DISCUSSION PAPER SERIES

IZA

IZA DP No. 8589

Does the Choice of Well-Being Measure Matter Empirically? An Illustration with German Data

Koen Decancq
Dirk Neumann

October 2014

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Does the Choice of Well-Being Measure Matter Empirically? An Illustration with German Data

Koen Decancq
CSB, University Antwerp, CES, KULeuven,
CORE, Université catholique de Louvain, ZEW and FMSH

Dirk Neumann
CORE, Université catholique de Louvain,
IZA and ZEW

Discussion Paper No. 8589
October 2014

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 8589
October 2014

ABSTRACT

Does the Choice of Well-Being Measure Matter Empirically? An Illustration with German Data*

We discuss and compare five measures of individual well-being, namely income, an objective composite well-being index, a measure of subjective well-being, equivalent income, and a well-being measure based on the von Neumann-Morgenstern utilities of the individuals. After examining the information requirements of these measures, we illustrate their implementation using data from the German Socio-Economic Panel (SOEP) for 2010. We find sizeable differences in the characteristics of the individuals identified as worst off according to the different well-being measures. Less than 1% of the individuals belong to the bottom decile according to all five measures. Moreover, the measures lead to considerably different well-being rankings of the individuals. These findings highlight the importance of the choice of well- being measure for policy making.

JEL Classification: D31, D63, I30

Keywords: income, composite well-being index, life satisfaction, equivalent income, von Neumann-Morgenstern utility function, worst off, Germany

Corresponding author:

Dirk Neumann
CORE
Voie du Roman Pays 34
bte. L1.03.01
1348 Louvain-la-Neuve
Belgium
E-mail: dirk.neumann@uclouvain.be

* Draft prepared for the *Oxford Handbook of Well-Being and Public Policy*, edited by Matthew D. Adler and Marc Fleurbaey. We thank the editors, Sabina Alkire, Kristof Bosmans, Tim Goedemé, François Maniquet, Alexander Michiels, Andrew Oswald, Andreas Peichl, Erik Schokkaert, and audi- ences in Antwerp, Boston, Ispra, Leuven, Louvain-la-Neuve, Paris, and Princeton for helpful discussions and suggestions. This research is part of the EquiRisk project supported by the Agence Nationale de la Recherche (ANR-12-INEG-0006-01). Neumann acknowledges support from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement n◦ 269831.

1 Introduction

A central question in the design of social and public policies is how to compare the well-being of different beneficiaries and how to identify the worst off. Indeed, the implementation of redistributive policies such as poverty alleviation programs and targeted social benefit schemes requires a method to make interpersonal comparisons of well-being. In practice, these comparisons rely almost exclusively on income as measure of individual well-being. The worst off are then identified as the persons with the lowest incomes.

This standard approach is attractive from a pragmatic and practical viewpoint, as the analyst only needs information on the income position of the individuals. Yet, many authors have pointed at the dark side of the practical appeal of income-based well-being comparisons: income alone may be a too narrow information basis to measure well-being attractively (see Stiglitz et al. (2010) for an extensive survey). Income-based well-being comparisons assume that all individuals with the same income can be considered equally well-off, irrespective of their non-monetary outcomes such as their health, educational achievements, labor market status, and so on. It is by now well-documented in the literature on subjective well-being that people care about these non-monetary dimensions of their life (see Benjamin et al. (2012), Clark (frthc.), and Graham (frthc.)). Various alternative approaches have therefore been proposed to perform well-being comparisons with a broader and richer information basis (see Decancq et al. (frthc.) for a recent overview).

A first alternative approach is to use an objective composite well-being index to aggregate the outcomes of the individuals in all different dimensions that constitute well-being (see Hurka (frthc.) and Arneson (frthc.)). A popular example of a composite well-being index is the Human Development Index (HDI). Similarly, most multidimensional poverty measures are built on an objective view on how to aggregate across the different dimensions of poverty (for a discussion of multidimensional poverty measurement, see Alkire (frthc.); Chakravarty and Lugo (frthc.); Duclos and Tiberti (frthc.); and Pogge and Wisor (frthc.)).

A second alternative is based on the idea that well-being comparisons should be based on the subjective well-being scores reported by the concerned persons themselves. These well-being scores can be obtained by asking respondents directly how well-off, happy or satisfied with their life they are (see Clark (frthc.) and Dolan and Fujiwara (frthc.)). In this approach, the problems of the selection of the relevant dimensions and the aggregation procedure are left entirely to the respondents.

A third alternative uses the individual preferences, i.e., the opinions on the trade-offs between the different aspects of well-being, to construct an interpersonally comparable measure of well-being. An example of such a well-being measure is the so-called equivalent income measure (see Fleurbaey (frthc.); Maniquet (frthc.); as well as Fleurbaey and Maniquet (2011)). Preference-based well-being measures require additional information on the ordinal preferences or indifference map of the concerned individuals, which typically is not readily available in existing data sets. As we will see below, indifference maps need therefore to be estimated.

A final alternative considered here is to use the so-called von Neumann-Morgenstern (vNM) utility function of the concerned individual to make well-being comparisons. As was shown by von Neumann and Morgenstern (1947), under some rationality conditions, the preferences of the individuals over risky outcomes or lotteries can be represented by the expected value of a vNM utility function. The idea of using this vNM utility function to make interpersonal well-being comparisons goes back to Harsanyi (1953, 1955) and has recently been advocated in the extended preferences approach of Adler (frthc.). Also this vNM utility function needs to be estimated or calibrated.

We compare these four alternative well-being measures to the standard approach of using income exclusively. We focus thereby on various issues that arise when operationalizing the measures with a single data set, namely the German Socio-Economic Panel (SOEP) of 2010. The SOEP is a widely used data set and contains detailed information on various aspects of life for a representative sample of about 20,000 German citizens. We compare the well-being measures in terms of the characteristics of the individuals in the data set who belong to the bottom decile and can thus be considered as worst off. We find substantial differences between the characteristics of the people at the bottom of the well-being distribution according to the different measures. In addition, we find a low degree of overlap between the groups of worst-off individuals according to the different well-being measures and a considerable amount of re-ranking.

Our empirical results confirm findings of other studies which are focusing on developing or transition countries. Laderchi et al. (2003), for instance, compare four different underlying well-being measures when measuring poverty: a standard monetary one, a measure based on the capability approach, a measure based on the notion of social exclusion and, finally, a measure based on participatory approaches. The authors find a significant lack of overlap between the approaches when measuring poverty in Peru and India. Fleurbaey et al. (2009) compare the bottom of the well-being distribution in Russia according to income, life satisfaction and equivalent income and also find low overlap between the different groups.

The paper is structured as follows. Section 2 introduces the well-being measures and discusses the required information to implement them empirically. Section 3 illustrates how the measures can be implemented based on an existing household survey, namely the

SOEP of 2010. In Section 4 we compare the characteristics of the worst off, quantify the degree of overlap between the worst off according to the different measures, and illustrate the extent to which the ranking of the individuals differs. Section 5 concludes.

2 Five well-being measures and their information requirements

In this section we discuss five operational well-being measures and their information requirements. To do so, it is useful to make a distinction between four different pieces of information that these measures can use as input.

First, let ℓ_i denote the vector of life aspects or attributes. This vector provides a full description of the life of individual i covering the diversity of outcomes, activities and possibilities enjoyed by the individual. In practice, however, the number of aspects of life that can actually be included in an empirical analysis of well-being depends on the data set at hand. In the next section we will illustrate that even with a state-of-the-art and rich data set such as the SOEP, the data limitations constrain the scope of empirical multidimensional analysis considerably.

Second, we assume that individuals have a well-informed opinion on what is a good life. We represent this opinion for each individual i by an ordinal preference ordering $\mathcal{R}_i$ over the vectors of life aspects ℓ_i. If individual i weakly prefers the life described by ℓ_i to the life described by ℓ_i', we write $\ell_i \mathcal{R}_i \ell_i'$. We denote strict preference by writing $\ell_i \mathcal{P}_i \ell_i'$, and use $\ell_i \mathcal{I}_i \ell_i'$ to refer to indifference between ℓ_i and ℓ_i'. It is important to stress from the outset that the preference ordering that we are considering here captures the well-informed opinions on a good life of the individuals. This notion of a preference ordering does not necessarily coincide with the preferences of the individuals that are revealed by their behavior and choices, as these may incorporate information limitations and mistakes (see, e.g., Hausman (2012)).

Third, individuals are equipped with a so-called von Neumann–Morgenstern (vNM) utility function, V_i. Such a *cardinal* vNM utility function V_i contains more information about the individual than her *ordinal* preference ordering $\mathcal{R}_i$, since the concavity or curvature of the vNM utility function embeds additional information on her attitude towards risk. We will see later that different approaches to well-being measurement disagree on whether this additional cardinal information should be taken into account or not.

Finally, individuals are assumed to be able to attach a particular satisfaction score s_i to their vectors of life aspects ℓ_i. The satisfaction function S_i is the individual-specific function that connects a satisfaction score to each ℓ_i, so that $s_i = S_i(\ell_i)$. The satisfaction function may—but is not required to—represent the ordinal preference ordering of the individual.

These four different pieces of well-being information are contained in the quadruple $(\ell_i, \mathcal{R}_i, V_i, S_i)$. A well-being measure WB compares the quadruples across individuals and attaches a numerical value to each of them. The numerical value $WB(\ell_i, \mathcal{R}_i, V_i, S_i)$ can then be interpreted as a well-being measure of individual i with vector of life aspects ℓ_i, preference ordering $\mathcal{R}_i$, vNM utility function V_i and satisfaction function S_i. Using such a well-being measure for all individuals, the worst off can be identified by simply looking at the bottom of the well-being distribution. Not all well-being measures use all the information available in the quadruple, and—as we will see below—different measures make different use of the same information. These differences reflect the inherently normative nature of well-being comparisons. In the remainder of this section we will compare five different well-being measures in detail.

2.1 Income

We assume that one of the outcomes in ℓ_i captures the command over material resources of the individual. We will refer to that dimension as "income" and denote it y_i, so that $\ell_i = (y_i, x_i)$, where the vector $x_i = (x_i^1, \ldots, x_i^m)$ contains all the m non-income aspects of life. Our first method for interpersonal comparisons takes income y_i as the only relevant measure of well-being. That is, the well-being of individual i is measured by her income:

$$WB^1(\ell_i, \mathcal{R}_i, V_i, S_i) = y_i. \tag{1}$$

We use the label "income" rather loosely as a placeholder for any interpersonal comparable measure capturing a person's command over material resources. The well-being comparisons implicit in most of the literature on poverty and inequality measurement, and in the design of (targeted) social policies are almost exclusively based on income in one of its empirical manifestations such as measures of wealth, lifetime income, disposable income or expenditures. Using disposable income, i.e., income after taxes and benefits, has become standard practice in European poverty and inequality measurement, for instance. Each of these manifestations raises its own practical measurement problems (see Cowell (2011) for a discussion).

A particularly pervasive measurement problem is that most operational income measures are collected at the household level, whereas one would like to make comparisons of well-being between individuals (see Lewbel and Pendakur (2008) and Chiappori (frthc.)). Two (heroic) assumptions are needed to move from the household level to the individual one. First, an assumption is usually made on how incomes are distributed within the household. Perfect income pooling is the standard assumption, although it is often challenged by empirical evidence (see Lundberg et al. (1997), for instance). Second, an assumption is needed on the extent to which resources held by the family can be commonly used by its members, which generates economies of scale of living together. In practice, the conversion from income information observed at the household level to the individual level is made by first dividing household incomes by an equivalence scale that is a concave function of the household size and then by attributing the resulting "equivalized incomes" equally to all members of the household.

Besides the measurement problems to collect the required information for this well-being measure, another—more fundamental—objection has been raised. Using income as measure of well-being assumes that all individuals with the same command of resources are necessarily equally well-off. This position reduces the vector of all aspects of life $\ell_i = (y_i, x_i)$ to a single measure of command over material resources y_i, which implies that all other aspects of life are deemed as irrelevant for well-being comparisons. Yet, people may differ in how they manage to convert their resources into whatever they care about in their life. Sen (1985), amongst others, has argued that income may indeed be used as a means to generate well-being, but that reducing well-being to command over material resources mixes up means and ends. Sen labelled this reductionist perspective as *resource fetishism*. Neglecting the non-income components of ℓ_i, even as an operational first-order approximation of well-being, may therefore be far from harmless.

2.2 Composite index of well-being

A natural alternative to measuring well-being by looking at income alone, consists in incorporating the other relevant aspects of life explicitly in the well-being analysis. By doing so, the well-being analysis becomes multidimensional. There is a recent and growing literature on multidimensional well-being and poverty measurement (see Weymark (2006) and Chakravarty and Lugo (frthc.), for surveys).

A popular way to make multidimensional well-being comparisons is by constructing a so-called "dashboard" of separate measures without aggregating them (see Ravallion (2011) for a discussion). This procedure has some intuitive appeal as it allows to treat the dimensions of life as incommensurable. Yet, these dimension-by-dimension comparisons of well-being remain blind to the correlation between the dimensions. Moreover, well-being comparisons between two individuals can only be robustly made when one individual is worse off in all dimensions of life, which we denote $\ell_i \ll \ell_j$. In reality, these cases of so-called *vector dominance* are likely to be rare, especially when many dimensions are included in the analysis. We consider this procedure therefore as a non-starter in our quest for an operational multidimensional well-being measure that can be used for policy making.

If we want to obtain a measure that is able to rank all persons with respect to their well-being, we have to specify how to trade-off a better score in one aspect of life (say income) for a worse score in another aspect (say health). In other words, we have to select a specific way to aggregate all the different aspects of ℓ_i. A *composite well-being index* is a function CI that does precisely that. It is the second well-being measure that we will discuss:

$$WB^2(\ell_i, \mathcal{R}_i, V_i, S_i) = CI(\ell_i). \tag{2}$$

As long as the composite well-being index CI is non-decreasing in each of its arguments, comparisons based on the composite index will be consistent with the coarse ordering imposed by vector dominance, i.e., if $\ell_i \ll \ell_j$ then we obtain that $CI\ (\ell_i) \leq CI(\ell_j)$. In practice, most composite well-being indices fit in the following general mathematical structure:

$$CI(\ell_i) = \left[w^0 \left(f^0 \left(y_i\right)\right)^\beta + w^1 \left(f^1 \left(x_i^1\right)\right)^\beta + \cdots + w^m \left(f^m \left(x_i^m\right)\right)^\beta\right]^{1/\beta}. \tag{3}$$

This structure reflects the three choices that have to be made to obtain an operational composite index of well-being (see Decancq and Lugo (2013)). First, as the different outcomes are often measured in different measurement units—such as, e.g., income in monetary units and schooling in years—the outcomes need first to be standardized to a common basis before they can be sensibly aggregated. The dimension-specific transformation functions $f^0, f^1, \ldots, f^m$ perform that standardization and, when desired, they may transform the outcomes further to capture decreasing returns to well-being, for instance. When the transformation functions are binary indicators taking 1 whenever the outcome is below an exogenous cut-off and 0 otherwise, CI becomes a counting measure, which is larger when the individuals are deprived in more dimensions of life (see Atkinson (2003) and Alkire and Foster (2011) on the counting approach to multidimensional poverty measurement).

Second, the weighting scheme $w = (w^0, w^1, \ldots, w^m)$ captures the relative weights given to each of the dimensions. Various methods have been proposed to select the most appropriate weighting scheme, ranging from explicitly normative approaches to completely data-driven ones (Decancq and Lugo (2013) provide a critical survey). In most applications the weights are set equally or are determined by statistical or mathematical algorithms such as Principal Component Analysis (PCA) that let "the data speak for themselves". Brandolini (2007, p. 10), however, warns that "we should be cautious in entrusting a mathematical algorithm with a fundamentally normative task".

The parameter β, finally, captures the degree of substitutability between the dimensions. Setting β equal to 1 leads to a simple additive aggregation which is most popular and allows for perfect substitutability between the transformed outcomes. When β is set equal to 1 and the weights are furthermore chosen to be equal, the composite index of well-being becomes an unweighted average of the standardized outcomes in all dimensions. Until 2010, the popular Human Development Index (HDI) followed this simple structure to aggregate average income, life expectancy and educational achievements for each country (see Ravallion (2012) for a critical appraisal of the trade-offs implicit in the HDI).

An important feature of composite indices is that the same index is used to aggregate the outcomes for all individuals. This means that the same opinion on how the different aspects of life should be traded-off is used to evaluate the well-being of all individuals. Various approaches can be followed to obtain a universal opinion on how the dimensions of life should be aggregated and traded-off. First, this opinion may reflect a single objective or perfectionist account of the good life, for which inspiration can be found in the work of Aristotle, for instance (see Hurka (frthc.) for a discussion). Alternatively, the common trade-offs may be determined in a democratic process through public reasoning (as advocated by Sen (2004)). Yet, whenever some disagreement exists on the precise content of the objective theory on the good life, or whenever a consensus cannot be reached on how to aggregate the different aspects of life, the well-being of some individuals is judged according to an opinion on the good life which is not their own. In other words, it can happen that $\ell_i \mathcal{P}_i \ell_i'$ whereas $CI(\ell_i) < CI(\ell_i')$. In these cases, the well-being comparison (and the policies relying on it) are said to be *paternalistic*.

2.3 Subjective well-being measure

An alternative to an objective and universal composite well-being index is to rely on a subjective well-being measure. A simple way to do that is to directly ask the concerned persons how well-off they consider themselves, by asking them about their life satisfaction, for instance. The responses are then immediately used as measure of well-being. In our framework, we say that the well-being of individual i is measured by:

$$WB^3(\ell_i, \mathcal{R}_i, V_i, S_i) = S_i(\ell_i), \tag{4}$$

where S_i is the individual-specific satisfaction function that attaches a particular satisfaction score s_i to each vector of life aspects ℓ_i. This is the third well-being measure that we discuss. Satisfaction scores are relatively easy to collect (see OECD (2013) for a detailed discussion) and hence they are becoming increasingly available for many countries of the world (see Helliwell et al. (2013) for a global comparison of subjective well-being data). A subjective well-being measure offers a practical and easily implementable tool for well-being comparisons. To evaluate its normative appeal, however, it is useful to open the black box of the satisfaction function a bit further.

We decompose the satisfaction function into an answer to two different problems (see also Fleurbaey and Blanchet (2013, p. 175)). First, the individual ranks her life described by ℓ_i relative to one or more reference lives (the ranking problem). Second, she translates this ranking into a verbal statement or numerical value on the scale of the survey (the calibration problem). This calibration may depend on so-called scaling factors, such as the interpretation of the verbal labels in the question, the expectations and aspirations of the respondents, the extent to which they have adapted to their situation, and various other personality traits.

Two important versions of the subjective well-being approach exist, depending on which information is used in the ranking problem. In the first version, the individuals are assumed to consider only their affects (feelings, emotions) when ranking different vectors of life aspects. Affects come in a permanent flow when individuals are awake. This version is the so-called hedonic variant of the subjective well-being approach, which argues that what matters for well-being comparisons are the affects of people (see Haybron (frthc.) for a discussion). It brings us to a modern version of the traditional Benthamite utilitarian calculus of the 18th century that aims to assure "maximum happiness to the maximum number of people" where happiness is understood as reflecting feelings of pleasure and pain. More recently, Layard (2005, p. 121), writes: "Ethical theory should focus on what people feel, rather than what other people think is good for them". In its hedonic variant, the function S_i is only sensitive to one aspect of ℓ_i, i.e., the affects of the individual. Relying exclusively on affects may seem again a rather narrow basis for interpersonal comparisons of well-being (see Nussbaum (2008)).

In the second version, the ranking depends on the cognitive valuations that are formed by the individuals themselves when they take some distance to formulate a judgment over their life. One may wonder whether this version of the subjective well-being approach respects the preferences of the concerned individuals. When satisfaction scores are aligned with the ordinal preferences (i.e., when the so-called *consistency assumption* holds) we have that $S_i(\ell_i) \geq S_i(\ell'_i)$ if and only if $\ell_i \mathcal{R}_i \ell'_i$, so that the subjective well-being approach indeed respects individual preferences when making *intrapersonal* comparisons. However, this result does not extend to the *interpersonal* comparisons that we are interested in. It may very well be the case that both individuals i and j have the same preference ordering and both strictly prefer the situation of individual i, whereas still $S_i(\ell_i) < S_j(\ell_j)$ holds. Indeed, both individuals may use a different calibration of their subjective well-being, so that individual j still reports a higher life satisfaction score even when they both agree that individual i is in a better situation (see also Fleurbaey et al. (2009) for a discussion).

The reason why these reversals may occur is reminiscent of Sen's (1985) argument of *physical condition neglect* against subjective well-being approaches. Person j may be satisfied with her life as she has adapted to her bad living conditions (Sen gives the example of a happy, but battered housewife). Person i, on the contrary, may have some "expensive tastes" which make her less satisfied even in an unanimously preferred situation. Arrow (1973, p. 254) famously refers to a person who has cultivated a taste for "prephylloxera claret and plover's eggs". See Graham (2009) and Di Tella et al. (2010) for some illustrations that these examples of adaptation and expensive tastes may occur in reality and are not mere theoretical curiosa.

2.4 Equivalent income

We have seen that objective composite indices of well-being are paternalistic since they may go against the individual's own preferences on what is a good life. On the other side of the spectrum, we have seen that subjective well-being measures respect preferences in intrapersonal comparisons, but may go against the unanimous preferences of the concerned individuals in interpersonal comparisons. A natural question at this point is whether an approach can be conceived that respects unanimous preferences in interpersonal comparisons.

The *equivalent income* approach is a method that aims to do that. Computing equivalent incomes has a long pedigree in economics (for more details see Fleurbaey (2009), and Fleurbaey (frthc.)). The equivalent income of an individual i is the hypothetical income y_i^* that, if combined with a reference value $\widetilde{x}$ on all non-income dimensions, would place the individual in a situation that she finds equally good as her initial situation. Formally, we say that:

$$WB^4(\ell_i, \mathcal{R}_i, V_i, S_i) = y_i^* \text{ such that } (y_i, x_i)\, \mathcal{I}_i\, (y_i^*, \widetilde{x})\,. \tag{5}$$

This is the fourth measure of well-being considered. The equivalent income measure respects unanimous preferences on what is a good life in interpersonal comparisons. When both individuals i and j have the same preference ordering and strictly prefer the situation

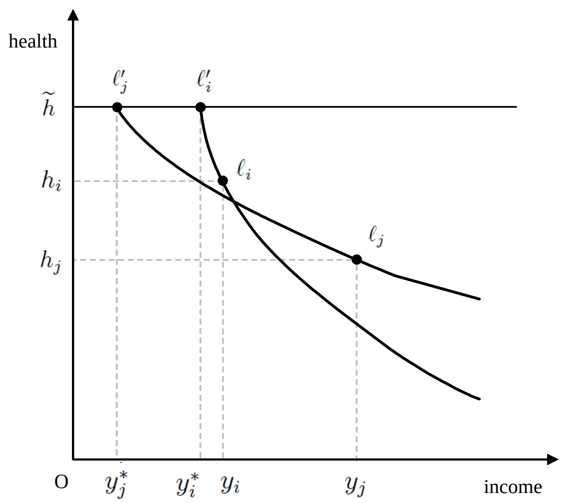

Figure 1: Measuring well-being with an equivalent income approach

of individual i, we obtain indeed that $y_i^* > y_j^*$.

Figure 1 illustrates the equivalent income measure for individuals i and j. For this illustration, we assume that there are only two life aspects: income and health. The vectors of life aspects of both individuals are depicted in the figure by means of ℓ_i and ℓ_j. Individual j is richer, but in worse health than individual i. For each individual, the indifference curve through her actual vector of life aspects is also shown. It is clear that both individuals have different opinions on the good life, as their indifference curves cross. Individual j has a "flatter" indifference curve, which means that individual j cares more about her health. The horizontal line in the graph represents the reference value for health, $\widetilde{h}$, the non-income dimension. The equivalent income of each individual can be obtained as the income level that, if combined with the reference value for health, is considered by the individual herself as equally good as her current vector of life aspects. These bundles are depicted by means of ℓ'_i and ℓ'_j in Figure 1.

From Figure 1 it is clear that the equivalent income depends on the choice of the reference values in the non-income dimensions (indeed, if the reference value was below the crossing of the indifference curves of i and j in Figure 1, their well-being ranking would have been reversed). The choice of the reference values depends on the value judgments of the social observer on how interpersonal comparisons of well-being should be made. Fleurbaey and Blanchet (2013, p. 250) provide some guidelines on how the reference values can be chosen. One observes that for all individuals who obtain the reference value in the non-income dimensions $x_i = \widetilde{x}$, the equivalent income y_i^* equals the actual income y_i and does not depend on the preferences at all (these cases correspond to ℓ'_i and

ℓ_j' in Figure 1). Inverting this reasoning, it follows that the reference values in the non-income dimensions should be chosen so that the well-being of individuals who obtain the reference value can be evaluated by only looking at their income. Consequently, when the non-income dimensions have a value that everybody aspires to—such as being in perfect health, or not being unemployed—it seems natural to take that value as reference value.

In the case where the respondent is not at the reference value in the non-income dimensions, the equivalent income level can be obtained from her actual income level, after adjusting for the gap between the outcomes for the non-income dimensions and the respective reference values. This gap in the non-income dimensions is valued in monetary units by using the personal *willingness-to-pay* (WTP). On the horizontal axis of Figure 1, the WTP of individual i to be in perfect health can be read as the distance between y_i and y_i^*. It is clearly smaller than the WTP of individual j who is in worse health and moreover cares more about health. Despite the use of a monetary measurement unit, the equivalent income method does not suffer from the resource fetishism critique formulated earlier because of the important role that is played by the outcomes in the non-income dimensions. Using a monetary measurement unit offers practical advantages as it allows to use existing results from the literature on income inequality and poverty measurement (see Cowell (frthc.)) when it comes to aggregating the well-being of different individuals in the society.

It is clear that well-being comparisons by means of the equivalent income method are not necessarily consistent with the idea that an individual who scores better on all aspects of life is considered to be better off (consider, for instance, a point on the indifference curve of individual i to the south-west of ℓ_j in Figure 1). In that case, we could have vector dominance $\ell_i \ll \ell_j$ while we still obtain that $y_i^* > y_j^*$. This is a consequence of the desire to respect individual preferences, which is central to the equivalent income method. According to this view, an individual who scores higher on all aspects of life may care more about his relative hardships than another individual who is worse off in all aspects. Choosing between the conflicting principle of respecting preferences, on the one hand, and the idea that vector dominance leads to higher well-being, on the other hand, appears to be one of the hard decisions to be made when evaluating multidimensional well-being (see Brun and Tungodden (2004)). Observers who prefer the latter principle of vector dominance, will have to rely on an objective and universal measure such as a composite well-being index, which, however, does not necessarily respect individual preferences and is therefore paternalistic.

2.5 von Neumann-Morgenstern utility

An equivalent income measure is aligned with the *ordinal* preferences of the individuals, irrespective of the *cardinal* values given by the individuals to their indifference curves. Some authors have argued that a method which only depends on ordinal preferences, does not incorporate enough information about the well-being of the respondents. When discussing so-called "economic quantity indices", the broad family of well-being measures to which equivalent income belongs, Sen (1979, p.11) writes: "A variation of one's intensities of pleasure or welfare cannot, therefore, find any reflection in this numbering system as long as the ordering remains unchanged" (see also Fleurbaey and Blanchet (2013, p.122) for a discussion). Besides information on the intensities of pleasure of the concerned individuals, one may also want to incorporate other cardinal information, such as the individual attitudes towards risk.

The extended preference approach which has been recently proposed by Matthew Adler allows to do that (see Adler (2012, 2014) and Adler (frthc.), for more details). The approach has its origin in Harsanyi's works (1953, 1955, 1977). Extended preferences are preferences of the "deliberator" or social observer k over histories, which are pairs consisting of a vector of life aspects and an individual who obtains this vector of life aspects. We will implement what Adler (frthc.) describes as the "sovereignty respecting" version of the simple case of the extended preference approach. In this case, the deliberator k respects $\mathcal{R}_i$ when evaluating ℓ_i for individual i. When formulating his measure of well-being, the deliberator k respects the preferences of the individuals over risky outcomes or lotteries, which under some rationality conditions can be expectationally represented by a vNM utility function V_i (see von Neumann and Morgenstern (1947)). Next, to obtain an interpersonally comparable well-being measure, we suppose that the deliberator normalizes the individual vNM utilities so that well-being for all individuals equals 0 and 1 in two selected calibration vectors ℓ^+, ℓ^{++}. Adler offers this as one approach to normalizing different vNM functions, but notes that other approaches are also possible. We obtain:

$$WB^5(\ell_i, \mathcal{R}_i, V_i, S_i) = \frac{V_i(\ell_i) - V_i(\ell^+)}{V_i(\ell^{++}) - V_i(\ell^+)}. \tag{6}$$

This measure is the fifth and final measure of well-being that we discuss. To implement it, one needs to know the individual-specific von Neumann-Morgenstern utility function V_i and, moreover, the deliberator needs to select the two calibration vectors ℓ^+ and ℓ^{++} where all individuals have the same well-being, irrespective of their preferences. The choice of the calibration points is again a value judgment. The origin or a bundle of "extreme" poverty seems a natural candidate for the first calibration vector ℓ^+, and, e.g., an "ordinary" poverty bundle or the maximum vector of life aspects can be selected as a second calibration vector ℓ^{++}. We set ℓ^{+} at the origin and ℓ^{++} at the maximum vector of life aspects.

Figure 2 presents a graphical example of the well-being measure based on the vNM utility function. In the left hand-panel, the same individuals i and j of Figure 1 are depicted again, with their vectors of life aspects ℓ_i and ℓ_j and indifference curves, as well as the two calibration vectors ℓ^{+} and ℓ^{++}. The well-being measure WB^5 is shown on the vertical axis in the right hand panel. The horizontal axis of the right hand panel corresponds to the dotted ray between ℓ^{+} and ℓ^{++} in the left hand panel. We depict also the vectors of life aspects ℓ_i' and ℓ_j', which are obtained at the intersection of this dotted line and the indifference curves. The gray curve in the right hand panel depicts the case of a risk-neutral vNM utility function along the ray between ℓ^{+} and ℓ^{++}, and the black curve a risk-averse vNM utility function along the same ray. One easily checks in Figure 2 that individual i is ranked as worse off compared to individual j when both individuals have the same risk aversion ($WB_i^{RA} < WB_j^{RA}$ as well as $WB_i^{RN} < WB_j^{RN}$). Note, however, that this ranking is opposite to the ranking according to the equivalent incomes in Figure 1, where we obtained $y_j^* < y_i^*$, even if exactly the same outcome bundles and ordinal preferences are used. Relying on a different cardinalization of the ordinal preferences, may indeed lead to a re-ranking of the individuals when their indifference curves cross. In the empirical part we illustrate the empirical relevance of this phenomenon. Moreover, assuming that individual i is risk averse and individual j risk neutral, it can be seen from the figure that the ranking of both individuals according to WB^5 is reversed again, with $WB_j^{RN} < WB_i^{RA}$. This observation highlights the important role of risk aversion when making well-being comparisons based on vNM utility functions. For a more extensive discussion about whether risk aversion is relevant to well-being comparisons, see the chapters by Adler (frthc.) and Fleurbaey (frthc.).

Finally, it is clear that preference-based methods such as equivalent incomes and the vNM utility function pose formidable data requirements: for each person in the data set, at least information is needed on the shape of her indifference curve. This information is typically not readily available. In the next section, we discuss various strategies how the necessary preference information can be elicited or estimated.

3 Implementation of the well-being measures with German data

We illustrate how the different measures can be implemented using a single data set, namely the German Socio-Economic Panel (SOEP). The SOEP is a representative house-

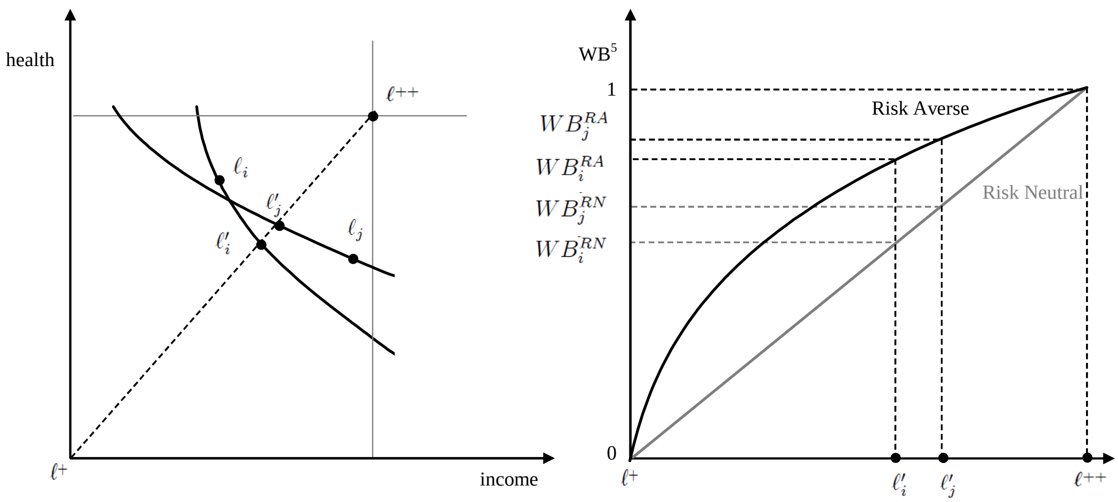

Figure 2: Measuring well-being with the extended preference approach

hold survey of the German population which is conducted yearly since 1984 (see Wagner et al. (2007) for more details). For the empirical illustration, we use the data of 2010 and focus on the subpopulation of individuals who are at least 25 years old and for whom we have data for all necessary variables. This leaves us with a sample of 14,027 individuals, living in 8,657 households. Various studies have used the SOEP to study subjective well-being and its determinants (see, e.g., Ferrer-i-Carbonell (2005), Ferrer-i-Carbonell and Frijters (2004), Frijters et al. (2004b, 2004a), and Layard et al. (2008)).

In this section, we first describe the information contained in the vector of life aspects ℓ_i, which can be readily obtained from the data set, and then we turn to the estimation of the preferences $\mathcal{R}_i$. Finally, we discuss the computation of the five well-being measures introduced in the previous section.

3.1 Observed outcomes in three dimensions of well-being

Given the data limitations imposed by our data set, we assume that the vector of life aspects ℓ_i that describes the life aspects of the individual i has three dimensions: income (y_i), health (h_i), and labor market status (u_i), so that $\ell_i = (y_i, h_i, u_i)$. These three dimensions belong to the set of eight dimensions suggested by Stiglitz et al. (2010). Clearly, it is only a limited set of dimensions of well-being, so that our analysis inevitably remains at a broad-brush level. However, the data set does not permit us to go much further and to include in a satisfactory way other potentially relevant dimensions such as social interactions, political freedom and liberties, safety and security, etc. We do not include education as a dimension of life, even though the data set would allow us to. We return later in this section to this issue.

Income is measured by monthly equivalized disposable income. This income measure is based on the total monthly income of a household after including the received benefits and subtracting the taxes paid. As discussed in the previous section, disposable income is observed at the household level, whereas individual information is desired. To account for household size, we divide total income by the so-called modified OECD equivalence scale, which attaches a weight of 1 to the household head, 0.5 to each additional adult household member and 0.3 to children. The resulting equivalized total income is then equally attributed to all household members, assuming full income pooling. It is important to distinguish between the notion of *equivalent* income, defined in the previous section as a preference-based measure of well-being, and *equivalized* income, which is income corrected for household size by means of an equivalence scale (Decancq et al. (frthc.) discuss the relation between the two notions).

To measure the *health* of the individuals, the SOEP allows two strategies to be followed. First, respondents of the SOEP have been asked to assess their overall health level on a scale between 1 and 5. This so-called self-assessed health is documented to be a good predictor for overall health status or longevity in other surveys (see, e.g., van Doorslaer and Jones (2003)). Yet, similarly to subjective well-being scores, the self-assessed health status of an individual may depend on personality traits and other scaling factors (Jürges (2007)). We opt therefore for a second strategy which involves the construction of an objective health index. The SOEP data set contains a set of indicators of health problems such as disability status, the number of visits to the doctor, presence of constraints in daily and social life through the prevalence of pain, etc. We derive the weights of the indicators in the objective health index from the coefficients of a regression with self-assessed health as dependent variable and the indicators as explanatory variables (van Doorslaer and Jones (2003) provide a comparison of this procedure with other health indices). The index is then normalized to a scale between 0 and 100. Since only health problems are taken into account, the value of 100 can be interpreted as "perfect" health (which is reached by about 9% of the individuals in the considered sample).

Labor market status is measured by a binary variable taking 1 whenever the respondent is unemployed. Labor market status has been found to be an important determinant of subjective well-being (see, e.g., Clark and Oswald (1994), Clark et al. (2001) and Luechinger et al. (2010)). Yet, a binary employment indicator can be argued to be a rather crude measure of the labor market status of an individual (see Schokkaert et al. (2011) for a preference-sensitive multidimensional measure of job quality, for instance). In the considered sample (including pensioners), 6.4% of the sample is registered as unemployed. This number is a bit lower than the official unemployment rate of Germany in

2010, when 7.7% of the total labor force was unemployed. Unemployment is traditionally higher in East (12.0%) than in West Germany (6.6%), but was slightly lower for young people aged 15-25 (6.8%) in 2010 (Bundesagentur für Arbeit (2014)).

3.2 Estimating indifference maps with life satisfaction data

To compute the preference-sensitive well-being measures WB^4 and WB^5 introduced in the previous section, additional information on the preference ordering of the concerned individuals is necessary. In general, three approaches can be followed to retrieve information on preferences (see Decancq et al. (frthc.) and Adler (2013) for a discussion).

First, in a *stated-preference* approach, the respondents are asked directly about the trade-offs between income and the non-income dimensions of well-being, for instance by using contingent valuation techniques. This method has been used mainly in two-dimensional well-being comparisons involving income and health. Fleurbaey et al. (2013) have used the stated-preference approach to calculate equivalent incomes for income-health combinations with specific and tailored survey data in France. In general, stated-preference questions are only rarely asked in standard household surveys and are also not available in the SOEP. One reason is that contingent valuation techniques may be cognitively challenging for the respondents, especially when many dimensions are involved.

A second approach aims at inferring preferences from observed behavior and choices made by the respondents. Consequently, such a *revealed preference* method is only applicable to dimensions over which individuals do actually make choices. Bargain et al. (2013) and Decoster and Haan (frthc.) infer indifference maps from an econometrically estimated labor supply model in which agents choose optimally between various income-leisure combinations under a budget constraint. Revealed preference methods have the disadvantage that possible decision errors will be incorporated in the actual preferences of the individuals. Moreover, this method is not directly applicable on the SOEP for the dimensions of interest, because choices over these dimensions are not observed.

We opt therefore for a third method, namely to estimate the preferences using *self-reported life satisfaction*. There is a tradition in economics of estimating trade-offs between monetary and non-monetary dimensions of well-being based on life satisfaction data (see Clark and Oswald (2002) and the references therein; Dolan and Fujiwara (frthc.) provide a critical discussion). Fleurbaey et al. (2009), Decancq and Schokkaert (2013) and Decancq et al. (2014) provide other examples of using this method to measure well-being, social progress and poverty respectively. Benjamin et al. (frthc.) compare marginal rates of substitution between life satisfaction and revealed preference rankings for residency choices.

Under the consistency assumption, the satisfaction scores contain information on the preference ordering and shape of the indifference curves of the individuals. Yet, as we have seen, the reported satisfaction scores depend also on the individual-specific scaling factors such as the aspiration level and interpretation of the verbal labels in the question, which determine the label or cardinal level attributed to the indifference curves. As these scaling factors may be correlated with the variables of interest such as income and health, the estimation of the effect of these variables on life satisfaction may be biased. Therefore, the effect of the scaling factors must ideally be controlled for. Examples of such scaling factors are the effect of age and education on life satisfaction. The U-shaped effect of age is a widely documented finding in cross-section satisfaction studies and remains significant even after controlling for the usual determinants of life satisfaction such as income, health, marital status and unemployment (see Gwozdz and Sousa-Poza (2010) and Van Landeghem (2012) for a discussion of the age effect on life satisfaction in Germany). In our analysis with the control variables discussed below, we find that German respondents reach their lowest life satisfaction around the age of 47. This pattern is consistent with an intuition that individuals report life satisfaction scores based on a comparison of the realized outcomes with the outcomes to which they aspire. Around the middle of their adult life, respondents may adjust their aspiration level downwards (to a more realistic level) so that afterwards they report higher life satisfaction for the same or even worse vectors of life aspects. Contrarily, higher education may adjust the aspiration level upwards, and hence reduce life satisfaction for the same or even better vectors of life aspects. To filter out this and similar scaling factors, we make use of a regression model.

We start from a satisfaction regression with self-reported life satisfaction s_i as the explained variable and a series of usual explanatory variables, including the three dimensions of life: income y_i, health h_i, unemployment u_i. We also include a vector of observable socio-demographic characteristics z_i including age, education, marital status and regional dummies (at the "Länder" level). In addition, we include three additional control variables in the regression to capture personality traits. Personality traits are generally found to be important determinants of life satisfaction, constituting a large part of individual heterogeneity. If panel data are available, it is therefore a common practice to include individual fixed effects in the regression to control for time-invariant (unobserved) characteristics of the respondents such as their personality traits (see, e.g., Ferrer-i-Carbonell and Frijters (2004) and Frijters et al. (2004a, 2004b) for examples using the SOEP). Yet, there is a price to be paid for the inclusion of these fixed effects as they will control for all the time-invariant characteristics of the individuals, including the time-invariant dimensions that may be of interest. Furthermore, they might bias the estimated effect of variables with low within-person variation (Boyce (2010)). As we work here with a single cross section of the data set we can not include individual fixed effects, but rather use three self-reported personality traits as controls in the regression (Anand et al. (2009) and Boyce (2010) discuss this approach in detail). In 2010, a limited set of questions concerning personality traits was included in the SOEP. We include three aspects, namely the extent to which respondents i) consider to have control over their life, ii) think to have achieved what they deserve and iii) report a positive attitude towards themselves.

Bringing everything together, we obtain the following standard life satisfaction model:

$$s_i = \alpha + \beta y_i + \gamma h_i + \delta u_i + \zeta z_i + \varepsilon_i, \tag{7}$$

where $\alpha, \beta, \gamma, \delta, \zeta$ are coefficients to be estimated and ε_i an error term. The coefficients β, γ and δ capture the trade-offs or marginal rates of substitution between the dimensions in which we are interested. These coefficients provide us with the necessary information to plot a linear indifference map of each individual. Note that the constant, the error term, and the vector of socio-demographic characteristics are irrelevant for the marginal rates of substitution between the dimensions.

The method requires us to classify every variable as a dimension of life or as a scaling factor. As discussed above, a variable like education, however, can be argued to be a dimension of life and a scaling factor. Unfortunately, both effects cannot be disentangled econometrically. We choose here to treat education as a scaling factor and hence we cannot include it as a dimension of life.

Given the linear specification of the standard model, all estimated indifference curves will be linear and parallel, leading to constant trade-offs between the dimensions. This clearly is an unattractive feature of the standard model. A few refinements are therefore necessary to obtain more realistic preferences and to allow for heterogeneity in the estimated trade-offs.

First, we allow the β, γ and δ's to differ across groups of individuals with the same socio-demographic background. This can be achieved by including interaction effects between the well-being dimensions and some variables that capture the socio-demographic background of the individuals. In our illustration we allow for different coefficients for different socio-demographic groups depending on whether the respondent lives with a partner, is male, and belongs to the age group 45-60 or not. Let β_g, γ_g and δ_g refer to the coefficients of income, health and unemployment of the socio-demographic group g.

Second, the marginal rates of substitution between the dimensions may not be constant. Indeed, the increase in income that is necessary to keep life satisfaction unchanged after a decrease in health, may depend on the level of income and health of the individual. For that reason, the dimensions are transformed by means of a so-called Box-Cox transformation, $f_\eta(x) = (x^\eta - 1)/\eta$, with the parameter η capturing the concavity of the transformed function (see Box and Cox (1964) for an extensive treatment). For η equal to 1, the Box-Cox transformation is linear. The transformation function f_η is strictly concave when $\eta < 1$. The lower the parameter η is, the more concave the function becomes. As unemployment is a binary variable, we do not transform unemployment. The resulting, refined, satisfaction model then becomes:

$$s_i = \alpha + \beta_g \frac{y_i^{\eta_y} - 1}{\eta_y} + \gamma_g \frac{h_i^{\eta_h} - 1}{\eta_h} + \delta_g u_i + \zeta z_i + \varepsilon_i, \tag{8}$$

where η_y and η_h are the parameters of the Box-Cox transformation for income and health, respectively. All parameters of the model are jointly estimated by maximum likelihood estimation assuming a normal distribution of the error term. The estimation results according to this refined model can be found in Table 1. The table has five parts.

The first part shows the coefficients of the well-being dimensions y_i, h_i, and u_i. The three rows give the effect of the outcomes in the dimensions on life satisfaction for members of the reference socio-demographic group (who do not live with a partner, are female, and are between 45 and 60 years old). The coefficients have the expected sign and are statistically significant at the 5% level. In the interpretation of the magnitude of the coefficients of income and health, one has to take account of the Box-Cox transformation.

The next three rows show the interaction effects of income with the three binary indicators capturing the socio-demographic group to which the respondent belongs. We see that respondents living with a partner have a higher coefficient of income compared to the reference group (the income coefficient of those respondents equals $0.1747 = 0.150 + 0.0247$), while the coefficient of the male respondents is smaller. All three interaction effects are quantitatively modest but statistically significant at the 0.1% level, which indicates moderate but relevant heterogeneity in preferences. Interactions between these three binary socio-demographic indicators and health and unemployment are mostly found to be insignificant and have been dropped to obtain a parsimonious model, so that $\gamma_g = \gamma$ and $\delta_g = \delta$ for all groups g in equation (8).

In the next two rows, the parameters of the Box-Cox transformation of income and the health index are reported. They are found to be slightly lower for income than for the health index (0.114 and 0.233 respectively), indicating somewhat stronger decreasing returns for income than for health, which stands to reason (the magnitude of η_y is in line with the literature: Becker et al. (2005) calibrate it at a level of 0.2 and Layard et al. (2008) estimate it to be around 0.26).

[Table 1 about here]

The fourth part of the table highlights the effect of the control variables age, age squared, an indicator of higher education, and three marital status variables. The regional dummies are not shown for lack of space (7 out of the 15 dummies are statistically significant at the 5% level). All three personality variables have a positive and statistically significant effect on life satisfaction. The strongest effect can be observed for having a positive attitude.

One final remark needs to be made about the last part of Table 1. The pseudo R^2 of this and similar life satisfaction regressions tends to be rather low (see, Graham (2004) for other examples), which indicates that the actual preferences are still largely determined by unobserved factors which are not captured by our econometric model. However, it can be argued that these actual preferences of the individuals are too idiosyncratic to be normatively compelling and should be laundered before they can be used in the construction of a well-being measure (see Decancq et al. (2014)). By using only the average preferences of the socio-demographic group to which the individual belongs, the estimation procedure provides such a laundering procedure. If the sample size would have allowed a more fine-grained classification in subgroups, however, that would certainly have been desirable.

3.3 Computing the well-being measures

The first well-being measure WB^1 coincides with household equivalized disposable income, which can be readily obtained from the data set, after adjusting it for the household size as discussed above. We focus therefore on the implementation of the other well-being measures.

To compute the composite well-being index WB^2, choices have to be made on the dimension-specific transformation functions, on the degree of substitutability, and on the weighting scheme. Income is first transformed by means of a logarithmic transformation to capture the decreasing marginal effect of income on well-being. All dimensions have been normalized to a scale between 0 and 1 and then aggregated additively (i.e. we set $\beta = 1$ in equation (3)). We use principal component analysis (PCA) to obtain the weighting scheme from the data (see Ram (1982) and Nguefack-Tsague et al. (2011) for a defense of this method). The weighting scheme $w = (0.41, 0.24, 0.35)$ consists of the normalized loadings of the first principal component of the original data and provides the linear combination that captures the maximal variance of the original variables. We find that the PCA method attributes a larger weight to income and a lower weight to health compared to equal weighting. The resulting composite index has an average value of 0.83 and ranges between 0.25 and 0.99 in the sample.

For the subjective well-being measure WB^3, we use observed life satisfaction s_i (the same variable that we used as dependent variable in the regression discussed in the previous subsection). This variable is measured by the following question (after translation): "*To what extent are you satisfied with your life in general at the present time?*". Respondents attach a score from 0 ("not satisfied at all") to 10 ("completely satisfied") to the question. The wording of this question invites the respondent to make a cognitive valuation of her life. Yet, we do not know how the respondents have precisely interpreted the question. In 2010, the SOEP also contains some information on affects, namely respondents are asked how frequently they felt sad during the last 4 weeks and how frequently they felt happy. Respondents reply on a 1 to 5 scale in both cases. In order to compare life satisfaction scores and affects, we also construct an—admittedly crude—measure of the "hedonic flow" in a Benthamite spirit by subtracting the frequency of feeling sad from the frequency of feeling happy (see Kahnemann and Krueger (2006) for a discussion of more sophisticated methods). This constructed affect measure ranges on a scale from -4 to 4. The Spearman rank correlation between the life satisfaction variable and the affect measure is relatively high (0.50) and significantly different from 0. Whether this correlation suggests that positive affects are an important component of life satisfaction or that respondents do not distinguish clearly between questions about affects and cognitive valuations, is unclear however.

We compute for each individual her equivalent income WB^4 by solving equation (5) using the shape of the indifference maps estimated with the procedure described in the previous subsection. To do so, we need to select the reference value of the non-income dimensions. As discussed above, it is intuitive to take for each dimension the best possible or the value that people aspires to. We use the estimated indifference curves to look for the best possible value for all socio-demographic groups. All socio-demographic groups have well-behaved monotonic preferences and we select the highest value as reference, i.e., a value of 100 for the objective health indicator (meaning perfect health or absence of any health problem) and not being unemployed. The reference values are reached by about 9% of our sample for health, and by 94% for unemployment.

For the implementation of the final preference measure WB^5 we need additional information on the vNM utility function of the individuals. Estimating vNM utility functions requires information on the preferences of the individuals over lotteries of vectors of life aspects. Unfortunately, this information is not available in the data set. To illustrate the role of the risk attitude when using this well-being measure, however, we calibrate the vNM utility functions based on some other information on the risk attitude reported by the respondents. In its 2010 wave, the SOEP contains a question on the personal willingness to take risks, in which respondents classify themselves on a scale with 11 categories, ranging from "no willingness to take risks" to "very willing to take risks". With respect to our measure WB^5, we interpret the answers of the respondents as their willingness to take risks in lotteries of vectors of life aspects on a path between the reference vectors ℓ^+ and ℓ^{++}. We classify the respondents as either being risk neutral or risk averse through a median split of the risk attitude variable. The risk neutral respondents are then equipped with the following vNM utility function:

$$V_i(y_i, h_i, u_i) = \beta_g \frac{y_i^{\eta_y} - 1}{\eta_y} + \gamma_g \frac{h_i^{\eta_h} - 1}{\eta_h} + \delta_g u_i, \tag{9}$$

with the parameters as reported in Table 1. This vNM utility function equals the refined satisfaction regression net of the scaling factors (as one notices from comparing equations (8) and (9)). The underlying ordinal preferences are hence the same as those that are used for the derivation of the equivalent income measure WB^4. For the risk averse respondents, we take a mildly concave Box-Cox transformation of equation (9) with the concavity parameter calibrated at the value of 0.8. This procedure results in a situation similar to the right hand panel of Figure 2 where half of the sample is considered to be risk neutral (using the linear gray curve to attribute well-being level to the indifference curves) and half of the sample is risk averse (using the concave black curve). Using the appropriate vNM utility function for each individual, we compute WB^5 using equation (6), with ℓ^+ equal to the origin and ℓ^{++} equal to the "maximum" vector of life aspects with income equal to 25,000 euro per month, health equal to 100, and the employment status being not unemployed. As mentioned above, this is only one of the possible choices for ℓ^+ and ℓ^{++}.

4 Empirical comparison of the well-being measures

In this section we empirically compare the different well-being measures by taking three different perspectives. First, we look at the characteristics of the individuals who are identified as worst off according to the five measures, then we study the degree of overlap across the groups identified as worst off and, finally, we discuss the re-ranking between the different well-being measures taking into account the full sample. The first two perspectives are especially relevant from a policy perspective when it comes to targeting the worst off in a society. The last perspective is relevant for the implementation of (progressive) redistribution, which requires a ranking of all individuals in a society.

Before starting the analysis of the "worst off", we need to decide on the size of the group that we are referring to as worst off. Empirically, it seems not to make much sense to focus on a single worst off respondent from the sample or on a very small group. On the other hand, the group should be small enough to capture what is going on at the bottom of the distribution without being biased by the outcomes of the better off. Focussing on the bottom decile seems a reasonable compromise to us. To be precise, we focus on the bottom 9.2 % of the population (which corresponds to the percentage of respondents reporting a life satisfaction lower or equal than 4 on the scale between 0 and 10). It is important to stress that our analysis of the characteristics of the bottom tail of the well-being distribution is different from a standard poverty analysis in which an exogenous (absolute or relative) poverty line is used to identify the poor. In an extremely rich society there will be a group of people identified as worst off, and vice versa in a poor society there may be more people being identified as poor than the group we classify here as the worst off.

Various other recent studies have used the SOEP to study well-being and poverty: Clark et al. (2013) and d'Ambrosio and Frick (2007) investigate the link between subjective well-being and income poverty, Juhasz (2012) discusses so-called "satisfaction based poverty measures", and Rippin (frthc.) compares income and multidimensional poverty in Germany.

4.1 Portrait of the worst off

Table 2 is the central table of this section and shows a portrait of the worst off. The right hand columns of the table correspond to the five well-being measures. The rows are classified into three parts. The first four rows show life satisfaction as well as the three outcome variables. The second part of the table reports various socio-demographic variables. The final two rows show some additional information on shape of the indifference curves, namely the willingess-to-pay (WTP) for being in perfect health and for not being unemployed. The higher the WTP in a non-income dimension, the more the respondent "suffers" from her imperfect outcome in that dimension according to her own opinion on the good life. Individuals who obtain the best possible value in a non-income dimension, have a WTP of 0 in that dimension.

In each cell of the table we report the average of the variable mentioned in its row amongst the worst off according to the well-being measure mentioned in its column. As a reference, the unconditional sample averages are given in the first column. All averages are obtained after weighting with corresponding sample weights. Figures between brackets refer to non-statistically-significant differences between the full sample average and the average amongst the worst off according to a particular measure at the 5% significance level.

[Table 2 about here]

Several observations from Table 2 are worth noting. We start by comparing the worst off according to the income measure WB^1 with the sample averages (i.e., we compare column 1 and 2). We see, for instance, that the average satisfaction score of the full sample is 6.95 whereas the average score amongst the worst off according to the income measure score is only 5.80. Average household disposable income is only slightly above one third of the average income in Germany while the difference between the health index of the worst off according to WB^1 and the full sample is rather small. Unemployment figures, on the contrary, are about 5 times higher. When looking at the socio-demographic characteristics in the next part of the table, one sees that the worst off according to the income measure are less likely to be living together with a partner, while they are more likely to be divorced or to live in a single parent household. The worst off according to the income measure are more often non-German, and more often living in East Germany. A final finding is the low educational level of the worst off according to WB^1.

Next, we focus on the portrait according to the composite well-being index, WB^2. When comparing the first two measures of well-being, it can be seen that the worst off are richer, less healthy and more likely to be unemployed according to the composite index compared to income, which reflects the impact of including health and unemployment in the well-being measure. In particular the large share of unemployed among the worst off (about 70%) is striking. When investigating the socio-demographic variables, we find again an over-representation of the divorced and single parent families compared to the full sample. A much larger proportion of disabled people is identified as worst off compared to the income well-being measure.

The third measure relies on the subjective well-being assessment of the individuals, WB^3 and offers a quite different picture. The worst off are richer and less likely to be unemployed compared to the former two well-being measures. Yet, the average health of the worst off is comparable to the previous measure as is the share of people suffering a disability. The educational level of the worst off is better than when using the previous measures and only slightly worse compared to the full sample. With respect to the personality traits, the worst off according to self-reported life satisfaction report lower levels of control over their life and have a more negative attitude towards themselves compared to the former two well-being measures and also the full sample. This finding illustrates the importance of personality traits as determinants of life satisfaction. We also find considerable differences when it comes to the affect variables. Unsatisfied people feel more often sad and less often happy. As discussed above, however, it is hard to disentangle whether we obtain this result because the respondents give a large weight to affects as dimensions of their life, or whether respondents do not distinguish the questions about affects and cognitions.

Fourth, we discuss WB^4, the preference-based equivalent income measure. Again a different picture emerges. The worst off according to the equivalent income measure are in much worse health compared to the earlier measures. More than half of the worst off report to be disabled, which is almost a doubling with respect to the composite index and even more than three times as much as in the full sample. Remarkable is also the large number of widowed individuals and pensioners. The average age of the worst off is 62 years, more than 10 years older than the worst off according to the composite index. The reader may be wondering why the differences between the equivalent income measure and the composite index are so pronounced, as the same three dimensions are considered in the analysis. The penultimate row of Table 2 sheds some light on this puzzle. One sees that the worst off according to the equivalent income measure are willing to pay much more to be in perfect health compared to the worst off according to the composite index. This is intuitive as the equivalent income measure takes account of the preferences of the individuals. Given the large coefficient of health and the concave transformation of health (recall that the Box-Cox parameter of health is 0.233), unhealthy individuals are willing to give up a large part of their income to be in perfect health, which is reflected in the low health of the individuals identified as worst off. Many pensioners are identified as worst off due to a higher prevalence of health problems at older ages. The average WTP for not being unemployed among the worst off with the equivalent income measure is lower compared to the composite index, since pensioners cannot be unemployed and are on the reference value by definition.

Finally, we consider WB^5, the well-being measure based on the vNM utility function. This measure is based on the same ordinal preferences underlying the equivalent income measure. Most of the characteristics of the worst off according to both measures are therefore quite similar. Recall that we have used an additional concave transformation for the group of respondents that are not willing to take much risk. The role of this transformation is clear when looking at the bottom row of the socio-demographic characteristics. The worst off according to the last well-being measure are clearly more willing to take risks. This finding is in line with the effect shown in Figure 2, where (for the same vectors of life aspects) the risk neutral individuals are worse off compared to the risk averse ones (our robustness checks show that when choosing a more concave vNM utility function transformation for the risk averse respondents, the willingness to take risks among the worst off increases further).

In sum, our inspection of the characteristics of the worst off highlights that different socio-demographic groups are identified as worst off with the different measures. Broadly speaking, we find that the worst off according to the income measure are more likely to be low educated, unemployed individuals who are living without a partner. The composite index gives a relatively large weight to unemployment so that more than two-thirds of the worst off are unemployed. The individuals with a low life satisfaction are on average almost as rich as the full sample but have a more negative attitude towards themselves and are more often sad and less often happy. The equivalent income measure identifies unhealthy people as worst off, which results in an overrepresentation of pensioners and people with a disability among the worst off. Finally, the worst off according to the measure based on vNM utilities are similar to those under the equivalent income measure, but are willing to take more risks.

4.2 Overlap between the worst off

We have seen that substantial socio-demographic differences exist between the individuals identified as worst off according to the five different well-being measures. We complement these findings by discussing the extent to which there is overlap between the groups of individuals identified as worst off according to the different measures.

In Table 3, the five columns in the middle part correspond to the five well-being measures. According to each measure an individual can be classified either as belonging to the worst off or not. This gives us $2^5 = 32$ possible combinations, which are listed in the rows of the table. In each row a checkmark indicates the combinations in which an individual is considered as worst off. The last column shows the percentage of the sample for which this combination occurs (taking sample weights into account).

We start by computing the percentage of individuals who do not belong to the worst off according to any of the well-being measures, which amounts to 75.59% (not reported in the table). Contrarily, that implies that more than 24% of the sample are classified as worst off according to at least one of the five well-being measures. This relatively large number suggests a low degree of overlap between the well-being measures. In the first five rows of Table 3, we show the percentage of all individuals that are identified by at least one particular measure as being worst off. For all measures, this gives the selected percentage of worst off, i.e., 9.2%.

The next group of rows in Table 3 shows the number of people that are identified as worst off according to at least two specific well-being measures. If all measures identified exactly the same group of individuals as being worst off, we would observe also 9.2% in each of these rows. However, as one can see, percentages are considerably lower. In most cases, they are well below half, except for the pairwise overlaps between the composite index, the equivalent income measure and the measure based on the vNM utility function, with the highest percentage for the latter two. When looking further down the table at the overlap of three measures, we see again that the overlap is highest for these three measures (4.6%), which means that about half of the worst off according to any of these three measures are also identified by the other two measures as belonging to the worst off. The overlap between four measures is never larger than 2%. Finally, we observe that only 0.87% of the individuals are identified as worst off according to all measures.

[Table 3 about here]

4.3 Re-ranking between the well-being measures

In this section, we broaden the picture to the full sample and visualize the degree of re-ranking between the different well-being measures . We do so by providing in Figure 3 scatter plots of the rank of the individuals (the percentile to which they belong) according to each pair of well-being measures. To visualize overplotting of different observations, we use a so-called "sunflower" scatter plot in which each point has a number of leaves equal to the number of observations in that point. If all points are situated on the diagonal, there is no re-ranking. The more spread-out the plots are, the more re-ranking takes place. To quantify the overall degree of re-ranking, we provide also the Spearman rank correlation coefficient r below each scatter plot (all correlation coefficients are significant at the 1% level).

The top left hand panel of Figure 3 plots the rank according to income (WB^1) against the rank according to the composite well-being index (WB^2). One sees that the sunflowers are spread out on almost the entire graph, with some concentration around the diagonal. This highlights some re-ranking between these two well-being measures (the Spearman rank correlation coefficient equals 0.51). No observations are found in the top left hand part of the graph, however. This is intuitive, as it is impossible for income poor individuals to reach top ranks according to the composite well-being index that contains income as one of its three components.

The next panel in the top right hand corner of Figure 3 shows much more re-ranking. It plots the income ranks against the ranks according to life satisfaction (WB^3). To allow a graphical representation, we use the ranks of a latent continuous life satisfaction measure that is consistent with the observed discrete responses. As we saw in the previous section, the overlap between the worst off identified according to these measures is low (2.04%, see Table 3). In this panel, we see that this low degree of overlap between the measures extends to the entire sample (the Spearman rank correlation coefficient equals only 0.25).

The plots of the income rank against the rank according to equivalent income (WB^4) and the measure based on the vNM utility function (WB^5) on the next row of Figure 3, are again similar to the top left hand panel with comparable Spearman rank correlation coefficients. The next scatter plot, with the rank according to the composite well-being index on the horizontal axis and the life satisfaction rank on the vertical axis is again similar to the spread out plot in the top right hand corner between the income and life satisfaction ranks.

The next two scatter plots (between the composite index and both preference based well-being measures) have a distinctive shape, which resembles a "hockey stick". The amount of re-ranking is found to be low with a strong concentration around the diagonal (the Spearman rank correlation coefficients are larger than 0.9), except for the individuals found in the bottom percentiles according to the composite index, who are scattered in the first five deciles according to the equivalent income and the measure based on the vNM utility function. All these individuals are unemployed. As the weight of unemployment is larger in the composite index than in the other multidimensional measures, all unemployed individuals are bottom ranked according to the composite index, whereas this is much less the case for the preference-based measures which give more weight to health and income. This finding is consistent with the high unemployment rate among the worst off according to the composite index, found in Table 2. Yet, it suggests that there would be even more overlap between these three measures when a larger group of worst off was considered than is currently used to construct Table 3. A closer inspection of the scatter plot of the ranking according to the composite well-being index and the measure based on the vNM utility function shows that the right-hand side of the plot is somewhat separated into two parts, one above and one below the diagonal. This is due to our splitting of the sample in two groups based on their willingness to take risks. As shown in Figure 2, the risk averse individuals obtain a higher well-being than the risk neutral ones for a given vector of life aspects.

The next two graphs show the re-ranking when the rank according to life satisfaction is put on the horizontal axis and a preference based measure on the vertical axis. Both scatter plots look very similar to the top right hand graph. This finding illustrates once more that the ranking according to life satisfaction is substantially different from the other methods.

The final scatter plot, in the bottom right hand corner, shows the re-ranking between both preference-based well-being measures. Recall that both measures are based on the same ordinal preferences, but use a different way of cardinalizing them. The ranking coincides at the top and bottom of the well-being distribution, but there is considerable re-ranking around the median. This is intuitive, as the scope for re-ranking due to the different cardinalization of the heterogeneous ordinal preferences is largest around the median. Each of the stripes on the graph reflects a different combination of risk attitude and membership of a socio-demographic group.

Our findings in this final section highlight the substantial amount of re-ranking that occurs in the full sample between the different well-being measures. These findings are consistent with the earlier analyses focussing on the bottom part of the well-being distribution. In particular life satisfaction ranks the individuals considerably different compared to the other measures.

[Figure 3 about here]

5 Conclusion

We have illustrated—once more—that measurement matters. We have compared five well-being measures: income, a composite well-being index, subjective well-being, equivalent income, and a measure based on individual vNM utility functions. We discussed the information requirements of these measures and have illustrated how they can be applied on a cross-section of a single data set, the German SOEP.

Our results show that the measures identify individuals with substantially different socio-demographic characteristics as worst off. Moreover, the overlap between the worst off according to some of the measures is low. Clearly, more research is needed to understand better the empirical consequences of choosing a particular well-being measure. It remains an open question, for instance, how generalizable our illustrative findings are. Moreover, even with a state-of-the-art data set as the SOEP, our analysis remained rather broad-brushed, covering only three aspects of life. Collecting richer data sets of a high quality will be essential to foster further development in the field of empirical well-being measurement.

Yet, our findings have potentially far-reaching policy implications when it comes to the design of social policies that involve some kind of targeting or redistribution. A policy maker who wants to take the idea seriously that well-being is a multidimensional notion, and who wants to work with a single well-being measure to take account of the potential correlation between the dimensions, will sooner or later face the challenging question which well-being measure to use. We have illustrated some of the real-world implications of this question when it comes to identifying the worst off.

Some well-being measures were relatively straightforward to implement (life satisfaction and income information was readily available in the data set), other measures required econometric estimations (the indifference maps for the preference-based approaches), or parameter calibrations (the composite well-being index and the curvature of the vNM utility function). Whether or not these extra efforts are worthwhile depends largely on the normative appeal of the measures. Personally, we believe that the additional estimations to retrieve information on the opinions on the good life of the concerned individuals are not only worthwhile, but also necessary to arrive at attractive interpersonal well-being comparisons.

References

Adler, M. (frthc.): "Extended preferences," in *Oxford Handbook on Well-being and Public Policy*, ed. by M. D. Adler, and M. Fleurbaey. Oxford University Press.

Adler, M. D. (2012): *Well-Being and Fair Distribution: Beyond Cost-Benefit Analysis*. Oxford University Press.

——— (2013): "Happiness surveys and public policy: What's the use?," *Duke Law Journal*, 62, 1509–1601.

——— (2014): "Extended preferences and interpersonal comparisons: a new account," *Economics and Philosophy*, 30, 123–162.

Alkire, S. (frthc.): "Capabilities," in *Oxford Handbook on Well-being and Public Policy*, ed. by M. D. Adler, and M. Fleurbaey. Oxford University Press.

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

Anand, P., G. Hunter, I. Carter, K. Dowding, F. Guala, and M. Van Hees (2009): "The development of capability indicators," *Journal of Human Development and Capabilities*, 10(1), 125–152.

Arneson, R. (frthc.): "Does fairness require a multidimensional approach?," in *Oxford Handbook on Well-being and Public Policy*, ed. by M. D. Adler, and M. Fleurbaey. Oxford University Press.

Arrow, K. (1973): "Some ordinalist-utilitarian notes on Rawls' theory of justice," *Journal of Philosophy*, 70(9), 245–263.

Atkinson, A. B. (2003): "Multidimensional deprivation: contrasting social welfare and counting approaches," *Journal of Economic Inequality*, 1(1), 51–65.

Bargain, O., A. Decoster, M. Dolls, D. Neumann, A. Peichl, and S. Siegloch (2013): "Welfare, labor supply and heterogeneous preferences: evidence for Europe and the US," *Social Choice and Welfare*, 41, 789–817.

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

Benjamin, D., O. Heffetz, M. S. Kimbal, and A. Rees-Jones (frthc.): "Can Marginal Rates of Substitution Be Inferred from Happiness Data? Evidence from Residency Choices," *American Economic Review.*

Benjamin, D. J., O. Heffetz, M. S. Kimball, and N. Szembrot (2012): "Beyond Happiness and Satisfaction: Toward Well-Being Indices Based on Stated Preference," Discussion paper, Mimeo.

Box, G. E. P., and D. R. Cox (1964): "An analysis of transformations," *Journal of the Royal Statistical Society. Series B (Methodological)*, 26(2), 211–252.

Boyce, C. J. (2010): "Understanding fixed effects in human well-being," *Journal of Economic Psychology*, 31(1), 1–16.

Brandolini, A. (2007): "On synthetic indices of multidimensional well-being: Health and income inequalities in France, Germany, Italy, and the United Kingdom," Bank of Italy, Economic Research Department 668.

Brun, B. C., and B. Tungodden (2004): "Non-welfaristic theories of justice: Is "the intersection approach" a solution to the indexing impasse?," *Social Choice and Welfare*, 22, 49–60.

Bundesagentur für Arbeit (2014): *Statistik der Bundesagentur für Arbeit, Arbeitsmarkt in Deutschland - Zeitreihen bis 2010.*

Chakravarty, S., and M. A. Lugo (frthc.): "Multidimensional indicators," in *Oxford Handbook on Well-being and Public Policy*, ed. by M. D. Adler, and M. Fleurbaey. Oxford University Press.

Chiappori, P.-A. (frthc.): "Households," in *Oxford Handbook on Well-being and Public Policy*, ed. by M. D. Adler, and M. Fleurbaey. Oxford University Press.

Clark, A. (frthc.): "Happiness," in *Oxford Handbook on Well-being and Public Policy*, ed. by M. D. Adler, and M. Fleurbaey. Oxford University Press.

Clark, A. E., C. D'Ambrosio, and S. Ghislandi (2013): "Poverty and Well-Being: Panel Evidence from Germany," Working Papers 291, ECINEQ, Society for the Study of Economic Inequality.

Clark, A. E., Y. Georgellis, and P. Sanfey (2001): "Scarring: The Psychological Impact of Past Unemployment," *Economica*, 68(270), 221–241.

Clark, A. E., and A. J. Oswald (1994): "Unhappiness and Unemployment," *The Economic Journal*, 104, 648–659.

——— (2002): "A simple statistical method for measuring how life events affect happiness," *International Journal of Epidemiology*, 31(6), 1139–1144.

Cowell, F. (frthc.): "Income-based inequality and poverty measures," in *Oxford Handbook on Well-being and Public Policy*, ed. by M. D. Adler, and M. Fleurbaey. Oxford University Press.

Cowell, F. A. (2011): *Measuring Inequality*. Oxford University Press, Oxford, third edn.

D'Ambrosio, C., and J. R. Frick (2007): "Income Satisfaction and Relative Deprivation: An Empirical Link," *Social Indicators Research*, 81(3), 497–519.

Decancq, K., M. Fleurbaey, and F. Maniquet (2014): "Multidimensional poverty measurement with individual preferences," Discussion paper, Princeton University - William S. Dietrich II Economic Theory Center Working Paper 058-2014.

Decancq, K., M. Fleurbaey, and E. Schokkaert (frthc.): "Inequality, income and well-being," in *The Handbook on Income Distribution, volume 2*. Elsevier.

Decancq, K., and M. A. Lugo (2013): "Weights in Multidimensional Indices of Well-Being: An Overview," *Econometric Reviews*, 32(1), 7–34.

Decancq, K., and E. Schokkaert (2013): "Beyond GDP: measuring social progress in the EU," Discussion paper, Euroforum (4).

Decoster, A., and P. Haan (frthc.): "Empirical welfare analysis in random utility models of labour supply," *International Tax and Public Finance*.

Di Tella, R., J. H.-D. New, and R. MacCulloch (2010): "Happiness adaptation to income and to status in an individual panel," *Journal of Economic Behavior & Organization*, 76(3), 834 – 852.

Dolan, P., and D. Fujiwara (frthc.): “Happiness-based policy analysis,” in *Oxford Handbook on Well-being and Public Policy*, ed. by M. D. Adler, and M. Fleurbaey. Oxford University Press.

Duclos, J.-Y., and L. Tiberti (frthc.): “Multidimensional poverty indices: A critical assessment,” in *Oxford Handbook on Well-being and Public Policy*, ed. by M. D. Adler, and M. Fleurbaey. Oxford University Press.

Ferrer-i-Carbonell, A. (2005): “Income and well-being: an empirical analysis of the comparison income effect,” *Journal of Public Economics*, 89(5-6), 997–1019.

Ferrer-i-Carbonell, A., and P. Frijters (2004): “How important is methodology for the estimates of the determinants of happiness?,” *Economic Journal*, 114(497), 641–659.

Fleurbaey, M. (2009): “Beyond GDP: the quest for a measure of social welfare,” *Journal of Economic Literature*, 47(4), 1029–1075.

——— (frthc.): “Equivalent income,” in *Oxford Handbook on Well-being and Public Policy*, ed. by M. D. Adler, and M. Fleurbaey. Oxford University Press.

Fleurbaey, M., and D. Blanchet (2013): *Beyond GDP: Measuring Welfare and Assessing Sustainability*. Oxford University Press.

Fleurbaey, M., S. Luchini, C. Muller, and E. Schokkaert (2013): “Equivalent income and fair evaluation of health care,” *Health Economics*, 22(6), 711–729.

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

Fleurbaey, M., E. Schokkaert, and K. Decancq (2009): “What good is happiness?,” CORE discussion paper 2009/17.

Frijters, P., J. P. Haisken-DeNew, and M. A. Shields (2004a): “Investigating the patterns and determinants of life satisfaction in Germany following reunification,” *Journal of Human Resources*, XXXIX(3), 649–674.

——— (2004b): “Money Does Matter! Evidence from Increasing Real Income and Life Satisfaction in East Germany Following Reunification,” *American Economic Review*, 94, 730–740.

Graham, C. (2009): *Happiness around the world: the paradox of happy peasants and miserable millionaires.* Oxford University Press., Oxford.

——— (frthc.): "Subjective well-being in economics," in *Oxford Handbook on Well-being and Public Policy*, ed. by M. D. Adler, and M. Fleurbaey. Oxford University Press.

Graham, C., A. Eggers, and S. Sukhtankar (2004): "Does happiness pay?: An exploration based on panel data from Russia," *Journal of Economic Behavior and Organization*, 55(3), 319 – 342.

Gwozdz, W., and A. Sousa-Poza (2010): "Ageing, Health and Life Satisfaction of the Oldest Old: An Analysis for Germany," *Social Indicators Research*, 97(3), 397–417.

Harsanyi, J. C. (1953): "Cardinal Utility in Welfare Economics and in the Theory of Risk Taking," *Journal of Political Economics*, 62, 434–435.

——— (1955): "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility," *Journal of Political Economics*, 63, 309–321.

——— (1977): *Rational Behaviour and Bargaining Equilibrium in Games and Social Situations.* Cambridge University Press, Cambridge.

Hausman, D. M. (2012): *Preference, Value, Choice, and Welfare.* Cambridge University Press.

Haybron, D. (frthc.): "Hedonic views," in *Oxford Handbook on Well-being and Public Policy*, ed. by M. D. Adler, and M. Fleurbaey. Oxford University Press.

Helliwell, J., R. Layard, and J. Sachs (2013): *World happiness report 2013.* United Nations Sustainable Development Solutions Network.

Hurka, T. (frthc.): "Objective-list approaches," in *Oxford Handbook on Well-being and Public Policy*, ed. by M. D. Adler, and M. Fleurbaey. Oxford University Press.

Juhász, A. (2012): "A satisfaction-driven poverty indicator - A bustle around the poverty line," Discussion paper, SOEPpapers 461.

Jürges, H. (2007): "True health vs response styles: exploring cross-country differences in self-reported health," *Health Economics*, 16(2), 163–178.

Kahneman, D., and A. B. Krueger (2006): "Developments in the measurements of subjective well-being," *Journal of Economic Perspectives*, 20(1), 3–24.

Laderchi, C. R., R. Saith, and F. Stewart (2003): "Does it matter that we do not agree on the definition of poverty? A comparison of four approaches," *Oxford Development Studies*, 31(3), 243–274.

Layard, R. (2005): *Happiness: Lessons for a new Science*. Penguin, London.

Layard, R., G. Mayraz, and S. Nickell (2008): "The marginal utility of income," *Journal of Public Economics*, 92(8-9), 1846–1857.

Lewbel, A., and K. Pendakur (2008): "Equivalence Scales," in *The New Palgrave Dictionary of Economics*, ed. by S. Durlauf, and L. Blume. Palgrave Macmillan, New York.

Luechinger, S., S. Meier, and A. Stutzer (2010): "Why Does Unemployment Hurt the Employed?: Evidence from the Life Satisfaction Gap Between the Public and the Private Sector," *Journal of Human Resources*, 45(4), 998–1045.

Lundberg, S. J., R. A. Pollak, and T. J. Wales (1997): "Do Husbands and Wives Pool Their Resources? Evidence from the United Kingdom Child Benefit," *Journal of Human Resources*, 32(3), 463–480.

Maniquet, F. (frthc.): "Social Ordering Functions," in *Oxford Handbook on Well-being and Public Policy*, ed. by M. D. Adler, and M. Fleurbaey. Oxford University Press.

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

Nussbaum, M. C. (2008): "Who Is the Happy Warrior? Philosophy Poses Questions to Psychology," *The Journal of Legal Studies*, 37(S2), S81–S113.

OECD (2013): *OECD Guidelines on Measuring Subjective Well-being*. OECD publishing.

Pogge, T., and S. Wisor (frthc.): "Measuring poverty: A proposal," in *Oxford Handbook on Well-being and Public Policy*, ed. by M. D. Adler, and M. Fleurbaey. Oxford University Press.

Ram, R. (1982): "Composite indices of physical quality of life, basic needs fulfilment, and income: a principal component representation," *Journal of Development Economics*, 11(2), 227–247.

Ravallion, M. (2011): "On multidimensional indices of poverty," *Journal of Economic Inequality*, pp. 1–14.

——— (2012): "Troubling tradeoffs in the Human Development Index," *Journal of Development Economics*, 99(2), 201–209.

Rippin, N. (frthc.): "Multidimensional Poverty in Germany: A Capability Approach," *Forum for Social Economics*.

Schokkaert, E., L. Van Ootegem, and E. Verhofstadt (2011): "Preferences and Subjective Satisfaction: Measuring Well-being on the Job for Policy Evaluation," *CESifo Economic Studies*, 57(4), 683–714.

Sen, A. K. (1979): "The Welfare Basis of Real Income Comparisons: A Survey," *Journal of Economic Literature*, 17(1), 1–45.

——— (1985): *Commodities and Capabilities*. North-Holland, Amsterdam and Oxford.

——— (2004): "Capabilities, lists, and public reason: continuing the conversation," *Feminist Economics*, 10, 77–80.

Stiglitz, J. E., A. K. Sen, and J. P. Fitoussi (2010): *Mismeasuring Our Lives: Why GDP Doesn't Add Up*. New Press, New York.

van Doorslaer, E., and A. Jones (2003): "Inequalities in self-reported health: validation of a new approach to measurement," *Journal of Health Economics*, 22(1), 61–87.

Van Landeghem, B. (2012): "A test for the convexity of human well-being over the life cycle: Longitudinal evidence from a 20-year panel," *Journal of Economic Behavior & Organization*, 81(2), 571–582.

von Neumann, J., and O. Morgenstern (1947): *Theory of Games and Economic Behavior*. Princeton University Press, Princeton, NJ.

Wagner, G. G., J. R. Frick, and J. Schupp (2007): "The German Socio-Economic Panel Study (SOEP) - Scope, Evolution and Enhancements.," *Schmollers Jahrbuch*, 127, 139–169.

Weymark, J. A. (2006): "The normative approach to the measurement of multidimensional inequality," in *Inequality and Economic Integration*, ed. by F. Farina, and E. Savaglio. Routledge, London.

Table 1: Life satisfaction regression

income	0.150**	(0.0563)
health	0.674*	(0.319)
unemployment	-0.575***	(0.0680)
income × partner	0.0247***	(0.00617)
income × male	-0.0178***	(0.00390)
income × “non-midlife”	0.0202***	(0.00514)
Box-Cox parameter income	0.114*	(0.0459)
Box-Cox parameter health	0.233*	(0.115)
age	-0.0344***	(0.00750)
age squared	0.000366***	(0.0000692)
higher education	0.0941**	(0.0305)
divorced	0.00629	(0.0583)
separated	-0.0371	(0.0988)
widowed	-0.0411	(0.0680)
control over life	0.144***	(0.0107)
achieved what deserved	0.0745***	(0.00815)
positive attitude	0.332***	(0.0124)
constant	-2.331*	(0.980)
N	14,027	
pseudo R^2	0.1007	

Regional dummies included. Clustered standard errors in parentheses.
“non-midlife” denotes an age below 45 or above 60 years.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
Source: Own calculations based on the SOEP 2010.

Table 2: Portrait of worst off according to different well-being measures (bottom 9.2%)

	full sample ($N = 14,027$)	WB^1 income	WB^2 comp. ind.	WB^3 satisfaction	WB^4 equiv. inc.	WB^5 vNM util.
satisfaction (0-10)	6.95	5.80	5.36	2.98	5.16	5.39
income (EUR/mth)	1,705	619	980	1,317	1,125	1,166
health index (0-100)	76.15	71.00	59.04	58.26	37.20	37.18
unemployed (in %)	6.44	32.93	69.91	16.75	24.44	21.84
male (in %)	45.93	44.05	(44.55)	(45.55)	(44.99)	41.59
age (years)	54.35	53.83	52.65	55.14	62.77	63.45
living with partner (in %)	56.54	33.45	39.72	47.09	38.32	50.38
divorced (in %)	11.49	24.64	22.60	17.39	20.04	16.84
separated (in %)	2.35	3.80	(2.91)	(2.23)	(2.18)	(1.48)
widowed (in %)	10.47	(10.68)	(8.16)	11.36	22.45	18.30
single parent (in %)	4.89	11.92	11.27	6.78	5.36	(4.08)
non German (in %)	6.09	13.03	10.45	(6.24)	(7.19)	(7.48)
East Germany (in %)	20.88	30.83	30.94	27.01	28.92	28.20
low education (in %)	17.55	36.95	30.31	23.96	29.31	28.87
pension (in %)	32.93	(31.73)	23.55	(34.20)	56.31	58.86
disabled (in %)	16.01	18.36	32.44	29.82	59.12	57.22
control over life (1-7)	5.23	4.56	4.49	4.20	4.35	4.34
achieved what deserved (1-7)	4.62	3.93	3.81	4.06	3.99	4.03
positive attitude (1-7)	5.54	5.22	5.08	4.54	5.09	5.13
frequency sad	2.45	2.78	2.97	3.50	3.23	3.17
frequency happy	3.40	3.00	2.87	2.36	2.71	2.81
risk taking (0-10)	4.11	3.88	(3.99)	3.50	3.55	(4.25)
WTP perfect health (EUR/mth)	2,250	1,358	3,257	3,673	5,764	5,762
WTP not unemployed (EUR/mth)	97	350	1,050	236	317	285

(.) denotes non-significant differences from the full sample mean (5% significance level).
"midlife" denotes an age between 45 and 60 years. *Source*: Own calculations based on the SOEP 2010.

Table 3: Worst off overlap across the different well-being measures

poor according to:	WB^1 income	WB^2 comp. ind.	WB^3 satisfaction	WB^4 equiv. inc.	WB^5 vNM util.	% overlap
1 measure	✓					9.20
1 measure		✓				9.21
1 measure			✓			9.20
1 measure				✓		9.20
1 measure					✓	9.21
2 measures	✓	✓				3.79
2 measures	✓		✓			2.04
2 measures	✓			✓		2.77
2 measures	✓				✓	2.43
2 measures		✓	✓			2.72
2 measures		✓		✓		5.02
2 measures		✓			✓	4.78
2 measures			✓	✓		3.17
2 measures			✓		✓	2.77
2 measures				✓	✓	7.36
3 measures	✓	✓	✓			1.31
3 measures	✓	✓		✓		2.26
3 measures	✓		✓	✓		1.17
3 measures	✓	✓			✓	2.02
3 measures	✓		✓		✓	1.02
3 measures	✓			✓	✓	2.15
3 measures		✓	✓	✓		2.01
3 measures		✓	✓		✓	1.90
3 measures		✓		✓	✓	4.59
3 measures			✓	✓	✓	2.54
4 measures	✓	✓	✓	✓		1.01
4 measures	✓		✓	✓	✓	1.91
4 measures	✓	✓		✓	✓	1.85
4 measures	✓	✓	✓		✓	0.94
4 measures		✓	✓	✓	✓	0.91
all measures	✓	✓	✓	✓	✓	0.87

Source: Own calculations based on the SOEP 2010.

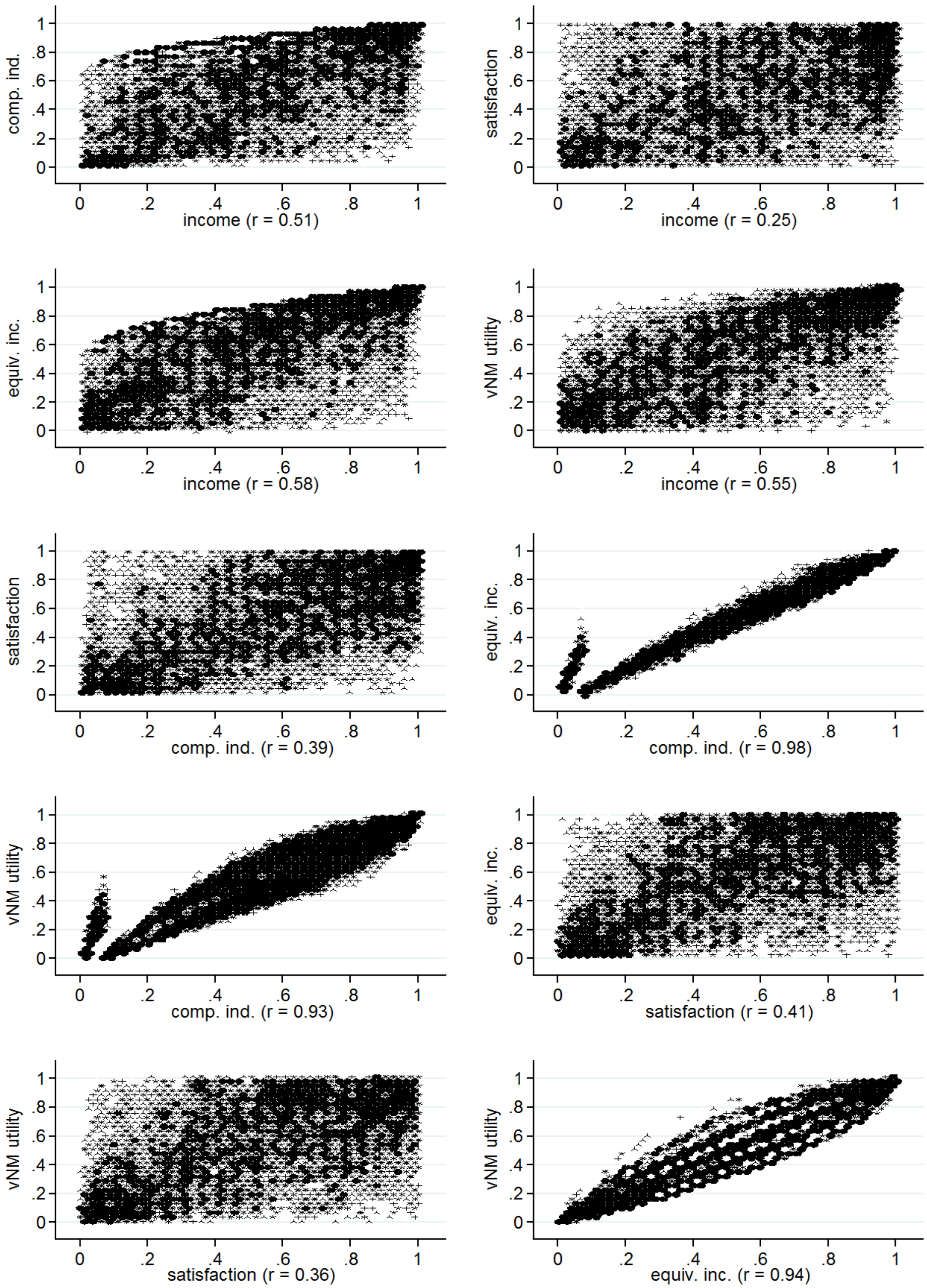

Figure 3: Re-ranking between each pair of considered well-being measures