start with the simplest method to estimate the fraction of multidimensionally poor individuals from two separate surveys. We first present the method and then show that it works surprisingly well when comparing Nigerian states in 2019.

For exposition purposes, assume (counterfactually) that the joint distribution is available. That is, there exists a comprehensive survey that collects the relevant outcomes in all dimensions, monetary as well as non-monetary. Assume that an individual is identified as multidimensionally poor in two cases: either she is below the monetary line or she is above the monetary line but cumulates enough deprivations in the other (non-monetary)

[41] Both income and health are among the dimensions most correlated with subjective well-being (Diener et al., 2018).

[42] The global MPI ignores incomes and the World Bank's MPM ignores health (see Appendix 6.1).

dimensions. Thus, this approach assumes two poverty standards $\underline{\hat{u}}_M$ and $\underline{\hat{u}}_{OD}$, respectively for the monetary dimension and the other (non-monetary) dimensions. These two standards $\underline{\hat{u}}_M$ and $\underline{\hat{u}}_{OD}$ are considered "equivalent", as in the World Bank's MPM.[43] The comprehensive survey allows computing the "true" fraction of multidimensionally poor individuals, denoted by $\hat{H}$.

We propose a method to construct an estimate $\hat{\hat{H}}$ of $\hat{H}$ when the joint distribution between monetary and non-monetary dimensions is missing. That is, there is no comprehensive survey but rather two separate surveys, say one LSMS household consumption survey and one DHS or MICS household survey. The former survey allows computing the fraction of monetary poor individuals, which we denote by $\hat{H}_M$. The latter survey allows computing the fraction of other-dimensions poor individuals, which we denote by $\hat{H}_{OD}$. Some individuals are simultaneously monetary poor and other-dimensions poor. The only information missing to compute the "true" fraction $\hat{H}$ can be called the coefficient of overlap (c) between the two types of poverty. This coefficient captures "by how much" monetary poor individuals are more likely than those who are not monetary poor to be other-dimensions poor. Formally, c is the computed as the fraction of monetary poor individuals who are other-dimensions poor divided by the fraction of individuals who are not monetary poor but are other-dimensions poor

$$c = \frac{\frac{\hat{H}_M + \hat{H}_{OD} - \hat{H}}{\hat{H}_M}}{\frac{\hat{H} - \hat{H}_M}{1 - \hat{H}_M}}.$$

In this setting, coefficient c summarizes all the relevant information on the missing joint distribution. If we knew the value for c, we could compute the "true" value for $\hat{H}$ through

$$\hat{H} = \frac{(c-1) * \hat{H}_M^2 + \hat{H}_M + \hat{H}_{OD} - \hat{H}_M * \hat{H}_{OD}}{1 + (c-1) * \hat{H}_M}.$$

The proposed method consists in assuming a value $\hat{c}$ for the unknown coefficient c. This allows estimating the "true" value for $\hat{H}$ through an indicator $\hat{\hat{H}}(\hat{c})$, computed by replacing c with $\hat{c}$ in last equation. By definition, we know that $c \geq 0$. We have $c = 0$ when no monetary poor individual is simultaneously other-dimensions poor. We have $c = 1$ when monetary poverty status is uncorrelated with other-dimensions poverty status. Finally, we have $c \to \infty$ when all other-dimensions poor individuals are simultaneously monetary poor. If other-dimensions poverty status is positively correlated with monetary poverty status, we have $c > 1$. For instance, the coefficients of overlap between the extreme IPL and the global MPI can be computed for the six developing countries studied

[43] Indeed, in the WB's MPM, individuals below the extreme IPL receive a weight of 1/3, which is the same identification threshold ($k = 1/3$) used to identify multidimensionally poor individuals who are not below the extreme IPL.

in Table 6 in Evans et al. (2023). All six coefficients of overlap lie in the range $(1.5, 3.6)$.[44]

We illustrate the proposed method's potential to improve on cross-regions comparisons (Objective 2). We consider the multidimensional poverty indicator constructed by Decerf and Fonton (2023) to compare the 36 Nigerian states in 2019. This indicator considers four dimensions, namely consumption (monetary), health, housing and security.[45] At national level, Decerf and Fonton (2023) estimate $\hat{H}$ at 52 percent, $\hat{H}_M$ at 40 percent and $\hat{H}_{OD}$ at 28 percent. There are large differences in multidimensional poverty rates across states, with $\hat{H}$ being around 20 percent for some states and around 90 percent for other states. The coefficients of overlap of all 36 states lie in the range $(0.8, 2.7)$, with median value 1.7.

Table 2 compares the cross-states comparisons of various indicators with those obtained with the "true" indicator $\hat{H}$. Most indicators considered correspond to $\hat{\hat{H}}(\hat{c})$ when assuming the same (counterfactual) value $\hat{c}$ for all Nigerian states. The deviations from $\hat{H}$ are quantified through a "bias" index, which sums over all 36 states the absolute difference in the state's rank for the indicator considered and for $\hat{H}$.[46] The larger the bias index, the less reliably the indicator considered yields cross-states comparisons in line with those of $\hat{H}$.

Table 2: Assuming the joint distribution yields better comparisons than wasting the data on one partial distribution.

	$\hat{H}_M$	$\hat{H}_{OD}$	$\hat{H}_M + \hat{H}_{OD}$ $\left(\sim \hat{\hat{H}}(0)\right)$	$\hat{\hat{H}}(0.8)$	$\hat{\hat{H}}(1.7)$	$\hat{\hat{H}}(2.7)$	$max(\hat{H}_M, \hat{H}_{OD})$ $\left(\sim \hat{\hat{H}}(\infty)\right)$
Bias	72	194	32	18	16	22	60

The key insight from Table 2 is that, in this context, the proposed method yields a smaller bias than the classical "wasteful" approaches *for all possible values of* $\hat{c}$. Recall the classic approach consists in using an indicator coming from a single survey, which at global level essentially means using either $\hat{H}_M$ or $\hat{H}_{OD}$ when the joint distribution is missing. Even when assuming the most extreme values for $\hat{c}$, either 0 or ∞, indicator $\hat{\hat{H}}$ yields a smaller bias than both $\hat{H}_M$ and $\hat{H}_{OD}$.[47] The method performs much better when assuming a value for $\hat{c}$ equal to the median value for c among Nigerian states.

[44] Also see Tran et al. (2015) on the overlap between monetary poverty and the global MPI.

[45] The baseline version of this indicator satisfies the necessary assumptions for this exercise. Individuals are considered monetary poor when their consumption is below the extreme IPL, in which case they receive a weight of $w_M = 1/4$. The identification threshold used is also $k = 1/4$. Note that the non-monetary dimensions of this indicator do not correspond to those of the global MPI.

[46] For instance, this difference is equal to 2 if a state is ranked 1 by the indicator considered and ranked 3 by $\hat{H}$.

[47] Formally, indicator $max(\hat{H}_M, \hat{H}_{OD})$ yields the same ranking as $\hat{\hat{H}}(\infty)$ when $\hat{c}$ tends to ∞. Also, indicator $\hat{H}_M + \hat{H}_{OD}$ yields the same ranking as $\hat{\hat{H}}(0)$ when $\hat{H}_M + \hat{H}_{OD} \leq 100$.

In this context, the missing joint distribution, i.e., the value for c, does not contain much of the information that is relevant to perform cross-regions well-being comparisons. Most of the relevant information is available in the two separate surveys. By ignoring one of the separate survey, the classical approach wastes relevant information and yields suboptimal well-being comparisons. Shockingly, "making up" (overlap) data improves well-being comparisons. The reason is that the "made up" data allows using more "true" data, i.e., from both partial distributions.

More generally, these results illustrate a sometimes overlooked point: when dealing with data constraint, the objective is *not* to approximate the ideal indicator we could construct in the absence of constraint. Rather, the objective is to construct an indicator that makes better well-being comparisons than the comparisons made by the alternative indicators that are currently used. In other words, what matters the most is not whether the value of $\hat{\hat{H}}$ is "close enough" to the value of $\hat{H}$. The relevant question to ask is whether the well-being comparisons by $\hat{\hat{H}}$ are better than the well-being comparisons by $\hat{H}_M$ (and those by $\hat{H}_{OD}$).

The results presented in Table 2 are valid only for one context using one particular definition for $\hat{H}$ and one way of quantifying the deviations from $\hat{H}$.[48] They call for more research to check that the proposed method can reliably improve on the classical "wasteful" approach. In particular, it would be interesting to investigate the performance of the proposed method for a multidimensional poverty measure constructed from the data available in a typical LSMS survey and a typical DHS or MICS survey. This performance can be assessed in contexts where the full multidimensional poverty measure can be constructed, e.g., those considered in Table 6 in Evans et al. (2023). Similarly, it could be interesting to study simple ways of extending the proposed method to multidimensional poverty measures for which the monetary dimension is not captured by a dichotomous status (i.e., monetary poor or not monetary poor) but by a trichotomous status (i.e., extremely, moderately or not monetary deprived) or by a continuous variable.

3.4.2 Imputing incomes into the non-monetary survey

The method mentioned here is a more elaborate variant of the method proposed in the Section 3.4.1. It also makes assumptions on the missing joint distribution between monetary and non-monetary outcomes. The idea is to impute consumption or income into the non-monetary survey, say DHS or MICS.[49] One way of doing it is to use survey-to-survey imputation techniques, as pioneered by Elbers et al. (2003).[50] These techniques build an imputation model based on the common variables in the two surveys, e.g., demographic variables and maybe some common outcomes. This model is trained in a similar context

[48] See Section 3.2 for one alternative quantification of deviations.

[49] Another possibility is to impute non-monetary outcomes in an LSMS survey.

[50] See Dang and Lanjouw (2023) for a recent review.

where the joint distribution is known. The underlying (strong) assumption is that the joint distribution is the same as in the context in which the model is trained. Another clear drawback is that imputations techniques are relatively complex to implement.

Such imputation methods has significant advantages over the method proposed in Section 3.4.1. First, it artificially generates a database where all outcomes are available for each household. As a result, the multidimensionally poor households can be identified with methods that are more elaborated than the AF identification implicit in Section 3.4.1. Indeed, the monetary dimension need not necessarily be dichotomized when identifying the poor.[51] Hence, imputations allow constructing more sophisticated multidimensional poverty measures. Second, and related, the "comprehensive" database generated by this imputation could potentially also be used for inter-personal well-being comparisons (Objective 1). This is unlike the method proposed in Section 3.2, which only allowed cross-regions well-being comparisons (Objective 2).

Of course, the added-value of these advantages depend on the relevance of the joint distribution assumed. And given the scarcity of contexts where such an imputation model can be trained, it is far from clear that the joint distribution assumed will often decently approximate the unknown joint distribution. Yet, one fundamental point to keep in mind is that this approximation needs not be decent. The key comparison is not how close is the assumed joint distribution from the unobserved joint distribution. Rather, what truly matters is whether the well-being indicators constructed on the "comprehensive" imputed database yield better well-being comparisons than those constructed using the classical approach, which "wastes" available data. This is the message illustrated in Table 2, where even very bad assumptions $\hat{c}$ on the unknown joint distribution c yield better well-being comparisons than those of $\hat{H}_M$ (or those of $\hat{H}_{OD}$).

The most discussed question in this area asks which particular imputation method yields most reliable results given the particular data constraints (Dang, 2021; Dang et al., 2023). Another question, which perhaps deserves more prominence given our purpose, is whether the use of an imputation method with limited reliability is better than no imputation at all. In the case of multidimensional well-being, an imputation method that is only semi-reliable could improve well-being comparisons by allowing to draw information from both a monetary and a non-monetary survey.

It would be interesting to study the reliability of well-being comparisons yielded by indicators constructed on a "comprehensive" database obtained by imputing the monetary dimension into a non-monetary survey (DHS or MICS). Studying this reliability requires at least two contexts for which the joint distribution between monetary and non-monetary outcomes is observed. The more dissimilar these two contexts, the more conservative the assessment. The imputation model is trained on one context and then tested on the second context. Let $\hat{u}$ and $\hat{P}$ denote the well-being indicators we would like to

[51] For instance, one could use the PX identification method described in Section 3.2.

construct if the joint distribution was known (it is in fact known in both contexts). Let $\hat{\hat{u}}$ and $\hat{\hat{P}}$ denote these well-being indicators constructed on the database generated by imputing income in the second context. Let $\hat{\hat{u}}^C$ and $\hat{\hat{P}}^C$ denote the well-being indicators constructed by the classical approach in the second context. Indicators $\hat{\hat{u}}^C$ and $\hat{\hat{P}}^C$ are either monetary indicators or non-monetary indicators like the global MPI. The reliability of the imputation method for *inter-personal* comparisons is assessed by quantifying the deviations between $\hat{\hat{u}}$ and $\hat{u}$ and comparing them to the deviations between $\hat{\hat{u}}^C$ and $\hat{u}$.[52]

3.4.3 Total consumption

A different way of circumventing the missing joint distribution consists in using a hybrid individual well-being index. The idea relies on the total consumption approach. An individual's total income is the sum of her monetary aggregate (consumption or income) plus a monetary valuation of her non-monetary outcomes. Conceptually, this monetary valuation could capture her willingess-to-pay to achieve these non-monetary outcomes (Hicks, 1942; Chipman and Moore, 1980). Oftentimes, this valuation is equated to the cost of providing the relevant public services (health, education, etc.) or the price of these services when privately offered (Barofsky and Younger, 2019). When one cannot observe the unequal access that different households have to these services, one natural solution is to assume an equal access at the local level. The individual well-being index is thus obtained by *hybrid* aggregation because a statistic on the distribution of non-monetary outcomes (say the average) is attributed to all individuals in the same district.

Formally, the total consumption t_i of household i can be defined as

$$t_i = \underbrace{\sum_{j \in markets} \frac{p_j x_{ij}}{Price\ index}}_{Monetary\ dimension} + \underbrace{\sum_{j \in health, educ, secur} v_j x_j^{district_i}}_{Non-Monetary\ dimensions}$$

where the first term captures monetary consumption and the second term captures the monetary valuation of non-monetary outcomes (or services). In particular, $x_j^{district_i}$ denotes the level of non-monetary outcome j in the district of i and v_j denotes the monetary value attributed to one unit of non-monetary outcome j.

The total consumption approach is relatively complex to implement, especially when $x_j^{district_i}$ capture services. One reason is the difficulty to estimate meaningful valuations v_j for non-monetary services, because the quality of these services matters for the non-monetary outcomes achieved and this quality is hardly observed (Simpson, 2009). At global level, this requires using national-level approximations of the service quality, like

[52]The reliability of the imputation method for *cross-regions* comparisons is assessed by quantifying the deviations between $\hat{\hat{P}}$ and $\hat{P}$ and comparing these to the deviations between $\hat{\hat{P}}^C$ and $\hat{P}$. This can be done for cross-regions comparisons if the second context has a large number of regions between which $\hat{P}$ can be compared.

is for instance done for schooling quality by the Human Capital Index (Kraay, 2019).

However, the total consumption approach shares one important advantage with the imputation method described in Section 3.4.2. Unlike the overlap method proposed in Section 3.4.1, the total consumption approach allows for both inter-personal well-being comparisons (Objective 1) and cross-regions well-being comparisons (Objective 2).

Yet, when total consumption t_i is taken to be the individual well-being index $\hat{u}$, the total consumption approach has one conceptual drawback. This approach does not account for the essentiality of the core dimensions of well-being. Indeed, the total consumption t_i assumes perfect substitution between monetary and non-monetary outcomes. This is problematic because, if i is not interested in the non-monetary services offered in her district, she cannot exchange them to buy more consumption goods. Importantly, alternative aggregations like those implicit in the AF or the PX identifications do *not* assume perfect substitution. The intuition can be easily grasped by contrasting Figures 5 and 6 in Section 3.2. The total consumption approach yields a linear achievement-based identification, which systematically makes exclusion errors on some extremors. In contrast, the PX linear deprivation-based identification avoids making these exclusion errors.

One potential solution to this drawback could be to apply a PX aggregation of the monetary valuations in each dimension, i.e.,

$$t_i^{PX} = \omega_{Mon} d_{iMon} + \omega_{Health} d_{iHealth} + \omega_{Edu} d_{iEdu} + \dots$$

where ω_j is the PX weight for dimension j, where $\sum \omega_j = 1$ and d_{ij} is the deprivation variable obtained from a continuous variable and a cutoff through Eq. (3). For the monetary dimension, the continuous variable is the monetary consumption and the cutoff is a poverty line. For a non-monetary dimension, say health, the continuous variable is the achievement $v_{health} x_{health}^{district_i}$ and the cutoff is a threshold value for this achievement.[53] When identifying the multidimensionally poor, comparing t_i^{PX} to some identification threshold $\underline{t}^{PX} > 0$ has conceptual advantages over comparing t_i to some identification threshold $\underline{t} > 0$ (Decerf, 2023b). Again, the empirical significance of such total consumption approach must be assessed by comparing its well-being comparisons to those of purely monetary indicators.[54]

There are many open questions related to the way in which total consumption indicators would be constructed in practice. To name just two: which data is used to capture the level of the non-monetary service $x_j^{district}$? which procedure is used to select its valuation v_j?[55] This is an active area of research (Gethin, 2023).

[53]The PX weight ω_j can be selected from one moderate and one extreme achievement thresholds in dimension j, as suggested for the monetary dimension in Section 3.2.

[54]The methodology for comparing the performance with purely monetary indicators would be similar to the methodology outlined at the end of Section 3.4.2.

[55]Different procedures for selecting valuation v_j yield in practice different results (Barofsky and Younger, 2019).

3.5 Taking advantage of individual-level variables

One key data constraint to global well-being measurement is that major surveys collect many outcomes at the household level rather than at the individual level. This implies that within-household inequalities are not observed even though well-being is typically not equally distributed within households. Importantly, this is a problem that affects not only inter-personal comparisons (Objective 1), but also cross-region comparisons (Objective 2) because the latter rely on the former.

Clearly, this is not a big issue for some dimensions that are almost equally distributed at household level, such as housing. However, key dimensions of well-being are typically unequally distributed within households. This is the case of the monetary dimension, as documented by a growing literature on intra-household decision making (Lise and Seitz, 2011).[56] Arguably, this is even more so the case for the health dimension, if only because health tends to decrease with age.

Yet, many household surveys contain some information on individual-level outcomes. This is typically the case for children's schooling information or for the BMI information underlying stunting and underweight variables. The current practice in global multi-dimensional poverty measurement treats these individual level outcomes as-if they are household level outcomes.[57] For instance, under the World Bank's MPM, a household is considered deprived in a schooling variable if one of its children is not enrolled in school (see Appendix 6.1). A natural alternative would be to switch the unit of analysis from the household level to the individual level. "Equal sharing" assumptions could generate individual level variables from the variables that are only available at household level, such as income of housing.[58] The advantage of such switch would be to allow accounting for the unequal individual outcomes observed in the data.

One disadvantage with such switch in unit of analysis is the risk of misinterpretation. Individual-level well-being indices obtained in this way may lead to improper inter-personal comparisons. To give a concrete example, DHS surveys provide information on underweight women but often do not provide information on underweight men. Taking advantage of this information on underweight women makes sense, but the resulting individual well-being indices cannot be compared across men and women.[59] However, they can certainly improve the inter-personal well-being comparisons across women. Impor-

[56] This literature relies on methods that are complex, even though some simplified methods have recently been proposed by Lechene et al. (2022) with limited additional data requirements. These methods are certainly worth further exploring.

[57] This need not be the case of children multidimensional poverty measures, where the unit of analysis is sometimes the individual child (Dirksen and Alkire, 2021; Alkire and Haq, 2023).

[58] Observe that such assumptions are already implicitly made when the unit of analysis is at household level. Indeed, a person living in a household identified as multidimensionally poor is herself considered to be multidimensionally poor.

[59] Observe that, similarly, interpersonal comparisons across men and women cannot be made using the World Bank's MPM or the global MPI, at least not without the qualifier "living in multidimensionally poor household".

tantly, they could also improve inter-personal comparisons across men. Such advantage is easier to grasp when thinking of another type of individual level data that depends less on household level circumstances. This is arguably the case of (self reported) disabilities. The extreme disability of one senior individual may drastically affect her well-being, but it needs not necessarily have the same impact on other household members. If such disability is accounted at the household level as a moderate deprivation, this leads to many identification mistakes. For instance, the ("extremely") disabled person may not be identified as poor if her household does not face other types of deprivations. Alternatively, her relatives may wrongly be identified as poor if the household's moderate disability deprivation pushes the household beyond the identification threshold. Interestingly, accounting at individual level for deprivations that affect one category of individual (e.g., seniors) may improve well-being comparisons across another category of individuals (say children).

3.6 Beyond ad-hoc weights

One key data constraint is that non-monetary data typically do not contain information on the relevant preference u. As a result, selecting appropriate values for the weights of a multidimensional indicator is challenging for practitioners. This explains the popularity of the "equal-weights" procedure, which allows accounting for non-monetary data without requiring any information on u.[60] The main risk is that equal weights deviate too much from the weights that would best reflect preference u (Ravallion, 2012).[61] Another risk is that the "equal-weights" procedure decreases the perceived legitimacy of multidimensional poverty measures because of this obviously ad-hoc procedure.

Improving on this issue requires some information on the relevant preference u. This information can potentially be taken from different sources (Decancq and Lugo, 2013). Subjective well-being data may be one source of information on u, as suggested by Decancq et al. (2015) or by in Decancq et al. (2019). Ideally, the survey that collects all relevant outcomes also collects information on life satisfaction or happiness. Preference elicitation surveys may provide another sources of information on u, as done in Decancq and Nys (2021). These surveys ask subjects to rank pairs of achievement vectors in one

[60] In spite of its paternalistic weights, it could be that the World Bank's MPM has a larger signal-to-noise ratio than the fraction of extremely monetary poor individuals. That is, the additional information coming from its non-monetary variables may more than compensate for the bias coming from using equal-weights $\hat{w}^{eq}$, which are bound to be different than weights w informed from individual preferences.

[61] There are other limitations associated with the "equal-weights" procedure. For instance, this procedure provides incentives to ignore relevant information. The reason is that the only way through which the practitioner can influence the weight attributed to a given variable is by changing the number of dimensions and/or changing the number of dichotomous variables considered within the same dimension. Under the equal-weights procedure, there is an incentive not to include one variable if another variable belonging to the same dimension is already included and has a clearly different importance for well-being. Indeed, the aggregation with equal weights would not look plausible if both variables were simultaneously included.

form or another, e.g., using vignettes or eliciting willingness-to-pay. There exist other methods less directly connected to our preference-based framework. For instance, asking directly subjects about their preferred weights (Al-Ajlani et al., 2020) or asking to poor individuals which dimensions matter most to poverty (Kanbur and Lustig, 2001; Bennett, 2021). These latter methods may guide the selection of dimensions[62] and maybe provide a source of legitimacy for weights, although the relationship between such weights and preference u is unclear. In practice, the information on preference u could be collected in a few contexts and then used to select a fixed set of weights to be applied to all contexts.[63]

Clearly, these different sources of information on u have important limitations. Conceptually, none of them provides unbiased information on well-being rankings. In practice, each also has its issues.[64] In spite of their important limitations, these sources of data nevertheless offer a promising way of improving on the ad-hoc "equal-weights". Again, we are looking for better weights, not perfect weights. Observe also that selecting the value of weights from information on preferences does not necessarily imply changing the weights across time and space. Indeed, changing the weights may compromise comparability. The advantage of selecting the values of weights from preferences is that these values are less likely to be implausible than "equal-weights".

In this section, we present one method to derive a set of weights from (partial information on) preference u. Before that, we argue that seemingly natural methods do not appropriately account for the meaning of the weights used in multidimensional poverty indices. As mentioned in Alkire and Foster (2011b), AF weights do not correspond to marginal rates of substitution. The intuition is illustrated for the two dimensions case, where i is identified as poor when $w_1 s_{i1} + w_2 s_{i2} \geq k$ (Eq. (2)). Assume without loss of generality that $w_1 > w_2 = 1 - w_1$. Figure 7 illustrates the resulting identification contour (in red) for two different cases. This contour corresponds to the union approach when $w_2 \geq k$ while it corresponds to the intersection approach when $w_1 < k$. Crucially, there is no achievement vector on the identification contour at which the slope of this

[62] The selection of dimension could also be guided by SWB data, as argued by Kingdon and Knight (2006).

[63] This is less of an issue when assuming that preference u is the same for all individuals. In any case, weights should not take equal values if preference data reliably shows that some variables should receive a larger / smaller weight.

[64] First, subjective well-being questions do not provide unbiased information on well-being. For instance, people make trade-offs between their income and their level of happiness, which is the component of happiness they tend to prefer (Adler et al., 2017). Also, choices need not align with anticipated subjective well-being (Benjamin et al., 2014a). The link between different aspects of SWB questions and different aspects of preferences is still imperfectly understood (Benjamin et al., 2017, 2023). This matters because inter-personal comparisons may change when we change the aspects of SWB from which preference u is elicited (Decancq et al., 2017). See also Ravallion (2014) for additional concerns with the use of SWB data in poverty measurement. However, subjective well-being data are easy to collect. Second, preference elicitation surveys may work relatively well for two or three variables at a time, but multidimensional poverty indicators often rely on more than 10 variables. Also, the ranking that subjects provide on achievement vectors that they did not experience themselves can be criticized as being not sufficiently well-informed (Karimi et al., 2017).

contour is equal to w_2/w_1. Equating the values of AF weights to the marginal rate of substitution of u at z (the slope of the dashed line in Figure 7) does not make sense. Instead, AF parameters (weights w, cutoffs z and identification threshold k) should be selected in a way that the identification contour best approximates indifference curve $\underline{u}$ (Decerf, 2023b). A similar point can be made for PX weights.[65]

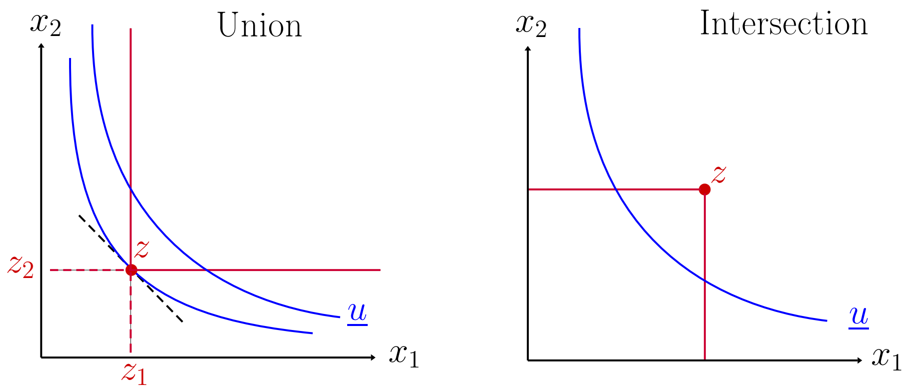

Figure 7: AF weights are not marginal weights of substitution

The fact that multidimensional poverty weights are not mere marginal rates of substitution casts serious doubts on seemingly natural methods to select their values. This is clearly the case for methods whose objective is precisely to uncover marginal rates of substitution. For instance, AF weights cannot directly be selected from the methodology proposed (in another context) by Benjamin et al. (2014b) because its aim is to uncover *average* marginal rates of substitutions for non-monetary dimensions. They can neither be selected directly from the coefficient values obtained from OLS regressions on subjective well-being data.[66] Also, it is unclear whether directly asking subjects for their preferred AF weights would yield meaningful results. Indeed, the link between AF weights and the shape of the identification contour is not trivial.

Rather, we outline a "weights selection" method whose objective is to yield interpersonal well-being comparisons that are best in line with preference u. This method is in line with the spirit of Maasoumi and Xu (2015). For Objective 1, we would like to select parameters such that the indicator most reliably yields $\underline{u} > \hat{u}(\hat{x}_i)$ whenever $\underline{u} > u(\hat{x}_i)$. Assume that $\underline{u}$ is defined from an arbitrarily selected reference vector $\underline{z}$, which may differ from the cutoff vector z.[67] Hence, we have $\underline{u} = u(\underline{z})$. The information on preference u

[65]It is possible to select PX weights such that the slope of the intermediate section of its identification contour (recall Figure 5) corresponds to the marginal rate of substitution of u at z. However, the shape of the identification contour must be a good approximation of *the whole* indifference curve $\underline{u}$. The PX parameters (weights w, cutoffs z and identification threshold k) must also be selected in a way that the implicit z_j^e defining its vertical and horizontal segments make sense.

[66]Such method could consist in running an OLS regression of subjective well-being data on the components of achievement vectors and then selecting AF weights so that they correspond to the relative values of the regression coefficients.

[67]For instance, for the WB's MPM, the reference vector $\underline{z}$ could be defined as living on the extreme IPL without suffering any additional deprivation. That is, $\underline{z}$ is defined from one of the extreme deprivation

can be used to identify (observed) achievement vectors as poor or non-poor, i.e., determine whether $u(\underline{z}) > u(\hat{x}_i)$.[68] Ideally, the proposed method selects the parameter values (w, k, z) that best approximate the inter-personal well-being comparisons following from the information on preference u. Formally, for the AF identification method, which relies on dichotomous status vectors s_i, these parameter values solve the problem

$$\min_{w,k,z} Bias$$

where $Bias$ quantifies the deviations from[69]

$$\sum_j w_j s_{ij} \geq k \qquad \text{when} \qquad u(\underline{z}) > u(\hat{x}_i).$$

In practice, some parameters value can be constrained by the data, which is typically the case for some cutoffs z_j. An equivalent problem can be defined for PX identification when replacing status s_i by deprivation d_i.[70] It could be interesting to develop such method so as to derive the weights of a global multidimensional poverty indicator from subjective well-being data and then compare the bias associated to the derived weights with the bias associated to equal-weights.

The method outlined above readily extends to more general well-being indicators, which compare any two achievement vectors.[71] One natural way of selecting the parameters defining the individual well-being indicator $\hat{u}$ is to select them in a way that maximizes the correlation between $\hat{u}(\hat{x}_i)$ and $u(\hat{x}_i)$, where u stands for the information on preference u (see footnote 68).

thresholds z_j^e.

[68] For instance, preference u can be assumed to be the preference that is most in line with the subjective well-being data, as is done in Decancq et al. (2019), or with the preference elicitation survey data, as is done in Decancq and Nys (2021).

[69] Examples of how to quantify "Bias" are proposed at the end of Section 3.2.

[70] Observe that both AF and PX methods are *linear* identification methods. This linearity constrains the substitutability-complementarity across different dimensions, i.e., the shape of their identification contour, as shown with an illustrative example in Maniquet (2023). Some authors proposed approaches for going beyond linearity. This is for instance the case of Jones (2022) when identifying the poor based on status s_i or the case of Maasoumi and Racine (2016) when relying on achievements vectors x_i (although the latter authors do not aim at aggregating dimensions as a function of preferences). Jones (2022) suggests a relatively simple identification method based on "limit vectors". One open question is how to efficiently elicit these limit vectors in practice when the number of variables is as large as a dozen. Such elicitation may require an algorithm more parsimonious than that proposed by Decancq and Nys (2021).

[71] Poverty identification methods merely compare any achievement vector to a fixed reference threshold $\underline{u}$.

3.7 Improving multidimensional analysis

Multidimensional policy analysis is another area where further progress could be made. After a multidimensional well-being indicator is constructed and its values are computed, there remains to decide what analyses to conduct with the indicator and what results should be reported to support evidence-based policy making. From this perspective, merely reporting that monetary poor individuals tend to also cumulate non-monetary deprivations falls short of what can be achieved with such indicator.

One key aspect of prioritarian policy making is to target individuals with low well-being. Another aspect is to identify the policy that delivers the largest increase in well-being at the lowest cost. These aspects could benefit from careful multidimensional policy analysis.[72]

Admittedly, policy making is routinely done dimension-by-dimension, with limited interractions between, say, the health ministry and the housing ministry. Such approach to policy making has little use for summary indicators of well-being. However, this dimension-by-dimension approach is bound to be suboptimal, if only because it does not guarantee that individuals with low well-being are prioritized. We argue it could be worth standardizing multidimensional poverty analyses so as to provide the necessary information for more efficient policy-making.

In this section, we merely illustrate with a toy-example the added value that multidimensional analysis could bring to prioritarian policy making. Assume for simplicity that individual well-being depends on two dimensions. Further assume that individual i has low well-being only when she cumulates deprivation in both dimensions, namely when $(s_{i1}, s_{i2}) = (1, 1)$. The country has three regions A, B and C, each with three individuals whose outcomes (s_{i1}, s_{i2}) are as follows:

- Region A: $(1, 0)$, $(1, 0)$ and $(0, 1)$,
- Region B: $(1, 1)$, $(1, 0)$ and $(0, 0)$,
- Region C: $(1, 1)$, $(0, 1)$ and $(0, 0)$.

Assume that the government can enact a policy j that can pull an individual out of deprivation j. This policy can be easily targeted to individuals deprived in dimension j. For instance, an electrification policy can easily be targeted at individuals without access to electricity.

To serve as benchmark, we first analyze this country using a dashboard approach, which summarizes outcomes in each dimension:[73]

- Region A: two individuals deprived in dim 1 and one individual deprived in dim 2,

[72] See for instance Azevedo and Robles (2013) on the multidimensional targeting of social protection programs.

[73] To be sure, dashboards can also contain information on cumulative deprivation (Decancq, 2022).

- Region B: two individuals deprived in dim 1 and one individual deprived in dim 2,
- Region C: one individual deprived in dim 1 and two individuals deprived in dim 2,

Provided that both dimensions are equally important to well-being, the dashboard analysis suggests that the populations in the three regions are equally well-offs. Also, if budget is limited, it would appear natural from this dashboard to implement a policy 1 in regions A and B and implement a policy 2 in region C.

We now contrast these policy recommendations with those coming from a multidimensional indicator. By assumption, we have

- Region A: no individual has low well-being
- Region B: one individual has low well-being
- Region C: one individual has low well-being

The first difference with the multidimensional analysis is that it allows prioritizing individuals with low well-being. Only in regions B and C do some individuals cumulate deprivation in the two dimensions. The policy maker should prioritize efforts towards regions B and C. Clearly, a profile of individuals with low well-being would help targeting.

The second difference is more subtle. If policy j can only be targeted to individuals who are deprived in dimension j, how can the policy maker efficiently allocate her efforts? The multidimensional analysis can provide useful information on the "targetability" of policy j. This targetability relates to the fraction of individuals deprived in j who have low well-being.

- Region B:
  - 50 percent individuals deprived in dim 1 have low well-being
  - 100 percent individuals deprived in dim 2 have low well-being
- Region C:
  - 100 percent individuals deprived in dim 1 have low well-being
  - 50 percent individuals deprived in dim 2 have low well-being

Thus, the most efficient action is to implement one policy 2 in region B and one policy 1 in region C. In that way, we perfectly target individuals with low well-being. Importantly, this policy recommendation is the exact opposite as the one that appears natural with a dashboard. This illustrates not only the relevance for policy making of multidimensional analysis. It also hints at the type of multidimensional analysis that can help improve policy making. The best practices for such analysis can take inspiration in the study by Santos et al. (2023) that evaluates the impact of policy on multidimensional poverty reductions.

4 Discussion of opportunities

We summarize in Table 3 the opportunities identified, the challenge to which they relate and whether they may help improve inter-personal comparisons (Objective 1) or cross-regions comparisons (Objective 2).[74]

Table 3: Improvement opportunities for multidimensional poverty measures.

Challenges	Opportunities	Objectives
Waste less available data	3.2 Do not needlessly dichotomize variables	
	(trichotomize categorical variables)	1 & 2
	(PX identification for consumption)	1 & 2
	Cover more dimensions through hybrid index	
	3.3 Synthetic integration of mortality-longevity	only 2
	3.4 Assume missing joint-distribution M vs OD	
	3.4.1 Assume overlap between H_M and H_{OD}	only 2
	3.4.2 Impute consumption into OD survey	1 & 2
	3.4.3 Total consumption	1 & 2
Better "filter" available data	3.5 Use most fined-grained unit of analysis	1 & 2
	(take advantage of individual-level data)	
	3.6 Select weights from information on preference	1 & 2
	(SWB data or preference-elicitation surveys)	
Beyond dim-by-dim policy making	3.7 Develop best practices for MPM analysis	1 & 2

Many opportunities presented in this review are complementary and are thus worth pursuing in parallel. Some opportunities are substitutes for one another, like the three alternative methods to deal with the missing joint distribution between monetary and non-monetary outcomes. Among these three methods, the first (overlap) is the simplest and the second (imputation) is conceptually the most promising. This does not disqualify the third method (total consumption), which could be very useful if the second method proves impractical or unreliable.

For what it is worth, our own expectation is that the largest increase in signal-to-noise ratio will come from the integration of additional data. This could in particular be the case for the methods to go beyond dichotomous variables (section 3.2), the integration of longevity (section 3.3) and the methods to bypass the missing joint distribution (section 3.4). It is unclear whether the methods suggested to go beyond equal-weights (section 3.6)

[74]Recall that Objective 1 aims at improving the targeting of policies across individuals or across households, while Objective 2 aims at improving the targeting of policies across regions or geographic areas (see Section 2.6).

can significantly increase the signal-to-noise ratio. However, going beyond equal-weights can also serve another goal, namely improving the perceived credibility and legitimacy of multidimensional poverty indicators. Yet, these methods come with their own issues, which limit the impact they may yield on the indicators' perceived credibility. There will always be doubts about the relationship between the preference data and "true" well-being (see footnote 64). Furthermore, applying preference information in another context than the one in which it is obtained is bound to raise questions on its external validity.

Finally, although the dimension-by-dimension approach to policy making is unlikely to change overnight, it might still be worth conducting policy analysis in a multidimensional way. Indeed, some of its insights may be productively used. And even if it is not immediately the case, these insights could raise awareness of the potential gains of an integrated approach to policy making.

5 Research papers

In this section, we outline a list of seven research papers based on the opportunities identified.

1. *Empirical implications of going beyond the dichotomous counting methodology*

   Dichotomizing non-binary variables "wastes" well-being relevant information. The solutions listed in section 3.2 provide ready-to-use conceptual improvements on the dichotomous counting methodology underlying mainstream multidimensional poverty measures. They will thus help improve the targeting of policies across households. One open question is the empirical significance of these solutions. By how much do they affect inter-household and cross-region well-being comparisons? How different are the households they identify as poor when compared to those identified under the counting methodology? This could be investigated both for the World Bank's MPM as well as for national multidimensional poverty measures.

2. *Harnessing separate surveys through hybrid well-being indices*

   Hybrid well-being indices, which are part-synthetic and part-objective, are one way of avoiding "wasting" part of the data available in separate surveys, like LSMS and DHS. Sections 3.3 and 3.4.1 list solutions for the construction of hybrid multidimensional indices. They may thus help improve cross-regions well-being comparisons as well as a region's well-being trend. The paper should study how to apply this kind of solutions in practice and quantify their impact on well-being comparisons.

3. *Impute or not impute monetary outcomes into DHS and MICS, and how?*

   Imputing monetary outcomes into a non-monetary survey is one way of avoiding "wasting" some of the data present in separate surveys, like LSMS and DHS. However, imputation methods have limitations, mainly that their underlying assumptions may not hold in practice. One open empirical question is whether they still improve (inter-households and/or cross-regions) well-being comparisons *when their underlying assumptions do not hold*. Hence, is it the case that using data from two separate surveys more than compensates for the errors made on their joint distribution? Another empirical question is which imputation method performs best to impute monetary outcomes into non-monetary surveys. The paper should study these questions and quantify how different are the well-being comparisons based on these imputations from currently used comparisons and from "ideal" comparisons made when the joint distribution is known (see sections 3.4.1 and 3.4.2).

4. *Accounting for local public services through the total consumption approach*

   LSMS consumption surveys often provide very limited information on households' outcomes in dimensions such as health, education or security. One pragmatic solution is to construct a household's "total consumption" from monetary valuations of the public services provided in their area. The paper should investigate the best approach to make these monetary valuations at scale given prevailing data constraints. It should then quantify the impact of such valuations on well-being comparisons. Conceptually, it remains an open question how to aggregate consumption aggregates with monetary valuations of public services. Indeed, households cannot freely allocate the monetary valuations from these services. As a result, assuming perfect substitution (i.e., summing these valuations) may not be conceptually ideal. The paper should provide a conceptual discussion of alternative aggregation methods.

5. *Multidimensional poverty weights selected from information on preferences*

   The "equal-weights" procedure is regularly used when designing national multidimensional poverty measures and always used for global measures. The use of this obviously ad-hoc procedure runs two risks. First, measures based on ad-hoc weights may yield implausible well-being comparisons in the sense that they would deviate too much from people's views. Second, ad-hoc weights may reduce the perceived legitimacy of these measures. The paper should propose a conceptually sound method to select values for these weights that is consistent with information on preferences (see section 3.6). It should also study empirically the extent to which the obtained weights yield different well-being comparisons than "equal-weights". The information can come from either subjective well-being data or from a one-off dedicated preference elicitation survey.

6. *A guide on best practices for multidimensional poverty analyses at the World Bank*

   A lot of attention goes into the design of multidimensional poverty measures. Yet, an important aspect for their policy relevance is the type of analyses conducted with these measures. Merely quantifying the extent to which poor individuals cumulate deprivations from different dimensions is far from exhausting their potential in guiding policies (see section 3.7). The Poverty and Equity GP may want to produce a guide for the types of multidimensional analyses that it deems most relevant for policy making.

7. *An improved design for the World Bank's Multidimensional Poverty Measure*

   The previous papers will lay the groundwork for an updated design for the World Bank's MPM, adressing several of the major limitations in its current design. Even if the MPM is unlikely to be computed with the same frequency and timeliness as monetary indicators, it still provides a minimal robustness check on the insights gathered with the latter. The end of the SDG agenda in 2030 provides a timeline for delivering such an updated design.

6 Appendix

6.1 The global MPI and the World Bank's MPM

In this appendix, we succinctly present the definition of the two main global multidimensional poverty measures. We start with the global MPI from UNDP-OPHI (Alkire et al., 2015). The global MPI considers three dimensions, namely health, education and living standards. Each dimension is captured by one or more variables, whose respective weights are summarized in Table 4. The global MPI identifies as (multidimensionally) poor any individual who lives in a household whose total deprivation is at least as large as the identification threshold $k = 1/3$.

Table 4: global MPI.

Dimension	Variables	Weight
Health	A member of the household is malnourished	$1/6$
	A child has died in the family	$1/6$
Education	No household member has completed five years of schooling	$1/6$
	A school-aged child is not attending school	$1/6$
Living Standards	The household has no access to electricity	$1/18$
	The household lacks access to sanitation facility	$1/18$
	The household lacks access to clean water	$1/18$
	The housing is built with inadequate construction material	$1/18$
	The household cooks with inadequate combustible	$1/18$

We continue with the World Bank's Multidimensional Poverty Measure (MPM) (World Bank, 2018). The MPM considers three dimensions, namely monetary poverty, education and access to basic infrastructure. Each dimension is captured by one or more variables, whose respective weights are summarized in Table 5. The MPM identifies as (multidimensionally) poor any individual who lives in a household whose total deprivation is at least as large as the identification threshold $k = 1/3$. This is for instance the case if her household is monetary poor. This is also the case if her household lacks access to two basic infrastructures and is deprived in one education variable, because its total deprivation is then equal to $1/9+1/9+1/6$, which is larger than k.

Table 5: World Bank's Multidimensional Poverty Measure (3-dimensions version) (WorldBank, 2020).

Dimension	Variables	Weight
Monetary poverty	Daily consumption is less than US$ 1.9 per person	1/3
Education	At least one school-age child is not enrolled in school	1/6
	No adult in household has completed primary education	1/6
Basic infrastructure	The household lacks access to limited-standard drinking water	1/9
	The household lacks access to limited-standard sanitation	1/9
	The household has no access to electricity	1/9

References

Adler, M. D., Dolan, P., and Kavetsos, G. (2017). Would you choose to be happy? tradeoffs between happiness and the other dimensions of life in a large population survey. *Journal of Economic Behavior & Organization*, 139:60–73.

Al-Ajlani, H., Van Ootegem, L., and Verhofstadt, E. (2020). Does well-being vary with an individual-specific weighting scheme? *Applied Research in Quality of Life*, 15:1285–1302.

Alkire, S., Conconi, A., and Seth, S. (2014). Multidimensional destitution: An ordinal counting methodology for constructing linked subsets of the poor. *OPHI Research in Progress*, 42:1–24.

Alkire, S. and Foster, J. (2011a). Counting and multidimensional poverty measurement. *Journal of Public Economics*, 95(7-8):476–487.

Alkire, S. and Foster, J. (2011b). Understandings and misunderstandings of multidimensional poverty measurement. *The Journal of Economic Inequality*, 9(2):289–314.

Alkire, S., Foster, J. E., Maria, F., Paola, E., and Seth, S. (2015). *Multidimensional Poverty Measurement and Analysis*. Oxford University Press.

Alkire, S. and Haq, R. U. (2023). Analyzing individual deprivations alongside household poverty to illuminate gendered and intrahousehold disparities.

Anand, S. and Sen, A. (1994). Human development index: Methodology and measurement.

Atkinson, A. (1970). On the measurement of inequality. *Journal of economic theory*, 2(3):244–263.

Atkinson, A. (2016). *Monitoring Global Poverty: Report of the commission on global poverty*. The World Bank, Washigton, DC.

Azevedo, V. and Robles, M. (2013). Multidimensional targeting: Identifying beneficiaries of conditional cash transfer programs. *Social Indicators Research*, 112:447–475.

Baland, J.-M., Cassan, G., and Decerf, B. (2021). “too young to die”: Deprivation measures combining poverty and premature mortality. *American Economic Journal: Applied Economics*, 13(4):226–257.

Baland, J.-M., Cassan, G., and Decerf, B. (2023). Poverty-adjusted life expectancy.

Bank, W. (2022). *Poverty and shared prosperity 2022: Correcting course*. The World Bank.

Bank, W., UNDP, and UNICEF (2021). A roadmap for countries measuring multidimensional poverty.

Barofsky, J. and Younger, S. (2019). The effect of government health expenditure on the income distribution: a comparison of valuation methods in ghana. Technical report, Tulane University, Department of Economics.

Becker, G. S., Philipson, T. J., and Soares, R. R. (2005). The quantity and quality of life and the evolution of world inequality. *American economic review*, 95(1):277–291.

Benjamin, D. J., Cooper, K. B., Heffetz, O., and Kimball, M. (2017). Challenges in constructing a survey-based well-being index. *American Economic Review*, 107(5):81–85.

Benjamin, D. J., Debnam Guzman, J., Fleurbaey, M., Heffetz, O., and Kimball, M. (2023). What do happiness data mean? theory and survey evidence. *Journal of the European Economic Association*, page jvad026.

Benjamin, D. J., Heffetz, O., Kimball, M. S., and Rees-Jones, A. (2014a). Can marginal rates of substitution be inferred from happiness data? evidence from residency choices. *American Economic Review*, 104(11):3498–3528.

Benjamin, D. J., Heffetz, O., Kimball, M. S., and Szembrot, N. (2014b). Beyond happiness and satisfaction: Toward well-being indices based on stated preference. *American economic review*, 104(9):2698–2735.

Bennett, F. (2021). The hidden dimensions of poverty: reflections on ngo/academic partnership. *Journal of Poverty and Social Justice*, 29(3):317–324.

Blanchflower, D. G. and Bryson, A. (2022). Taking the pulse of nations: A biometric measure of well-being. *Economics & Human Biology*, 46:101141.

Boarini, R., Fleurbaey, M., Murtin, F., and Schreyer, P. (2022). Well-being during the great recession: new evidence from a measure of multi-dimensional living standards with heterogeneous preferences. *The Scandinavian Journal of Economics*, 124(1):104–138.

Boarini, R., Murtin, F., Schreyer, P., and Fleurbaey, M. (2016). Multi-dimensional living standards: A welfare measure based on preferences.

Browning, M. and Chiappori, P.-A. (1998). Efficient intra-household allocations: A general characterization and empirical tests. *Econometrica*, pages 1241–1278.

Cherchye, L., De Rock, B., Lewbel, A., and Vermeulen, F. (2015). Sharing rule identification for general collective consumption models. *Econometrica*, 83(5):2001–2041.

Chi, G., Fang, H., Chatterjee, S., and Blumenstock, J. E. (2022). Microestimates of wealth for all low-and middle-income countries. *Proceedings of the National Academy of Sciences*, 119(3):e2113658119.

Chipman, J. S. and Moore, J. C. (1980). Compensating variation, consumer's surplus, and welfare. *The American Economic Review*, 70(5):933–949.

Dang, H.-A., Kilic, T., Abanokova, K., and Carletto, C. (2023). Poverty imputation in contexts without consumption data: a revisit with further refinements.

Dang, H.-A. H. (2021). To impute or not to impute, and how? a review of poverty-estimation methods in the absence of consumption data. *Development Policy Review*, 39(6):1008–1030.

Dang, H.-A. H. and Lanjouw, P. F. (2023). Regression-based imputation for poverty measurement in data-scarce settings. In *Research Handbook on Measuring Poverty and Deprivation*, pages 141–150. Edward Elgar Publishing.

Deaton, A. and Muellbauer, J. (1980). An almost ideal demand system. *The American economic review*, 70(3):312–326.

Decancq, K. (2022). Cumulative deprivation: identification and aggregation.

Decancq, K., Fleurbaey, M., and Maniquet, F. (2019). Multidimensional poverty measurement with individual preferences. *Journal of Economic Inequality*, 17(1):29–49.

Decancq, K., Fleurbaey, M., and Schokkaert, E. (2015). Happiness, equivalent incomes and respect for individual preferences. *Economica*, 82:1082–1106.

Decancq, K. and Lugo, M. A. (2013). Weights in multidimensional indices of wellbeing: An overview. *Econometric Reviews*, 32(1):7–34.

Decancq, K. and Neumann, D. (2014). Does the choice of well-being measure matter empirically? an illustration with german data.

Decancq, K. and Nys, A. (2021). Non-parametric well-being comparisons. *European Economic Review*, 133:103666.

Decerf, B. (2023a). Mortality and poverty measurement. *Chapters*, pages 572–580.

Decerf, B. (2023b). A preference-based theory unifying monetary and non-monetary poverty measurement. *Journal of Public Economics*, 222:104898.

Decerf, B. and Fonton, K. (2023). Reconceptualizing global multidimensional poverty measurement, with illustration on nigerian data.

Decerf, B., Friedman, J., Mendes, A., Pennings, S., and Yonzan, N. (2024). Lives, livelihoods and learning. a global perspective on well-being impacts of the covid-19 pandemic.

Defloor, B., Verhofstadt, E., and Van Ootegem, L. (2017). The influence of preference information on equivalent income. *Social Indicators Research*, 131:489–507.

Diener, E., Oishi, S., and Tay, L. (2018). Advances in subjective well-being research. *Nature Human Behaviour*, 2(4):253–260.

Diewert, W. E. (1998). Index number issues in the consumer price index. *Journal of Economic Perspectives*, 12(1):47–58.

Dirksen, J. and Alkire, S. (2021). Children and multidimensional poverty: Four measurement strategies. *Sustainability*, 13(16):9108.

Dunbar, G. R., Lewbel, A., and Pendakur, K. (2013). Children's resources in collective households: identification, estimation, and an application to child poverty in malawi. *American Economic Review*, 103(1):438–471.

Elbers, C., Lanjouw, J. O., and Lanjouw, P. (2003). Micro-level estimation of poverty and inequality. *Econometrica*, 71(1):355–364.

Evans, M., Nogales, R., and Robson, M. (2023). Monetary and multidimensional poverty: Correlation, mismatches, and a combined approach. *The Journal of Development Studies*, pages 1–24.

Fleurbaey, M. (2009). Beyond gdp: The quest for a measure of social welfare. *Journal of Economic literature*, 47(4):1029–1075.

Fleurbaey, M. and Maniquet, F. (2011). *A Theory of Fairness and Social Welfare*. Cambridge University Press, Cambridge.

Gethin, A. (2023). Revisiting global poverty reduction: Public goods and the world distribution of income, 1980-2022.

Hicks, J. R. (1942). Consumers' surplus and index-numbers. *The Review of Economic Studies*, 9(2):126–137.

Jolliffe, D. M., Mahler, D. G., Lakner, C., Atamanov, A., and Tetteh Baah, S. K. (2022). Assessing the impact of the 2017 ppps on the international poverty line and global poverty.

Jones, C. I. and Klenow, P. J. (2016). Beyond gdp? welfare across countries and time. *American Economic Review*, 106(9):2426–2457.

Jones, S. (2022). Extending multidimensional poverty identification: from additive weights to minimal bundles. *The Journal of Economic Inequality*, 20(2):421–438.

Kanbur, R. and Lustig, N. (2001). Voices of the poor. *World Development Report 2000/2001*.

Kanbur, R. and Mukherjee, D. (2007). Premature mortality and poverty measurement. *Bulletin of economic Research*, 59(4):339–359.

Karimi, M., Brazier, J., and Paisley, S. (2017). Are preferences over health states informed? *Health and quality of life outcomes*, 15(1):1–11.

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

Kniesner, T. J. and Viscusi, W. K. (2019). The value of a statistical life. *Forthcoming, Oxford Research Encyclopedia of Economics and Finance, Vanderbilt Law Research Paper*, (19-15).

Kraay, A. (2019). The world bank human capital index: a guide. *The World Bank Research Observer*, 34(1):1–33.

Kraay, A., Lakner, C., Ozler, B., Decerf, B. M. A., Jolliffe, D. M., Sterck, O. C. B., and Yonzan, N. (2023). A new distribution sensitive index for measuring welfare, poverty, and inequality. *World Bank Working Paper Series*.

Lechene, V., Pendakur, K., and Wolf, A. (2022). Ordinary least squares estimation of the intrahousehold distribution of expenditure. *Journal of Political Economy*, 130(3):681–731.

Lefebvre, M., Pestieau, P., and Ponthière, G. (2018). Fgt old-age poverty measures and the mortality paradox: Theory and evidence. *Review of Income and Wealth*, 64(2):428–458.

Lise, J. and Seitz, S. (2011). Consumption inequality and intra-household allocations. *The Review of Economic Studies*, 78(1):328–355.

Maasoumi, E. and Racine, J. S. (2016). A solution to aggregation and an application to multidimensional well-being frontiers. *Journal of Econometrics*, 191(2):374–383.

Maasoumi, E. and Xu, T. (2015). Weights and substitution degree in multidimensional well-being in china. *Journal of Economic Studies*, 42(1):4–19.

Mancini, G. and Vecchi, G. (2022). On the construction of a consumption aggregate for inequality and poverty analysis. *Washington, DC: The World Bank*.

Maniquet, F. (2023). 37. multidimensional poverty measurement and preferences. *Research Handbook on Measuring Poverty and Deprivation*, page 401.

Nussbaum, M. C. (2009). Creating capabilities: The human development approach and its implementation. *Hypatia*, 24(3):211–215.

Pattanaik, P. K. and Xu, Y. (2018). On measuring multidimensional deprivation. *Journal of Economic Literature*, 56(2):657–72.

Piacquadio, P. G. (2017). A fairness justification of utilitarianism. *Econometrica*, 85(4):1261–1276.

Ravallion, M. (2011). On multidimensional indices of poverty. *The Journal of Economic Inequality*, 9(2):235–248.

Ravallion, M. (2012). Mashup indices of development. *The World Bank Research Observer*, 27(1):1–32.

Ravallion, M. (2014). Poor, or just feeling poor? on using subjective data in measuring poverty. *Happiness and economic growth*, pages 140–178.

Ravallion, M. (2016). *The Economics of Poverty: History, Measurement, and Policy.* Oxford University Press.

Ravallion, M. and Chen, S. (2011). Weakly relative poverty. *Review of Economics and Statistics*, 93(4):1251–1261.

Riumallo-Herl, C., Canning, D., and Salomon, J. A. (2018). Measuring health and economic wellbeing in the sustainable development goals era: development of a poverty-free life expectancy metric and estimates for 90 countries. *The Lancet Global Health*, 6(8):e843–e858.

Santos, M. E., Lustig, N., and Zanetti, M. M. (2023). Counting and accounting: Measuring the effectiveness of fiscal policy in multidimensional poverty reduction.

Sen, A. (1999). *Development as freedom*. New York: Anchor Books.

Simpson, H. (2009). Productivity in public services. *Journal of Economic Surveys*, 23(2):250–276.

Steele, J. E., Sundsøy, P. R., Pezzulo, C., Alegana, V. A., Bird, T. J., Blumenstock, J., Bjelland, J., Engø-Monsen, K., De Montjoye, Y.-A., Iqbal, A. M., et al. (2017). Mapping poverty using mobile phone and satellite data. *Journal of The Royal Society Interface*, 14(127):20160690.

Steptoe, A., Deaton, A., and Stone, A. A. (2015). Subjective wellbeing, health, and ageing. *The Lancet*, 385(9968):640–648.

Stiglitz, J. E., Sen, A., Fitoussi, J.-P., et al. (2009). Report by the commission on the measurement of economic performance and social progress.

Szende, A., Janssen, B., and Cabases, J. (2014). Self-reported population health: an international perspective based on eq-5d.

Tran, V. Q., Alkire, S., and Klasen, S. (2015). Static and dynamic disparities between monetary and multidimensional poverty measurement: Evidence from vietnam. In *Measurement of poverty, deprivation, and economic mobility*, pages 249–281. Emerald Group Publishing Limited.

World Bank (2018). *Poverty and Shared Prosperity 2018: Piecing Together the Poverty Puzzle*. World Bank, Washington, DC.

WorldBank (2020). Poverty and shared prosperity 2020: Reversals of fortune.