Network for Studies on Pensions, Aging and Retirement

Netspar DISCUSSION PAPERS

Danilo Cavapozzi, Wei Han and Raffaele Miniaci

Alternative Weighting Structures for Multidimensional Poverty Assessment

DP 11/2013-082

Alternative weighting structures for multidimensional poverty assessment

Danilo Cavapozzi[a,b], Wei Han[c], Raffaele Miniaci[d]

[a] Ca' Foscari University of Venice, Department of Economics, Fondamenta San Giobbe, Cannaregio 873,30121 Venezia, Italy; Phone: +39 041 234 9180; E-mail: danilo.cavapozzi@unive.it
[b] Netspar, P.O. Box 90153, 5000 LE Tilburg, The Netherlands
[c] Blavatnik School of Government, University of Oxford, Albion House, Littlegate Street, OX1 1AN, Oxford, UK; E-mail: wei.han@bsg.ox.ac.uk
[d] University of Brescia, Department of Economics and Management, Via S.Faustino 74/b, 25122 Brescia, Italy; Phone: +39 030 298 8830; E-mail:miniaci@eco.unibs.it; (corresponding author)

November, 2013

Abstract

A multidimensional poverty assessment requires a weighting scheme to aggregate the well-being dimensions considered. We use Alkire and Foster's (2011a) framework to discuss the channels through which a change of the weighting structure affects the outcomes of the analysis in terms of overall poverty assessment, its dimensional and subgroup decomposability and policy prescriptions. We exploit the Survey on Health, Ageing and Retirement in Europe to evaluate how alternative weighting structures affect the measurement of poverty for the population of over 50s in ten European countries. Further, we show that in our empirical exercise the results based on hedonic weights estimated on the basis of life satisfaction self-assessments are robust to the presence of heterogeneous response styles across respondents.

Keywords: multidimensional poverty measurement, weights, life satisfaction, anchoring vignettes

JEL codes: I3, I32, C43, D63

1. Introduction

During the past decades it has been gradually recognized that the concept of well-being cannot be comprehensively captured by any conventional unidimensional indicator based on income, consumption or expenditure (Nussbaum, 2001; Sen, 1985). Focusing on a unique dimension keeps blind of the information about the overall life quality, of which it might be worthwhile for policy-makers to keep track given that pursuing well-being rather than wealth itself appears to be the ultimate goal of human society (Ruger 2010).

Although the multidimensional perspective on well-being measurement moves beyond the focus on a single indicator, it is still far from reaching an agreement on how to translate this perspective into practice. One of the complex and highly debatable issues emerging in a multidimensional context of well-being research lies in how to set the relative weights across the dimensions. Summarizing the achievements with respect to different well-being dimensions in a single indicator is needed to measure the diffusion of poverty, defined as pronounced deprivation in well-being, within a population.

This paper is aimed at showing how the adoption of different weighting schemes affects the outcomes of a multidimensional poverty study. We frame our analysis in the Alkire and Foster's multidimensional poverty framework (Alkire and Foster, 2011a). According to this approach, well-being dimensions are described by a set of one or more indicators. The achievements with respect to the whole battery of indicators can be aggregated into a single well-being score according to a weighting structure specified a priori. Poor households are those whose well-being scores fail to reach a minimum threshold. Alkire and Foster (2011a) propose a poverty measure, the adjusted headcount ratio, that reflects prevalence of poverty in the population and the intensity of the poverty among the poor. This measure can be decomposed in order to assess both the contribution of each dimension to overall poverty and how poverty varies across subgroups.

We analytically show that the variation in the adjusted headcount ratio induced by a change in the weighting scheme depends on the difference between weighting schemes, on the achievements of the households whose poverty status varies with the weighting structure adopted, and on the achievements of those that are classified as poor regardless of the weighting structure. The flows in and out of poverty can play a major role in explaining variation in multidimensional poverty measure. This result emphasizes a key characteristic of Alkire and Foster's approach. As the intensity of poverty is determined by the poverty level of the households classified as poor, the dimensional and subgroup decomposition reflects the achievements of this set of households by dimension or group. Different sets of poor households do not necessarily have the same achievements by dimension or group. Then, the change in the weighting scheme might have a direct effect on the results of dimensional and subgroup decompositions via a change in the weights but also an indirect effect induced by the change in the set of households classified as poor.

Finally, we provide a formal result showing that policy makers who want to reduce poverty by providing support in specific domains should not set their priorities merely accordingly to the weighting scheme adopted, but they should also take into account the distribution of the achievements in the population.

How poverty assessments and the choice of key-indicators for anti-poverty policies are affected by the weighting scheme remains an empirical issue. To this end, we draw data from the second wave of the Survey of Health, Ageing and Retirement in Europe (SHARE) to analyse the poverty of the elderly by a multidimensional poverty assessment exercise under alternative weighting schemes. The socio-economic inclusion of the elderly in Europe is an issue receiving an increasing attention by policy makers given the current and future demographic trends that are boosting the proportion of older individuals in European societies. Our analysis will investigate the socio-economic inclusion of the elderly in Europe according to a multidimensional approach explicitly designed at identifying the elderly individuals living in poverty and assessing the factors that contribute most to their status. SHARE is a cross-country dataset suited to analyze such issues since it is based on a representative sample of individuals aged 50 or more living in Europe and administers a multi-disciplinary questionnaire covering important aspects expected to be relevant determinants of socio-economic exclusion, such as health, employment, financial situation, social and family networks.

Decanq and Lugo (2013) surveyed three main classes of weights: normative, data-driven and hybrid. Normative weights are based on an explicit value judgment of analysts about the trade-offs between the well-being dimensions. Data-driven weights are based on the actual distribution of the achievements in the society with respect to the indicators of interest. Hybrid weights combine value judgements and statistical facts. They lie in the middle between defining weights by arbitrary decisions of analysts and letting data distribution be the only criterion used.

In our exercise, we follow the classification by Decanq and Lugo (2013) and choose one example for each of these classes. As for normative weights, we use equal weighting, which is the weighting scheme most widely used in measuring multidimensional well-being due to its simplicity[1]. We follow the Human Development Index and Multidimensional Poverty Index to assign equal weights to each dimension and equal weights to each indicator

[1] It has been employed in approximate 50% of literatures according to the summary provided by Decancq and Lugo (2013).

in each dimension (UNDP, 2011).[2] Within the class of data-driven weights, we adopt he frequency weights, which are motivated by the idea that, when assessing well-being, individuals put a high value on the shortfalls where the majorities do not fall short. We follow Desai and Shah (1988) to set the weight of a given indicator as the corresponding proportion of the non-deprived in the society. Finally, within the hybrid class, we choose the hedonic weights. Fleurbaey et al. (2009) propose to derive hedonic weights using life satisfaction self-assessments of respondents as a source of information for value judgments about well-being. Many social-science surveys ask respondents to rate their satisfaction with life according to a predetermined scale usually spanning from "very dissatisfied" to "very satisfied". Life satisfaction self-assessments have been widely used in the applied research focusing on well-being determinants (see for instance Frey and Stutzer, 2002 and Dolan et al., 2008). We follow Fleurbaey et al. (2009) and derive two alternative sets of hedonic weights by first regressing the life satisfaction self-assessments of respondents on a set of variables representing the indicators involved in the well-being score and, secondly, allowing for household and individual characteristics. The estimates of the coefficients on the indicators are used to derive their corresponding weights. Unlike equal weighting, in this approach the value judgement about dimension trade-offs is not set a priori by researchers but comes from the opinions of the individuals in the population of interest.

When dealing with self-reported life satisfaction data it is important to recognize that as a subjective measure, its variability across socioeconomic groups can be ascribed to genuine differentials in well-being (Schokkaert, 2007) as well as heterogeneity in the way in which individuals with different characteristics interpret the scale used to provide self-assessments. As an example, two individuals might have different expectations about the conditions that should realize to self-define as satisfied with their lives. Then, even if they experience the same level of well-being, they might produce different self-assessments due to their different reporting styles. Neglecting such heterogeneity when studying life satisfaction determinants might end up with assigning to an explanatory variable a biased role coming from the combination between its relationship with the reporting style used in life-satisfaction self-assessments and its actual role in explaining genuine differences in well-being (see King et al., 2004 and Angelini et al. 2012 and in press).

The second wave of SHARE provides us with a survey-instrument designed to take into account heterogeneity in the reporting styles used in collecting subjective data on life satisfaction. This approach is based on anchoring vignettes. A representative subsample of SHARE respondents is asked to report their own life satisfaction self-assessments along with

[2] It can be set equally either at the dimension level or at the indicator level. Inherently, it is an arbitrary approach regardless how we make it equal.

their assessments about the life satisfaction of hypothetical individuals described in vignettes kept constant across respondents. Differences in vignette-evaluations provided by respondents can then be of use to identify heterogeneity in reporting styles and disentangle such variability from actual differentials in well-being. We estimate a third set of hedonic weights based on the estimates of the relationship between indicators and life satisfaction self-assessments once heterogeneity in reporting styles is formally controlled for. Our results will be of help to shed light on the robustness of the results based on hedonic weights to the relaxation of the assumption of invariance of reporting styles in the population.

The paper is organised as follows. Section 2 describes Alkire and Foster's multidimensional poverty framework that our analysis is built on. In Section 3, we describe analytically the effects of changing the weighting scheme on the overall poverty measurement and on the prescriptions for the design of anti-poverty interventions. Section 4 is devoted to describe the approach we follow to derive the hedonic weights used in our analysis. Section 5 describes the data used in our empirical exercise and all the ingredients, but weights, involved in our application of the Alkire and Foster's multidimensional framework. Results are reported in Section 6. Section 7 presents our conclusions and policy implications.

2. A framework for the multidimensional poverty assessment

Sen (1976) concisely summarized two problems that must be faced in the poverty measurement: (1) the identification problem, i.e. how to choose the criterion of poverty and then distinguish those who fall into that criterion and those who do not; and (2) the aggregation problem, i.e. how to construct a poverty index using the available information on the poor. Dealing with these two issues is particularly challenging in a multidimensional framework. Alkire and Foster (2011a) tackle the identification problem by defining indicator-specific thresholds – which refer to specific achievements – and an overall threshold, which refers to a comprehensive well-being score based on the achievements. Moreover, they adopt the Foster, Greer and Thorbecke (FGT, 1984) framework to handle the aggregation problem and deliver a methodology satisfying desirable properties for poverty measurement. Compared with other multidimensional poverty frameworks, the Alikire and Foster's strategy is superior since it allows exploiting the information coming from achievements measured even on ordinal and categorical scales.

In this paper we investigate the sensitivity of the multidimensional poverty index to alternative weighting structures in Alkire and Foster's framework, which consists of five key ingredients: the dimensions, the indicators, the indicator-specific thresholds, the overall threshold and the weighting structure. The overall well-being is assessed with respect to a set of dimensions. As far as which dimensions should be considered in the research and how to choose them, Sen (2004) gave the following guidelines: (1) focus on those that are of special importance to the society or people in question; (2) focus on those that are an appropriate focus for public policy, rather than a private good or a capability that cannot be influenced from outside. For each dimension, one or more indicators are selected to describe the achievements of the households with respect to particular aspects of the dimension in hand. The nature and the number of the indicators used in each dimension depend on data availability and research purposes. Each indicator has its own threshold that determines whether the household meets the minimum standard with respect to that indicator; if not, we say the household is deprived in that indicator. Every household is therefore characterized by an achievement vector, each entry of which records the achievement status (dichotomised) of a given indicator. These entries are aggregated into a well-being score defined as the weighted sum of the single achievement statuses and measuring the household's overall well-being. Consequently, a household is identified as poor if and only if its overall well-being score does not reach a given threshold. As far as the identification problem got solved, a FGT poverty measure can be applied to generate a multidimensional poverty index. Evidently, poverty assessment depends on the (questionable) choice of these five components. In this paper we investigate to what extent the poverty assessment can be influenced by the choice of the weighting structures.

The units of our analysis are n households. Each dimension, $j = 1,2 \dots d,$ is described by d_j indicators, y_{jk}^h denotes the achievement of household h on the k-th indicator of the dimension j and every indicator has its own threshold, z_{jk}, to indicate the minimum standard. Let w_{jk} be the weight of the k-th indicator belonging to dimension j, the weights sum up to 1, i.e. $\sum_{j=1}^{d} \sum_{k=1}^{d_j} w_{jk} = 1$. The Alkire and Foster's three-step identification procedure works as follows.

Step 1: dichotomise achievement with respect to each indicator for every household

The indicator-specific achievement status of household h in terms of the k-th indicator belonging to dimension j, a_{jk}^h, is defined as

$$a_{jk}^h = \mu\left(y_{jk}^h > z_{jk}\right) \tag{1}$$

where $\mu(\cdot) = 1$ if expression is true; otherwise $\mu(\cdot) = 0$.

Step 2: define a household well-being score as the aggregation of the achievement statuses

Denote the overall well-being score of household h as s^h:

$$s^h = \sum_{j=1}^{d} \sum_{k=1}^{d_j} a_{jk}^h \cdot w_{jk} \tag{2}$$

Step 2 solves the problem of aggregation across indicators as well as dimensions by defining a scalar well-being function as weighted average of $\sum_{j=1}^{d} d_j$ achievement statuses. We thus refer to the debate in the well-being indices literature for the choice, interpretation and estimation of the weighting schemes (e.g. Decancq and Lugo, 2013; Schokkaert, 2007). For our purposes, it is relevant to stress that $a_{jk}^h = \mu\left(y_{jk}^h > z_{jk}\right)$ is not a derivable function of the achievement y_{jk}^h, and therefore the ratio of two weights cannot be interpreted as the marginal rate of substitution between two indicators. Nevertheless, the ratio of the weights has a natural interpretation. Considering the achievement statuses of the household h with respect to the indicator *m* of dimension *p* and the indicator *l* of dimension *q* and name them with a_{pm}^h and a_{ql}^h, respectively. Household h's well-being score can be written as

$$s_{(a_{pm}^h, a_{ql}^h)}^h = a_{pm}^h w_{pm} + a_{ql}^h w_{ql} + \sum_{jk} a_{jk}^h w_{jk} \tag{3}$$

where $\forall jk \neq pm, ql$. For instance, $s_{(1,0)}^h$ is the well-being score of household h if it achieved the minimum standard in the indicator *p* of the dimension *m* but not in the indicator *q* of the dimension *l*. Then,

$$\frac{w_{pm}}{w_{ql}} = \frac{s_{(1,1)}^h - s_{(0,1)}^h}{s_{(1,1)}^h - s_{(1,0)}^h} = \frac{s_{(1,0)}^h - s_{(0,0)}^h}{s_{(0,1)}^h - s_{(0,0)}^h} \tag{4}$$

The ratio of the weights between pm and ql is therefore equal to the ratio of the changes in the well-being score due to obtaining the achievement in the former indicator rather than the latter, all the other indicators being constant.

Step3: identify the poor

Those households whose well-being score is lower than an arbitrarily chosen well-being threshold φ, with $\varphi \in (0,1)$, are identified as poor. The poverty status of household *h*, P^h, is defined as

$$P^h = \mu\left(s^h < \varphi\right) \tag{5}$$

Until now, we have solved the identification problem. The standard method to overcome the aggregation problem is to refer to the poverty incidence measured by the headcount ratio:

$$H = \frac{1}{n}\sum_{h=1}^{n} P^h \tag{6}$$

However, this index violates the dimensional monotonicity property, that is, other things being constant, if the shortfall[3] of those identified as 'poor' varies, the headcount index remains unchanged. As early as in the Sen's well-known paper on the ordinal approach to poverty measurement (Sen, 1976), monotonicity has been listed as one of the most important axioms that a valid poverty index should satisfy. The adjusted headcount ratio M, proposed by Alkire and Foster (2011a) in the multidimensional framework, satisfies the monotonicity axiom by combining information on the incidence of the poor in the population with the degree of poverty among the poor. The former is measured by the headcount ratio H; the latter is measured by the average shortfall among the poor:

$$A = \frac{\sum_{h=1}^{n} P^h (1 - s^h)}{\sum_{h=1}^{n} P^h} \tag{7}$$

Formally, the adjusted headcount ratio is

$$M = HA = \frac{1}{n}\sum_{h=1}^{n} P^h \,(1 - s^h) \tag{8}$$

which also represents the total shortfall experienced by the poor ($\sum_{h=1}^{n} P^h (1 - s^h)$) divided by the maximum shortfall that could be experienced by the entire population. In fact, when none of the households meet any minimum standard with respect to any indicator, $s^h = \sum_{j=1}^{d} \left(\sum_{k=1}^{d_j} a_{jk}^h \cdot w_{jk} \right) = 0$ and thus $\sum_{h=1}^{n} P^h (1 - s^h) = n$.

The adjusted headcount ratio satisfies a range of desirable properties for poverty indexes[4] (see Alkire and Foster, 2011a). In particular, the decomposability by dimension and country is of significant importance to policy makers.

Subgroup decomposability: The overall poverty is the weighted average of subgroup poverty, where weights are subgroup population shares. Formally, suppose population can be divided into q groups. Let θ_g be the population share of subgroup g. Denote M_g as the adjusted headcount ratio of subgroup g, so we have $M = \sum_{g=1}^{q} \theta_g M_g$. The contribution of subgroup g to the overall poverty $RG_g = \theta_g M_g / M$ takes the population share into account.

[3] Note that, in this paper, shortfall refers to the gap between the actual well-being score s^h and the full well-being score (i.e. when all the minimum standards are met, and $s^h = 1$) rather than to the gap with respect to the well-being threshold φ.

[4] It satisfies the following properties: replication invariance, symmetry, poverty and deprivation focus, weak and dimensional monotonicity, non-triviality, normalization, weak rearrangement and decomposability.

Dimensional decomposability: The overall poverty is the weighted average of the *censored* deprivation regarding each indicator, where the weights are the indicator-specific weights. The term "censored" is used to emphasize the fact that the indicator-specific deprivations of those who are classified as non-poor are not taken into account. Let I_{jk} be the censored deprivation of indicator k belonging to dimension j,

$$I_{jk} = \frac{1}{n}\sum_{h=1}^{n} P^h \,(1 - a_{jk}^h) \tag{9}$$

the censored deprivation with respect to dimension j is

$$V_j = \sum_{k=1}^{d_j} I_{jk} \cdot w_{jk} \tag{10}$$

and the overall adjusted headcount ratio can be written as

$$M = \sum_{j=1}^{d} V_j \tag{11}$$

Therefore, the proportional contribution of deprivations in dimension p to the overall poverty is $RD_p = V_p/M$. This offers a decomposition of M by dimension and thus can shed light on the sources of poverty.

3. The influence of weights on multidimensional poverty assessment

One of the main ingredients of the Alkire and Foster's approach is the weighting structure adopted for the aggregation across different dimensions. Changing the weights potentially has influence on who is identified as poor, on the headcount and the adjusted headcount ratios, and on the dimensional as well as subgroup decompositions.

In this section we analytically investigate how the variation of the weighting scheme affects the adjusted headcount ratio and show that the effects of such changes cannot be unambiguously predicted in advance by looking only at the differences in the weights and at the outcomes of the households originally classified as poor. The analytical analysis shows that the households whose poverty status varies with the weighting scheme play a significant role in such changes: (1) their population share matters; (2) their achievement statuses of all the indicators are relevant. Besides, the households who are always identified as poor play an active role. Both their share and their achievement statuses in the indicators with varying weights affect the variation in the adjusted headcount ratio.

Finally, from a policy perspective, we analogously show that a policy eliminating the deprivation with respect to the indicator associated with the highest weight is not necessarily the most effective anti-poverty strategy. Rather, policy makers should base interventions considering not only the weight of indicators but also the number of households who can exit poverty after the intervention as well as their outcomes.

3.1 Effects of weighting scheme variations on the adjusted headcount ratio

All other things being held constant, assume the weight of indicator p increases from w_p to w_p'. We further assume, without losing generality, the weight of indicator q decreases from w_q to w_q':[5]

$$w_p' - w_p = w_q - w_q' = \Delta w_{pq} > 0 \tag{12}$$

For the generic household h, the well-being scores before (s^h) and after ($s^{h\prime}$) the change of weights, respectively, can be written as

$$\begin{aligned} s^h &= a_p^h w_p + a_q^h w_q + \sum_{j \neq p,q} a_j^h w_j \\ s^{h\prime} &= a_p^h w_p' + a_q^h w_q' + \sum_{j \neq p,q} a_j^h w_j \end{aligned} \tag{13}$$

We name M the adjusted headcount ratio with the original set of weights and M' its counterpart obtained with the new weighting scheme. Our aim is to analyze the determinants of the variation in the adjusted headcount ratio $\Delta M = M' - M$.

Let us consider a generic household h. In terms of the achievement status of indicator p and q, there are four possible combinations: $\left(a_p^h, a_q^h\right) = \{(0,0), (0,1), (1,0), (1,1)\}$. For those who meet the minimum standard in both or neither of the indicators, i.e. $\left(a_p^h, a_q^h\right) = (1,1)$ or $\left(a_p^h, a_q^h\right) = (0,0)$, we have $s^h = s^{h\prime}$. That is to say, the change of weights in indicator p and q has no effect on the well-being score, let alone the aggregated poverty indices. Therefore, we only need to consider those households who meet the minimum standard in only one of the two indicators, i.e. $\left(a_p^h, a_q^h\right) = (0,1)$ or $\left(a_p^h, a_q^h\right) = (1,0)$. The households relevant for our purposes can be grouped into the following three cases.

<u>*Case 1: households entering poverty only after the change in the weighting scheme*</u>

This case can be observed only when $\left(a_p^h, a_q^h\right) = (0,1)$. By contrast, if a household h is not poor according to the original set of weights and has $\left(a_p^h, a_q^h\right) = (1,0)$, its well-being

[5] As the dimensions do not play any significant role in the current section, in what follows we suppress the index referring to them in order to simplify the notation.

score will increase by construction with the new weighting scheme and it will keep on being not poor.

The change in the contribution of household h to the adjusted headcount ratio is

$$\begin{aligned} P^{h'}\left(1-s^{h'}\right)-P^{h}\left(1-s^{h}\right)=P^{h'}\left(1-s^{h'}\right)&=1-\left(a_p^h w_p'+a_q^h w_q'+\sum_{j\neq p,q} a_j^h w_j\right) \\ &=1-\left(w_q'+\sum_{j\neq p,q} a_j^h w_j\right) \end{aligned} \tag{14}$$

It is worth noting that Eq.(14) does not depend just on the new weight $\left(w_q'\right)$ on indicator q but on the achievements and the corresponding weights for all the other indicators whose weights did not change across weighting schemes $\left(\sum_{j\neq p,q} a_j^h w_j\right)$. This amounts to say that looking at the achievement and at the new weight for indicator q is not sufficient to predict the variation in the contribution of household h to the overall poverty measure.

Let us suppose that the number of households in this group is equal to $n^{NP\to P}$. If we focus on this group only, the variation in the adjusted headcount ratio is equal to

$$\begin{aligned} &\frac{1}{n^{NP\to P}}\left(\sum_{h\in(NP\to P)} P^{h'}\left(1-s^{h'}\right)-\sum_{h\in(NP\to P)} P^{h}\left(1-s^{h}\right)\right) \\ &=\frac{1}{n^{NP\to P}}\sum_{h\in(NP\to P)} P^{h'}\left(1-s^{h'}\right) \\ &=\frac{1}{n^{NP\to P}}\left(\sum_{h\in(NP\to P)}\left(1-w_q'\right)-\sum_{h\in(NP\to P)}\sum_{j\neq p,q} a_j^h w_j\right)=M'_{NP\to P} \end{aligned} \tag{15}$$

Finally, the aggregate contribution of all the households in this group to the overall variation in the adjusted headcount ratio due to change in the weighting scheme is

$$\frac{n^{NP\to P}}{n} M'_{NP\to P} \tag{16}$$

Case 2: households exiting poverty after the change in the weighting scheme

This case is possible only when $\left(a_p^h, a_q^h\right)=(1,0)$. By contrast, if a household is originally classified as poor and the indicator in which it achieves the minimum standard is appreciated less, i.e. $\left(a_p^h, a_q^h\right)=(0,1)$, it will be still classified as poor according to the new weighting scheme.

The change of the contribution of household h to the adjusted headcount ratio is

$$
\begin{aligned}
P^{h'}\left(1-s^{h'}\right)-P^{h}\left(1-s^{h}\right) &= -P^{h}\left(1-s^{h}\right) \\
&= -\left(1-a_p^h w_p - a_q^h w_q - \sum_{j\neq p,q} a_j^h w_j\right) \\
&= -\left(1-w_p-\sum_{j\neq p,q} a_j^h w_j\right)
\end{aligned} \tag{17}
$$

Analogously to case 1, this variation is not only affected by the old weight on indicator p $\left(w_p\right)$ but it is also affected by the achievements and corresponding weights for indicators whose weights do not change $\left(\sum_{j\neq p,q} a_j^h w_j\right)$.

If this group includes $n^{P\to NP}$ households and we focus on the variation in the adjusted headcount ratio for this group only, it can be defined as

$$
\begin{aligned}
&\frac{1}{n^{P\to NP}}\left(\sum_{h\in(P\to NP)} P^{h'}\left(1-s^{h'}\right)-\sum_{h\in(P\to NP)} P^{h}\left(1-s^{h}\right)\right) \\
&\qquad = -\frac{1}{n^{P\to NP}}\left(\sum_{h\in(P\to NP)} P^{h}\left(1-s^{h}\right)\right) \\
&\qquad = -\frac{1}{n^{P\to NP}}\left(\sum_{h\in(P\to NP)}\left(1-w_p\right)-\sum_{h\in(P\to NP)}\sum_{j\neq p,q} a_j^h w_j\right) = -M_{P\to NP}
\end{aligned} \tag{18}
$$

The aggregate contribution of the households in the group to the overall variation in the adjusted headcount ratio is

$$
-\frac{n^{P\to NP}}{n} M_{P\to NP} \tag{19}
$$

<u>*Case 3: households who are poor under both weighting schemes*</u>

The change of the contribution of household h to the adjusted headcount ratio is

$$
\begin{aligned}
P^{h'}\left(1-s^{h'}\right)-P^{h}\left(1-s^{h}\right) &= \left(a_p^h w_p + a_q^h w_q\right)-\left(a_p^h w_p' + a_q^h w_q'\right) \\
&= a_p^h\left(w_p-w_p'\right)+a_q^h\left(w_q-w_q'\right) = \Delta w_{pq}\left(a_q^h-a_p^h\right) \\
&= \begin{cases} \Delta w_{pq}, & if\ \left(a_p^h, a_q^h\right)=(0{,}1) \\ -\Delta w_{pq}, & if\ \left(a_p^h, a_q^h\right)=(1{,}0) \end{cases}
\end{aligned} \tag{20}
$$

Unlike the former two cases, the variation in the contribution of the households in this group depends only on the achievements in the indicators whose weights vary and on the change in their weights.

Let us name $n^{P\to P}$ the sample size of this group, its variation in the adjusted headcount ratio variation is equal to

$$\frac{1}{n^{P\to P}}\left(\sum_{h\in(P\to P)} P^{h'}\left(1-s^{h'}\right)-\sum_{h\in(P\to P)} P^{h}\left(1-s^{h}\right)\right)=M'_{P\to P}-M_{P\to P} \qquad (21)$$

The aggregate contribution of the households in this group to the overall variation in the adjusted headcount ratio is

$$\frac{n^{P\to P}}{n}\left(M'_{P\to P}-M_{P\to P}\right) \qquad (22)$$

For sake of completeness, we point out that the contribution of all the households who are not poor according to both weighting schemes is null since $P^{h'}\left(1-s^{h'}\right)=P^{h}\left(1-s^{h}\right)=0$.

Now we are in the position of providing an analytical expression describing the determinants of the overall variation in the adjusted headcount ratio ΔM due to a change in the weighting scheme. Summarizing the results obtained so far in the case-by-case analysis, we can write

$$\Delta M=\left(\frac{n^{NP\to P}}{n}M'_{NP\to P}-\frac{n^{P\to NP}}{n}M_{P\to NP}\right)+\frac{n^{P\to P}}{n}\left(M'_{P\to P}-M_{P\to P}\right)=\Delta M_1+\Delta M_2 \qquad (23)$$

Even if we consider the simplest case in which only two weights change, the induced variation in the adjusted headcount ratio is not a matter involving only the weights that vary and the associated achievements $\left(w_p, w_p', w_q, w_q', a_p^h, a_q^h\right)$. Indeed, it depends on the summations of the achievements on the indicators whose weights are invariant calculated over the households who enter or exit poverty when passing from one weighting structure to the other ($\sum_{h\in(NP\to P)}\sum_{j\neq p,q} a_j^h w_j$ and $\sum_{h\in(P\to NP)}\sum_{j\neq p,q} a_j^h w_j$, respectively). Moreover, the proportion of households changing poverty status ($\frac{n^{NP\to P}}{n}$ and $\frac{n^{P\to NP}}{n}$) as well as the proportion of those staying in poverty ($\frac{n^{P\to P}}{n}$) matter. The ΔM_1 terms then reflects a composition effect driven by the households who change their poverty status from one weighting scheme to the other. The term ΔM_2 is instead driven by the changes in the weighting scheme and the outcomes of the households who are poor regardless of the weighting scheme used. Consequently, it is impossible to predict the sign of ΔM without knowing the underlying distribution of all the indicators and identifying the set of households remaining poor and the sets of those switching their poverty status.

Finally, there are two more remarks on the Eq.(23). Firstly, it holds regardless of the relationship between w_p and w_q. Put differently, even if we further assume $w_p > w_q$, i.e. we shift weight away from the indicator that has already weighted relatively less to the one that has already weighted more, the sign of ΔM will still be determined by Eq.(23). Secondly, it holds even when more than two weights are allowed changing and then the entire vector of weights changes from w to w'. Although the structure of $M'_{NP\rightarrow P}, M_{P\rightarrow NP}, M'_{P\rightarrow P}, M_{P\rightarrow P}$ will change according to the variation in the weighting structure object of study, these factors will still be combined according to Eq.(23) in order to identify the variation in the adjusted headcount ratio.

3.2 Effects of the weighting scheme on the design of anti-poverty interventions

Policy makers might be interested in selecting the indicators, or the set of indicators, on which they should intervene to set-up the most efficient anti-poverty intervention. We show that in a multidimensional framework choosing the set of such indicators being based on indicator weights only might lead to inefficient programs. It turns out that providing relief for the deprivation in an indicator with a higher weight is not necessarily more effective to reduce the adjusted headcount ratio than providing relief in an indicator with a lower weight.

Let us consider the simplest case in which the intervention is focused on a single indicator p. The policy consists of making all households in the population able to achieve the minimum threshold. Our aim is to predict the variation in the adjusted headcount ratio due to this policy. If we name M^0 the adjusted headcount ratio before the intervention and M^1 its counterpart after the intervention, $\Delta M^{int} = M^1 - M^0$ is the variation of interest.

For the generic household h, the well-being scores before (s^{h^0}) and after (s^{h^1}) the intervention in indicator p, respectively, can be written as

$$\begin{aligned} s^{h^0} &= a_p^{h^0} w_p + \sum_{j \neq p} a_j^h w_j \\ s^{h^1} &= a_p^{h^1} w_p + \sum_{j \neq p} a_j^h w_j \end{aligned} \tag{24}$$

The change in the contribution to the variation in the adjusted headcount ratio is null if the household, albeit poor before and after the intervention, meets the minimum standard in indicator p before the intervention since its condition is unchanged by the intervention. Also, the contribution to the adjusted headcount ratio of the households which are not poor before the intervention is null, regardless of the original achievement status in the indicator p. By construction, the intervention cannot make them poor. To sum up, only those households who did not meet the minimum standard of indicator p and were identified as poor before the intervention are able to contribute to reduce the adjusted headcount ratio after intervention. They can be grouped into two cases.

Case 1: households still in poverty after the intervention

Assume household h meets the minimum standard with respect to the indicator p after the intervention, but this is not sufficient to make it escape from poverty. The change of household h's contribution to the adjusted headcount ratio is

$$P^{h^1}(1-s^{h^1}) - P^{h^0}(1-s^{h^0}) = -w_p \tag{25}$$

If this group consists of $m^{P\to P}$ households, their contribution to the overall variation in the adjusted headcount ratio will be equal to

$$-\frac{m^{P\to P}}{n}w_p \tag{26}$$

Case 2: households exiting poverty after the intervention

The change of the contribution of household h in this group to the adjusted headcount ratio is:

$$\begin{aligned} P^{h^1}\left(1-s^{h^1}\right) - P^{h^0}\left(1-s^{h^0}\right) = -\left(1-s^{h^0}\right) = -\left(1-a_p^h w_p - \sum_{j\neq p} a_j^h w_j\right) \\ = -\left(1-\sum_{j\neq p} a_j^h w_j\right) \end{aligned} \tag{27}$$

Suppose the number of households in this group is $m^{P\to NP}$, their variation in the adjusted headcount ratio is

$$-\frac{1}{m^{P\to NP}}\sum_{h\in(P\to NP)}\left(1-\sum_{j\neq p} a_j^h w_j\right) = -M_{P\to NP}^p \tag{28}$$

The aggregated contribution of this group to the overall variation in the adjusted headcount ratio is

$$-\frac{m^{P\to NP}}{n}M_{P\to NP}^p \tag{29}$$

Therefore, the overall change in the adjusted headcount ratio due to the policy support in indicator p is

$$\Delta M^{int} = -\frac{1}{n}\left(m^{P\to P}w_p + m^{P\to NP}M_{P\to NP}^p\right) \tag{30}$$

If the government is aimed at reducing the adjusted headcount ratio, it should focus on the indicator maximizing the whole term in brackets, that this is not necessarily the one with the highest weight. Clearly, the second term in the brackets is larger the higher is the number of households exiting poverty thanks to the intervention. This suggests that policy maker should also be concerned with identifying those indicators whose relief maximizes the exits from poverty and not restrict the attention to the weighting scheme.

4 Deriving weights with the hedonic approach

The hedonic approach of deriving a weighting scheme is hybrid since it combines value judgements about trade-offs among dimensions, as it is typical in the normative weighting, with statistical facts. We follow Fleurbaey et al. (2009) and use life satisfaction self-assessments of respondents to elicit value judgements about trade-offs between well-being dimensions. A widely used approach to measure well-being in applied research is to ask individuals to evaluate their life satisfaction according to a predetermined scale, e.g. by answering the question "How satisfied are you with your life in general?". Self-assessments are measured according to an ordinal scale, such as "Very dissatisfied", "Dissatisfied", "Neither satisfied nor dissatisfied", "Satisfied", "Very satisfied". Frey and Stutzer (2002) and Dolan et al. (2008) survey the main findings of the empirical research on the determinants of life satisfaction self-assessments.

We estimate a first set of hedonic weights by running an ordered probit regression having the life satisfaction self-assessments of individuals as dependent variables and their achievements with respect to the well-being indicators considered as explanatory variables. We indicate the standardized weight for indicator k in dimension j with w_{jk}. This is retrieved from the corresponding estimated coefficients $\hat{\beta}_{jk}$ in the ordered probit equation by

$$w_{jk} = \frac{\hat{\beta}_{jk}}{\sum_{i=1}^{d} \sum_{l=1}^{d_i} \hat{\beta}_{il}} \quad j = 1, \dots, d; k = 1, \dots, d_j \tag{31}$$

Second, we estimate a second set of standardized hedonic weights by enriching the explanatory variables showing up in the ordered probit regression with household and individual characteristics. Controlling for all these factors "is necessary to 'clean' the happiness measure to separate the 'ethically' relevant information from the irrelevant noise" (Shokkaert, 2007).

On the one hand, life satisfaction self-assessments have the advantage of summarizing in a single index all the factors that individuals consider relevant determinants of their well-being. On the other hand, a recent research vein (Angelini et al. 2012 and in press, Kapteyn et al., 2009) has shown that the benchmarks used to self-evaluate life satisfaction are not invariant across individuals but depend on their own characteristics. Even if individuals are asked to self-evaluate their own life satisfaction according to the same survey question, they might provide different evaluations due to inter-personal and inter-cultural heterogeneity in the interpretation of the response scale. Furthermore, a phenomenon of adaptation might be at work. In fact, individuals may adjust their aspiration levels to their realistic opportunities (Schokkaert, 2007). In psychometrics such heterogeneity has been called differential item functioning (DIF). If DIF is an issue, life satisfaction self-assessments fail to be comparable across individuals or socioeconomic groups since their differences might not reflect actual differences in well-being but only differences in the reporting styles adopted by respondents. Individuals with the same actual level of well-being might provide different life satisfaction self-evaluations because they have in mind different concepts about what being satisfied with their life means. As a consequence, the presence of DIF implies that a welfare analysis based on the comparison of life satisfaction self-evaluations should take into account heterogeneity in reporting styles in order to provide meaningful results.

This paper takes advantage of the SHARE data to control for DIF by a vignette methodology. After having provided life satisfaction self-assessments, a subsample of SHARE respondents are asked to evaluate the life satisfaction of two hypothetical individuals described in particular situations (anchoring vignettes), which are reported below.

1. John is 63 years old. His wife died 2 years ago and he still spends a lot of time thinking about her. He has 4 children and 10 grandchildren who visit him regularly. John can make ends meet but has no money for extras such as expensive gifts to his grandchildren. He has had to stop working recently due to heart problems. He gets tired easily. Otherwise, he has no serious health conditions. How satisfied with his life do you think John is?
2. Carry is 72 years old and a widow. Her total after tax income is about € 1,100 per month[6]. She owns the house she lives in and has a large circle of friends. She plays bridge twice a week and goes on vacation regularly with some friends. Lately she has been suffering from arthritis, which makes working in the house and garden painful. How satisfied with her life do you think Carry is?

[6] This values is PPP-adjusted to account for cross-country differentials in price levels.

Respondents' evaluations of vignettes are recorded according to the same response scale used for their self-assessments ("Very dissatisfied", "Dissatisfied", "Neither satisfied nor dissatisfied", "Satisfied", "Very satisfied").

The situations described in the vignettes do not vary across respondents, who are also explicitly asked to evaluate the vignettes according to their own preferences. Differences in the evaluations of the anchoring vignettes can be ascribed to the heterogeneity in the reporting styles of respondents and be of use to filter the life satisfaction self-assessments of respondents from DIF as long as respondents use the same reporting style when assessing the life satisfaction of themselves and of the hypothetical individuals described in the vignettes (response consistency) and the life satisfaction of the hypothetical individuals in the vignettes is on average perceived by respondents in the same way (vignette equivalence).

More specifically, we analyze the determinants of life satisfaction and control for the presence of DIF by the hierarchical ordered probit (Hopit) model introduced by King et al. (2004). This econometric specification consists of two components modeling self-assessments and vignette evaluations as ordered variables.

Self-assessment component

Let Y_i^* be the life satisfaction perceived by individual $i = 1, \dots, I$ and assume that it comes from a linear combination of individual characteristics stored in the vector X_i and an error term $\varepsilon_i \sim N(0,1)$ independent of X_i,

$$Y_i^* = X_i \beta^H + \varepsilon_i \tag{32}$$

where β^H is a vector of unknown parameters. In our case, the vector X_i includes the well-being indicators as well as the household and individual characteristics used to derive the second set of hedonic weights.

Although Y_i^* cannot be observed, we know individual's life satisfaction self-evaluation Y_i, which is coded as an ordered discrete variable spanning from 1 ("Very dissatisfied") to 5 ("Very satisfied"). We can write

$$Y_i = j \quad \text{if} \quad \tau_i^{j-1} \leq Y_i^* \leq \tau_i^j, \quad j = 1, \dots, 5 \tag{33}$$

The thresholds τ_i^j are individual-specific and depend on the individual characteristics X_i

$$\tau_i^0 = -\infty, \quad \tau_i^5 = \infty \tag{34}$$

$$\tau_i^1 = X_i \gamma^1 \tag{35}$$

$$\tau_i^j = \tau_i^{j-1} + exp\left(X_i \gamma^j\right), \quad j = 2,3,4 \tag{36}$$

where γ^j are vectors of unknown parameters. The set of thresholds τ_i^j formally allows individuals with different characteristics to provide different self-evaluations Y despite the same perceived level of life satisfaction Y^*. The Hopit model can then be seen as a generalization of the standard ordered probit specification, which restricts the thresholds to be invariant across individuals and implicitly assumes that reporting styles adopted by individuals do not depend on their own characteristics.

The information conveyed by life satisfaction self-evaluations is not sufficient to disentangle the effect of the individual characteristics X_i on Y_i^* and their effect on the thresholds τ_i^j. To achieve this goal, we make use of vignette evaluations.

Vignette evaluation component

Let Z_{il}^* be the life satisfaction of the hypothetical person in vignette $l = 1{,}2$ perceived by individual i. We assume that

$$Z_{li}^* = \theta_l + v_{il} \tag{37}$$

$v_{il} \sim N(0, \sigma_l^2)$ and v_{il} is independent of ε_i and X_i. The parameter θ_l is assumed to be vignette-specific and invariant across individuals. This restriction follows from the vignette equivalence assumption, according to which respondents have the same perception of the life satisfaction of the hypothetical person in the vignette, up to an individual idiosyncratic error term.

Again, we cannot observe the perception Z_{il}^* but we know the evaluation Z_{il}, defined as

$$Z_{il} = j \quad \text{if} \quad \tau_i^{j-1} \le Z_{il}^* \le \tau_i^j, \quad j = 1, \dots, 5 \tag{38}$$

The thresholds τ_i^j are those used to derive the life satisfaction self-assessments. This results from imposing the response consistency assumption, which ensures that respondents use the same reporting style when evaluate themselves and the hypothetical persons in the vignettes.

The self-assessment and vignette evaluation components are connected by the use of the same set of individual-specific thresholds τ_i^j. This implies that we can combine the information relevant to estimate these two specifications in order to identify all the parameters of interest in the Hopit model. Along the lines of King et al. (2004), the joint estimation can be carried out by maximum likelihood techniques. Life satisfaction self-evaluations serve to identify the parameters in β^H, as it would happen if we were estimating a standard ordered probit equation, whereas the vignette evaluations are needed to identify $\theta_l, l = 1,2$ and $\gamma_j, j = 1,2,3,4$.

Once the Hopit model has been estimated, we derive a third set of standardized hedonic weights for the well-being indicators being based on the estimates of their coefficients in the vector β^H. This vector of hedonic weights is expected to reflect the relationships between achievement in the indicators and well-being once their effect on reporting styles in life satisfaction self-assessments has been filtered out.

5 Data, dimensions, indicators and thresholds

In this paper we use data from the 2006 wave of the Survey of Health, Ageing and Retirement in Europe (SHARE). SHARE is an interdisciplinary survey on ageing that is run every two years and collects extensive information on health, socioeconomic status and family interactions of individuals aged 50 and over in a host of European countries. The choice of using SHARE rather than other well established surveys (e.g. EU SILC for the European countries) is dictated by the fact that SHARE collects self-assessments and anchoring vignette evaluations on life satisfaction and makes it possible to implement the formal econometric framework discussed in the previous section in order to evaluate to what extent the results coming from hedonic weights based on the life satisfaction self-assessments of respondents are robust to the presence of heterogeneity in response styles.

Data are collected by face-to-face, computer-aided personal interviews (CAPI), supplemented by a self-completion paper and pencil questionnaire, which collects self-assessments and vignette evaluations on life satisfaction. We select only those respondents who provide both the self-evaluation and at least one vignette evaluation. Our final estimation sample for the hedonic weights is composed by 3,804 households, corresponding to 5,545 individuals living in Sweden, Denmark, Germany, The Netherlands, Belgium, France, Greece, Italy, Spain and Czech Republic[7].

Different multidimensional poverty indexes consider alternative sets of dimensions due to differences in theoretical perspectives, reference population and data limitation. Material deprivation, health conditions, educational attainments, empowerment, labour market participation, environmental quality, safety from violence, and social relationships are all relevant domains and their relevance has been assessed for the European Union population (Eurostat 2012).

[7] We restrict our sample to the countries in which vignette data have been collected with the exception of Poland, for which some of the data used for the analysis show some inconsistency with respect to the rest of the sample.

In our illustrative exercise we focus on a representative sample of elderly respondents living in ten European countries and we consider three dimensions to represent the main drivers of their well-being: economic situation, housing and health conditions. The economic dimension is meant to describe the monetary resources available to the household. It includes two indicators: per-capita net income and per-capita net wealth. The thresholds for income and wealth indicators are set equal to 60% of the country specific median values. By doing so, we follow Stiglitz's Commission suggestions to consider both income and wealth. The housing dimension has one indicator, a measure of accessibility of the dwelling given by the number of steps people have to climb up and/or down to the entrance of their home. The architectural barriers of the accommodation are potentially relevant for the population we consider, as ageing is often accompanied by limitation to the mobility. In our sample, 44% of respondents report limitations with mobility, arm functions and fine motor function. This percentage is higher than 50% for Belgium, Czech Republic, Greece and Italy. We considered also some overcrowding indicators, but, once controlled for others dimensions, none of them proved to have any significant effect on the self-assessed life satisfaction of the households. Unfortunately we do not have information on the quality of the neighbourhood for most of the estimation sample. Finally, we use three indicators for the health domain: the presence of chronic diseases (in a list of 17 diseases) and the number of limitations with the activities of daily living (ADL, that is dressing, walking across a room, bathing or showering, eating, getting in and out of bed, using the toilet) to take into consideration physical health, and the presence of depression symptoms (EURO-D caseness, see Prince et al. 1999) for what concern mental wellbeing.

Table 1 summarizes the details about the dimensions, the indicators and the corresponding thresholds used to define the presence of deprivation. Rephrasing Alkire and Foster's (2011b) words, the aim of our empirical exercise is not to suggest that this set of indicators, dimensions and cut-offs is appropriate in every application. Rather, the aim of our illustrative exercise is twofold. On the one hand, we aim at describing the effects of changes in weighting schemes on the outcomes of the multidimensional poverty analysis run according to the Alkire and Foster's methodology on a sample of elderly individuals living in ten European countries. On the other hand, we want to assess the robustness of the results obtained with hedonic weights based on respondents' life satisfaction self-assessments once the heterogeneity in reporting styles is formally taken into account.

Table 1: Dimensions, indicators and thresholds

Dimensions	Indicators	Thresholds (meet the minimum standard if)	Percent meeting the minimum standards
Economic	per-capita net income	equal or above 60% of median (country specific)	78.80%
	per-capita net wealth	equal or above 60% of median (country specific)	66.70%
Housing	dwelling accessibility	less than 16 steps to climb up/down to entrance	82.60%
Health	chronic disease	none of household members have more than two chronic diseases	44.50%
	ADL	none of household members have ADL problem	86.30%
	EURO-D	none of household members have EURO-D caseness	66.80%

Note: Percentage meeting the minimum standard of a given indicator is a weighted average for the entire sample of households[8].

The proportion of households who meet the minimum standards with respect to single indicators ranges between 44.5% for the presence of chronic diseases to 86.3% for the presence of impediments with the activities of daily living. It is important to notice that the indicators are only weakly correlated between them (Table 2). This suggests that the information conveyed by the dimension considered is not redundant.

Table 2: Tetrachoric correlation coefficients among indicators $\left(a_{jk}^{h}\right)$

	per-capita net income	per-capita net wealth	dwelling accessibility	chronic disease	ADL	EURO-D
per-capita net income	1.0000					
per-capita net wealth	0.3565	1.0000				
dwelling accessibility	-0.0164	0.2384	1.0000			
chronic disease	0.0253	0.2000	0.0476	1.0000		
ADL	0.1155	0.2067	-0.0500	0.5238	1.0000	
EURO-D	0.0588	0.1458	0.0618	0.3755	0.4414	1.0000

In addition to the thresholds of the indicators (z_{jk}), the overall well-being threshold φ plays an important role in this multidimensional poverty framework. Unlike the thresholds of the indicators that can be mostly determined by convention, the choice of φ seems more arbitrary and less grounded since it works across the dimensions where general understanding is hard to be applied. One possible method is to choose it on the basis of the specific policy goals or interests of evaluation (Alkire and Foster, 2011a). We take 0.6 as the default well-

[8] All the descriptive statistics and findings are based on the weighted average for the entire sample of households, unless otherwise stated.

being threshold in all the analysis and conduct a sensitivity analysis of the robustness of the results with respect to the choice of this parameter.

6 Results

6.1 Life satisfaction and hedonic regression

To set the hedonic weights, we exploit the individual question about life satisfaction in general. The top panel of Table 3 shows that about 77% of the interviewed individuals declared to be satisfied or very satisfied, while 5.4% declared to be dissatisfied or very dissatisfied. The lower panel provides a first insight on the relation between achievements and life satisfaction. For each indicator, we compute the risk ratios for each level of life satisfaction, that is, the ratio of the probability to reach a given level of satisfaction for individuals living in households falling below the minimum standard, over the same probability for those living in households reaching this standard. It is therefore possible to appreciate that the percentage of income-poor individuals declared to be very satisfied with their life is only 3/4 of those whose income is above the income threshold. Viceversa, the percentage of dissatisfied among the income-poor is the double of those with higher income. The differences between being below and above the thresholds of indicators are even more striking when focusing on the health indicators, suggesting a prominent role played by the health dimension on the overall life-satisfaction of the individuals.

Table 3: Distribution of the answers to the life-satisfaction question and risk ratios for each life satisfaction level by indicator

	Very dissatisfied	Dissatisfied	Neither satisfied nor dissatisfied	Satisfied	Very satisfied
Percentage of respondents	0.63	4.8	17.48	56.77	20.32
	Percentage of respondents below the minimum standard of a given indicator/ percentage of respondents above the minimum standard of a given indicator				
per-capita net income	2.74	2.02	1.36	0.90	0.74
per-capita net wealth	1.82	1.79	1.47	0.97	0.61
dwelling accessibility	1.63	1.54	1.6	0.95	0.57
chronic disease	5.84	2.24	1.29	0.97	0.72
ADL	6.62	2.97	1.72	0.85	0.42
EURO-D	4.40	5.85	1.89	0.85	0.45

Note: Sample of 5545 individuals used to estimate the hedonic weights.

As explained in the Section 4, the Hopit model can be seen as a generalization of the standard ordered probit model in that the thresholds used to provide self-assessments will be allowed varying with respondents' characteristics. This is made possible by exploiting the information provided by the vignette evaluations. Table 4 shows the distribution of respondents evaluations of the life satisfaction of the hypothetical individuals described in the vignettes, which are crucially kept constant across respondents. While about 44% of respondents rate John (the person in vignette 1) as very dissatisfied or dissatisfied with his life, only 15% of them think that John is at least satisfied. Also, while 13% of respondents rate Carry (the person in the second vignette) as dissatisfied or very dissatisfied, 55% of the sample think she should be at least satisfied. Although the same vignettes about John and Carry have been administered to all the respondents, their evaluations show considerable variability and suggest the presence of heterogeneity in the way they report life satisfaction. If this is an issue, comparisons of life satisfaction self-assessments neglecting this source of heterogeneity might bring about misleading results.

Table 4: Distribution of the answers to the vignette evaluation

	Very dissatisfied	Dissatisfied	Neither satisfied nor dissatisfied	Satisfied	Very satisfied
Vignette 1 (John) Percentage of respondents	5.78	38.59	40.35	14.32	0.96
Vignette 2 (Carry) Percentage of respondents	1.3	11.83	30.58	48.65	7.63

Note: Sample of 5545 individuals used to estimate the hedonic weights.

We estimate the hedonic weights using three different specifications: (1) a standard ordered probit model with only the indicators referring to the dimensions introduced above; (2) we augment (1) by including a full set of observable household and individual characteristics: country of residence, gender, age, presence of a cohabiting partner, children, and grandchildren, employment status, involvement in social activities, education, home ownership, type of the area the accommodation is located in, season at the time of the interview (see Table A.1 in Appendix for the descriptive statistics); and (3) we use the additional information coming from the anchoring vignettes in our dataset and we estimate an Hopit model.

The regression results (Table A.2 in Appendix) for all the specifications show that the indicators are strongly correlated with the self-reported life satisfaction and that demographic characteristics play a significant role. Life satisfaction exhibits remarkable cross-country heterogeneity, it is higher for women, it is at its minimum among individuals aged between 50 and 55, which is consistent with Blenchflower and Oswald (2008) and de Ree and Alessie (2011). Further, life satisfaction increases with the presence of a cohabiting partner, the involvement in social activities and with being at work or retiree instead of out of work due to reasons other than retirement.

In general, these estimates confirm the necessity to "clean" the well-being function in order to identify the relevant components as suggested by Shokkaert (2007).

When we use the Hopit model to control for the possible effect of the heterogeneity in response styles due to differential item functioning (DIF), we have two relevant results. First, the effects of the achievement indicators on life satisfaction change with respect to the previous standard ordered probit. Given the non linearity of the models, such changes can be better appreciated looking at the variation of the probability of being satisfied or very satisfied when the indicators switch from deprivation to achievement.

Table 5: Average percentage change of the probability of being satisfied or very satisfied with life in general due to the variation in the status from non-achievement to achievement

	Ordered probit models		Hopit model
	Indicators only	Indicators + demographics	Indicators + demographics + vignettes
per-capita net income	7.73	3.41	3.98
per-capita net wealth	8.05	6.99	6.96
dwelling accessibility	13.07	5.90	9.90
chronic disease	4.72	7.86	8.39
ADL	15.96	15.72	21.59
EURO-D	22.00	19.98	29.85

Note: Sample of 5545 individuals used to estimate the hedonic weights.

Table 5 shows that, according to the basic ordered probit model, the average effect of enjoying a sufficient level of net income is an increase of 7.73%, the same quantity is reduced to less than the half when the demographic variables are introduced and it equals 3.98% if the heterogeneity in response styles is taken into account. Similar remarkable differences are present also for all the indicators, with the exception of per-capita net wealth.

The second result concerns the response style heterogeneity. This is correlated with country and seasonal dummies, age, the presence of a cohabiting partner, employment status, home-ownership, the type of area in which the accommodation is located and some of the achievement indicators. Overall, our results confirm the evidence provided by Angelini et al. (2012 and in press) that there is heterogeneity in the response styles. Therefore, the estimation of the hedonic weights can be biased if such heterogeneity is neglected.

Table 6: Summary of the weights derived from different approaches (w_{jk})

	Equal weights	Frequency weights	Hedonic weights		
			Indicators only	Indicators + demographics	Indicators + demographics + vignettes
per-capita net income	0.1667	0.1851	0.1081	0.0569	0.1092
per-capita net wealth	0.1667	0.1567	0.1126	0.1167	0.1629
dwelling accessibility	0.3333	0.1940	0.1827	0.0986	0.0751
chronic disease	0.1111	0.1046	0.0659	0.1314	0.0788
ADL	0.1111	0.2026	0.2231	0.2626	0.2359
EURO-D	0.1111	0.1570	0.3075	0.3338	0.3380

6.2 Comparing alternative weighting schemes

The five sets of weights are presented in Table 6. We set the equal weights mimicking the Human Development Index and Multidimensional Poverty Index, that is, the three dimensions have the same relevance and the indicators share the same weight within each dimension (UNDP, 2011). For the frequency weights, we follow Desai and Shah (1988) to set the weight of every indicator as the proportion of the non-deprived households in the sample[9].

As compared with the equal weighting scheme, the frequency one reduces the weight of the housing domain from 1/3 to 0.194 in favor of the health conditions, whose weight goes from 1/3 to 0.464. The weights attached to the economic conditions remain almost unchanged. As compared with frequency weights, the hedonic approach doubles the weight of the EURO-D indicator to about 30%, and in general it increases the prominence of the health domain. The overall importance of the economic domain is reduced between 17.3% and 27.2%, mainly due to a sharp decrease of the weight associated with the per-capita income indicator. When the observable characteristics of the respondents are taken into account (last two columns) the accessibility of the accommodation loses its significance.

Each weighting scheme gives origin to a different well-being score, with its own empirical distribution and therefore potentially with different incidence of the poverty for any given poverty threshold φ.

[9] Frequency weights are standardized in order to sum up to one. Equal weights are standardized by construction.

Table 7: Relevant percentiles of the distribution of the multidimensional well-being score under alternative weighting schemes (s^h)

Percentile	Equal weights	Frequency weights	Hedonic weights: Indicators only	Hedonic weights: Indicators + demographics	Hedonic weights: Indicators + demographics + vignettes
5th	0.3333	0.3593	0.2953	0.2154	0.2631
10th	0.4444	0.3967	0.4034	0.3612	0.3472
15th	0.5000	0.5358	0.4718	0.4181	0.4202
20th	0.5556	0.5447	0.5139	0.4780	0.4990
25th	0.6111	0.5818	0.5942	0.5349	0.5831
30th	0.6111	0.6493	0.6265	0.5676	0.5831
35th	0.6667	0.6864	0.6643	0.6533	0.6528
40th	0.7222	0.7103	0.7047	0.6662	0.6831
45th	0.7222	0.7384	0.7133	0.7519	0.7279
50th	0.7778	0.7388	0.7793	0.7700	0.7582
55th	0.8333	0.8149	0.8215	0.8264	0.8371
60th	0.8333	0.8430	0.8874	0.8686	0.8460
65th	0.8889	0.8433	0.8919	0.8686	0.9212
70th	0.8889	0.8954	0.9341	0.8686	0.9212
75th	0.8889	0.8954	0.9341	0.9014	0.9212
80th	0.8889	0.8954	0.9341	0.9431	0.9249
85th	1.0000	1.0000	1.0000	1.0000	1.0000

Table 7 shows some relevant percentiles of the distribution of the multidimensional well-being score obtained under the alternative weighting schemes considered. The percentiles of the well-being score distribution under equal weights are barely distinguishable from those under frequency weights. Also, the lowest percentiles under equal and frequency weighting are almost always higher than their counterparts obtained under hedonic weighting, at least up to the 40th percentile. This implies that for almost any sensible value of the poverty threshold φ, the headcount ratio H will be lower with the equal and the frequency weight schemes than with hedonic weights. Our choice of setting $\varphi = 0.6$ as benchmark implies that we will classify as poor approximately one third of the households in our sample under all weighting schemes.

The well-being scores are highly correlated to each other. In particular, the pairwise correlations of the scores based on the three sets of hedonic weights lies around 98%. The pairwise correlations between the frequency weight score and the hedonic ones are around 90%, and those between the equal weight score and the hedonic ones range between 76% and 85% (Table 8).

Table 8: Correlation of score using different weight schemes (s^h)

		Equal weights	Frequency weights	Hedonic weights: Indicators only	Hedonic weights: Indicators + demographics	Hedonic weights: Indicators + demographics + vignettes
Equal weights		1.0000				
Frequency weights		0.9485	1.0000			
Hedonic weights	Indicators only	0.8481	0.9374	1.0000		
	Indicators + demographics	0.7600	0.8910	0.9757	1.0000	
	Indicators + demographics + vignettes	0.7807	0.9153	0.9811	0.9853	1.0000

We are now in the position of comparing the aggregated poverty index. By setting the poverty threshold $\varphi = 0.6$, the commonly used headcount ratios vary between 23.8% of the equal weights and 31.7% of the hedonic weights with the vignettes, whereas the corresponding adjusted headcount ratio M ranging between 13.6% and 18.8% (see Table 9).

Table 9: Dimensional decomposition (poverty threshold $\varphi = 0.6$)

	Equal weights	Frequency weights	Hedonic weights: Indicators only	Hedonic weights: Indicators + demographics	Hedonic weights: Indicators + demographics + vignettes
H	0.2380	0.2759	0.2661	0.3112	0.3165
M	0.1364	0.1496	0.1532	0.1881	0.1848
Relative contribution of dimension j to the overall adjusted headcount index M $(RD_j = V_j/M)$					
Economic (%)	33.13	35.46	19.48	12.34	19.42
Housing (%)	36.03	12.67	10.75	4.14	3.11
Health (%)	30.84	51.86	69.77	83.52	77.48

Despite the fact that the level of the adjusted headcount indices is similar across alternative weighting structures, the relative contribution of the dimensions is remarkably different. Consider the health dimension: its contribution to the overall level of the adjusted headcount index is measured by the ratio $RD_{Health} = V_{Health}/M$, which is 51.9% for the frequency weights and 77.5% for the hedonic weights taking the heterogeneity in response styles into account (Table 9). As for the economic condition, its relative contribution is 35.5% with frequency weighting and reduces to 12.3% with hedonic weighting based on an ordered probit model allowing for household and individual characteristics but neglecting response heterogeneity. This variability magnifies for the housing dimension: it explains more than one third of poverty under equal weighting but only 3% under the hedonic weighting based on the Hopit model.

Besides, the differences between subgroups can be affected by the weighting structure too (Table 10). If we look at the Panel A of the table, although Italy always results to have the highest level of adjusted headcount ratio, Germany has an adjusted headcount ratio that is only 8% lower than the Italian one under equal weighting, but more than 30% lower under hedonic weighting. For France the adjusted headcount ratio is 55% lower than that of Italy under equal weighting but this differential shrinks to 31% when hedonic weighting based on Hopit estimates is considered. When turning our attention to Panel B, we also find that the households where the oldest member is aged 55 or less experience the lowest level of poverty, whereas those in which the oldest member is aged 76 or more the highest. However, while the adjusted headcount ratio of the youngest households is 40% lower than the one of oldest households under equal weighting, this reduction lies around 60% for the remaining weighting schemes. In Panel C we looked at the poverty variation by household size, which is measured according to the OECD modified equivalence scale. We find analogous inconsistencies. The poorest households are on average always those consisting of one adult only. Still, if we consider the households consisting of two (equivalent) adults, we find that their adjusted headcount ratio is 35% lower under equal weighting and only 6% lower under hedonic weighting based on Hopit estimates.

It is worth pointing out that, once household and individual characteristics have been taken into account in the estimation of hedonic weights, allowing for heterogeneity in response styles does not lead to dramatic variations in either the *H* and *M* indexes or the poverty decomposition by dimension and group. Although controlling for reporting styles has been proved to be an issue in modeling life satisfaction self-assessments, in our case the results related to hedonic weighting structures seem to be overall robust to the way individuals interpret life satisfaction response scale.

Table 10: Group decomposition, by weighting scheme (poverty threshold $\varphi = 0.6$)

	Population share (%) (θ_g)	The adjusted headcount ratio of the subgroup (M_g)				
				Hedonic weights		
		Equal weights	Frequency weights	Indicators only	Indicators + demographics	Indicators + demographics + vignettes
Panel A: by country						
- SE	3.49	0.0790	0.0873	0.0809	0.1127	0.1127
- DK	2.14	0.0708	0.0946	0.0808	0.1199	0.1178
- DE	29.31	0.1685	0.1759	0.1471	0.1737	0.1722
- NL	5.33	0.0777	0.0835	0.0653	0.0874	0.0865
- BE	3.35	0.0600	0.0976	0.1224	0.1829	0.1807
- FR	18.02	0.0823	0.1068	0.1400	0.1738	0.1759
- ES	10.88	0.1137	0.1376	0.1590	0.2089	0.2026
- IT	20.61	0.1834	0.1995	0.2233	0.2630	0.2561
- GR	2.87	0.1690	0.1344	0.1325	0.1536	0.1364
- CZ	3.98	0.1690	0.1519	0.1409	0.1803	0.1686
Panel B: by age						
- 55-	17.93	0.1040	0.0888	0.0997	0.1207	0.1266
- 56-60	18.64	0.1353	0.1345	0.1433	0.1709	0.1696
- 61-65	15.88	0.1269	0.1287	0.1279	0.1635	0.1589
- 66-75	28.22	0.1388	0.1412	0.1397	0.1770	0.1730
- 76+	19.34	0.1718	0.2500	0.2529	0.3034	0.2917
Panel C: by household size (equivalent adult)						
- 1	21.78	0.1782	0.1828	0.1737	0.1978	0.1971
- 1.1-1.5	51.79	0.1247	0.1435	0.1494	0.1947	0.1887
- 1.6-2	14.80	0.1154	0.1324	0.1534	0.1848	0.1845
- 2.1+	11.64	0.1370	0.1365	0.1316	0.1447	0.1446

6.3 Decomposing the variation of *M* due to different weighting schemes

The Eq.(23) in Section 3 shows that the variation in the adjusted headcount index originated by the change in the weighting scheme can be decomposed in two components: ΔM_1, which depends on a composition effect driven by the sets of households switching their poverty status with the weighting scheme used, and ΔM_2, which depends on the mere variation in the weights vector and it is based on the set of households classified as poor regardless of the weighting scheme used.

Table 11: Effects of weighting scheme variations under different weighting schemes (poverty threshold $\varphi = 0.6$)

Weighting schemes		Relative group size (%)			Group specific adjusted headcount index				7 The difference	
w	w'	$\frac{n^{P\to NP}}{n}$	$\frac{n^{NP\to P}}{n}$	$\frac{n^{P\to P}}{n}$	$M_{P\to NP}$	$M'_{NP\to P}$	$M_{P\to P}$	$M'_{P\to P}$	$M' - M$	ΔM_1
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
1	2	4.98	8.77	18.82	0.4636	0.4480	0.6023	0.5861	0.0132	0.0162
	3	8.06	10.88	15.74	0.4959	0.5039	0.6129	0.6254	0.0168	0.0149
	4	9.19	16.51	14.61	0.5126	0.5456	0.6114	0.6706	0.0516	0.0430
	5	9.43	17.29	14.36	0.5137	0.5157	0.6124	0.6658	0.0484	0.0407
2	3	4.43	3.45	23.16	0.4530	0.4460	0.5593	0.5951	0.0036	-0.0047
	4	5.2	8.73	22.39	0.4668	0.4629	0.5597	0.6594	0.0385	0.0161
	5	5.44	9.5	22.15	0.4683	0.4342	0.5604	0.6480	0.0352	0.0158
3	4	1.9	6.41	24.71	0.4086	0.4647	0.5886	0.6405	0.0348	0.0220
	5	1.37	6.41	25.24	0.4159	0.4171	0.5844	0.6262	0.0316	0.0210
4	5	0.25	0.77	30.88	0.4926	0.4472	0.6052	0.5872	-0.0033	0.0022

Note: 1- equal weighting, 2 - frequency weighting; 3 - hedonic weighting based on ordered probit allowing for well-being indicators only; 4 - hedonic weighting based on ordered probit allowing also for household and individual characteristics; 5- hedonic weighting based on the Hopit model.

Table 11 summarizes the elements of Eq.(23) to decompose the variations $M' - M$ when the weights go from w (column 1) to w' (column 2). Columns (3) and (4) inform that abandoning the equal weights for any other weighting scheme is associated with a substantial change of the set of households regarded as poor. If we consider the change from equal weighting (indexed by 1 in the table) to hedonic weighting based on the Hopit estimates (indexed by 5), 9.43% of the households in the sample exit poverty, 17.29% of the households enter poverty and 14.36% remain poor. As a consequence, despite the fact that $M_{P\to NP}$ and $M'_{NP\to P}$ are quite similar (see columns 6 and 7), the component ΔM_1 is dominant and explains 84% of the variation in the adjusted headcount ratio. This also provides empirical evidence that the households whose poverty statuses vary with the weighting scheme play a significant role. However, the sets of poor households are more stable for changes between the other weighting schemes. In particular, the set of poor households identified using the hedonic weights obtained by taking into consideration the role of the covariates (scheme 4) is almost identical to the group of poor families identified when also response style heterogeneity is taken into account (scheme 5): only around 1% of the household change the poverty status. This finding, combined with the small changes in the estimated weights reported in Table 6, clearly shows the robustness of the results obtained under hedonic weighting to relaxing the assumption of invariance of reporting styles in life satisfaction self-assessments.

6.4 Differences in policy prescriptions due to different weighting schemes

Table 12 summarizes the results of a counterfactual simulation in which we consider one indicator at once and make all households able to meet the minimum standard with respect to such indicator. We calculate the resulting adjusted headcount ratios under all the weighting schemes considered. There are at least two findings. Firstly, policy makers should not base interventions only considering the weight of indicators. For instance, in terms of frequency weights (recall Table 6), the most weighted indicator is ADL; however, the simulation shows that intervening on ADL has the least effectiveness since it is associated with the highest level of poverty after intervention. Also, if we consider hedonic weighting derived from the Hopit model the weights for the per-capita net wealth and chronic diseases indicator were 0.1629 and 0.0788, respectively. Still, intervening on the chronic disease indicator reduces the adjusted headcount ratio by 25.87%, whereas the reduction associated with an intervention on per-capita net wealth is only 18.72%. Secondly, the variability in the effectiveness of the intervention across weighting schemes is evident. Intervening on dwelling accessibility reduces the adjusted headcount ratio by 57% according to the equal weighting but only by 6% under the hedonic weighting based on the Hopit estimates. An analogous gradient is found for an intervention on income: it reduces poverty by almost 30% under equal weighting and only by 9% under vignette-based hedonic weighting. Instead, in the hypothetical scenario in which the governments is in the position of making all households not deprived with respect to mental health, the poverty decreases by 21% under equal weighting and by more than 70% under all the hedonic weighting approaches considered. As noted before, allowing for heterogeneity in reporting styles in the calculation of hedonic weighting does not lead to remarkable differences in the results.

Table 12: Effect of intervention on adjusted headcount ratio under different weighting schemes (poverty threshold $\varphi = 0.6$)

	Equal weights	Frequency weights	Hedonic weights: Indicators only	Hedonic weights: Indicators + demographics	Hedonic weights: Indicators + demographics + vignettes
M (before intervention)	0.1364	0.1496	0.1532	0.1881	0.1848
M' (after intervention on indicator *p*)					
p=per-capita net income	0.0981	0.1107	0.1288	0.1794	0.1681
p=per-capita net wealth	0.0803	0.0890	0.1073	0.1597	0.1502
p=dwelling accessibility	0.0585	0.1154	0.1210	0.1731	0.1729
p=chronic disease	0.0904	0.0739	0.1378	0.1210	0.1370
p=ADL	0.1179	0.1168	0.1138	0.1526	0.1539
p=EURO-D	0.1081	0.0880	0.0417	0.0435	0.0395

6.5 Sensitivity analysis

Any assessment of the effect of the change of the weighting scheme on the headcount and adjusted headcount ratios depends on the choice of the poverty threshold. In order to appreciate to what extent our empirical conclusions are robust to changes in φ, we repeat the analysis setting $\varphi = 0.47$, that is the 60% of the median of the well-being score computed with the equal weights. As expected, the reduction in the poverty threshold causes a decrease in both the headcount and the adjusted headcount ratios (see Table 13). More interestingly, the dimensional decomposition delivers results very similar to our benchmark case.

Table 13: Dimensional decomposition (poverty threshold $\varphi = 0.47$)

			Hedonic weights		
	Equal weights	Frequency weights	Indicators only	Indicators + demographics	Indicators + demographics + vignettes
H	0.1348	0.1210	0.1489	0.1903	0.1741
M	0.0881	0.0797	0.0990	0.1316	0.1206
Relative contribution of dimension *j* to the overall adjusted headcount index M ($RD_j = V_j/M$)					
Economic (%)	32.89	38.52	15.44	13.78	23.12
Housing (%)	36.94	13.32	10.81	4.29	2.36
Health (%)	30.17	48.16	73.75	81.93	74.52

Table 14: Effects of weighting scheme variations under different weighting schemes (poverty threshold $\varphi = 0.47$)

Weighting schemes		Relative group size (%)			Group specific adjusted headcount index				The difference	
w	w'	$\frac{n^{P\to NP}}{n}$	$\frac{n^{NP\to P}}{n}$	$\frac{n^{P\to P}}{n}$	$M_{P\to NP}$	$M'_{NP\to P}$	$M_{P\to P}$	$M'_{P\to P}$	$M' - M$	ΔM_1
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
1	2	4.30	2.92	9.18	0.5840	0.6220	0.6863	0.6701	-0.0085	-0.0070
	3	4.65	6.06	8.83	0.6393	0.6476	0.6613	0.6759	0.0108	0.0095
	4	4.04	9.58	9.44	0.6537	0.6962	0.6537	0.6870	0.0434	0.0403
	5	5.97	9.89	7.51	0.6242	0.6673	0.6772	0.7267	0.0325	0.0287
2	3	2.52	5.31	9.58	0.6021	0.5834	0.6734	0.7094	0.0193	0.0158
	4	1.91	8.84	10.19	0.5797	0.6252	0.6733	0.7492	0.0519	0.0442
	5	1.99	7.29	10.11	0.5839	0.6148	0.6732	0.7494	0.0409	0.0332
3	4	0.12	4.26	14.77	0.5983	0.5794	0.6650	0.7240	0.0326	0.0240
	5	2.06	4.57	12.84	0.5596	0.5830	0.6812	0.7321	0.0217	0.0151
4	5	1.93	0.31	17.10	0.5632	0.6101	0.7061	0.6944	-0.0110	-0.0090

Note: 1- equal weighting, 2 - frequency weighting; 3 - hedonic weighting based on ordered probit allowing for well-being indicators only; 4 - hedonic weighting based on ordered probit allowing also for household and individual characteristics; 5- hedonic weighting based on the Hopit model.

For what concern the components of the change $M' - M$, the role of ΔM_1 is even more relevant than in the benchmark case (see Table 14).

Finally, in the estimation of the hedonic weights, the different number of indicators within each dimension may potentially affect the overall relevance of each dimension, and therefore the conclusions of the decomposition analysis. In particular, the overall weight of a dimension may increase with the number of indicators considered. We thus run a robustness check in which we estimate the weight of the entire dimension and constrain each indicator to have the same weight within the dimension. This is done by replacing the indicators in the ordered probit and Hopit specifications with their averages by dimension. By doing this, we estimate the same number of parameters (i.e. one) for each dimension. The estimated coefficients on the dimension averages of indicators deliver the dimensional unstandardized weights. Then, all the indicators within a given dimension equally split the dimensional weight and are finally standardized. The estimated vectors of weights can therefore be interpreted as a constrained version of the benchmark hedonic weights. The new vectors of weights are shown in Table 15, with the corresponding headcount and adjusted headcount ratios for $\varphi = 0.6$ in Table 16.

Table 15: Summary of the hedonic weights derived from different approaches

			Hedonic weights		
	Equal weights	Frequency weights	Indicators only	Indicators + demographics	Indicators + demographics + vignettes
per-capita net income	0.1667	0.1851	0.1180	0.0928	0.1412
per-capita net wealth	0.1667	0.1567	0.1180	0.0928	0.1412
dwelling accessibility	0.3333	0.1940	0.1865	0.0988	0.0734
chronic disease	0.1111	0.1046	0.1925	0.2385	0.2147
ADL	0.1111	0.2026	0.1925	0.2385	0.2147
EURO-D	0.1111	0.1570	0.1925	0.2385	0.2147

Table 16: Dimensional decomposition (poverty threshold $\varphi = 0.6$)

			Hedonic weights		
	Equal weights	Frequency weights	Indicators only	Indicators + demographics	Indicators + demographics + vignettes
H	0.2380	0.2759	0.2759	0.3528	0.3528
M	0.1364	0.1496	0.1608	0.2080	0.2030
Relative contribution of dimension *j* to the overall adjusted headcount index M ($RD_j = V_j/M$)					
Economic (%)	33.13	35.46	23.22	13.26	20.68
Housing (%)	36.03	12.67	11.33	4.19	3.19
Health (%)	30.84	51.86	65.45	82.55	76.13

The constrained versions of the hedonic weights confirm the emphasis on the health dimension as well as the results coming from the dimensional decomposition of the adjusted headcount ratio.

7 Conclusions

Using multidimensional poverty measures instead of simple monetary poverty indicators is now a standard practice. The increase in the number and heterogeneity of the dimensions makes the weighting scheme a key ingredient of the poverty assessment. In this paper we carry out a multidimensional poverty assessment framed in the approach proposed by Alkire and Foster (2011a) to show to what extent the outcomes of a multidimensional poverty analysis are robust to the change in the weighting scheme adopted.

We analytically show that a change in the weighting structure has an effect on the overall poverty assessment that cannot be unambiguously predicted by looking at the change in the weights and at the outcomes of the households originally classified as poor. In fact, the effect depends on the sample size and the outcomes of the households changing their poverty status from one weighting scheme to the other and of those that are classified as poor according to both weighting schemes. This result stresses that in Alkire and Foster’s framework, everything else constant, different weighting schemes vary the set of poor households and then they lead to different results in terms of dimension and subgroup decompositions. Consistently, we show that different weighting schemes deliver different prescriptions for anti-poverty policies whose effectiveness is assessed with respect to the reduction of the adjusted headcount index. Policy makers should choose the dimensions to target by considering not only the weights of indicators but also the number of households who could exit poverty thanks to the intervention.

To analyze empirically how alternative weighting schemes impact on the results of a multidimensional poverty analysis, we draw data from the second wave of SHARE, a multi-country survey administering a multidisciplinary standardized questionnaire to a representative sample of individuals aged 50 or over and living in Europe. We apply the Alkire and Fosters’ approach to our data and consider three dimensions. We measure the economic condition by looking at income and wealth, the housing condition by a measure a dwelling accessibility based on the number of steps respondents have to climb up and/or down to enter their accommodation and the health condition by looking at the presence of chronic diseases, difficulties in activities of daily living (ADL) and mental health.

Following the classification by Decanq and Lugo (2013), we consider equal weighting, data-driven weighting and hedonic weighting as examples of the three main classes of weights: normative, statistical and hybrid. Building upon Fleurbaey et al. (2009), we estimate three sets of hedonic weights by means of ordered probit regressions having respondents' life satisfaction self-assessments as dependent variable. In a first case the explanatory variables include the well-being indicators of interest only. In a second case, we enriched the right-hand-side variables by including a set of household/individual characteristics. However, there is an increasing literature investigating the effects on the comparability of self-assessments across individuals produced by the heterogeneity in the way they interpret the scales according to which such self-assessments are provided. As long as individuals with the same level of well-being interpret the life satisfaction scale differently and provide different evaluations, the comparison between their self-assessments may give a picture of genuine well-being differentials blurred by heterogeneity in reporting styles. We take advantage of the second wave of SHARE to develop a third set of hedonic weights based on Hopit regressions of life satisfaction self-assessments, which formally take into account the variability of response styles across individuals by means of an anchoring vignette methodology (King et al., 2004).

Our results show that changes in the weighting scheme produces substantial differences in the set of households classified as poor. In particular, households who enter or exit poverty when passing from equal weighting to hedonic weighting explain most of the variation in the overall poverty assessments. In addition, when we look at the contribution of each dimension to the overall poverty level, we find that they widely change across weighting structures. For instance, health explains 51.8% of the overall poverty with frequency weights and more than 83.5% with hedonic weighting based on regressions allowing for the heterogeneity in reporting styles. This variability is also confirmed when looking at the other dimensions. Changing weighting scheme has an effect on the comparisons of poverty levels by country, age and household-size. Our estimates show also that omitting to condition on observable characteristics when estimating the hedonic weights from a life satisfaction regression equation can lead to a distorted weighting structure. Although our empirical exercise confirms that the heterogeneity in response styles is an important issue in modeling life satisfaction self-assessments, it does not highlight significantly differences in neither the level nor the decomposition of the poverty index based on the hedonic weights.

Finally, for each indicator considered, we carry out counterfactual simulations based on our data and on the weighting schemes considered to analyze the effectiveness of hypothetical anti poverty interventions consisting of making all households not deprived for such indicator. We find that changes in the weighting scheme significantly affect the results of these hypothetical policies. As an example, we find that according to frequency weighting the ADL indicator receives the highest weight, still an intervention removing ADL deprivation in the population turns out to be the least effective. Also, removing problems of dwelling accessibility in the whole population reduces poverty by 57% according to the equal weighting but only by 6% under the hedonic weighting allowing for heterogeneity in reporting styles. Although our empirical results can be affected by the selected sample we use and by the settings we adopt to implement the general Alkire and Foster's approach, they clearly warn us that the choice of the weighting schemes is not innocuous for the outcomes of a multidimensional poverty analysis. Comparisons of poverty across groups and policy evaluations based on this framework should then take into account this issue in order to provide meaningful and reliable conclusions.

Acknowledgements

This paper uses data from SHARE wave 4 release 1.1.1, as of March 28th 2013 or SHARE wave 1 and 2 release 2.5.0, as of May 24th 2011 or SHARELIFE release 1, as of November 24th 2010. The SHARE data collection has been primarily funded by the European Commission through the 5th Framework Programme (project QLK6-CT-2001-00360 in the thematic programme Quality of Life), through the 6th Framework Programme (projects SHARE-I3, RII-CT-2006-062193, COMPARE, CIT5- CT-2005-028857, and SHARELIFE, CIT4-CT-2006-028812) and through the 7th Framework Programme (SHARE-PREP, N° 211909, SHARE-LEAP, N° 227822 and SHARE M4, N° 261982). Additional funding from the U.S. National Institute on Aging (U01 AG09740-13S2, P01 AG005842, P01 AG08291, P30 AG12815, R21 AG025169, Y1-AG-4553-01, IAG BSR06-11 and OGHA 04-064) and the German Ministry of Education and Research as well as from various national sources is gratefully acknowledged (see www.share-project.org for a full list of funding institutions). We would like to thank the participants at the Italian SHARE Users' Conference (2012), 12th Journées Louis-André Gérard-Varet Conference in Public Economics (2013), SIS 2013 Statistical Conference and at seminars at the University of Groningen and CEMFI.

References

Alkire, S., Foster, J., 2011a. Counting and multidimensional poverty measurement. J. Public Econ. 95, 476-487.

Alkire, S., Foster, J., 2011b. Understandings and misunderstandings of multidimensional poverty measurement. J. Econ. Inequal. 9, 289-314.

Angelini, V., Cavapozzi D., Corazzini, L., Paccagnella, O., Do Danes and Italians rate life satisfaction in the same way? Using vignettes to correct for individual-specific scale biases. Oxf. Bull. Econ. Stat. In press.

Angelini, V., Cavapozzi D., Corazzini, L., Paccagnella, O., 2012. Age, health and life satisfaction among older Europeans. Soc. Indic. Res. 105, 293-308.

Blenchflower, D.G., Oswald, A.J., 2008. Is well-being U-shaped over the life cycle?, Soc. Sci. Medicine 66, 1733-1749.

Decancq, K., Lugo, M.A., 2013. Weights in multidimensional indices of well-being: An overview. Econom. Rev. 32, 7–34.

Desai, M., Shah, A., 1988. An econometric approach to the measurement of poverty. Oxf. Econ. Pap. 40, 505-522.
Dolan, P., Peasgood, T., White, M., 2008. Do we really know what makes us happy? A review of the economic literature on the factors associated with subjective well-being. J. Econ. Psychol. 29, 94–122.
Eurostat, 2012. Measuring material deprivation in the EU — Indicators for the whole population and child-specific indicators, Publications Office of the European Union, Luxembourg.
Fleurbaey, M., Schokkaert, E., Decancq, K., 2009. What good is happiness?. Université Catholique de Louvain, Center for Operations Research and Econometrics (CORE) Discussion Paper 2009/17.
Foster, J.E., Greer, J., Thorbecke, E., 1984. A class of decomposable poverty measures. Econom. 52, 761–766.
Frey, B. S., Stutzer, A., 2002. What can economists learn from happiness research?. J. Econ. Lit. 40, 402–435.
United Nation Development Programme, 2011., Human Development Report 2011. Sustainability and Equity: A Better Future for All. United Nations Development Programme, New York.
Kapteyn, A., Smith, J. P., Van Soest, A.H.O., 2009. Comparing life satisfaction. RAND Working Paper WR-623-1.
King, G., Murray, C., Salomon, J., Tandon, A,. 2004. Enhancing the validity and cross-cultural comparability of measurement in survey research. Am. Polit. Sci. Rev. 98, 567–583.
Nussbaum, M. C., 2001. Women and Human Development: The Capabilities Approach, Cambridge University Press, Cambridge, UK.
Prince, M. J., Reischies, F., Beekman, A. T., Fuhrer, R., Jonker, C., Kivela, S. L., Lawlor, B. A., Lobo, A., Magnusson, H., Fichter, M., van Oyen, H., Roelands, M., Skoog, I., Turrina, C., Copeland, J. R., 1999. Development of the EURO-D scale - a European, Union initiative to compare symptoms of depression in 14 European centres. Br. J. Psychiatry 174, 330-338.
de Ree, J., Alessie, R., 2011. Life satisfaction and age: Dealing with underidentification in age-period-cohort models. Soc. Sci. Medicine 73, 177-182.
Ruger, J. P., 2010. Health and Social Justice, Oxford University Press, New York.
Schokkaert, E., 2007. Capabilities and satisfaction with life. J. Hum. Dev. 8, 415-430.
Sen, A., 1976. Poverty: An ordinal approach to measurement. Econom. 44, 219-31.
Sen, A., 1985. Commodities and Capabilities, North-Holland, Amsterdam.
Sen, A., 2004. Capabilities, lists, and public reason: Continuing the conversation. Fem. Econ. 10, 77-80.

Appendix

Table A.1: Descriptive statistics of the covariates used in the hedonic regression

	Mean	Std. dev.
SE	0.0772	0.2669
DK	0.1717	0.3771
NL	0.0833	0.2764
BE	0.0931	0.2905
FR	0.0637	0.2442
GR	0.0895	0.2854
IT	0.1068	0.3088
ES	0.0828	0.2756
CZ	0.1473	0.3545
male	0.4449	0.4970
aged 55 or less	0.2341	0.4235
aged 56-60	0.2076	0.4056
aged 61-65	0.1729	0.3782
aged 66-75	0.2530	0.4348
living with a cohabiting partner	0.7803	0.4141
have children	0.9039	0.2948
have grand children	0.6388	0.4804
retired from work	0.4956	0.5000
employed or self-employed	0.3203	0.4666
not involved in social activity	0.4923	0.5000
low education	0.4855	0.4998
middle education	0.2819	0.4500
house owner	0.7482	0.4341
residing in city	0.3234	0.4678
residing in town	0.4227	0.4940
interviewed in winter	0.4074	0.4914
interviewed in spring	0.3627	0.4808
interviewed in summer	0.0294	0.1689

Note: Sample of 5545 individuals used to estimate the hedonic weights.

Table A.2: Hedonic weights estimation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Ordered probit model				Hopit model		
	Indicators only	Indicators + demographics	main equation	/cutoff1	/cutoff2	/cutoff3	/cutoff4
per-capita net income	0.1970***	0.0876*	0.1662***	0.0358	0.0295	-0.0043	-0.0111
per-capita net wealth	0.2051***	0.1797***	0.2480***	0.0963	-0.0505	0.0604	-0.0171
dwelling accessibility	0.3328***	0.1518***	0.1143*	-0.1555*	0.0426	-0.0033	0.0775**
chronic disease	0.1201***	0.2022***	0.1200***	-0.1327**	0.0467	-0.0123	-0.0125
ADL	0.4065***	0.4043***	0.3591***	-0.1222	0.0262	0.0384	0.0111
EURO-D	0.5603***	0.5138***	0.5145***	-0.1899***	0.0403	0.0702**	0.0806***
SE	-	0.3422***	0.3828***	0.5189***	-0.0508	-0.1528**	-0.2432***
DK	-	0.5840***	0.0268	-0.2729	-0.0260	-0.1010	-0.1474***
NL	-	0.2203***	0.5862***	0.6714***	-0.2989***	0.0175	0.0311
BE	-	0.0184	-0.0292	0.5334***	-0.0797	-0.2115***	-0.2839***
FR	-	-0.1265	0.2492*	0.7092***	-0.1762	0.0335	-0.0560
GR	-	-0.5224***	-0.5964***	0.5848***	-0.1854	-0.0299	-0.6128***
IT	-	-0.3165***	-0.0312	0.7414***	-0.1196	-0.2477***	0.0136
ES	-	0.0061	0.0636	0.5812***	-0.0416	-0.4044***	-0.0927
CZ	-	-0.2729***	-0.5580***	-0.0003	-0.1265	-0.0038	-0.0682
male	-	-0.1078***	-0.0807*	0.0152	0.0034	0.0050	0.0107
aged 55 or less	-	-0.2000***	-0.4066***	-0.1798	0.0188	-0.0290	-0.0366
aged 56-60	-	-0.0541	-0.1977**	-0.0669	-0.0300	-0.0042	-0.0422
aged 61-65	-	-0.0098	-0.0918	-0.2059*	0.0928	-0.0147	-0.0177
aged 66-75	-	0.0011	0.0419	0.1722*	-0.0904	0.0254	-0.0433
living with cohabiting partner	-	0.4119***	0.4279***	-0.0182	0.0111	-0.0242	0.0700**
have children	-	0.0051	-0.0167	-0.0778	0.0934	-0.0384	-0.0491
have grand children	-	0.0386	0.0737	-0.0023	-0.0018	0.0256	0.0178
retired from work	-	0.2053***	0.2827***	-0.0393	0.0450	0.0257	0.0527
employed or self-employed	-	0.2458***	0.4288***	-0.0913	0.0889	0.0581	0.1136***
not involved in social activity	-	-0.2266***	-0.2536***	0.0989	-0.0808*	0.0443	-0.0643***
low education	-	-0.0233	-0.0491	-0.1170	0.0525	-0.0016	0.0183
middle education	-	-0.0059	-0.0110	-0.0245	-0.0232	0.0095	0.0473
house owner	-	-0.0388	-0.0558	-0.0772	0.0194	-0.0261	0.0668**
residing in city	-	0.0643	0.0739	0.0640	-0.0205	-0.0650*	0.0476
residing in town	-	0.1142***	0.0600	-0.0401	0.0171	-0.0733**	0.0516*
interviewed in winter	-	-0.0045	-0.0991	0.1904*	-0.1983	0.0284	0.0189
interviewed in spring	-	0.0624	0.028	0.2265*	-0.1574**	0.0145	-0.0075
interviewed in summer	-	-0.0202	-0.1626	0.1554	-0.2489*	0.0787	0.0346
/cutoff1	-1.3303	-1.4208	-	-	-	-	-
/cutoff2	-0.3495	-0.3970	-	-	-	-	-
/cutoff3	0.6024	0.6308	-	-	-	-	-
/cutoff4	2.2886	2.4940	-	-	-	-	-
vignettes question 1	-	-	-0.4821***	-	-	-	-
vignettes question 2	-	-	0.6794***	-	-	-	-
constant	-	-	-	-2.0703***	0.4395**	0.1647	0.4631***

Note: *P<0.1; **p<0.05; ***p<0.01. Sample of 5545 individuals used to estimate the hedonic weights.