Dealing with Increasing Dimensionality in Wellbeing and Poverty Measurement, Some Problems and Solutions

Gordon Anderson*  Teng Wah Leo†
University of Toronto  St. Francis Xavier University

Paul Anand‡
The Open University

January 10, 2014

*Department of Economics, University of Toronto. Email Address: anderson@chass.utoronto.ca
†Department of Economics, St. Francis Xavier University. Email Address: tleo@stfx.ca
‡Department of Economics, The Open University. Email Address: paul.anand@open.ac.uk
Abstract

A recent literature has fomented considerable interest in multi-dimensional measurement of wellbeing and deprivation. However, increasing the number of dimensions over which various aspects of wellbeing are measured raises concerns regarding the robustness of such measures, the extent to which they relate to individual and societal preferences, and their diminishing ability to discriminate between alternative states of the world. Here these problems are investigated in a multi-dimensional context, and some solutions are proposed using a satisfaction survey as a test bed for exploring agent preferences over a range of wants and needs. An alternative way of viewing the robustness issue is proposed and ideas, based upon various notions of separable functions (borrowed from consumer choice theory and non-parametric statistics) are shown to facilitate measurement of wellbeing, and increase our ability to discriminate between alternative states of the world. The result is a modification to the classic AlkireFoster multi-dimensional impoverishment index which admits notions of substitutability and complementarity between the various aspects of impoverishment. The modified version of the index proved to have superior robustness properties compared to more restrictive versions of the index.

JEL Code:

Key Words: Wellbeing Index
1 Introduction

A considerable literature (Sen (1995), Anand and Sen (1997), Atkinson (2003), and essays in Grusky and Kanbur (2006), Stiglitz et al. (2011)) has refocused attention away from an aggregate of goods and services towards a collection of individual capabilities as a measure of wellbeing, thereby stressing the multi-dimensional nature of individual wellbeing. The argument is in essence about what should appear as representing individual wellbeing, since simple consumption/income measures have been deemed inadequate due to doubts in the linkage between possessions and wellbeing, compared to a collection of various abilities that describes the tools within an individual that empowers her and allows her to actively shape her environment and achieve her desires. This has consequently resulted in a substantial expansion in the number of dimensions over which wellbeing or deprivation (as the case may be) has been deemed to be within the purview of an accurate depiction and analysis. An early example of Multi-dimensional Wellbeing measurement is the Human Development Index (United Nations Development Programme (2008)), first published in 1990, and commonly referred to as the HDI. A more recent and very popular multivariate Poverty or Deprivation (wellbeing) index that is illustrative of these multi-dimensional indices\(^1\) is that of Alkire and Foster (2011). The common concern in all of these indices are questions regarding aspects of wellbeing that should be included in the analysis, how they should be aggregated, and how robust they may be to alternative assumptions on parameterizations.

The HDI index which has a country as its basic agent is principally based on a equally weighted aggregation across countries of the three dimensions of education (a combination of literacy and school enrolment rates), life expectancy, and GDP per capita. On the other hand, the Alkire-Foster index, as an advancement over the former, has an individual as its basic agent, thus providing a better reflection of the true level of wellbeing within a society. The Alkire-Foster index is a “counting” based measure, like its predecessor, but incorporates elements of the Foster-Greer-Thorbecke (FGT) index (Foster et al. 1984) that allows it to reflect the complexity of wellbeing, through the number of dimensions that can be considered. These dimensions of individual wellbeing are themselves weighted and they need not be the same, while each individual receives the same weight towards the calculation of the index. Some methods used in determining the weights for each dimension of wellbeing include statistical, survey-based, normative-participatory,

\(^1\)See also Bossert et al. (2009), and Chakravarty and D’Ambrosio (2006).
frequency-based, or a combination of these (See inter alia Atkinson et al. (2002), Brandolini and D’Alessio (1998), Decancq and Lugo (2008) and Sen (1996, 1997)). However, there are several concerns with the measure. Firstly, insofar as the various dimensions of wellbeing are weighted sums, there is an implicit assumption of separability (see for example de la Vega and Urrutia (2011)) which has implications on the structure of the underlying preferences, and may be too restrictive to accurately reflect the true level of wellbeing. Secondly, insofar as much of the parameters within the index are subject to investigator’s choice, these choices may impinge on the index’s ability to correctly identify the set of individuals that constitute the poor (see for example Ravallion (2010)). Finally, the breadth and depth granted by its ability to incorporate the multi-dimensional nature of individual wellbeing suffers from the “curse of dimensionality” as the number of dimensions considered rises, so that the results may not be robust (see for example Yalonetzky (2012)).

This paper presents a constructive method to address these three issues of choice of weight and robustness to parameter choices, thereby enhancing the Alkire-Foster index, and its broad application and appeal. Firstly, by realising and acknowledging the implicit relationship between the index and consumer theory, we propose the estimation of the weights. Insofar as the subjective wellbeing (SWB) responses used in developing a wellbeing/poverty index typically have several sub-categories/sub-dimensions, we propose the use of the nonparametric welfare bounds obtained from Anderson et al. (2011) to dispense with the weights applied to this sub-categories in calculating the sub-utility functions, thereby reducing the multidimensional burden on the index. Secondly, we propose the use of the overlap measure (Anderson et al. 2010; Anderson et al. 2012) to examine the ability of the original Alkire-Foster index vis-a-vis one which estimates the weights applied to accurately identify the relevant sets of individuals, in examining the robustness of the index. To demonstrate the force of the issues and the potential value of the approach taken here, we use data from a recent survey of adult life quality in the U.S., and U.K. (Anand, Gray and Roupe (2013 in preparation)). The data affords an opportunity to explore these issues and permits the proposal of some practical solutions.

In the following section, the relationship between the generalized Alkire-Foster is elaborated on. Section 3 discusses the robustness issues pertaining to the choice of weights used, and the informational burden imposed on the index with increasing dimensions considered, and a viable solution and enhancement to the Alkire-Foster index and robustness
test is provided. Essentially, the approach provides a way of engaging agent preferences in the calculus by estimating the agent’s utility function, and implementing it in the Alkire-Foster index. Section 4 provides the detailed approach to the problem in relation to the data set used for a better grounding of the method. The results from this method then provides a baseline for evaluating the effectiveness of the Alkire-Foster index. Section 5 reports the results and comparisons, and some conclusions are drawn in section 6.

2 Consumer Theory Based Weighting of Alkire-Foster Wellbeing Index

The Alkire and Foster (2011) index has the individual as its basic agent. To see the relationship between the Alkire-Foster index and Consumer Theory, let there be a sample of $N$ agents with $D$ dimensions that constitute potential deprivation (poverty/wellbeing). In addition, let $K \in \{0, 1, 2, \ldots, D - 1\}$ denote the number of dimensions in which the agent fails to meet the threshold for minimum wellbeing. Then the generalized version of the index for the $K^{th}$ level of deprivation may be written as:

$$M^\alpha_K = \frac{1}{N} \sum_{n=1}^{N} \prod_{c_n > K} \sum_{d=1}^{D} \frac{w_d}{D} \left( \frac{x_{nd}}{x_d} < 1 \right) \left( \frac{x_d - x_{nd}}{x_d} \right)^\alpha$$  \hspace{1cm} (1)

where $\prod(\cdot)$ is an indicator function (which is 1 when $c_n > K$, and 0 when $c_n \leq K$), $c_n$ is the count of the number of dimensions in which agent $n \in \{1, \ldots, N\}$ is deprived, $w_d$ is a weight attached to the $d^{th}$, $d \in \{1, \ldots, D\}$, deprivation dimension, $x_{n,d}$ is the deprivation level experienced by the $n^{th}$ agent in the $d^{th}$ dimension, $x_d$ is the deprivation threshold on the $d^{th}$ dimension, and $\alpha = \{0, 1, \ldots\}$ is the parameter for the degree of aversion to poverty (deprivation/wellbeing) in the generalized Foster et al. (1984) (FGT) index. The parameter $\alpha$ is most frequently set to zero in practice (and for the empirical analysis that follows), which implies that the index involves an average across agents of a weighted count of the number of dimensions in which an individual has not met the deprivation metric particular to each dimension. Some of the variety of approaches to setting dimension weights include statistical, survey-based, normative-participatory, frequency-based, or a combination of these (See inter alia Atkinson et al. (2002), Brandolini and D’Alessio (1998), Decancq and Lugo (2008) and Sen (1996, 1997)).

---

2This incorporates a notion of the Foster et al. (1984) class of impoverishment indices.
On closer examination of the Alkire-Foster index of equation (1), one begins to observe the measure as similar to the aggregation of additively separable utility functions (see for example de la Vega and Urrutia (2011)), where the weights could be construed as the marginal effect of each respective dimension. Consider the familiar complete union (failure to meet at least one metric, \( c_n = K = 0 \)) – complete intersection (failure to meet all metrics \( c_n > K = D - 1 \)) approaches are extreme versions of the Alkire-Foster index. The former can be interpreted as the case where all goods are perfect complements in deprivation, so that deprivation in one good implies deprivation in all, and the latter can be interpreted as the case where all goods are perfect substitutes for each other, so that deprivation will only arise when there is deprivation in all.

In other words, this suggests that the choice of weights for equation (1) is much like approximating the deprivation/wellbeing boundary of an individual’s preference function \( U(x) \) over \( x \), which we will denote as \( U^* \) (in the present context the vector \( x \) represents the various dimensions of functionings, and capabilities of agents). Indeed as will be seen in the following section, equation (1) implies a very restrictive form of \( U(.) \), which may not represent the true preferences or needs of agents well due to the separability assumptions. However, it is possible to avoid this issue and generate parameter free wellbeing indices. Anderson et al. (2011) showed how two-sided bounds can be placed on a welfare index for each agent using only the assumptions that wellbeing (as measured by the welfare index) be non-decreasing, and weakly quasi-concave with respect to its many indicators/dimensions. Unfortunately, from a policy maker’s perspective, these Anderson et al. (2011) bounds are frequently considerably far apart. Furthermore, it is not always clear from such an index what the role or importance of a particular dimension of capability or functioning is, so that some sort of parameterization will inevitably be necessary.

### 3 Robustness and the Dimensionality Problem

The issue of robustness of the orderings in cross society comparisons using aggregating measures such as the Alkire-Foster index is principally due to the parametric choices of the weights, the wellbeing/poverty cut-offs, and the critical number of dimensions in which an agent is below the cutoff before it is deemed to constitute impoverishment. The issue arises from the parameters being chosen subjectively by the investigator, or policy
maker (such concerns could equally apply to the HDI). Although the issue with subjective choice of weights is addressed by the estimation and linear programming method that will be discussed in detail in the following section, the issue of dimensionality diluting and confounding results as the dimensionality of wellbeing/impoverishment being considered increases remain.

These cross society comparisons are most commonly performed using stochastic dominance techniques. The classic multi-dimensional stochastic dominance comparisons (Atkinson and Bourguignon (1982), Duclos et al. (2006), and Anderson (2008) in a continuous framework, and Anderson and Hachem (2012) in a continuous/discrete framework) can be interpreted as providing conditions (the order and direction of dominance) under which certain classes of wellbeing/poverty measures, such as the union and intersection restricted cases of index (1), will be robust in the sense that if the conditions prevail, all measures in the class will unambiguously reveal a consistent ordering of the distributions. In all cases, they end up comparing some notion of distance between two surfaces. However, when dimensionality increases substantially, these techniques run into some practical difficulties not unlike the “curse of dimensionality” in nonparametric economics. The issue results from the rapidly increasing demands placed upon sample size for effective estimation and testing of multi-dimensioned functions as dimensions considered increases. Here, as is the case for nonparametric econometrics, as dimensionality increases, probability distributions underlying the surfaces being compared lose mass at the center of their distribution\(^3\). This results in the flattening of the distributions over their support, and the surfaces of the distributions inevitably begin to resemble each other, thereby reducing the distance between them, and making it more difficult to discriminate between them. Another way to consider the effect is to note that notionally similar points in \(D\) dimensional space grow further apart as \(D\) increases. As an example, consider \(\phi(0)\) the peak at the center of the joint distribution of \(D\) i.i.d. standard normal variables (here \(0\) is the \(D\) dimensional null vector), \(\phi(0) = 1/(2\pi)^{D/2}\) which goes to 0 as \(D\) increases, and the Euclidean distance between this null vector and the unit vector in \(D\) dimensioned space is \(\sqrt{D}\), which obviously increases with \(D\). This issue thus fundamentally impinges on the primary ability of the Alkire-Foster index to identify poverty, diluting its insights as the dimensions increase.

\(^3\)This problem is familiar to neural networks researchers, and is referred to as the empty space phenomenon (Verleysen 2003)
Yalonetzky (2012) examines this by developing multivariate stochastic dominance techniques to provide full robustness conditions on $f_A(x)$ and $f_B(x)$ for such “counting” type measures in multi-dimensional settings. He demonstrates that when the number of dimensions considered are greater than 2, generalizations of the Atkinson and Bourguignon (1982) continuous distribution stochastic dominance framework are usually not applicable (an early sign that increasing dimensionality creates complications). Nonetheless, he does obtain some results by restricting the family of measures to particular subclasses (typically the aforementioned union and intersection type measures). His empirical application (based upon the SILC dataset) illustrating the use of these conditions for ordinal variables, found robustness in only 45 of the 325 pairwise comparisons in the two variable case (13.85%), and 38 of the 325 pairwise comparisons (10.8%) in the three variable case (Note the poor and deteriorating robustness when moving from two to three dimensions).

The lack of robustness found in Yalonetzky (2012) may be due to some technical, and practical statistical problems associated with increasing dimensionality in the joint distributions of wellbeing factors. Increasing dimensionality, or the “Curse of Dimensionality”, is a problem familiar to non-parametric econometricians. It concerns the rapidly increasing demands placed upon sample size for effective estimation and testing of multi-dimensional functions as dimensions considered increases. Essentially as dimensionality increases, probability distributions underlying the surfaces being compared lose mass at the center of the distribution\(^4\), the distributions flatten over the support, and the surfaces of the distributions inevitably begin to resemble each other, thereby reducing the distance between them, and making it more difficult to discriminate between them. Therefore, to adequately address the robustness issue, there is a need to examine both the type of functional adjustments that needs to be adopted for the Alkire-Foster index, and other alternative methods of robustness checks in lieu of the issues that increasing dimensionality creates.

3.1 “Curse of Dimensionality” & the Functional Form of Preferences

Yatchew (1998, 2008) nicely describes the curse of dimensionality problem, and points to a potential solution. He considers the general equation $y = f(x_1, x_2) + e$, where

\[^4\]This problem is familiar to neural networks researchers, and is referred to as the empty space phenomenon (Verleysen 2003)
it is supposed that \( f(\ldots) \) is a two dimensioned function on the unit square. In order to approximate the function, it is necessary to sample throughout its domain. If \( T \) points are distributed uniformly on the unit square, each will “occupy” an area \( 1/T \), and the typical distance between points will be \( 1/\sqrt{T} \), so that the approximation error is now \( O(1/\sqrt{T}) \). Repeating this argument for functions of \( k \) variables, the typical distance between points becomes \( 1/(T^{1/k}) \), and the approximation error is \( O(1/(T^{1/k})) \). In general, this method of approximation yields error proportional to the distance to the nearest observation. For \( T = 100 \), the approximation error is 10 times larger in two dimensions than in one, and 40 times larger in five dimensions.

However, if \( f(\ldots) \) can be assumed to be additively separable on the unit square so that \( f(x_1, x_2) = f_1(x_1) + f_2(x_2) \), and \( 2T \) observations are taken (\( T \) along each axis), then \( f_1(\cdot) \) and \( f_2(\cdot) \) can each be approximated with error \( O(1/T) \) so that the approximation error for \( f(\ldots) \) is also \( O(1/T) \). In other words, additive separability maintains the approximation error to a multi-dimensioned function within a single-dimensioned framework, so that the assumption of additive (sometimes referred to as Strongly) separability buys much in terms of information requirements. However it is a very strong assumption which needs significant theoretical and empirical justification. Some indication of how such reasoning could be pursued may be gleaned from the theory of consumer behavior.

Strongly or Additively Separable structures for Utility functions have played an important role in the empirical development of the theory of consumer behavior, largely as simplifying assumptions for the purposes of facilitating estimable demand equations (for an extensive discussion see Deaton and Muellbauer (1980)) in the context of very limited data sets. Perhaps the best known strongly separable utility function is the Stone-Geary utility function underlying the Linear Expenditure System (Stone 1954) which, for goods \( q_d, d = \{1, \ldots, D\} \), may be written as:

\[
U(q_1, q_2, \ldots, q_d) = \sum_{d=1}^{D} \beta_d \ln(q_d - \gamma_d) \tag{2}
\]

Here, for \( U(\cdot) \) to be appropriately concave, \( \beta_d \geq 0, q_d > \gamma_d \) for all \( d = \{1, \ldots, D\} \), and \( \sum_{d=1}^{D} \beta_d = 1 \). Usually the \( \gamma_d \)'s are interpreted as minimum subsistence requirements defining the lower boundary of the feasible consumption set. Note should be taken here of the mathematical similarity of this structure to (1), indeed Alkire-Foster is in essence a strongly separable representation of preferences for (or aversion to) deprivations.
Stone-Geary is frequently referred to as a “want independent” model, since $\partial^2 U / \partial q_1 \partial q_2 = 0$ for $i \neq j$ (note the same is true for (1)). In a similar fashion, the subsistence levels can be considered independent in this system of preferences since the cross partials with respect to the gamma’s are also zero. Its attraction is that it requires estimation of only $2n - 1$ parameters, much like the deprivation part of $M_K^{\alpha=1}$ in (1) (an unrestricted demand system would require $n^2 - 1$ parameters to be estimated). The problem with (2) in the present context is that it does not admit inferior goods ($\beta_d < 0$ violates concavity), and unfortunately sustaining the concavity assumption also precludes the presence of complementary goods. This is instructive when thinking about the Alkire-Foster index, especially when considering the complementarity arguments for using the set intersection ($K = 0$) version of (1). This is because should an investigator consider estimating a utility function in order to derive the requisite weights for use in the Alkire-Foster index, a Stone-Geary utility function would not suffice. What is required is a more general description of the relationship between various functionings and capabilities of an agents. Furthermore if (1) is seen as an index representing deprivations across groups of goods, with each dimension representing a subgroup, then additively separable preferences over such subgroups imposes extremely strong restrictions over the substitutability between the components across subgroups, especially at the level of subsistence.

To put it more succinctly, the Alkire-Foster index does not admit tradeoffs between the domains of capability at the margin of deprivation. For instance, it does not admit the possibility that an agent may be willing to trade some domestic satisfaction for some work environment satisfaction at the margin. If one wished to accommodate such inter-relationships in an Alkire-Foster index, one could contemplate a modification of (1) to the following form:

$$M_K^\alpha = \frac{1}{N} \sum_{n=1}^{N} \mathbb{I}(c_n > K) \sum_{d=1}^{D} \frac{w_d}{D} \left\{ \mathbb{I} \left( \frac{x_{nd}}{x_d} < 1 \right) \left( \frac{x_d - x_{nd}}{x_d} \right)^\alpha \times \left[ 1 + \sum_{h=d+1}^{D} \frac{w_{ch}}{D} \mathbb{I} \left( \frac{x_{nh}}{x_h} < 1 \right) \left( \frac{x_h - x_{nh}}{x_h} \right)^\alpha \right] \right\}$$

(3)

where $w_{ch}$’s are chosen to reflect the complementary ($w_{ch} > 0$) or substitutability ($w_{ch} < 0$) nature of capabilities $x_d$ and $x_h$, $d \neq h$, $d, h \in \{1, \ldots, D\}$.

\footnote{Stone (1954) only had 19 observations on the prices of and expenditures on 6 commodities!}
3.2 Empirical Examination of the Robustness Issue

Taking a step back from the traditional way of considering robustness through examining ordering across populations, it is proposed here that one should consider instead the “identity” of observations classified as belonging to a group. In other words, to consider the extent to which alternative parameterizations of index (1) identify the same underlying impoverishment groups rather than focus on the index itself, since robust parameterizations should largely identify very similar groups. In other words, to focus on the robustness of the impoverished groups identified under differing parameterizations, rather than on the robustness of the index. This idea in two dimensions is illustrated in diagram 1 below. $U^*$ denotes the deprivation frontier estimated under a relevant set of parameterization, while the lines denoted by “Alkire-Foster” represents the alternative frontier according with Alkire and Foster (2011) under another relevant set of parameterization (the Alkire-Foster frontiers are horizontal and vertical lines due to the fact that it is essentially a count measure for a particular level of deprivation). Robustness in this context is a question then of the extent to which sets $(A \cup C \cup D)$, and $(C \cup B \cup D)$ are equivalent, which thus reduces the question to whether sets $A$ and $B$ are indeed empty sets. This may be examined by a combination of overlap measures (using Anderson et al. (2010), and Anderson et al. (2012)), one which reports the proportion of agents defined as poor by $U^*$ that are reported poor by Alkire-Foster (a value of 1 indicates Alkire-Foster fully covers $U^*$), and one which reports the proportion of agents defined as poor by Alkire-Foster that are reported poor by $U^*$ (a value of 1 indicates $U^*$ fully covers Alkire-Foster). In other words, the two respective measures can be written as,

$$\begin{align*}
\text{OV}(U^* \text{ vs. A-F})|A-F &= \sum_{\delta \in \Delta_{A-F}} \min \left\{ f_{A-F}(z_\delta), \frac{n_{U^*}}{n_{A-F}} f_{U^*}(z_\delta) \right\} \quad (4) \\
\text{OV}(U^* \text{ vs. A-F})|U^* &= \sum_{\delta \in \Delta_{U^*}} \min \left\{ \frac{n_{A-F}}{n_{U^*}} f_{A-F}(z_\delta), f_{U^*}(z_\delta) \right\} \quad (5)
\end{align*}$$

where $n_j \in \{n_{A-F}, n_{U^*}\}$ is the number of observations classified as poor under the respective method $j$, $z_\delta$ is the $K$ vector of variable realizations, and $\Delta_j$ is the set of cells

---

Note that when wellbeing is measured by a single variable, the difference between equation (1) and another alternative is in the weight applied, so that either one distribution will first order stochastically dominate the other depending on the weight, and one of these sets will always be empty. This is another sign of the complications associated with increasing dimensionality, since the likelihood of such sets arising increases as the number of dimensions increases.
classified as poor. Further, the overlap index of Anderson et al. (2010) is used to examine the overlap or joint coverage of the two poverty samples generated.

\[
OV_{cov} = \sum_{\delta \in \Delta_{U^*\cup A-F}} \min \{f_{A-F}(z_\delta), f_{U^*}(z_\delta)\}
\] (6)

A robustness index\(^7\) on the \([0, 1]\) interval is provided by taking an average of the two overlap measures, with 1 implying absolute robustness, and 0 implying no connection between the two poverty groups.

---

\(^7\)Generally these Overlap measures are approximately \(N(p, (p(1-p)/n))\), which facilitates inference.
4 Incorporating a Flexible Representation of Agent Preferences into the Calculus

Since the Alkire-Foster index can be thought of as an aggregation across individuals of various aspects of their respective wellbeing functions, and the extent to which they are below a particular wellbeing level, the index’s relationship to the utility or welfare function should thus be grounded in wellbeing or utility theory. Since the self-reported satisfaction questionnaire data set employed here records responses on satisfaction with respect to overall wellbeing, as well as responses to the various dimensions of wellbeing, it presents an opportunity to examine the importance of the various dimensions of wellbeing that constitute overall happiness, $U(x)$. To be precise, overall happiness $U(x)$ has $K$ dimensions, in other words $U(x) = U(U_1(x_1), U_2(x_2), \ldots, U_K(x_K))$, where $x_k$, $k \in \{1, \ldots, K\}$ are each $M_k$ vectors of responses to questions on wellbeing on the $k$th dimension, so that $x = [x_1, \ldots, x_K]'$ is a $\sum_{k=1}^{K} M_k$ dimensional vector. Unlike conventional empirical approaches to Consumer Theory (see for example Deaton and Muellbauer (1980)), where the nature of preferences (the structure of $U(.)$) has to be inferred from agents’ observed incomes, expenditures on goods, and the prices they face, here direct observations on $U(x)$, and estimates of its various sub-components $U_k(x_k)$, $k = \{1, \ldots, K\}$, are available. Given these, the structure of $U(.)$ can be estimated directly, which will in turn provide information on the structure of indices such as (1), or provide alternative indices.

A data set useful in affording an opportunity of studying the relationships between various categories of wellbeing in the context of an overall measure of wellbeing is available (Anand, Gray and Roupe (2013 in preparation)). In 2011, nationally representative populations were surveyed in the U.S.A., U.K. and Italy. The panel of respondents was drawn equally from four divisions of each country, representative of working age adults in terms of age, gender and social class. The U.S. data is used for demonstrative purposes, with 1061 observations. However, due to non-response to question 4 regarding opportunities and constraints (appendix A for details of question), the final data used in the analysis was culled to 725. Although unreported here, results from the full data set with question 4 excluded instead, as well as a smaller random set were investigated, and the results found here are robust. Aside from an overall index of wellbeing ($Y$) (based upon 4

\textsuperscript{8}For discussions of the satisfaction, happiness and utility nexus see Clark et al. (2008), Clark et al. (2009), Kahneman and Deaton (2010), Kimball and Willis (2006), Layard (2005).
indicators, though only 3 will be used in the study here), satisfaction with respect to six categories of wellbeing were surveyed (respondents were asked to report satisfaction on a 0 – 10 scale, with 0 denotes “strongly disagree”, and 1 denotes “strongly agree”). There were 6 general subcategories of Wellbeing (Details of specific questions are in Appendix A), each with a number of sub-dimensions. The subcategories were as follows:

1. $X_1 \sim$ Social Relationships External to the Family (3 dimensions)
2. $X_2 \sim$ Social Relationships Internal to the Family (7 dimensions)
3. $X_3 \sim$ Work opportunities and constraints (6 dimensions)
4. $X_4 \sim$ Local Social opportunities and constraints (4 dimensions)
5. $X_5 \sim$ Environmental Concerns (5 dimensions)
6. $X_6 \sim$ Access to services (7 dimensions)

4.1 Structuring the Estimating Equations

Estimating an unrestricted wellbeing function over all of the sub-components within each dimension of wellbeing in this data would require 496 parameters, which given a sample size of 725 leaves very few degrees of freedom, so the following approach is proposed. Begin with $U(X) \equiv U(U_1(X_1), U_2(X_2), \ldots, U_6(X_6))$, where $X$ is the vector of all of the sub-components in all of the dimensions, so $X = [X_1, X_2, X_3, X_4, X_5, X_6]'$ with $X_k, k = \{1, \ldots, 6\}$, being the vector of sub-components of category $k$. Thus $U$ is assumed to be weakly separable in the sub-components of each dimension of wellbeing. For these sub-component functions, $U_i(.)$, Anderson et al. (2011) is employed to provide an aggregating wellbeing index of the sub-components in each dimension. It is shown that two-sided bounds can be placed on each dimension of wellbeing for each observation, using only the assumptions that each wellbeing component be non-decreasing and weakly quasi-concave with respect to the sub-components. The bounds encompass the entire set of wellbeing sub-components consistent with monotonicity and quasi-concavity, and as such

---

9This is tantamount to assuming a weakly separable wellbeing function with respect to the sub-components, whilst admitting more standard substitutability and complementarity properties between their aggregates (Gorman (1968), and Kannai (1980)).
they provide a region within which any parametric index with these properties can be expected to lie\textsuperscript{10}, here the midpoint of the bounds is used. In other words,

\[ U_i(X_i) = ||X_i|| \left( 1 - \frac{D_i + D_i}{2} \right) \]

The two dimensional representation of the technique is depicted in figure 2. The approach is applied directly to the data, and is fully nonparametric in the sense that it does not require any further assumptions on the functional form of the welfare function, nor does it require the estimation of any functions of the data. Indeed the method can be applied to very small datasets (as well as to large ones) where statistical techniques — and especially nonparametric statistical techniques — could not be relied upon. A useful feature is that the methodology is easily replicable requiring nothing more complex than standard linear programming techniques.

The wellbeing index \( U_i(X_i) \) for an observation \( X_i \) in the \( i^{th} \) wellbeing component may be written as:

\[ U_i(X_i) = \left( 1 - \frac{D_i + D_i}{2} \right) ||X_i|| \]

\[ = \left( 1 - \frac{\bar{D}_i + D_i}{2} \right) \sqrt{\sum_{j=1}^{J_i} x_{ij}^2} \]

where \( \bar{D}_i \) and \( D_i \) are the upper and lower bounds respectively of the distance function, illustrated for two dimensions above in figure 2. In essence, this is one minus the distance function which measures the amount by which one has to scale the Euclidean norm of the component vector, \( ||X_i|| \), of an observation, so that it is increasing in the sub-components, and decreasing in the reference welfare level, \( W \). In other words, it can be thought of as measuring the “size” of \( X_i \) relative to the reference welfare level \( W \), which is notionally midway between the two bounds. Insofar as \( W \) is arbitrary, we define it as follows:

\[ W = U_i(X_i) \left( \frac{\bar{D}_i + D_i}{2} \right) \]

\[ \Rightarrow U_i(X_i) = \frac{W}{\left( \frac{\bar{D}_i + D_i}{2} \right)} \]

\textsuperscript{10}The bounds’ span can be shown to diminish with the wellbeing index (bounds are tighter on high wellbeing, and loose on low wellbeing), and generally increase with dimensionality.
To understand the impacts of the various components note that:

\[
\begin{align*}
\partial U_i(X_i) &= \frac{x_{ij}}{\sqrt{\sum_{j=1}^{J_i} x_{ij}^2}} \left( 1 - \frac{D_i + D_i^2}{2} \right) \\
&= \frac{x_{ij}}{\sum_{j=1}^{J_i} x_{ij}^2} U_i(X_i) \\
&= \left[ \sum_{j=1}^{J_i} \frac{x_{ij}^2}{(D_i + D_i)^2} \right] W > 0
\end{align*}
\]

where \(J_i\) denotes the total number of sub-components within component \(i\). Let overall wellbeing satisfaction (\(U\)) be quadratic in the components of wellbeing satisfactions (\(X_i\)'s), thus for the \(k^{th}\) agent, her vector of component wellbeing outcomes is \(X_k = [X_{k,1}, X_{k,2}, \ldots, X_{k,6}]^T\), and her overall satisfaction or level of felicity is:

\[
U_k = X_k'AX_k
\]
Further, let $|A| < 0$, and $U$ be monotonically non-decreasing, and concave in $X$ (essentially what consumer behaviour theorists require of preference structures). Note the caveat that concave preferences are usually only required in consumer theory because budget sets are weakly convex, and concavity facilitates constrained optimization solutions. Technically all that is required is for preferences to be less convex than the budget sets (or more concave than the budget sets if they are concave). Further, adopting Young’s theorem, $A$ will be symmetric, (indeed all that can be identified in a regression version of (10) is a symmetric version of $A$), and a version of $A$ consistent with an Alkire-Foster structure would have $A$ be a diagonal matrix.

Further, let the measure of marginal wellbeing with respect to the $j^{th}$ sub-component of the $i^{th}$ wellbeing component, and the marginal rate of substitution between sub-components of the $i^{th}$ and $h^{th}$ wellbeing components be given by:

$$\frac{\partial U}{\partial x_{ij}} = 2 \sum_{m=1}^{6} a_{im} X_m \frac{x_{ij}}{\sum_{j=1}^{J} \left( \frac{x_{ij}^2}{Y_{ij} + D_{ij}} \right)} W > 0 \text{ if } \sum_{m=1}^{6} a_{im} X_m > 0$$

and

$$\frac{\partial^2 U}{\partial x_{ij} \partial x_{hn}} = 2a_{ih} \frac{x_{ij}}{\sum_{j=1}^{J} \left( \frac{x_{ij}^2}{Y_{ij} + D_{ij}} \right)} \frac{x_{hn}}{\sum_{h=1}^{H} \left( \frac{x_{hj}^2}{Y_{hj} + D_{hj}} \right)} W^2$$

where the sign of the latter is governed by the sign of $a_{ih}$, the $\{i, h\}$ element of the matrix $A$.

The above representation of preferences requires only 21 estimated parameters (a substantial improvement over 496!), but still admits a measure of substitutability between sub-components of the different categories. Deprivation in functionings and capabilities may now be considered in terms of a basic standard, such as some percentile of the population who have satisfaction wellbeings below some particular metric (for example those below a wellbeing measure reflecting some percentage of the median level or average level in each dimension). Since $U(.)$ can be computed, standard univariate deprivation, inequality and polarization measures can all be employed using this calculation. The attraction of the approach is that it provides maximal flexibility in the representation of wellbeing component deprivations, whilst admitting the possibility of sub-component substitutability/complementarity in the overall index, and retaining the ability to measure the impact of improvements/worsenings of sub-components within each category.
5 Empirical Analysis

5.1 Summary Statistics

At the heart of the analysis are summary indices for overall wellbeing, and indices for each of the components of wellbeing for each agent, based upon Anderson et al. (2011). Tables 1 (the indices) and 2 (their bounds) report their statistical properties. From table 1, the distribution of each component across agents is left skewed (that is their means are less than medians, implying the distribution is dense in the lower range of the index, but sparse in the upper range), and are similar across the domains in terms of their summary statistics.

<table>
<thead>
<tr>
<th>Summary Stat</th>
<th>( Y(4) )</th>
<th>( X_1(3) )</th>
<th>( X_2(7) )</th>
<th>( X_3(6) )</th>
<th>( X_4(4) )</th>
<th>( X_5(4) )</th>
<th>( X_6(7) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimums</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>Maximums</td>
<td>10.0000</td>
<td>10.0000</td>
<td>10.0000</td>
<td>10.0000</td>
<td>10.0000</td>
<td>10.0000</td>
<td>10.0000</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>1.0933</td>
<td>1.1949</td>
<td>1.5053</td>
<td>1.7341</td>
<td>1.2395</td>
<td>1.5192</td>
<td>1.3774</td>
</tr>
</tbody>
</table>

As noted above, the ACL bounds encompass the entire set of possible welfare indices consistent with monotonicity and quasi-concavity, and can be used as a computationally convenient robustness check on parametric methods. Thus the difference between the high and low bounds provides an idea of the range within which the true welfare function lies. As can be seen from table 2, the distribution of bounds are right skewed, and negatively correlated with the indices themselves, implying that the bounds on high levels of wellbeing are tight (i.e. narrow), whereas bounds on low levels of wellbeing are loose.

The domain indices were then employed to estimate a symmetric version of \( A \) in the regression equation:

\[
Y^2_n = \sum_{i=1}^{6} \sum_{j=1}^{6} a_{ij} X_{in} X_{jn} + e_n \quad \text{for} \quad n = 1, \ldots, N
\]

Here \( Y^2(.) \) is set to \( Y^2 \) for scaling reasons, the standard regression assumptions on \( e \) and the \( X \)'s are employed, and \( n \) is the number of agents in the sample (725 in this case).
Table 2: Bounds for ACL Indices (High – Low)

<table>
<thead>
<tr>
<th>Index (Dimensions)</th>
<th>Average Difference</th>
<th>Median Difference</th>
<th>Standard Deviation</th>
<th>Index/Bounds Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y(4)</td>
<td>0.5385</td>
<td>0.2222</td>
<td>1.2209</td>
<td>-0.6594</td>
</tr>
<tr>
<td>X1(3)</td>
<td>1.0077</td>
<td>0.3493</td>
<td>1.9795</td>
<td>-0.8858</td>
</tr>
<tr>
<td>X2(7)</td>
<td>1.8707</td>
<td>0.7334</td>
<td>2.7586</td>
<td>-0.9415</td>
</tr>
<tr>
<td>X3(6)</td>
<td>2.1042</td>
<td>0.6112</td>
<td>3.1338</td>
<td>-0.9258</td>
</tr>
<tr>
<td>X4(4)</td>
<td>1.1063</td>
<td>0.375</td>
<td>2.0585</td>
<td>-0.8224</td>
</tr>
<tr>
<td>X5(4)</td>
<td>1.5594</td>
<td>0.375</td>
<td>2.8352</td>
<td>-0.9284</td>
</tr>
<tr>
<td>X6(7)</td>
<td>1.395</td>
<td>0.4584</td>
<td>2.6192</td>
<td>-0.9533</td>
</tr>
</tbody>
</table>

Based on these results\textsuperscript{11}, table 3 reports the “marginal wellbeing” at various values of the component indices (means, medians and minimums). These evaluated derivatives (where \( \hat{a}_{ij} \) indicates estimated \( a_{ij} \) and mean, median and minimum \( X \) respectively) may be written as:

\[
\frac{\partial Y^2}{\partial X_j} = 2 \sum_{i=1}^{6} \hat{a}_{ij} \hat{X}_i \quad \text{for} \quad j = 1, \ldots, 6
\]

They may be construed as reflecting respectively the “average”, the “median”, and the poorest possible person’s marginal benefit from each domain, and would provide a basis for the weights in an Alkire-Foster (A-F) index of the components. Notice that these are not “need independent” equations (i.e. the incremental benefit accruing to an incremental change in the \( i \text{th} \) component depends upon the status of the \( j \text{th} \) component).

From the results of table 3, note first the substantial differences in the weights across the \( X \)’s, arguing for something other than equal weighting (Atkinson et al. (2002)) across the domains. Since the determinant of \( A \) is negative, the wellbeing function is concave (albeit slightly). However note the problem with the fifth factor (Local Social Opportunities and Constraints), since it is negative across all average, median and poorest agents, implying that \( U \) is not monotonic non-decreasing in this particular domain. Thus, based upon the elimination of the fifth factor, a 5 factor model is considered in the following.

\textsuperscript{11}The results are not reported here but are available from the authors on request.
Table 3: Marginal Benefits from the Six Domains

<table>
<thead>
<tr>
<th>Derivative</th>
<th>Sample Means</th>
<th>Sample Medians</th>
<th>Sample Minimums</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\partial U/\partial X_1$</td>
<td>3.0859</td>
<td>3.0771</td>
<td>0.3253</td>
</tr>
<tr>
<td>$\partial U/\partial X_2$</td>
<td>2.5161</td>
<td>2.6878</td>
<td>0.2945</td>
</tr>
<tr>
<td>$\partial U/\partial X_3$</td>
<td>1.0160</td>
<td>1.2512</td>
<td>0.1306</td>
</tr>
<tr>
<td>$\partial U/\partial X_4$</td>
<td>1.6600</td>
<td>1.6744</td>
<td>0.1680</td>
</tr>
<tr>
<td>$\partial U/\partial X_5$</td>
<td>-0.1846</td>
<td>-0.3052</td>
<td>-0.0338</td>
</tr>
<tr>
<td>$\partial U/\partial X_6$</td>
<td>0.9950</td>
<td>1.1880</td>
<td>0.1205</td>
</tr>
</tbody>
</table>

$\text{det}(A) = -0.0971$

5.2 5 Factor Model & Optimal Weights

Table 4 presents the full set of estimated coefficients for the 5 factor model, which is essentially the 6 factor model with the fifth factor omitted. A test of the restrictions implied by omitting the fifth domain ($\chi^2(5) = 8.2267$ ($P = 0.1442$) or $F(5, 254) = 2.1367$ ($P = 0.0616$)), fails to reject the hypothesis that the domain (Local Social Opportunities and Constraints) is not relevant in describing overall satisfaction. Recall that what is being estimated is a symmetric version of $A$, and a model consistent with a “needs independent” structure of the basic A-F model would require a diagonal $A$. The test of the 15 restrictions that this implies yields $\chi^2(15) = 70.277$ ($P = 0.0000$) or $F(15, 244) = 12.3800$ ($P = 0.0000$), strongly rejecting the restrictions that “need independence” implies.

The most important point here is the existence of statistically significant substitution trade-offs between the domains. $\partial^2 U/\partial X_1 \partial X_3 = -1.2134$, (Social Opportunities External to the Family versus Work Opportunities are substitutes), $\partial^2 U/\partial X_1 \partial X_6 = -0.5993$ (Social Opportunities External to the Family versus Access to Services are substitutes) and finally $\partial^2 U/\partial X_3 \partial X_5 = -0.7297$ (Work Opportunities versus Environmental Concerns are substitutes), all of which make a great deal of sense. Interestingly there are hints of some complementarities (Social Opportunities External versus Internal to the Family, Work versus Access to Services, and Environmental Concerns versus Access to Services), but none of which are significant at usual levels of significance. Note also that these tradeoffs could not be accommodated in an A-F type index (1) which essentially reflects “needs independent” preferences on the part of agents.
Table 4: The Five Variable Model Coefficients

| Variable | Coefficient | Std. Error | |Z| 1 - F(|Z|) |
|----------|-------------|------------|-----------|--------------|
| $X_1 \times X_1$ | 1.5013 | 0.1495 | 10.0428 | 0.0000 |
| $X_1 \times X_2$ | -0.5085 | 0.2079 | 2.4455 | 0.0072 |
| $X_1 \times X_3$ | -0.5307 | 0.2020 | 2.6270 | 0.0043 |
| $X_1 \times X_5$ | -0.0617 | 0.2199 | 0.2805 | 0.3896 |
| $X_1 \times X_6$ | -1.2253 | 0.2476 | 4.9477 | 0.0000 |
| $X_2 \times X_2$ | 1.3284 | 0.1552 | 8.5573 | 0.0000 |
| $X_2 \times X_3$ | -0.4169 | 0.1763 | 2.3655 | 0.0090 |
| $X_2 \times X_5$ | -0.7205 | 0.2016 | 3.5734 | 0.0002 |
| $X_2 \times X_6$ | -0.3155 | 0.1755 | 1.4639 | 0.0716 |
| $X_3 \times X_3$ | 0.3811 | 0.1475 | 2.5835 | 0.0049 |
| $X_3 \times X_5$ | 0.2396 | 0.1536 | 1.5598 | 0.0594 |
| $X_3 \times X_6$ | 0.1760 | 0.1984 | 0.8871 | 0.1875 |
| $X_5 \times X_5$ | 0.2535 | 0.1260 | 2.0129 | 0.0221 |
| $X_5 \times X_6$ | 0.1692 | 0.1736 | 0.9745 | 0.1649 |
| $X_6 \times X_6$ | 0.7359 | 0.1320 | 5.5730 | 0.0000 |

$R^2 = 0.4651$, and $\sigma^2 = 117.8034$.

Table 5 reports the marginal benefits in this model. As may be observed, this preference structure is convex ($|A| > 0$), though only marginally so, but now overall satisfaction is monotonic non-decreasing in all domains. Note again the very different marginal effects, which again provides a counter argument against equal weighting in an A-F type index.

To examine the effects of various specifications of deprivation, an arbitrary cutoff of 0.7× median is considered. Poverty counts for each domain would be for $X_1 \sim 0.1055$, $X_2 \sim 0.1545$, $X_3 \sim 0.0927$, $X_5 \sim 0.1127$ and $X_6 \sim 0.1164$. Table 6 reports an Intersection Rule ($c_n = K+1$) poverty measure, and three A-F measures, the first of which is with equal weighting, the second with needs independent utility weighting based upon weights from table 5, which are essentially Union Rule measures (since $c_n > K$), and the final one based on the full utility weighting from table 4 using equation (3). Note the substantial change in the measures when the weighting scheme changes from equal to “utility” weightings, which to a considerable degree supports the concerns expressed in Ravallion (2010). It is
Table 5: Marginal Benefits from the Five domains

<table>
<thead>
<tr>
<th>Derivative</th>
<th>Sample Means</th>
<th>Sample Medians</th>
<th>Sample Minimums</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\partial U / \partial X_1$</td>
<td>3.3140</td>
<td>3.1962</td>
<td>0.3382</td>
</tr>
<tr>
<td>$\partial U / \partial X_2$</td>
<td>2.9188</td>
<td>3.1849</td>
<td>0.3477</td>
</tr>
<tr>
<td>$\partial U / \partial X_3$</td>
<td>0.9298</td>
<td>1.1106</td>
<td>0.1151</td>
</tr>
<tr>
<td>$\partial U / \partial X_4$</td>
<td>0.6415</td>
<td>0.6932</td>
<td>0.0669</td>
</tr>
<tr>
<td>$\partial U / \partial X_6$</td>
<td>1.2617</td>
<td>1.3991</td>
<td>0.1381</td>
</tr>
</tbody>
</table>

$\det(A) = 0.0324$

also important to point out the robustness of the fully weighted A-F measure ($A-F_{WF}$) as the number of impoverished domains that constitute poverty increases, relative to all other versions of the index. This feature combined with its robustness in identifying poverty, evident in the following discussion, makes it the preferred measure.

Table 6: Comparison Between Intersection Rule vs. Alkire-Foster (A-F) Poverty Measures

<table>
<thead>
<tr>
<th>Minimum Number of Domains $(K + 1)$ in the A-F index</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intersection Rule – Equal Weight</td>
<td>0.1917</td>
<td>0.0455</td>
<td>0.0262</td>
<td>0.0166</td>
<td>0.0069</td>
</tr>
<tr>
<td>Alkire-Foster – Equal Weight (A-F)</td>
<td>0.2869</td>
<td>0.0952</td>
<td>0.0497</td>
<td>0.0234</td>
<td>0.0069</td>
</tr>
<tr>
<td>Alkire-Foster – Utility Weight Needs (A-F&lt;sub&gt;W&lt;/sub&gt;)</td>
<td>0.7251</td>
<td>0.5266</td>
<td>0.4376</td>
<td>0.4297</td>
<td>0.2000</td>
</tr>
<tr>
<td>Independent (A-F&lt;sub&gt;W&lt;/sub&gt;)</td>
<td>0.4566</td>
<td>0.3347</td>
<td>0.3698</td>
<td>0.3539</td>
<td>0.2000</td>
</tr>
</tbody>
</table>

5.3 Comparison of the Alkire-Foster and Utility Based Index

The enhancements to the Alkire-Foster index, manifested in the utility based index is best understood in the following comparison of the extent to which the utility based index ($U$), and A-F cover each other. For this exercise the “utility based” ($U$) measure considers the wellbeing levels enjoyed by individuals at 0.7, 0.8, 0.9, and 1.0 of the domain medians as
metrics for poorness. The deprivation levels of wellbeing (associated with the $U$ measure), and deprivation rate (associated with the A-F measure) pairings at 70%, 80%, 90%, and 100% of domain medians are respectively (44.2568, 0.0400), (57.8048, 0.0618), (73.1591, 0.1382) and (90.3199, 0.4309).

Table 7: Coverage of $U$ & an Equally Weighted Alkire-Foster (A-F) Index

<table>
<thead>
<tr>
<th>Minimum Number of Domains $(K + 1)$ in the A-F Index (Equal Weighting)</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coverage ($U$ vs. A-F)</td>
<td>A-F</td>
<td>0.0817</td>
<td>0.1739</td>
<td>0.3055</td>
<td>0.4706</td>
</tr>
<tr>
<td></td>
<td>$U(0.7 \times$ Domain Medians)</td>
<td>0.1010</td>
<td>0.2029</td>
<td>0.3333</td>
<td>0.5294</td>
</tr>
<tr>
<td></td>
<td>$U(0.8 \times$ Domain Medians)</td>
<td>0.2404</td>
<td>0.4348</td>
<td>0.5833</td>
<td>0.6471</td>
</tr>
<tr>
<td></td>
<td>$U(0.9 \times$ Domain Medians)</td>
<td>0.5529</td>
<td>0.7101</td>
<td>0.8056</td>
<td>0.8824</td>
</tr>
<tr>
<td></td>
<td>$U(1.0 \times$ Domain Medians)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Coverage ($U$ vs. A-F)</th>
<th>1.0000</th>
<th>0.7059</th>
<th>0.6471</th>
<th>0.4706</th>
<th>0.1765</th>
</tr>
</thead>
<tbody>
<tr>
<td>$U(0.7 \times$ Domain Medians)</td>
<td>0.7778</td>
<td>0.5185</td>
<td>0.4444</td>
<td>0.3333</td>
<td>0.1111</td>
</tr>
<tr>
<td>$U(0.8 \times$ Domain Medians)</td>
<td>0.7042</td>
<td>0.4225</td>
<td>0.2958</td>
<td>0.1549</td>
<td>0.0423</td>
</tr>
<tr>
<td>$U(0.9 \times$ Domain Medians)</td>
<td>0.3734</td>
<td>0.1591</td>
<td>0.0942</td>
<td>0.0487</td>
<td>0.0130</td>
</tr>
<tr>
<td>$U(1.0 \times$ Domain Medians)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Joint Coverage: Utility Based vs. Equal Weighting †

| | $U(0.7 \times$ Domain Medians) | 0.5409 | 0.4399 | 0.4763 | 0.4706 | 0.3882 |
|---|---|---|---|---|---|
| $U(0.8 \times$ Domain Medians) | 0.4394 | 0.3607 | 0.3889 | 0.4314 | 0.3556 |
| $U(0.9 \times$ Domain Medians) | 0.4723 | 0.4287 | 0.4396 | 0.4010 | 0.3211 |
| $U(1.0 \times$ Domain Medians) | 0.4631 | 0.4346 | 0.4499 | 0.4655 | 0.4065 |

†An index of the commonality of the groups being covered, where proximity to 1 implies identical groups are covered, while proximity to 0 implies groups are completely segmented.

Tables 7 and 8 report the various degrees to which the A-F, and the $U$ deprivation index proposed here cover the same group of agents using equations (4), (5), and (6). Each column reports the comparison of the two types of measures under the assumption that impoverishment in $K$ domains constitute poverty. The 2 dimensional figure 1 in section 3.2 illustrates the issue, agents in areas $C$ and $D$ are recorded as poor by both
Table 8: Coverage of $U$ & a Utility Weighted Alkire-Foster (A-F$_W$) Index

<table>
<thead>
<tr>
<th>Minimum Number of Domains ($K + 1$) in the A-F Index (Needs Independent Utility Weighting)</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coverage ($U$ vs. A-F$_W$)</td>
<td>Coverage ($U$ vs. A-F$_W$)</td>
<td>A-F$_W$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$U(0.7 \times$ Domain Medians)</td>
<td>0.1354</td>
<td>0.2075</td>
<td>0.2903</td>
<td>0.6364</td>
<td>0.6000</td>
</tr>
<tr>
<td>$U(0.8 \times$ Domain Medians)</td>
<td>0.1667</td>
<td>0.2453</td>
<td>0.3226</td>
<td>0.6364</td>
<td>0.6000</td>
</tr>
<tr>
<td>$U(0.9 \times$ Domain Medians)</td>
<td>0.3750</td>
<td>0.4906</td>
<td>0.5484</td>
<td>0.8182</td>
<td>0.6000</td>
</tr>
<tr>
<td>$U(1.0 \times$ Domain Medians)</td>
<td>0.7083</td>
<td>0.7358</td>
<td>0.8065</td>
<td>0.9091</td>
<td>0.8000</td>
</tr>
<tr>
<td>Coverage ($U$ vs. A-F$_W$)</td>
<td>Coverage ($U$ vs. A-F$_W$)</td>
<td>$U$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$U(0.7 \times$ Domain Medians)</td>
<td>0.7647</td>
<td>0.6471</td>
<td>0.5294</td>
<td>0.4118</td>
<td>0.1765</td>
</tr>
<tr>
<td>$U(0.8 \times$ Domain Medians)</td>
<td>0.5926</td>
<td>0.4815</td>
<td>0.3704</td>
<td>0.2593</td>
<td>0.1111</td>
</tr>
<tr>
<td>$U(0.9 \times$ Domain Medians)</td>
<td>0.5070</td>
<td>0.3662</td>
<td>0.2394</td>
<td>0.1268</td>
<td>0.0423</td>
</tr>
<tr>
<td>$U(1.0 \times$ Domain Medians)</td>
<td>0.2208</td>
<td>0.1266</td>
<td>0.0812</td>
<td>0.0325</td>
<td>0.0130</td>
</tr>
<tr>
<td>Joint Coverage: Utility Based vs. Needs Independent Weighting $†$</td>
<td>Joint Coverage: Utility Based vs. Needs Independent Weighting $†$</td>
<td>$U$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$U(0.7 \times$ Domain Medians)</td>
<td>0.4501</td>
<td>0.4273</td>
<td>0.4099</td>
<td>0.5241</td>
<td>0.3882</td>
</tr>
<tr>
<td>$U(0.8 \times$ Domain Medians)</td>
<td>0.3796</td>
<td>0.3634</td>
<td>0.3465</td>
<td>0.4478</td>
<td>0.3556</td>
</tr>
<tr>
<td>$U(0.9 \times$ Domain Medians)</td>
<td>0.4410</td>
<td>0.4284</td>
<td>0.3939</td>
<td>0.4725</td>
<td>0.3211</td>
</tr>
<tr>
<td>$U(1.0 \times$ Domain Medians)</td>
<td>0.4646</td>
<td>0.4312</td>
<td>0.4438</td>
<td>0.4708</td>
<td>0.4065</td>
</tr>
</tbody>
</table>

$†$An index of the commonality of the groups being covered, where proximity to 1 implies identical groups are covered, while proximity to 0 implies groups are completely segmented.

indices, agents in areas $B$ are recorded as poor by $U^*$ but not by A-F, and agents in areas $A$ are recorded as poor by A-F but not by $U^*$. These results highlight the fact that there are profound differences in who are identified as impoverished by the $U$ approach, and the A-F approach. Similarly the weighting system in the A-F index has a profound effect on who is deemed poor. In other words, in the context of figure 1, sets $A$ and $B$ are non-trivial in terms of the sample considered. Ultimately, the choice of measure between the A-F index or the utility based $U$ is dependent on the researcher’s assumptions regarding whether the variables or components used are a true reflection of individual utility. Tables
7 and 8 simply highlights that implications derived from one index need not generalize to the other.

The main contribution here is that the utility based weighting of the A-F index that incorporate notions of complementarity and substitutability of wellbeing dimensions is significantly more robust and consistent at identifying poverty groups. Robustness as far as A-F indices are concerned is best examined by comparing how well different versions of the indices cover each other. Table 9 reports the results of these comparisons. As may be seen, the Full Utility Weights index (which includes the substitutability components) fully covers both the more restrictive equally weighted, and needs independent utility weighted versions of the index, whereas the reverse is not generally true (“needs independent” and equally weighted versions do a reasonable job of covering each other). To put it another way, with the full utility weighting augmented A-F measure, regardless of the number of impoverishment dimensions that constitute poverty considered, the utility weighted A-F is very robust in its ability to cover the other versions of A-F, or other definitions of poverty, which combined with the fact that the utility based weights include notions of substitutability between the components of wellbeing, makes the augmented A-F measure very versatile, and therefore the preferred measure.
Table 9: Pairwise Coverage Between Different Versions of Alkire-Foster

<table>
<thead>
<tr>
<th>Pairwise Comparisons</th>
<th>Minimum Number of Domains ($K + 1$) in A-F Index</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Coverage (A-F vs. A-F&lt;sub&gt;W&lt;/sub&gt;)</td>
<td>1.0000</td>
</tr>
<tr>
<td>Coverage (A-F vs. A-F&lt;sub&gt;W&lt;/sub&gt;)</td>
<td>0.4615</td>
</tr>
<tr>
<td>Joint Coverage: Equal vs. Needs</td>
<td>0.7308</td>
</tr>
<tr>
<td><strong>Independent Weighting</strong></td>
<td></td>
</tr>
<tr>
<td>Coverage (A-F&lt;sub&gt;W&lt;/sub&gt; vs. A-F&lt;sub&gt;WF&lt;/sub&gt;)</td>
<td>0.4615</td>
</tr>
<tr>
<td>Coverage (A-F&lt;sub&gt;W&lt;/sub&gt; vs. A-F&lt;sub&gt;WF&lt;/sub&gt;)</td>
<td>1.0000</td>
</tr>
<tr>
<td>Joint Coverage: Needs Independent vs.</td>
<td>0.7308</td>
</tr>
<tr>
<td><strong>Full Weighting</strong></td>
<td></td>
</tr>
<tr>
<td>Coverage (A-F vs. A-F&lt;sub&gt;WF&lt;/sub&gt;)</td>
<td>1.0000</td>
</tr>
<tr>
<td>Coverage (A-F vs. A-F&lt;sub&gt;WF&lt;/sub&gt;)</td>
<td>1.0000</td>
</tr>
<tr>
<td>Joint Coverage: Equal vs. Full Weighting</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

Note: A-F $\sim$ equally weighted Alkire-Foster, A-F<sub>W</sub> $\sim$ needs independent weighted Alkire-Foster, and A-F<sub>WF</sub> $\sim$ Full utility weights Alkire-Foster (3)

6 Conclusions

The move toward understanding wellbeing (and the lack of it) in terms of functionings and capabilities has fostered considerable interest in multi-dimensional wellbeing measurement with increasing dimensionality in terms of the domains of deprivation being high on the agenda. This in turn has lead to proposals for measures of deprivation in the context of the functionings and capabilities approach which “count” the numbers of dimensions or domains in which agents are deprived and those agents who reach at least a prespecified number of domain deprivations are considered poor. These so-called “Mashup” indices have not been without their critics whose concern has largely been with the robustness of these indices to vagaries in the choice of parameters which are largely at the behest of investigators. There are also technical concerns with increasing dimensionality which have hitherto been overlooked both in terms of the informational requirements placed upon data and the over-simplistic view of the inter-relationships between the domains of satisfaction inherent in “Mashup” type indices.
Here a new approach is proposed which addresses some of these issues. It is facilitated by a unique data set which records overall satisfaction on the part of agents as well as their satisfaction with sub domains of functioning and capability which permits estimation of the inter-relationship between sub-domain satisfactions and overall satisfactions. In this work a subsample of a panel of respondents was drawn equally from four divisions of the USA which are representative of working age adults in terms of age, gender and social class in 2011. The attraction of the approach is that it provides as much flexibility and generality as possible in the representation of sub category deprivations whilst admitting the possibility of reflecting subgroup substitutability/complementarity in the overall index and retaining the ability to measure the impact of improvements/worsenings of sub-category sub components on measures of overall deprivation. Application of these ideas using these data sets revealed significant inter-relationships between sub domain satisfactions and considerable differences between the groups of individuals captured by the various measures as deprived or impoverished. This prompted the development of a modified version of the Alkire-Foster index which admits notions of substitutability and complementarity between the various aspects of functioning and capability impoverishment in the index. This modified index was found to have superior robustness properties to other more restrictive versions of the index.

**References**


A Appendix

A.1 Wellbeing Variables Used in this Study

1. Please could you say how you would rate each of the following aspects of your life? Please rate on a scale of 0 to 10, where 0 indicates the lowest rating you can give and 10 the highest.

   (a) Overall, how satisfied are you with your life nowadays?
   (b) Overall, how happy did you feel yesterday?
   (c) Overall, how anxious did you feel yesterday?
   (d) Overall, to what extent do you feel that the things you do in your life are worthwhile?

2. Please indicate how satisfied, or dissatisfied you are with the following on a scale of 0 to 10, where 0 indicates you are very dissatisfied and 10 that you are very satisfied.

   (a) Friendships
   (b) Relationships with colleagues at work
   (c) The neighborhood in which you live

3. Here are some questions about the opportunities and constraints that you face. For each of the following statements, please indicate how much you agree, or disagree on a scale of 0 to 10, where 0 indicates you strongly disagree and 10 that you strongly agree.

   (a) I am able to share domestic tasks within the household fairly
   (b) I am able to socialise with others in the family as I would wish
   (c) I am able to make ends meet
   (d) I am able to achieve a good work-life balance
   (e) I am able to find a home suitable for my needs
   (f) I am able to enjoy the kinds of personal relationships that I want
   (g) I have good opportunities to feel valued and loved

4. Here are some questions about the opportunities and constraints that you face. For each of the following statements, please indicate how much you agree, or disagree on a scale of 0 to 10, where 0 indicates you strongly disagree and 10 that you strongly agree.

   (a) I am able to find work when I need to
(b) I am able to use my talents and skills at work
(c) I am able to work under a good manager at the moment
(d) I am always treated as an equal (and not discriminated against) by people at work
(e) I have good opportunities for promotion or recognition at work
(f) I have good opportunities to socialize at work

5. For each of the following statements, please indicate how much you agree, or disagree on a scale of 0 to 10, where 0 indicates you strongly disagree and 10 that you strongly agree.

   (a) I have good opportunities to take part in local social events
   (b) I am treated by people where I live as an equal (and not discriminated against)
   (c) I am able to practice my religious beliefs (including atheism/agnosticism)
   (d) I am able to express my political views when I wish

6. For each of the following statements, please indicate how much you agree, or disagree on a scale of 0 to 10, where 0 indicates you strongly disagree and 10 that you strongly agree.

   (a) I am able to walk in my local neighborhood safely at night
   (b) I am able visit parks or countryside whenever I want
   (c) I am able to work in an environment that has little pollution from cars or other sources
   (d) I am able to keep a pet or animals at home with ease if I so wish
   (e) I am able to get to places I need to without difficulty

7. Moving on to think about access to services. Again please indicate how much you agree, or disagree on a scale of 0 to 10, where 0 indicates you strongly disagree and 10 that you strongly agree.

   (a) Make use of banking and personal finance services
   (b) Get my garbage cleared away
   (c) Get trades people or the landlord to help fix problems in the house
   (d) Be treated by a doctor or nurse
   (e) Get help from the police
   (f) Get help from a solicitor
   (g) Get to a range of shops