Composite indices and preference-based measures of wellbeing

Sustainable and Inclusive Wellbeing Series

HIGHLIGHTS

→ Wellbeing is a multidimensional concept that yields measurement challenges, such as how to aggregate different dimensions and which relative weights to give them.

→ Composite indices are widely used to measure different aspects of wellbeing. While they are relatively easy to understand and construct, they often contain underlying trade-offs between dimensions that are difficult to discern.

→ Preference-based measures are rooted in economic theory and derived from the trade-offs people make between different dimensions of wellbeing. However, they are data-intensive and pose measurement challenges.

→ Investments to better measure outcomes and preferences at the individual level could yield a more nuanced view of well-being across the EU.

INTRODUCTION

Promoting the wellbeing of citizens is a goal enshrined within article 3 of the Treaty of the European Union.[1] Wellbeing measures are key to evaluate progress towards this goal across member states and over time. Traditionally, societal wellbeing has been often equated with indicators of economic activity, such as income, or GDP per capita. However, such measures fail to capture the value of other aspects that people care about, such as health, social interactions, safety, or the environment, among others. In addition, they do not incorporate a concern for the underlying distribution of income or wellbeing in a society. Accordingly, there has been an impetus among academics and policymakers to move ***beyond GDP*** towards multi-dimensional indicators of wellbeing (Stiglitz et al. 2009). For instance, the EU Commission's 2023 Strategic Foresight Report lists the adjustment of GDP in terms of sustainability and inclusivity as a key priority area.

THE POLICYMAKING DILEMMA

Moving towards a multidimensional perspective introduces new challenges for the development of wellbeing indicators. To demonstrate these, consider the following hypothetical society in Table 1 below, which consists of two individuals that differ in their income and health. For simplicity, let's assume that health is measurable on a scale from 0 to 100, where 0 represents worst possible health state and 100 the

[1] Treaty on European Union - TITLE I: COMMON PROVISIONS - Article 3 (ex Article 2 TEU), accessed at link [07/11/24].

best possible.[2] Let's imagine now that a policymaker has to decide who is better off. Individual A is richer than individual B, but the latter has better health. If one focuses on income alone, as has traditionally been the case, individual A would be ranked better off. Bringing health into the picture, however, introduces a dilemma for the policymaker, since each individual is now better off in at least one dimension. Hence, to make interpersonal comparisons of wellbeing, one must consider the relative importance of each dimension.

Table 1 – Moving to a multidimensional perspective

Individual	Income	Health
A	50,000	70
B	40,000	100

Opinions differ in the research and policy-making spheres on the ways to address this issue. One may argue against the explicit aggregation of any dimensions at all, preferring instead to rely on **dashboard approaches** which display outcomes in each dimension separately. Yet, this approach does not allow for the assessment of inequalities in wellbeing, which, as demonstrated above, would require some form of aggregation across the dimensions.[3] One approach is to aggregate the dimensions by means of a mathematical procedure known as a **composite index**. This aligns with an objective viewpoint of wellbeing, in which an analyst decides on a common scale for the different dimensions (i.e., normalisation), their relative importance, and how they contribute to wellbeing. One of the most well-known and widely applied examples is the United Nations Development Programme's (UNDP) Human Development Index (HDI), which was first launched in 1990 (UNDP, 1990). The EU Regional Social Progress Index is a more recent example within the European context (de Dominicis et al., 2024).

Another strand of the literature proposes an aggregation approach based on trade-offs people make between different dimensions of wellbeing, known as **preference-based measures**.[4] By preferences, we mean the relative value that people place on different dimensions of life for their wellbeing. These values can be inferred from the trade-offs individuals are willing to make, either via observed or hypothetical choices. Preference-based measures typically augment income per capita to incorporate the value of other non-monetary dimensions, such as health, in monetary terms. They have recently been applied within the EU's 'Sustainable and Inclusive Wellbeing' (SIWB) framework (see Benczur et al., 2023).

A final stream of the literature argues that measures of **subjective wellbeing**, e.g., self-reported happiness and life satisfaction scores, should be used to resolve the policymaker's dilemma of assigning weights to the different dimensions of wellbeing. For instance, the World Happiness Report, published annually since 2012, reports average levels of life satisfaction by country. While such data are often readily available and easy to collect, they may be problematic for making intra and interpersonal comparisons of wellbeing (Fleurbaey and Blanchet, 2013). Individuals with the same circumstances and preferences may rate their life satisfaction differently, for instance, due to different interpretations of the rating scale and/or adaptations to one's situation.

This policy brief discusses the relative pros and cons of two of these approaches within the EU policymaking context: composite indices and preference-based measures. It also provides an overview of recent advances within the literature and the key challenges for constructing such measures. We do not cover subjective wellbeing measures in this note as they do not involve aggregation of dimensions on the part of an analyst, an issue we will explore in the next section. We also omit the capabilities approach (see Sen, 1985) due to the empirical challenges associated with its operationalisation, i.e., assessing capability sets.

COMPOSITE INDICES

Composite indices are a mathematical aggregation of several dimensions into a single number, which can be used to rank different assessment units (e.g., individuals, regions, countries). Typically, these

[2] Note, however, that the ability to measure health on a cardinal scale is controversial. See Hausman (2015) for a discussion of the ethical challenges associated with measuring health.

[3] Dashboard approaches only allow one to identify the worst off through union approaches, for instance, an individual or country that is below a specific level on one or more dimensions of wellbeing. Although this may be sufficient for poverty measurement but does not allow us to compare differences in wellbeing.

[4] We focus on preference-based measures that utilise the concept of indifference to identify equivalent situations to one's own in a unidimensional space. That is, measures which rely on ordinal information about preferences. This contrasts with cardinal approaches (i.e., subjective well-being) that assume well-being can be measured on a specific scale.

dimensions are chosen to reflect an underlying concept, such as wellbeing or, more broadly, social progress. For example, the EU Social Progress Index aggregates 53 indicators across three core dimensions: Basic Human Needs, Foundations of Wellbeing, and Opportunity. In recent years, the use of such indices has accelerated, with mapping exercises suggesting the existence of over 150 indices of human wellbeing and progress (Yang, 2014).[5]

The construction of a composite index usually follows several stages (see the toolkit developed by JRC Competence Centre on Composite Indicators for more details). Firstly, the researcher selects the dimensions and the relevant indicators to capture the underlying concept being measured (e.g., wellbeing). Next, the indicators are re-scaled (or normalised) between 0 and 1. A common approach, the min-max method, consists of scaling each dimension relative to a minimum and maximum level. The normalised dimensions are then aggregated using a set of relative weights and a specific aggregation function. While a variety of weighting methods exist, many indices rely on so-called "equal weights", expert judgements, or statistical approaches (e.g., principal component analysis, PCA).[6]

The aggregation function represents the way in which the dimensions are combined to yield changes in the wellbeing. Commonly used functions include the arithmetic or geometric mean of the normalised indicators. The form of aggregation also influences the substitutability between the dimensions. For instance, in a linear aggregation, such as the arithmetic mean, shortfalls in one transformed dimension can be perfectly compensated by gains in another. Meanwhile, a geometric mean reduces the degree of substitutability between dimensions by allowing the contribution of each dimension to depend on the level of the others. This makes it more difficult for high performance in one dimension to mask poor performance in another. Finally, researchers often carry out sensitivity analyses to test whether rankings of countries vary substantially according to different assumptions within the index, such as the weights and functional form.[7]

Table 2 demonstrates how a composite index could resolve the policymaker's dilemma in the introduction. First, we re-scale the dimensions using the min-max method. For the sake of the exercise, let's assume that the minimum and maximum income levels are 0 and 100,000 euros. Health, in this case, has lower and upper bounds of 0 and 100. To aggregate the normalised dimensions, we assume each dimension has equal weights and take a weighted average of the rescaled dimensions for each individual. The composite index indicates that, under these assumptions, individual B is better off than A.

Table 2 – A simple composite index

Individual	Income	Health	Index
A	0.5	0.7	**0.6**
B	0.4	1.0	**0.7**

Advocates of composite indices argue that they are transparent and easy to understand, which makes them effective for quickly communicating vast amounts of data and complex issues to policy makers.[8] This also helps to raise awareness of the multidimensional nature of specific concepts, such as wellbeing, among different audiences (e.g., stakeholders, policymakers, and the general public). In addition, they can be tailored to various objectives or contexts. This flexibility makes them useful for addressing the ever-changing needs or priorities of policymakers.

Nevertheless, a growing literature has begun to contend some of these claims. We will focus on two issues in particular. Firstly, it is not evident that composite indices are as transparent as they first seem. While the "explicit" weights and functional form of the index are easy to understand, the "implicit" trade-offs between the dimensions contained within such indices are sometimes obscure.[9] This is because they depend not only on the explicit weights given to each dimension, but also on the functional form used for aggregation and re-scaling of the dimensions. Ravallion (2012a), for instance, demonstrates that the

[5] See also the COIN Explorer, for a list of different composite indicators: https://composite-indicators.jrc.ec.europa.eu/explorer

[6] See Decancq and Lugo (2013) for an overview of different weighting approaches.

[7] See the COINr tool for the example of a comprehensive statistical programme that builds and analyses composite indicators: https://knowledge4policy.ec.europa.eu/composite-indicators/toolkit_en/coinr_en

[8] See Nardo et al., (2005) for a detailed discussion of the relative pros and cons of composite indices. They also cover a range of issues related to data processing (i.e., outliers) that are relevant for the construction of an index but not covered here.

[9] 'Non-compensatory' approaches have also been utilised for composite indicators. These are inspired by Condorcet voting, in which countries (or alternatives) are ranked based on pairwise comparisons across multiple indicators. See, for example, the JRC SOCRATES tool: https://web.jrc.ec.europa.eu/policy-model-inventory/explore/models/model-socrates/

arbitrary assumptions contained within the HDI, such as equal weights, a logarithmic transformation of income, and geometric aggregation, can lead to ethically questionable trade-offs between health, income, and education. In our simple example above, a one unit of change in health yields the same change in the overall wellbeing index as a 1,000 euro increment in income.[10] This value would change if we selected different normalisation bounds or weights.

Second, many composite indices rely on a single aggregation function with weights that are derived from expert opinion or statistical methods, such as principal component analysis (PCA). This is problematic for at least two reasons. Firstly, the implied trade-offs within these indices may be at odds with the preferences of the population under consideration. More specifically, experts may not be the best judges of an individual's wellbeing because they have imperfect information on what individuals' care about. Meanwhile statistical approaches, such as PCA, assign weights based on the underlying variance structure of the data. Again, this may not necessarily reflect the relative importance of each dimension for individual wellbeing. Secondly, people often differ in their views on what matters in life. For instance, some may place a lot of importance on generating income for a good life, while others would prefer to work less and spend more time with their families. Applying a single index (with identical weights and functional form) across the population fails to respect these differences.[11]

PREFERENCE-BASED MEASURES

Preference-based measures of wellbeing provide an alternative framework in which to aggregate various dimensions. This approach is grounded in welfare economic theory, which equates wellbeing with the satisfaction of individual preferences. Underlying this theory is the assumption that individuals are in the best position to judge about what is good for them (individual sovereignty) and that their preferences can (under certain conditions) serve as indicators of their wellbeing.[12] By utilising these preferences, it is possible to transform a multidimensional measurement problem into a unidimensional equivalent.[13]

In this section, we will focus on a broad class of preference-based measures known as equivalence approaches. These consist of evaluating wellbeing by searching for reference situations that an individual deems as good as (or equivalent to) their actual situation.

A BRIEF HISTORY

Economists have been proposing preference-based measures of wellbeing since the early 1970s. Dan Usher (1973) was the first to incorporate the value of changes in mortality risk within a national income accounting framework. His longevity adjusted consumption measure was defined as the hypothetical consumption level that if enjoyed by an individual in all remaining years of life, while facing the mortality rates of a reference year, would make them indifferent to their actual situation, defined by their current consumption level and mortality rates. A novelty of his approach was to weight the relative achievements in each dimension based on the trade-offs between consumption and mortality observed in different settings (e.g., the labour market). While imperfect, the empirical nature of these weights, and their theoretical underpinnings, made the measure popular among economists who subsequently applied and expanded upon the approach. Nordhaus (2003), for instance, demonstrated that the value of life expectancy gains over the twentieth century in the USA were comparable to total economic growth. Becker et al. (2005) found that incorporating the value of life expectancy gains within GDP per capita, reversed trends in global inequality based on income alone, particularly due to the large gains in life expectancy observed in low-income countries. Fleurbaey and Gaulier (2009) incorporated adjustments for leisure time, unemployment risk and household composition. They found that the rankings of OECD countries change substantially based using their measure vis-à-vis GDP per capita. Jones and Klenow (2016) extended the measure to incorporate the value of leisure time and inequalities in consumption. They found that while

10 In this simple framework, the trade-off is calculated as the marginal rate of substitution between the two dimensions, given by the ratio: $\frac{h^{max}-h^{min}}{y^{max}-y^{min}}$.

11 While some statistical approaches (e.g., data envelopment analysis) yield a set of weights for each individual that maximises their wellbeing, there is no a-priori reason why these will align with their preferences.

12 As noted by Hausman (2011), preferences may serve as an indicator of an individual's wellbeing, provided they are self-interested, based on full information and correct beliefs about different alternatives.

13 While there are a variety of approaches used in economics to build well-being measures, the most prominent methods tend to rely on some concept of equivalence (see Da Costa et al. 2024).

GDP per capita and their measure of wellbeing are highly correlated, differences between the two are often important.[14]

The aforementioned measures tend to rely on average outcomes at the country level and impose a single set of preferences across the population. This is limiting because it ignores within country inequalities in wellbeing and differences in preferences at the individual level. The equivalent income approach (Fleurbaey, 2009) seeks to overcome these limitations by measuring wellbeing at the *individual level* in a way that respects the relative values that people place on different non-monetary dimensions of life. More specifically, it is the hypothetical income that, if combined with a reference level of some non-monetary dimension (e.g., perfect health), would yield the same level of wellbeing as the initial situation. By placing individuals in the same reference situation it also allows for interpersonal comparisons of wellbeing when preferences differ.

Empirically, equivalent incomes have been applied in various contexts (see Decancq, Schokkaert and Fleurbaey, 2015a). More recently, they have been utilized to measure wellbeing across the EU. Using EU-SILC data, Decancq and Gamage (2023) construct equivalent incomes that incorporate five dimensions of wellbeing: income, employment, crime, pollution and health. They find that wellbeing inequality increased during the Great Recession (2008-15) with health being a key driver. They also demonstrate that preference heterogeneity is an important contributor to wellbeing inequality between European citizens and countries. Boarini et al. (2022) also assess the effect of the Great Recession across 28 OECD countries using equivalent incomes. They capture preference heterogeneity in terms of the welfare loss due to unemployment and changes in life expectancy. They find that while economic growth slightly increased over the period 2008-2013, societal wellbeing declined substantially.

EXAMPLE

To demonstrate the use of a preference-based measure, we will focus on the equivalent income (Fleurbaey, 2009). Let us return to the hypothetical society from before in Table 3 below. The left-hand side of the table displays the actual outcomes for the two individuals. Now suppose we could place both individuals in a hypothetical situation (right hand side of the table) where they both have perfect health, represented by a score of 100. Since both individuals have the same levels of health in this hypothetical situation, we can now compare their wellbeing in terms of a singular dimension, the equivalent income. To calculate this, we ask each individual which level of income, combined with perfect health, would make them indifferent between the hypothetical situation and the actual one. For simplicity, assume that both individuals have the same preferences and are willing to give up 1,000 euros (as implied by the composite index above) for each additional unit of health. Individual A would therefore be indifferent between their actual situation and a hypothetical situation with an (equivalent) income of 20,000 euros and perfect health. Individual B already has perfect health meaning that the hypothetical situation is identical to their actual situation. Their equivalent income is simply their current income. Comparing these values, we can state that individual B has a higher level of wellbeing than individual A.

Table 3 – Equivalent incomes

	Actual situation		Hypothetical situation	
Individual	**Income**	**Health**	**Eq.Income**	**Ref.health**
A	50,000	70	20,000	100
B	40,000	100	40,000	100

Note: Eq.Income = equivalent income; Ref.health = reference health. We assume that both individuals are willing to give up 1,000 euros for each additional unit of health.

ADVANTAGES AND LIMITATIONS

Preference-based measures are rooted in a theory of wellbeing which implies that they do not rely on arbitrary assumptions regarding how to aggregate different dimensions. In addition, they allow policy makers to remain agnostic about the relative weights of different dimensions of wellbeing by deferring to individual's judgements. By respecting individual sovereignty, policymakers can avoid the critique of paternalism. Measures such as the equivalent income are also more transparent about the underlying trade-offs between dimensions, which allows one to observe the direct contribution of different dimensions to wellbeing. Moreover, their cardinal nature enables the construction of meaningful 'growth rates' over time, or an analysis of rates of return or costs and benefits.

[14] Despite income and life expectancy being highly correlated, the former fails to capture the value that individual's place on living longer and healthier lives. Jones and Klenow (2016) demonstrate that this value may also increase as people get richer due to diminishing returns to income.

A common criticism against equivalent incomes is that they are money-centric. Accordingly, there may be some aversion among policymakers to placing monetary values on different dimensions of wellbeing, such as years of life. There two important points to note here. Firstly, income is an arbitrary numeraire used to quantify wellbeing. One could re-specify the equivalent income in terms of a non-market dimension, as Ponthiere et al. 2023 do using an equivalent lifetime measure. Secondly, such trade-offs are regularly used within EU policymaking. For instance, value of statistical life calculations are employed within cost benefit analyses of the REACH regulation. Regardless, making these trade-offs explicit may be a first step to better understanding what constitutes wellbeing.

A more fundamental critique relates to the dependence of equivalent incomes on reference values for the non-market dimensions. For instance, D'Albis and Bonnet (2018) demonstrate that the results of Becker et al. (2005) depend on the chosen reference value of life expectancy. Reference level dependency is not necessarily a limitation. This is because reference values provide policymakers with the flexibility to accommodate different ethical principles (Fleurbaey, 2016). The approach can also be expanded to incorporate individual specific reference points (ibid). Note, however, that difficult choices such as these arise within most measures of wellbeing. For instance, a similar reference dilemma is faced when choosing the minimum and maximum levels for each dimension within composite indices.

Lastly, there may be reasons why preferences are not good indicators of wellbeing. For instance, individuals may have false beliefs about the contribution of different dimensions to their own wellbeing (Hausman, 2012). Their preferences may also be malleable and easily influenced by subtle but irrelevant factors, such as the ordering of questions within a survey. Another strand of the literature argues that true underlying preferences may not exist at all and that individuals instead construct their preferences on the fly (Lichtenstein and Slovic, 2006), an issue we will return to in later sections.

RECONCILING THE TWO APPROACHES

It is important to recognise that many preference-based measures are special cases of composite indices, with a functional form based on standard representations of preferences from welfare economics (see Decancq et al. 2009).[15] The main difference between the two approaches is that preference-based measures are guided and constrained by a substantive theory of wellbeing. Meanwhile composite indices, traditionally understood, are only limited by the availability of data and information on relative weights. Although some composite indices allude to an underlying theoretical framework, the connections with theories of wellbeing are often loose.[16]

Another point of departure is the assessment unit. Preference-based measures, such as equivalent income, focus on measuring wellbeing at the individual level. Many composite indices, however, have followed the traditional path of the HDI by aggregating wellbeing dimensions measured at the country or regional level. This is problematic because such approaches overlook correlations between dimensions of wellbeing. To demonstrate this, let us return to our hypothetical society once again and label it society X. Now let us introduce a second society, labelled Y, in which individual B is now worse off across both dimensions. Consequently, outcomes in society Y are more correlated across individuals. We also add a final row in each table denoting average income and health.

Table 4 – Correlations between dimensions

Society X

Individual	Income	Health
A	50,000	70
B	40,000	100
Average	45,000	85

Society Y

Individual	Income	Health
A	50,000	100
B	40,000	70
Average	45,000	85

[15] This is also the case for GDP where the relative weights of different goods are provided by the prices

[16] For example, the HDI was intended to capture the notion of Sen's capabilities approach. However, as Ravallion (2012b) highlights, it is unclear how this link is made, given that GDP per capita would be considered a functioning rather than a capability.

Unlike our first example, which aggregated at the individual level, let's follow the traditional HDI like approach and aggregate by average outcomes in each society. Since both societies have the same average income and health, their composite index values are identical, in this case 0.65. Yet, society Y is inherently more unequal than society X since individual B is worse off in both dimensions.[17] However, this correlation is overlooked by the aggregation process. The key take-away here is that if one wants to introduce a concern for underlying correlations between dimensions, one must aggregate first at the individual level, using a composite index or equivalent income, and then aggregate again to capture a concern for inequality. Nevertheless, there have been very few applications of composite indices at the individual level, as proposed in Table 1.

More recently, there has been a move within the field of composite indices towards incorporating individual preferences within the aggregation process. For instance, some researchers have utilised discrete choice experiments to elicit the weights respondents give to dimensions of the HDI, usually finding that these diverge from the assumption of equality (Deyshappriya and Feeny, 2021; McGillivray et al., 2023). Others have utilised life satisfaction data, budget allocation methods or importance scores to similar ends (Yang, 2018; Esposito and Chiappero-Martinetti, 2019). The OECD Better Life Index is an example of the latter, which allows individuals to select weighting schemes for 11 different dimensions of wellbeing via an interactive web application. Nevertheless, by focusing solely on the weights, all of these approaches fail to address the broader problem of the implicit trade-offs contained within the indices.[18]

Benjamin et al. (2014) propose a two-step method for creating a personal wellbeing index that addresses some of the issues highlighted above. In the first step (rating task), respondents indicate their levels of different wellbeing dimensions on a scale from 0 to 100. In the second step (trade-off task), respondents make choices between different options, which can be used to infer the relative weights given to each dimension. The personal wellbeing index is then constructed by taking a weighted sum of the dimensions.[19] By eliciting these indices at the individual level, the index captures correlations between dimensions. In addition, it sidesteps the issue of how to scale each dimension by relying on self-reported ratings. Yet, the approach faces several challenges, as outlined by the authors (see Benjamin et al., 2017). The key issue is that there may be large differences in scale use across individuals and over time when answering the rating question. For instance, an individual may use the scale differently as time goes by to deal with ceiling or floor effects. While the authors propose the use of calibration questions to capture differences in scale use, the effectiveness of such approaches is an area of ongoing research.

Decancq (2025, forthcoming) proposes an individual level HDI, which incorporates preferences over income, health and education. He sets out a framework for eliciting the measure using a stated-preference experiment by asking respondents to compare their lives relative to different hypothetical lives that are balanced across all dimensions. Their choices are used to estimate personalised wellbeing levels, which can then be aggregated using a social welfare function.

CHALLENGES FOR USING PREFERENCE-BASED MEASURES IN POLICY-MAKING

As discussed above, the equivalent income can be viewed as a composite indicator that aggregates dimensions based on the preferences of individuals, using income as a numeraire. By measuring wellbeing at the individual level it allows us to capture correlations between dimensions that may arise. Yet, moving to the individual level requires additional data on their outcomes and preferences. There are two key challenges for the implementation of preference-based measures in policy making: i) preference elicitation and ii) constructed preferences. We now discuss each of these in turn.

PREFERENCE ELICITATION

Preference-based measures are data-intensive, requiring the estimation of trade-offs people make between different dimensions of life. Several methods are available in this regard. One option is to recover preferences from the choices people make in the

[17] Note that the inequality adjusted version of the HDI does not avoid this issue, given that it still aggregates by dimension.

[18] An exception is provided by Decancq and Watson (2019) who elicit trade-offs within the HDI using a discrete choice experiment. Da Costa and Decancq (2025, forthcoming) utilise a similar experiment in Tanzania. Yang (2018) also estimates an underlying utility function from which trade-offs can be derived.

[19] Note that this index mimics the construction of GDP, which relies on quantities (self-reported ratings) and prices (i.e., weights from trade-offs).

market or real-life settings (revealed preference, RP). For instance, the value of life expectancy gains can be inferred from the trade-offs people make in the labour market between income and mortality risk, known as the value of statistical life. Becker et al. (2005) and others utilize this information to construct their longevity adjusted income measures. Moreover, such trade-offs are regularly employed within EU policymaking, for example, within cost benefit analyses related to regulation of chemicals and their impacts on human health (see Georgiou et al., 2018). Still, many dimensions of wellbeing are not traded directly or indirectly within the market, limiting the use of RP methods. In addition, the elicitation of preferences using RP techniques relies on several strong assumptions, principally that individuals are fully informed about the consequences of their choices and there are no market constraints.

To address these challenges, welfare economists are increasingly turning to alternative methods to elicit preferences. Stated preference (SP) methods recover trade-offs from hypothetical choices, usually within a survey or experimental setting. They can therefore be applied to almost any context, a key advantage with respect to RP methods. In recent years, several studies have elicited equivalent incomes using contingent valuation questions (e.g., Samson et al 2018, Capéau et al., 2020). Typically, these ask respondents directly how much they would be willing to pay to obtain a reference level of a non-monetary dimension, e.g., perfect health. Dichotomous choice variations of the approach have also been in which respondents choose between hypothetical lives until a point of indifference is reached (see Decancq and Nys, 2021). Nonetheless, such methods have traditionally been met with scepticism from economists. Put simply, there are many reasons what people say they will do would differ from their behaviour in the actual situation. Moreover, SP surveys are costly to run and responses may be influenced by subtle aspects of the survey design, leading to context dependence.

More recently, economists have been exploring whether variations in responses to simple questions about life satisfaction can be used to infer trade-offs between different dimensions. An advantage of this approach is that life satisfaction data is easy to collect and readily available across most EU countries, for example through EU-SILC. Decancq and Gamage (2023), for instance, exploit this approach to estimate preferences at the country level across the EU. Nevertheless, the approach has several limitations. First, the estimation of trade-offs relies on assumption that life satisfaction scores are a viable representation of individual's underlying preferences, which is often untestable with available data. Second, people tend to interpret and use response scales in different ways when answering life satisfaction questions. While this can be controlled to some extent (e.g., through anchoring vignettes)[20] it often poses difficulties for the estimation of trade-offs between dimensions.[21] Lastly, since it relies on regression analysis, the approach can only capture preference heterogeneity at the group level through the inclusion of interaction terms for different socio-demographic groups (e.g., sex, age, location, etc.)

CONSTRUCTED PREFERENCES

The field of behavioural economics introduces a significant challenge for eliciting equivalent incomes. Accumulating evidence suggest that respondents may not have well-defined preferences from the outset that they can access when answering a survey question. Instead, they might construct their preferences as they go through a survey, especially for goods with which they have little experience of trading in the marketplace (Lichtenstein and Slovic, 2006). Additionally, this constructive process seems to be highly susceptible to contextual cues within the survey, such as anchoring points or the framing of a question (ibid). Consequently, the trade-offs elicited from RP and SP methods may be context-dependent.

Still, the literature does not take the stance that *all* preferences are constructed. On the contrary, there may exist a continuum along which the degree of preference construction lies. For instance, individuals are likely to have well-defined preferences for choices that are familiar, simple and experienced regularly. Difficulties arise, however, when choices are more complex and hypothetical, such as those contained within stated preference surveys.[22]

[20] This approach rescales individual's self-assessments based on their responses to anchoring vignettes (King and Wand, 2007). Variation in the latter allows one to capture differences in scale use as long as: i) all respondents use the response scale in the same way when assessing their own situation and the hypothetical vignettes (consistency); and ii) all respondents perceive the variables contained within the vignette in the same way (equivalence).

[21] There are various statistical issues related to the estimation of life satisfaction regression models. For a more detailed discussion, see Fujiwara and Campbell (2011).

[22] In such situations, individuals responses may reflect an aggregation process of their beliefs about the properties and consequences of difference alternatives as well as the underlying values – a process which is highly susceptible to contextual cues (Bernheim, 2016; Hausman, 2012).

At the more fundamental level, the constructive preferences literature raises the issues of whether preferences should be trusted at all, and if so, under what conditions. For instance, some argue that economists should abandon the concept of preferences altogether as a basis for wellbeing measurement given their malleability (Sugden, 2018). Others hold that preferences can serve as indicators of wellbeing when individuals act in their self-interest and hold correct beliefs and full information about the consequences of their choices (Hausman, 2012, Bernheim, 2021). If these conditions are satisfied, context-dependence can then be accommodated through variations of preference based measures that capture uncertainty in respondents' trade-offs (see Fleurbaey and Schokkaert, 2013).[23]

FUTURE AVENUES FOR POLICY USE

Moving beyond GDP and constructing broader wellbeing measures is an important challenge for the researchers and policymakers. This brief has covered some of the pros and cons of two different measures of wellbeing: composite indices and preference based measures. It has shown that, while these measures share similarities, they have moved in different methodological directions over the years.

Nevertheless, the two strands of the literature could be reconciled by considering hybrid approaches that offer the simplicity of composite indicators alongside the more nuanced understanding of wellbeing provided by preference-based measures. Existing composite indicators of wellbeing could, for instance, be extended i) to measure multidimensional wellbeing at the individual level thereby capturing correlations between dimensions and ii) consider the preferences of the population when constructing the individual wellbeing function (Benjamin et al, 2014; Fleurbaey, 2018). This would result in measures that are comprehensive and capture what individuals value the most. For instance, studies have shown that measures of well-being incorporating preferences yield very different distributions of well-being relative to other commonly used indicators, such as income per capita and life satisfaction (see, for example, Decancq et al, 2015a). Of course, the equivalent income approach is one type of composite indicator that already fulfils these criteria while being rooted in welfare economic theory.

Moving in this direction would require investments in data infrastructure to measure preferences and well-being at the individual level. While information on the latter is readily available to some extent across the EU (e.g., SILC), there are no dedicated surveys assessing preferences over different dimensions of wellbeing at the individual level. Moreover, as discussed above, efforts are still ongoing to determine which form an adequate preference elicitation survey should take and how behavioural factors, such as constructed preferences, can be accounted for. Future work in this direction requires the testing of various aspects of survey design and elicitation methods.

This note has shown that some first steps could be taken with the data at hand. Despite its limitations, life satisfaction data could offer some insights on preferences across the EU at the country level (see Decancq and Gamage, 2023). Revealed preference studies also offer considerable evidence on trade-offs for some dimensions of life, such as life expectancy. Such information has already been used to construct equivalent incomes at the EU-level (see Benczur et al., 2023). Nonetheless, such data only allows us to apply a broad-brushed approach to measuring wellbeing at the country level. It would need to be extended to incorporate differences between individuals on what matters for wellbeing, which, as this brief has highlighted, could be a core determinant of wellbeing.

REFERENCES

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, 277–291.

Benczúr, P., Kvedaras, V., and Preziosi, N. (2023). *Health-adjusted income: complementing GDP to reflect the valuation of life expectancy*. Luxembourg: Publications Office of the European Union.

Benjamin, D. J., Cooper, K. B., Heffetz, O., and Kimball, M. S. (2017). *Challenges in Constructing a Survey-Based Wellbeing Index.* American Economic Review 107 (5): 81–85.

Benjamin, D. J., Heffetz, O., Kimball, M. S., and Szembrot, N. (2014). *Beyond Happiness and Satisfaction: Toward Wellbeing Indices Based on Stated Preference.* American Economic Review 104 (9): 2698–735.

[23] These issues are at the core of a lively ongoing debate among welfare economists and philosophers, alike. If one accepts that preferences can serve as a basis for wellbeing measurement the task for researchers is to identify the best possible context or conditions in which to elicit them. More precisely, economists could view themselves as 'architects', helping respondents to construct defensible preferences, rather than 'archaeologists' seeking to uncover some underlying preference ordering (Gregory et al., 1993).

Bernheim, B. D. (2021): *In defense of behavioral welfare economics*. Journal of Economic Methodology, 28, 385–400.

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.

Capéau, B., Cherchye, L., Decancq, K., Decoster, A., Rock, B. D., Maniquet, F., Nys, A., Périlleux, G., Ramaekers, E., and Rongé, Z. (2020). *Who Has the Lowest Levels of Well-Being?* Well-being in Belgium: Beyond Happiness and Income, 175–181.

d'Albis, H., and Bonnet, F. (2018). *Inequalities in life expectancy and the global welfare convergence.* Economics Letters, 168, 49-51.

de Domincis, L., Cabeza Martinez, B., Kovacic, M., and Papadimitriou, E. (2024). *The EU regional Social Progress Index 2.0, 2024 Edition.* Directorate-General for Regional and Urban Policy Working Paper WP01/2024.

Da Costa, S., K. Decancq, M. Fleurbaey, and E. Schokkaert (2024). *Preference Elicitation Methods and Equivalent Income: An Overview*. PSE Working Paper 2024-60. https://shs.hal.science/halshs-04840652v1

Da Costa, S. and Decancq, K. (2025, forthcoming). *Trade-offs within the Human Development Index: evidence from a discrete choice experiment in Tanzania.* CSB Working Paper.

Decanq, K. (2025, forthcoming). *Towards a Micro-Founded Framework for Measuring Human Development.* UNDP Discussion Paper.

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

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

Decancq, K., and P Gamage, S. (2023). *Well-being convergence in the European Union*. EuSocialCit working paper.

Decancq, K., Decoster, A., and Schokkaert, E. (2009). *The evolution of world inequality in well-being.* World Development, 37(1), 11-25.

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

Decancq, K., Fleurbaey, M., and Schokkaert, E. (2015b). *Inequality, income, and well-being.* Handbook of income distribution. Vol. 2. Elsevier, 2015. 67-140.

Deyshappriya, N. R., and Feeny, S. (2021). *Weighting the dimensions of the multidimensional poverty index: findings from Sri Lanka*. Social Indicators Research, 156, 1-19.

Esposito, L., and Chiappero-Martinetti, E. (2019). *Eliciting, applying and exploring multidimensional welfare weights: Evidence from the field.* Review of Income and Wealth, 65, S204-S227.

European Commission (2023). *The 2023 Strategic Foresight Report: Sustainability and people's wellbeing at the heart of Europe's Open Strategic Autonomy*. COM(2023) 376 final, https://commission.europa.eu/strategy-and-policy/strategic-planning/strategic-foresight/2023-strategic-foresight-report_en

Fleurbaey, M. (2009). *Beyond GDP: The Quest for a Measure of Social Welfare.* Journal of Economic Literature, 47 (4): 1029–75.

Fleurbaey, M. (2016). *Equivalent income.* In Matthew D. Adler, and Marc Fleurbaey (eds.), "The Oxford Handbook of Well-Being and Public Policy", Oxford Handbooks.

Fleurbaey, M. (2018). *On human development indicators.* United Nations Development Programme Human Development Report Discussion Paper.

Fleurbaey, M. and Gaulier, G. (2009). *International comparisons of living standards by equivalent incomes.* Scandinavian Journal of Economics, 111, 597–624.

Fleurbaey, M., and Blanchet, D. (2013). *Beyond GDP: Measuring welfare and assessing sustainability.* Oxford University Press.

Fleurbaey, M., and Schokkaert, E. (2013). *Behavioral welfare economics and redistribution.* American Economic Journal: Microeconomics, 5(3), 180-205.

Fujiwara, D. and Campbell, R. (2011). *Valuation Techniques for Social Cost-Benefit Analysis: Stated Preference, Revealed Preference and Subjective Well-Being Approaches. A Discussion of the Current Issues.* HM Treasury and Department for Work and Pensions.

Georgiou, S., Rheinberger, C. M., and Vainio, M. (2018). *Benefit-cost analysis in EU chemicals legislation: Experiences from over 100 REACH applications for authorisation.* Journal of Benefit-Cost Analysis, 9(1), 181-204.

Gregory, R., Lichtenstein, S., and P. Slovic (1993). *Valuing environmental resources: a constructive approach*, Journal of Risk and Uncertainty, 7, 177–197Hausman, D. M. (2012). *Preference, value, choice, and welfare.* Cambridge University Press.

Hausman, D. M. (2015). *Valuing health: Well-being, freedom, and suffering.* Oxford University Press.

Jones, C. I., and Klenow, P. J. (201). *Beyond GDP? Welfare across countries and time.* American Economic Review, 106, 2426–57.

King, G., and Wand, J. (2007). *Comparing Incomparable Survey Responses: New Tools for Anchoring Vignettes.* Political Analysis, 15, 46-66.

Lichtenstein, S. and Slovic, P. (2006). *The construction of Preference.* Cambridge University Press.

McGillivray, M., Feeny, S., Hansen, P., Knowles, S., and Ombler, F. (2023). *What are valid weights for the human development index? A discrete choice experiment for the United Kingdom.* Social Indicators Research, 165(2), 679-694.

Nardo, M., Saisana, M., Saltelli, A., Tarantola, S., Hoffman, A., and Giovannini, E. (2005). *Handbook on constructing composite indicators: methodology and user guide.*

Nordhaus, W.D. (2003). *The Health of Nations: The Contribution of Improved Health to Living Standards.* In: Measuring the gains from medical research: an economic approach. Ed. by Murphy, K. M., and Topel, R. H. University of Chicago Press.

Ravallion, M. (2012a). *Troubling tradeoffs in the human development index*. Journal of Development Economics, 99(2), 201-209.

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

Samson, A.-L., Schokkaert, E., Thébaut, C., Dormont, B., Fleurbaey, M., Luchini, S. and Van de Voorde, C. (2018). *Fairness in cost-benefit analysis: A methodology for health technology assessment.* Health economics, 27, 102–114.

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

Stiglitz, J., Sen, A., and Fitoussi, J-P. (2009). *The measurement of economic performance and social progress revisited.* Reflections and overview. Commission on the Measurement of Economic Performance and Social Progress, Paris.

Sugden, R. (2018). *The community of advantage: A behavioural economist's defence of the market.* Oxford University Press.

UNDP (1990). *Human Development Report.* Tech. rep., United Nations Development Programme, New York

Usher, D. (1973). *An imputation to the measure of economic growth for changes in life expectancy.* In: The measurement of economic and social performance, ed. by M. Moss, NBER, 193–232.

Yang, L. (2018). *Measuring well-being: A multidimensional index integrating subjective well-being and preferences.* Journal of Human Development and Capabilities, 19(4), 456–476.

ACKNOWLEDGEMENTS

I would like to thank Begoña Cabeza Martínez, Peter Benczur and Lorenzo Pirelli for their useful comments and suggestions on previous versions of this brief.

AUTHORS OF THE REPORT

Shaun M. Da Costa

DISCLAIMER

The views expressed are purely those of the author and may not in any circumstances be regarded as stating an official position of the European Commission.

COPYRIGHT

CONTACT INFORMATION

Shaun-Mark.Da-Costa@ec.europa.eu

EU Science Hub — joint-research-centre.ec.europa.eu

JRC141107

PDF ISBN 978-92-68-25015-0 doi:10.2760/9276064 KJ-01-25-112-EN-N

Publications Office of the European Union