EXPLOITING ORDINAL DATA FOR SUBJECTIVE WELL-BEING EVALUATION

Marco Fattore[1], Filomena Maggino[2], Alberto Arcagni[3]

ABSTRACT

The evaluation of subjective well-being, and of similar issues related to quality of life, is usually addressed through composite indicators or counting procedures. This leads to inconsistencies and inefficiency in the treatment of ordinal data that, in turn, affect the quality of information provided to scholars and to policy-makers. In this paper we take a different path and prove that the evaluation of multidimensional ordinal well-being can be addressed in an effective and consistent way, using the theory of partially ordered sets. We first show that the proper evaluation space of well-being is the partially ordered set of achievement profiles and that its structure depends upon the importance assigned to well-being attributes. We then describe how evaluation can be performed extracting information out of the evaluation space, respecting the ordinal nature of data and producing synthetic indicators without attribute aggregation. An application to subjective well-being in Italy illustrates the procedure.

Key words: subjective well-being, multidimensional ordinal data, partial order.

1. Introduction

The aim of the paper is to show how the evaluation of subjective well-being can be addressed in a consistent and effective way, using tools from partial order theory and overcoming the limitations of composite and counting paradigms. The topic is valuable for two main reasons. In a "beyond GDP" perspective, the measurement of multidimensional well-being is progressively gaining importance for social scientists and policy-makers. In particular, the evaluation of subjective well-being proves relevant, since it makes clear that personal satisfaction is not just a matter of "objective achievements" and reveals that quality of life eludes

---

[1] Department of Statistics and Quantitative Methods, University of Milan – Bicocca, Italy. E-mail: marco.fattore@unimib.it.

[2] Department of Statistics, Informatics, Applications "G. Parenti" (DiSIA), University of Florence, Italy. E-mail: filomena.maggino@unifi.it.

[3] Department of Statistics and Quantitative Methods, University of Milan – Bicocca, Italy. E-mail: alberto.arcagni@unimib.it.

reductionist approaches. At the same time, available evaluation procedures are not suitable for well-being assessments; they are in fact mainly designed to deal with numerical variables and not with ordinal attributes, so common in well-being studies. Answering the needs of social scientists and policy-makers thus requires developing new statistical strategies, namely alternative procedures to exploit the information power of multidimensional ordinal data. Starting from this consideration, a new evaluation procedure has been recently proposed by the Authors (Fattore, Brueggemann and Owsinski, 2011; Fattore, Maggino and Colombo, 2012; Fattore and Maggino, 2015), capable to deal directly with ordinal data and to compute synthetic indicators without attribute aggregation. The procedure follows the classical steps of any evaluation process in the social field, namely identifying well-being attributes, choosing a well-being threshold and computing individual and overall statistics. All of these steps, however, are accomplished in "purely ordinal terms", without introducing any artificial transformation of ordinal degrees into numerical scores. Here, we extend the procedure introducing a formal and consistent way to insert into computations exogenous information pertaining to the relevance of well-being attributes. The main focus of the paper is on the concept of evaluation space, which is not simply the set of selected evaluation dimensions, but more profoundly the mathematical structure determining which kind of computations can be consistently performed on the data. It is shown that the natural evaluation space of well-being is a partially ordered set, whose specific structure depends upon the different relevance assigned to attributes. Once the proper evaluation space is set, it is outlined how the evaluation process proceeds and which overall indicators may be computed.

The paper is organized as follows. In Section 2 we briefly discuss the statistical problem of evaluation in a multidimensional setting, mainly focusing on the composite indicator approach and on the Counting Approach of Alkire and Foster. In Section 3, we introduce partial order theory, discuss the construction of the evaluation space and how to account for attribute relevance. In Section 4 we sketch the evaluation procedure. In Section 5, we apply the evaluation procedure to real data pertaining to subjective well-being in Italy, comparing different attribute relevance patterns. Section 6 concludes. The aim of the paper is mainly methodological, so artificial and real examples are meant to introduce the procedure, rather than to provide a deep study of subjective well-being. At the same time, examples aim at showing how the methodology may be, at least in principle, straightforwardly applied, despite its not trivial mathematical foundations. The mathematics involved in the evaluation procedure may seem abstract, at first. Indeed, it is the "mathematics of order", which is algebraic and combinatorial in nature. Working out all of the technical details behind the evaluation procedure would take too much space, so we describe it in a synthetic and rather informal way (other information can be found in cited references). Only the technical tools employed in accounting for attribute relevance are formally detailed, being the first time they are presented to scholars.

2. Evaluation in multidimensional systems of ordinal attributes

The central role of multidimensional ordinal data in current social studies neatly emerges, when inspecting the structure of national and international social surveys pertaining to well-being and quality of life. Consistently with a beyond GDP perspective, most of questionnaires' items is devoted to aspects of personal or familiar life that can be meaningfully described only in yes/no or ordinal terms. This poses statistical problems in data synthesis and, more specifically, in evaluation studies, since usual dimension reduction tools prove scarcely effective, if not inconsistent.

2.1. The composite indicator approach

Under more or less sophisticated forms, the main road to synthesis in social evaluation is the computation of composite indicators. These may be built using simple weighted averages, directly computed on variable scores, or can be obtained as outcomes of more complex procedures, e.g. of structural equation models under reflective or formative schemes, as is the case. Independently of the adopted statistical tools, an aggregative-compensative approach is followed and synthesis is basically achieved through weighted sums of original variable scores. Although not pertaining to well-being, an instructive and prototypical example of the difficulties involved in the composite indicator approach is provided by the Regional Competitiveness Index (RCI), developed by the Joint Research Centre (Annoni and Dijkstra, 2013). With the aim of producing a ranking of the economic attractiveness of 262 regional European areas, 73 different and heterogeneous indicators are aggregated in a single index. The aggregation proceeds by steps: first 11 so-called "sub-pillars" are built; these are in turn aggregated into three "pillars" (Basic pillar, Efficiency pillar and Innovation pillar); finally, RCI is computed as a weighted average of regional pillar scores (see Fattore, Arcagni and Barberis, 2014, for a graphical scheme of the aggregation process). Although the aim of comparing European regions may be sensible, it is legitimate to ask which kind of information is really gained by such a mixing of different dimensions and whether the resulting European attractiveness map is a faithful reproduction of reality. In its essence, the key problem is that socio-economic issues like well-being (or territorial attractiveness, in the case of RCI) are inherently multidimensional and complex. The complexity of a concept refers to the impossibility of capturing it through compensative approaches, ultimately based on dimensional reduction tools. A complex concept comprises many different dimensions, logically related, but not necessarily statistically correlated. Indeed, a major problem in social evaluation is that dimension reduction approaches often fail, since "logical components" of the concept are not enough correlated and thus cannot be "composed" in a satisfactory way. In the case of multidimensional well-being, the inadequacy of aggregative-compensative procedures is further made evident by the nature of the

attributes involved in the evaluation exercise. Well-being attributes are naturally expressed on ordinal scales, ruling out any consistent way to manipulate them using classical statistical analysis. The "vector space" approach (i.e. considering attributes as vectors that can be summed and multiplied by scalars) is useless and inconsistent. Even attempts to scale ordinal attributes into numerical variables are not satisfactory. If a concept is conceived in ordinal terms (e.g. naturally expressed through adverbial forms), why is one to force it into numbers? In addition, scaling procedures may lead to arbitrary and counterintuitive results (Madden, 2010), revealing that ordinal degrees should not be naively seen as rough manifestations of underlying truly continuous scores.

So we are left with an apparently unsolvable problem. On the one hand, we must produce synthetic views of well-being, out of complex systems of multidimensional attributes; on the other hand, aggregative procedures cannot be employed, for both conceptual and technical issues. Partial order theory provides a way out to this problem. Before showing how this is achieved, however, we briefly discuss the Counting Approach of Alkire and Foster (Alkire and Foster, 2011), which tries a different approach to evaluation and which has been gaining an increasing popularity, since its appearance in 2007.

2.2. The Alkire-Foster counting approach

The Counting Approach has been originally designed for deprivation measurement, but actually it provides a general framework for multidimensional evaluation studies and can be consistently applied to systems of ordinal data. It has the merit to realize that ordinal attributes cannot be handled like numerical variables and to provide a procedure that does not introduce any scaling tools. However, the Counting Approach still sticks to an aggregative paradigm and achieves consistency in the treatment of ordinal attributes at the cost of dichotomizing them, losing a great deal of information. With reference to well-being, the Counting Approach is composed of two main steps: (i) the **identification step**, where satisfied or dissatisfied (it depends upon the focus of the study) individuals are identified and (ii) the **measurement step**, where aggregate satisfaction/dissatisfaction indicators are computed. Suppose the focus is on dissatisfaction. Dissatisfied individuals are identified based on their achievements on a set of k ordinal satisfaction variables $v_1, \ldots, v_k$. The so-called **identification function** is computed through a **dual cutoff procedure**. First, a set of k dissatisfaction thresholds $c_1, \ldots, c_k$ is exogenously selected: individuals whose achievement on the i-th attribute is equal to or less than c_i are classified as dissatisfied on v_i. Then an overall cutoff c is defined: individuals whose number of "dissatisfactions" equals or exceeds c are classified as definitely dissatisfied. The Alkire-Foster identification function is thus a 0-1 function, classifying individuals as either non-dissatisfied or dissatisfied in a crisp way. Once dissatisfied individuals have been identified, three aggregate measures can be computed. Let n be the number of individuals in the population and m the number

of dissatisfied, identified through the dual cutoff procedure. The **Head Count Ratio** H is defined as $H = m/n$, i.e. as the fraction of dissatisfied individuals within the population. The **Average Deprivation Share** is instead the average fraction of "dissatisfactions" suffered by dissatisfied individuals, in formulas, $A = T/mk$ where T is the sum of the number of "dissatisfactions" over dissatisfied units. H and A can be combined together, as $M = HA = T/nk$. M can be interpreted as the share of "dissatisfactions" over the maximum number of possible "dissatisfactions" in the population (which is achieved when all individuals are dissatisfied on all of the k attributes).

As clear by the above description, the Counting Approach reduces to computing the number of dimensions an individual is dissatisfied upon, declaring it globally dissatisfied, if this number is equal or exceeds a pre-determined threshold. In this respect, it follows an aggregative approach, where summing over attributes is made equivalent to counting dissatisfactions, by virtue of the dichotomization process. Correspondingly, it also appears that the Counting Approach is not a truly "ordinal procedure", since ordinal information is not effectively exploited, being lost in simpler binary classifications. This leads to a crisp view of personal satisfaction/dissatisfaction, whereas well-being has a vague nature (Sen, 1992,. pp. 48-49), usually accounted for by means of fuzzy set theory. Moreover, when applied to multidimensional ordinal data (which is the case of interest here), there is no natural way to introduce attribute relevance into computations. In summary, the Counting Approach does not provide a completely satisfactory solution to multidimensional ordinal evaluation, being inefficient in the treatment of ordinal data and conceptually similar to the composite indicator paradigm.

3. Partial orders in social evaluation

Any multidimensional evaluation study, and the well-being case is no exception, starts by identifying the so-called **evaluation space**, i.e. the set of dimensions against which evaluation is performed. For example, in the application illustrated later in the paper, the evaluation space comprises four subjective well-being dimensions pertaining to satisfaction on personal health, personal economic situation, familiar relationships and leisure time. Not only the input space settles the conceptual framework of the evaluation process but, even more important, it also determines which information can be extracted from the data and how. What turns a set into a "space", in fact, is the mathematical structure put on it, which defines the set of meaningful operations that can be performed on its elements. In the Counting Approach, the focus is on attributes and the structure underlying the evaluation space is implicitly that of a vector space, with the drawbacks and inconsistencies already mentioned. Which is, instead, the proper structure of the input space associated with subjective well-being (and with multidimensional ordinal data in general)? Given k well-being

attributes $v_1,\ldots,v_k$, with each statistical unit a **well-being** (or **achievement**) **profile** (i.e. a sequence of ordinal scores on the well-being attributes) is naturally associated. Two profiles **p** and **q** may be ordered in terms of well-being, when achievements of (say) **p** are not worse than those of **q** and at least one of them is better. In this case, **p** is "better than" **q**, written **q** < **p**, and the two profiles are said to be **comparable**. On the contrary, if **p** and **q** have conflicting scores, i.e. if **p** is better than **q** on some attributes and is worse on others, than they cannot be ordered and are said to be **incomparable** (written **p** ‖ **q**). So some pairs of achievement profiles may be ordered, others cannot. The input space for well-being evaluation is therefore the set of well-being profiles, naturally structured as a **partially ordered set**. Notice that the focus here is not on ordinal attributes, but on multidimensional profiles, which are the entities that actually characterize individual well-being. Correspondingly, profiles are seen as elements of a "partial order space" and not as elements of a vector space. Although partially ordered sets may seem rather "poor" mathematical structures, they in fact prove very powerful, in view of evaluation. To show this, we must first introduce some basic notions of partial order theory.

3.1. Elements of partial order theory

Given a set Q, a **partial order** $\leq$ on Q is a **reflexive**, **antisymmetric** and **transitive** binary relation defined on it (Davey and Priestley, 2002); the pair (Q, $\leq$) is called a **partially ordered set** or a **poset**, for short. Let Π be the set of well-being profiles and let **p** = $(p_1,\ldots,p_k)$ and **q** = $(q_1,\ldots,q_k)$ be two elements of Π. The set of profiles with partial order defined by

$$\textbf{q} \leq \textbf{p} \text{ if and only if } q_i \leq p_i \text{ for each } i = 1,\ldots,k \qquad (1)$$

is called the **basic achievement poset**. Definition (1) is a purely mathematical condition, identifying the minimal set of comparabilities "anyone would agree upon". This is why Π is qualified as "basic". Posets defined on finite sets may be conveniently depicted *via* **Hasse diagrams**, a kind of directed acyclic graph written according to the following two rules: (i) if **q** $\leq$ **p**, then the node corresponding to **q** is put below the node corresponding to **p** and (ii) if **p** covers **q** (i.e. if **q** $\leq$ **s** $\leq$ **p** implies either **q** = **s** or **p** = **s**), then an edge is drawn from **p** to **q**. A subset of Π whose elements are mutually comparable is called a **chain**. On the contrary, a subset of mutually incomparable elements is called an **antichain**. A partially ordered set which is also a chain is called a **complete** (or **linear**) order. If each variable is conceived as a linear order over the set of its degrees, the basic achievement poset is the so-called **product order** of $v_1,\ldots,v_k$, i.e. it is the partial order defined by (1) over the Cartesian product of (the degrees of) $v_1,\ldots,v_k$. A subset D of Π such that "if **p** is in D and **q** < **p**, then **q** belongs to D" is called a **down**-set. Analogously, a subset U of Π such that "if **p** is in U and **p** < **q**, then **q** belongs to U" is called an **up**-set.

Example. A simple, yet useful, example of poset (the so-called "Cube") and of its Hasse diagram, is given by the set of all the profiles on three binary attributes v_1, v_2 and v_3, partially ordered according to the product order. The attributes may be thought of as expression of satisfaction (1) or dissatisfaction (0) on some well-being dimensions. The poset comprises $2^3 = 8$ profiles and is depicted in Figure 1.
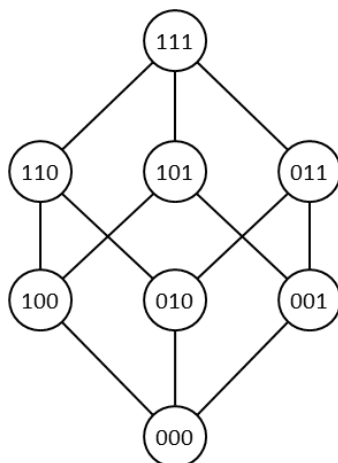

Figure 1. Hasse diagram of the product order of three binary variables ("Cube"). The diagram is to be read from top to bottom; for this reason, directed edges are drawn as simple lines and not as arrows. The subset {111, 110, 100} is a chain, while the subset{110, 001} is an antichain. The subset {011, 101, 100, 010, 001, 000} is a down-set; the subset {111, 110, 011, 010} is an up-set.

The link between partially ordered structures and well-being evaluation can be intuitively guessed: any measurement process is a comparison to a benchmark and partially ordered sets are the natural setting to perform multidimensional comparisons. With reference to Example 1, if one identifies profile 110 as dissatisfied, one can immediately conclude that profiles 100, 010 and 000 are dissatisfied as well, since they represent worse situations than 110. Similarly, one can assert that profile 111 is better than 110 and better than all of the other dissatisfied profiles. However, nothing can be said about profiles 101, 011 and 001, which are incomparable with 110. This represents a difference with respect to the Counting Approach: according to the Alkire-Foster procedure, two profiles with the same number of dissatisfactions are equivalent and two profiles with a different number of dissatisfactions may always be compared. The existence of incomparabilities, however, is deeply consistent with the intrinsic multidimensionality and vagueness of well-being (Qizilbash, 2006) and posets do account for them. The way partially ordered sets will be used to derive concrete evaluations of personal well-being, will be outlined later. What is of concern here, is to stress that (i) the input evaluation space is a partially ordered set and (ii) its

structure is the primary source of information pertaining to well-being. In fact, while achievement profiles describe the personal status of individuals, it is the network of comparabilities/incomparabilities they are embedded in, to determine the "social meaning" of their achievement configurations.

3.2. Attribute relevance and the structure of the achievement poset

The definition of the basic achievement poset does not incorporate any information about possible differences in attribute relevance. In real applications, this is not appropriate and attribute relevance should be accounted for (or, at least, the evaluation procedure should provide this opportunity), in order for the input space to better fit the system of social values under investigation. The basic idea is to inject information on attribute relevance into computations, properly modifying the structure of the basic input space. Before showing how to achieve this in practice, we must deepen the mathematical study of the basic achievement poset.

3.2.1. Decomposition of the basic achievement poset by linear extensions

Let $(\Pi, \leq)$ be the basic achievement poset. An **extension** $\Pi^{ext} = (\Pi, \leq^{ext})$ of $(\Pi, \leq)$ is a poset, such that $\mathbf{p} \leq \mathbf{q}$ implies $\mathbf{p} \leq^{ext} \mathbf{q}$. In practice, Π^{ext} comprises the same comparabilities of Π and adds some more; in this sense, the set of comparabilities of Π^{ext} extends that of Π. An extension λ which is also a linear order is called a **linear extension** of Π. The set of linear extensions of Π is denoted by $\Omega(\Pi)$. In view of accounting for attribute relevance, we are interested in a particular subset of $\Omega(\Pi)$, namely the set Lex(Π) of linear extensions **lexicographically ordered**.

Definition. Given a permutation π of indices $(1,\ldots,k)$, a linear extension λ_π is called lexicographically ordered along π if its order relation $\leq_\pi$ is defined by (here "<" refers to the ordering of attribute degrees):

$$\mathbf{q} \leq_\pi \mathbf{p} \text{ if } q_{\pi(1)} < p_{\pi(1)} \text{ , or } q_{\pi(h)} = p_{\pi(h)} \text{, for h < s and } q_{\pi(s)} < p_{\pi(s)} \text{ (s} = 2,\ldots,k).$$

(where $\pi(i)$ stands for the i-th component of the permuted vector of indices). In practice, the sequence of attributes is permuted and profiles are ordered in an "alphabetic" fashion, according to the permutation. Figure 2 depicts an extension, a linear extension and a lexicographic linear extension of the Cube. Linear extensions can be considered as those complete profile rankings that are compatible with the structure of the basic achievement poset, since no comparabilities of Π are violated in them. Clearly, the map L defined by

$$L(v_{\pi(1)},\ldots, v_{\pi(k)}) = \lambda_\pi$$

is a bijection between Lex(Π) and the set of possible permutations of attributes, so that the cardinality of Lex(Π) is k!. There is a deep link between $\Omega(\Pi)$ and the

partial order structure of the basic achievement poset, in fact one can reconstruct Π from its linear extensions. More precisely, it can be proved that (see Neggers and Kim, 1998):

Proposition 1. The basic achievement poset Π is the intersection of its linear extensions:

$$\Pi = \cap \Omega(\Pi).$$

Explicitly, this means that the set of comparabilities of Π coincides with all of the comparabilities common to its linear extensions. This fact is of central importance for the evaluation procedure. In view of attribute relevance, however, the following proposition is even more relevant.

Proposition 2. The basic achievement poset Π is the intersection of its lexicographic linear extensions:

$$\Pi = \cap \text{Lex}(\Pi).$$

Proof. Since $\text{Lex}(\Pi)$ is a subset of $\Omega(\Pi)$, then $\cap \Omega(\Pi)$ (which is equal to Π) is a subset of $\cap \text{Lex}(\Pi)$, i.e. $\cap \text{Lex}(\Pi)$ is an extension of Π. On the other hand, if **p** and **q** are incomparable in Π, then they have at least two conflicting scores in, say, position i and j. Then **p** and **q** are ordered differently in at least two lexicographic linear extensions whose corresponding permutations differ in the order v_i and v_j are listed. This proves that **p** and **q** are also incomparable in $\cap \text{Lex}(\Pi)$. Therefore, $\Pi = \cap \text{Lex}(\Pi)$. q.e.d.
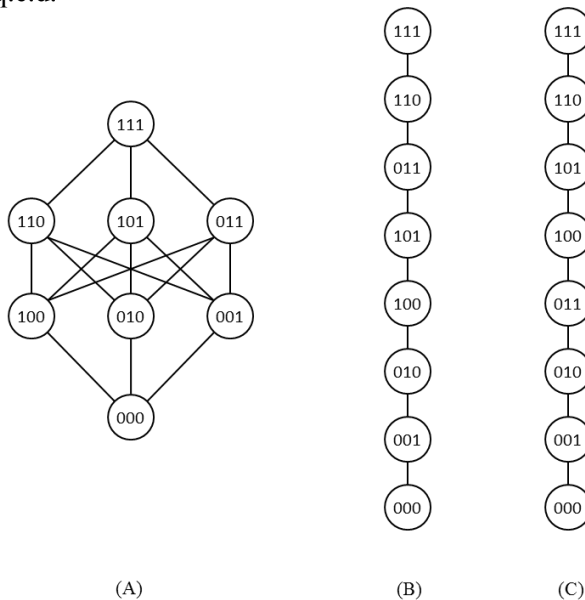

(A)   (B)   (C)

Figure 2. (A) an extension of the Cube; (B) a linear extension of the Cube; (C) a lexicographic linear extension of the Cube (along the identity permutation).

Lexicographic linear extensions are intuitively the "representatives" of attribute rankings in $\Omega(\Pi)$. Since $\text{Lex}(\Pi)$ generates Π by intersection, we see that a link can be made between attribute rankings and the structure of Π. In the next paragraph, we show how this link can be exploited to inject into the structure of the achievement poset exogenous information on attribute relevance.

3.2.2. The attribute poset

To show how lexicographic linear extensions of the achievement poset Π are involved in accounting for attribute relevance, we must first introduce a new poset $(\Lambda, \leq_\Lambda)$ on the set Λ of well-being attributes.

Definition. Let v_i and v_j be two well-being attributes, we write $v_i \leq_\Lambda v_j$ if and only if v_j is more relevant than v_i. The set Λ partially ordered by $\leq_\Lambda$ is called **attribute poset**.

Poset Λ represents a formal yet easy way to define the "relevance pattern" of well-being attributes. It is in fact more natural to assign partial orderings of relevance among attributes, rather than introducing weights (which would also be inconsistent with an ordinal setting), as if one could realistically state a set of precise "attribute equivalences". In addition, not any partially ordered relevance pattern can be reproduced through weighting schemes. To realize this, consider the attribute poset depicted in Figure 3, on the three attributes of Example 1. According to the Hasse diagram, attribute v_1 is more important than attribute v_2, but nothing is stated about v_3. If we were to adopt an equivalent weighting scheme, a greater weight should be attached to v_1 than to v_2; but then it is impossible to find a weight to attach to v_3 which makes it "indifferent" to both of the other attributes.
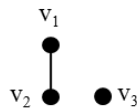

Figure 3. Attribute poset on three attributes. As it can be seen, v_3 is incomparable with both v_1 and v_2. It is impossible to emulate this relevance pattern attaching weights to each attribute.

As any finite partial order, according to Proposition 1 the attribute poset is equivalent to the set of its linear extensions $\Omega(\Lambda)$, i.e. to the set of attribute rankings, **compatible** with $\leq_\Lambda$ (or **admissible** with respect to Λ). These linear extensions, in turn, can be seen as attribute permutations. With a little notational abuse, we can thus associate lexicographic linear extensions of Λ to linear extensions of Π through the map L. The image of $\Omega(\Lambda)$ in $\text{Lex}(\Pi)$ through L identifies the set $\text{Lex}(\Pi; \Lambda)$ of lexicographic linear extensions of Π that are consistent with the attribute poset Λ (i.e. **Λ–admissible** lexicographic linear extensions). When Λ is an antichain, which means that attributes share the same relevance, $L(\Omega(\Lambda)) = \text{Lex}(\Pi)$ and all lexicographic linear extensions of Π are Λ-admissible. On the contrary, when Λ is not an antichain, some elements of

Lex(Π) are not Λ-admissible. This implies that Π itself is not consistent with Λ. To restore consistency, Π must be turned into a new poset Π*(Λ) = (Π, ≤*) such that Lex(Π*(Λ)) = L(Ω(Λ)). The only way to achieve this is to define Π*(Λ) as the intersection of all Λ-admissible linear extensions of Π:

$$\Pi^*(\Lambda) = \cap\ L(\Omega(\Lambda)).$$

Π*(Λ), as defined above, is a proper extension of Π if and only if Λ is not an antichain (in that case, it coincides with Π) and it is easily proved to be the smallest extension of Π consistent with Λ.

In summary, we have the following logic chain

$$\Lambda \rightarrow \Omega(\Lambda) \rightarrow \mathrm{Lex}(\Pi;\ \Lambda) \rightarrow \cap\mathrm{Lex}(\Pi;\ \Lambda) = \Pi^*(\Lambda)$$

that turns the attribute poset into the desired extension of the achievement poset, incorporating the exogenous information on attribute importance. Figure 4 gives an example of the above chain for the Cube. Two final remarks are in order. First, it is indeed expected that introducing attribute relevance leads to extending the basic achievement poset; additional information conveys in fact new criteria to resolve "comparison ambiguities", reducing the number of incomparabilities. Secondly, notice that the extension procedure is **purely ordinal**: no numerical weights enters the computations and the nature of the attributes is fully respected. The introduction of attribute relevance concludes the process of evaluation space definition. One can thus proceed to evaluating well-being, through the evaluation procedure briefly described in the next paragraph.

| (A) | (B) | (C) | (D) |
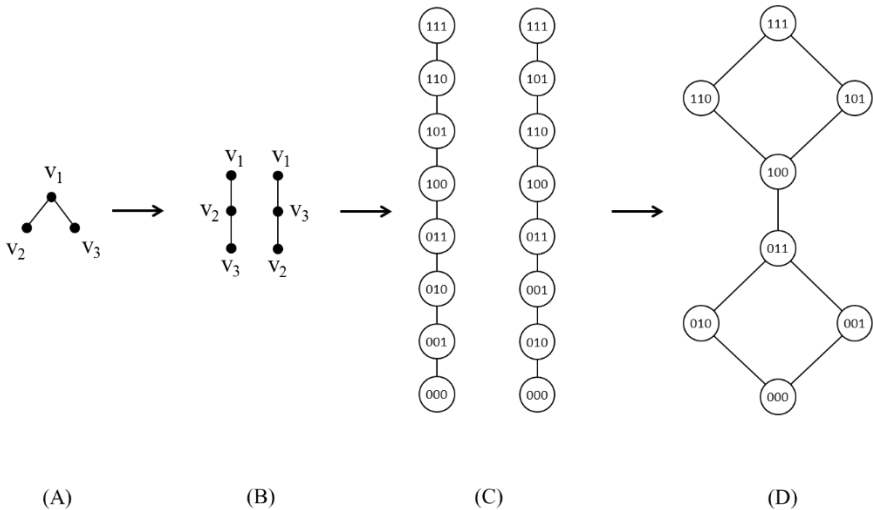
Figure 4. (A) attribute poset Λ on three binary attributes; (B) linear extensions of Λ; (C) Λ-admissible lexicographic linear extensions of the Cube; (D) extended achievement poset Π*(Λ). The Cube has 48 linear extensions; among them, 6 are lexicographically ordered. The selection of the relevance pattern Λ reduces them to two. The final extended achievement poset has 4 linear extensions and comprises only two incomparabilities (110‖101 and 010‖001).

4. Evaluating subjective well-being from the achievement poset

Previous paragraph has been devoted to show how the evaluation space of subjective well-being can be properly structured, focusing on achievement profiles and accounting for attribute relevance. Here we show how it can be used to evaluate well-being. Since the fundamentals of the evaluation procedure have been already introduced in other papers (Fattore, Brueggemann and Owsinski 2011; Fattore, Maggino and Colombo, 2012), in the following we limit ourselves to a brief outline.

The aim of the evaluation procedure is to assign subjective well-being scores to statistical units in the population. This will be achieved associating scores directly to profiles of the achievement poset; statistical units then inherit the scores of their profiles. Due to multidimensionality, however, we must distinguish between well-being **vagueness** and **intensity**. With reference to personal dissatisfaction, the former pertains to ambiguities in the classification of a statistical unit as dissatisfied or not; the latter to the severity of dissatisfaction, i.e. intuitively to the "distance" from satisfaction. Both concepts are distinctly accounted for and measured by the evaluation procedure, through a couple of evaluation functions, namely the **identification function** $\mathrm{idn}(\cdot)$, that quantifies the ambiguity of profile classification, and the **severity function** $\mathrm{svr}(\cdot)$, that quantifies dissatisfaction intensity. To build these functions, however, the preliminary identification of a dissatisfaction threshold is needed.

4.1. Setting the dissatisfaction threshold

The achievement poset Π^* conveys no explicit information on subjective well-being. To turn it into a direct input to evaluation, a well-being **threshold** τ must be introduced. The threshold must be conceived as a minimal set of exogenous information, which identifies profiles "on the edge of dissatisfaction", leaving to the evaluation procedure to spread such information across the poset. Due to multidimensionality, more than one profile may be "on the edge", so that the threshold must be in general chosen as an antichain of well-being profiles, whose elements describe **alternative dissatisfaction patterns**, to be considered as reference benchmarks. As always in evaluation studies, the choice of the threshold is a delicate step. Given the methodological aim of the paper, here we do not discuss this issue further. Notice, however, that the threshold is directly specified in terms of profiles, without any explicit reference to attribute cut-offs, as in the Counting Approach.

4.2. The identification function

Since there is no natural scale against which to assess subjective well-being, we address identification as a problem of multidimensional comparison between achievement profiles and threshold benchmarks. Differently from the unidimensional case, due to partial ordering not any well-being profile may be

unambiguously compared with the threshold. A profile whose scores are worse than those of an element of the threshold represents a dissatisfied condition (since it is worse than a "dissatisfied profile"). But in many cases, ambiguities arise and some profiles cannot be classified as below or above the threshold, due to conflicting scores. The identification function must account for such ambiguities; to this aim it is defined in such a way that:

- elements of the threshold are scored 1, i.e. they are classified as dissatisfied profiles;
- profiles below an element of the threshold in the achievement poset, i.e. profiles in the down-set generated by the threshold, are similarly scored 1;
- profiles above **any** element of the threshold in the achievement poset, i.e. profiles in the intersection of the up-sets generated by threshold elements, are scored 0 (they are classified as "non-dissatisfied", since they represent situations that are better than any "completely" dissatisfaction pattern identified in the threshold);
- all other profiles are scored by idn($\cdot$) in (0,1), i.e. they are scored as "ambiguously" or "partly" dissatisfied profiles.

To define the identification function in practice, we start by considering the set of linear extensions of the input achievement poset. In a linear extension, a well-being profile is either below (or coincide with) an element of the threshold or it is above all threshold elements, so that it can be unambiguously identified as "dissatisfied" or as "non-dissatisfied" in **that** linear extension. Therefore, on each linear extension λ one can define a 0-1 identification function $\mathrm{idn}_\lambda(\cdot)$ assigning value 1 to profiles classified as dissatisfied in λ and 0 to all of the others. Linear extensions are thus seen as 0-1 classificators. In different linear extensions, profiles classified as dissatisfied are in general different (only profiles below elements of the deprivation threshold in the input achievement poset are scored 1 in each linear extension and only profiles above all of the elements of the deprivation threshold in the input achievement poset are always scored 0). Classification ambiguities in Π^* thus reflect in different linear extensions classifying profiles differently. Counting the proportion of linear extensions where a profile is scored 1, one quantifies such ambiguities and gets a non-linear identification function assigning scores in [0,1] to well-being profiles. Formally, idn($\cdot$) is defined as:

$$\mathrm{idn}(\mathbf{p}) = \frac{1}{|\Omega(\Pi^*)|} \sum_{\lambda \in \Omega(\Pi^*)} \mathrm{idn}_\lambda(\mathbf{p}).$$

From the above definition, it follows easily that idn($\cdot$) is **order-preserving**, i.e. if $\mathbf{q} \leq \mathbf{p}$ in Π^* then $\mathrm{idn}(\mathbf{q}) \leq \mathrm{idn}(\mathbf{p})$. Notice that the poset approach to subjective well-being evaluation is, in a sense, a counting approach, but differently from other counting methodologies (Alkire and Foster, 2011, Cerioli and Zani, 1990) it counts over linear extensions and not over attributes.

4.3. The severity function

A faithful picture of subjective well-being requires taking into account both its vagueness and intensity. If intensity would not be assessed, profiles with same identification scores could be considered as equivalent, while they can actually correspond to very different self-perceived situations. Obviously, dissatisfaction intensity, or **severity**, can be meaningfully assessed only on the subset of completely or partly dissatisfied profiles. Analogously to the identification function, for each linear extension and for each dissatisfied profile in it, we define a severity function $svr_\lambda(\cdot)$ and compute the severity function $svr(\cdot)$ averaging on $\Omega(\Pi^*)$. Formally, let λ be a linear extension of Π^* and let $\mathbf{p}$ be a profile dissatisfied in it. The satisfied profile $\mathbf{q}$ nearest to $\mathbf{p}$ in λ is the "first" profile ranked above all the elements of the threshold, in λ. Dissatisfaction severity of $\mathbf{p}$ in λ, i.e. $svr_\lambda(\mathbf{p})$, is thus defined as the graph distance of $\mathbf{p}$ from $\mathbf{q}$ in the Hasse diagram of λ, i.e. as the number of edges between $\mathbf{p}$ and $\mathbf{q}$ ($svr_\lambda(\mathbf{p})$ is instead set to 0 for non-dissatisfied profiles in λ). Finally, we put:

$$\text{svr}(\mathbf{p}) = \frac{1}{|\Omega(\Pi^*)|} \sum_{\lambda \in \Omega(\Pi^*)} \text{svr}_\lambda(\mathbf{p}).$$

A relative measure of profile dissatisfaction severity can be obtained dividing $svr(\mathbf{p})$ by its maximum over the achievement poset, i.e. by the severity of the bottom profile of Π^*. Like the identification function, also the severity function is order-preserving.

4.4. Synthetic indicators

Identification and severity scores assigned to profiles are inherited by statistical units. Synthetic indicators may then be computed, averaging on the population or on suitable subpopulations. Three overall indicators are of particular interest.

1. **Population dissatisfaction degree** (H), defined as the average of $idn(\cdot)$ over the population.

2. **Specific dissatisfaction degree** (D), defined as the average of $idn(\cdot)$ over the subpopulation of completely or partly dissatisfied statistical units (i.e. with profiles $\mathbf{p}$, such that $idn(\mathbf{p}) > 0$). D measures the vagueness of dissatisfaction of individuals with a non-null dissatisfaction degree. High values of this indicator may reveal that the population is polarized into two groups, the non-dissatisfied and the (almost) completely dissatisfied.

3. **Population severity degree** (S), defined as the average of $svr(\cdot)$ over the subpopulation of statistical units with non-null identification scores. S can be turned into a relative measure S*, dividing it by its maximum over Π^*.

Other indicators may be indeed defined and computed, starting from the distribution of evaluation scores on the statistical population; nevertheless, H, D and S suffice to provide a synthetic and comprehensive view of subjective well-being. As anticipated in the Introduction, they are computed without attribute aggregation.

4.5. Computational aspects

The evaluation procedure is combinatorial in nature and draws upon the computation of linear extensions of the achievement poset. In real cases, it is unfeasible to list all of them and one must rely on sampling algorithms, computing the evaluation functions on a subset of linear extensions. The sampling procedure and some basic functions to manipulate partial orders and to compute the identification and severity functions together with the overall indicators are implemented in the R package PARSEC (available from the Authors). Details on computational aspects and the use of the package can be found in Fattore and Arcagni (2014).

5. Subjective well-being in Italy

In this section, we provide some examples of the evaluation process on data pertaining to subjective well-being in Italy, for the year 2012. The dataset comes from the "Multipurpose survey about families: aspects of daily life", held by the Italian National Statistical Bureau on a yearly basis. We focus on satisfaction on Health, personal Economic status, Family relationships and Leisure time. In the original dataset, satisfaction is expressed on a 4-degree scale: 1 – "very", 2 - "enough", 3 - "little" and 4 - "not at all". In the following, scores have been reversed, so that 1 stands for "not at all" and 4 stands for "very". The achievement poset is composed of 256 profiles and will not be displayed. The original dataset comprises 46464 records; among these, 6893 have missing values. Since they are not systematic and given the exemplificative purpose of this analysis, we have simply deleted them, reducing the input dataset to 39571 records. To provide examples of applications of the procedure, we consider three different attribute relevance patterns, comparing the results in terms of evaluation functions and overall indicators. Computations have been performed using the programming language R and the package PARSEC. The three attribute posets Λ_1, Λ_2 and Λ_3 are depicted in Figure 5. In the first pattern, all attributes share the same relevance; in the second, Health, Economic status and Family relationships are equivalent and more relevant than Leisure time; in the third, Leisure time is dominated by Economic status which in turn is dominated by Health, while Family relationships is incomparable with all of the other attributes (this is the only pattern not reproducible by numerical weighting schemes). To the three relevance patterns, there correspond three achievement posets Π_1 (coinciding with the basic achievement poset Π), Π_2 and Π_3. The threshold has been set to $\tau =$

(1223, 2123) (first score refers to Health, second to Economic status, third to Family relationships and fourth to Leisure time). The threshold is not symmetric with respect to attributes, but emphasizes the relevance of Health and Economic status. In principle, there should be consistency between attribute relevance implicit in threshold selection and the definition of the attribute poset. Given the exemplificative purpose of this section, we do not stress this aspect here.

Figure 5. Three alternative attribute posets on Health (Hea), personal Economic status (Eco), Family relationships (Fam) and Leisure time (Lei). Λ_1, Λ_2 and Λ_3 have 4!=24, 3!=6 and 4 linear extensions respectively.

Figure 6 depicts the identification and relative severity scores of the well-being profiles for each input achievement poset. The values of indicators H, D and S for the three cases are reported in Table 1.

Figure 6. Evaluation functions implied by the attribute posets depicted in Figure 5. Profiles are sorted on the x-axes in increasing order of identification (upper panels) or severity (lower panels) scores (for graphical purposes, the relative severity scores of completely non-dissatisfied profiles have been set to 0, although the severity function is not defined over them).

As it can be seen, the shapes of both evaluation functions strongly depend upon the relevance patterns of the attributes. Interestingly, the identification function is not linear and when the attribute poset comprises comparabilities it assumes a sigmoid shape. The severity functions allow for dissatisfied profiles to be further distinguished in terms of dissatisfaction intensity. As for the identification scores, also these functions are non–linear, but here non-linearities tend to disappear, when the attribute poset increases the number of comparabilities. As the attribute poset gets similar to a linear order, dissatisfaction severity increases, i.e. on average partly or completely dissatisfied statistical units increase their distance from non-dissatisfaction. The Head Count Ratio, on the contrary, has a more complex behaviour, revealing that the two indicators do measure different aspects of deprivation. In fact, H is lower under Λ_1 than under Λ_2. The Specific deprivation degree, instead, reveals that passing from Λ_1 to Λ_3, partly/completely dissatisfied units move from non-dissatisfaction towards dissatisfaction, and Italian society would appear as basically polarized into two groups, a bigger one of non-dissatisfied individuals and a smaller one of highly dissatisfied.

Table 1. Overall indicators (expressed on a 0-100 scale) for the patterns of attribute relevance depicted in Figure 5.

Attribute poset	Population dissatisfaction degree (H)	Specific dissatisfaction degree (D)	Population Severity degree (S)
Λ_1	6.7	23.3	10.7
Λ_2	2.7	31.9	15.7
Λ_3	12.1	60.8	23.3

Much more information could be extracted from the data, e.g. considering socio-demographic covariates or comparing subjective well-being patterns at territorial level. What has been reported, however, should be enough to prove the flexibility and the effectiveness of the evaluation procedure. The shapes of the identification function show how the procedure is capable to account for the nuances of subjective well-being, distinguishing among well-being patterns that counting approaches would have scored identically. Sensitivity of final results to the choice of the attribute poset, in turn, shows how it is important to properly specify the evaluation input space, whose structure, together with the threshold, determines the final evaluation scores.

6. Conclusion

In this paper, we have outlined a comprehensive procedure to address evaluation problems in a multidimensional ordinal setting. The procedure follows the general logic of any evaluation study: identification of evaluation dimensions, addition of attribute relevance, threshold selection, computation of evaluation scores at the statistical unit and population level; noticeably, however, all of these steps are defined and performed in a purely ordinal way. Partial order theory, the mathematics of order, allows for that and makes it possible to overcome the inconsistencies of composite indicators and the inefficiency of counting approaches. We have applied the procedure to data pertaining to subjective well-being in Italy for year the 2012, comparing different patterns of attribute relevance, to show how the evaluation process is flexible, simple and straightforward to apply. A major problem with the procedure is indeed its computational burden, that currently limits its application to small or medium size posets (typically, with up to about 250 profiles). Currently, a simplification of the procedure is under development, so as to reduce the number of linear extensions to evaluate, so as to virtually remove any computational issue in real applications. The primary aim of the paper, however, is to show that a sound conceptual and formal setting can be indeed defined, where subjective well-being evaluation and, more generally, evaluation on multidimensional systems of ordinal data can be properly addressed. Hopefully, this may help social scientists to innovate their methodological and statistical toolbox, so as to capture the complexity and nuances of human experiences in a more effective and realistic way.

REFERENCES

ALKIRE, S., FOSTER, J., (2011). Counting and multidimensional poverty measures. Journal of Public Economics, 95, pp. 476–487.

ANNONI, P., DIJKSTRA, L., (2013). EU Regional Competitiveness Index - RCI 2013. RCI 2013, JRC Scientific and Policy Reports. Luxembourg: Publications Office of the European Union.

CERIOLI, A., ZANI, S., (1990). A Fuzzy Approach To The Measurement Of Poverty. In: Dagum, C., Zenga, M. Income and Wealth distribution, Inequality and Poverty. Berlin Heidelberg: Springer-Verlag. pp. 272–284.

DAVEY, B. A., PRIESTLEY, B. H., (2002). Introduction to lattices and order. Cambridge: Cambridge University Press.

FATTORE, M., BRUEGGEMANN, R., OWSIŃSKI, J., (2011). Using poset theory to compare fuzzy multidimensional material deprivation across regions. In: Ingrassia S., Rocci R. and Vichi, M. New Perspectives in Statistical Modeling and Data Analysis. Berlin: Springer-Verlag.

FATTORE, M., MAGGINO, F., COLOMBO, E., (2012). From composite indicators to partial orders: evaluating socio-economic phenomena through ordinal data. In: Maggino F., Nuvolati G. Quality of life in Italy: researches and reflections. Berlin: Social Indicators Research Series.

FATTORE, M., ARCAGNI, A., (2014). PARSEC: an R package for poset-based evaluation of multidimensional poverty. In: Bruggemann R., Carlsen L. and Wittmann J. Multi-Indicator Systems and Modelling in Partial Order. Berlin: Springer.

FATTORE, M., ARCAGNI, A., BARBERIS, S., (2014). Visualizing Partially Ordered Sets for Socioeconomic analysis. Revista Colombiana de Estadistica, 34(2), pp. 437–450.

FATTORE, M., MAGGINO, F., (2015). A new method for measuring and analyzing suffering – Comparing suffering patterns in Italian society. In: Anderson R. E. World Suffering and the Quality of Life. New York: Springer.

MADDEN, D., (2010). Ordinal and cardinal measures of health inequality: an empirical comparison. Health Economics, 19, pp. 243–250.

NEGGERS, J., KIM, S. H., (1998). Basic posets. Singapore: World Scientific.

QIZILBASH, M., (2006). Philosophical Accounts of Vagueness, Fuzzy Poverty Measures and Multidimensionality. In Lemmi A. and Betti G. Fuzzy Set Approach to Multidimensional Poverty Measurement. New York: Springer.

SEN, A., (1992). Inequality reexamined, Harvard University Press.