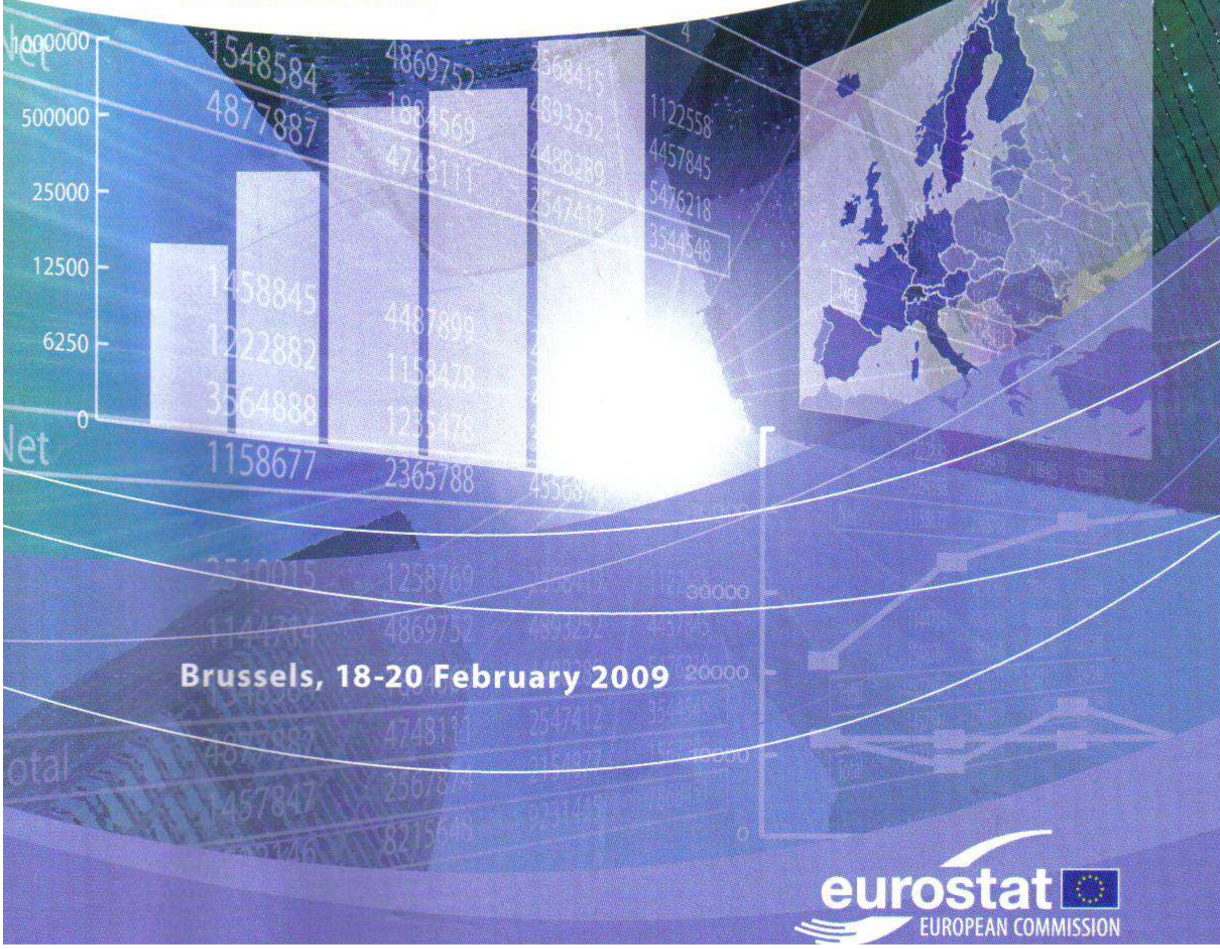
Obtaining weights: from objective to subjective approaches in view of more participative methods in the construction of composite

eurostat

NTTS - Conferences on New Techniques and Technologies for Statistics

The meeting place for Research in Official Statistics

POSTER SESSIONS
FULL PAPERS



Brussels, 18-20 February 2009

eurostat
EUROPEAN COMMISSION

Obtaining weights: from objective to subjective approaches in view of more participative methods in the construction of composite indicators

Filomena Maggino[1], Elena Ruviglioni[2]

[1] Università degli Studi di Firenze (Italy), email: filomena.maggino@unifi.it

[2] Università degli Studi di Firenze (Italy), email: elena.ruviglioni@unifi.it

Abstract

As known, the consolidated methodology aimed at the construction of composite indicators states particular approaches allowing differential *importance* weights to be determined and to be assigned to the indicators composing the synthesis. In this ambit, it is always asserted that the choice of weights should be preferably derived from objective principle. In recent works (Hagerty & Land, 2007) further views were introduced about weighting in the context of composite indicators construction, which should take into account the agreement among citizens concerning the importance to be assigned to each indicator. The final composite should maximize this agreement. Hagerty & Land (2007) provide a framework to jointly consider weights and social indicators as part of the research problem of constructing a composite indicator. Our purpose is to identify procedures that provide a framework allowing differential subjective weights to be determined and managed (*subjective weighting* procedure). In particular, the framework must clarify how to obtain importance weights at individual-subjective level through subjective judgments and to assign the weights to the corresponding subjective scores. This work intends

1. to introduce the general underlying principles in obtaining weights, with special attention to subjective weights
2. to identify and analyze the approaches for obtaining weights (a) Statistical approaches (for obtaining "objective" weights), generally applied in the ambit of composite indicators construction, (b) Multi-Attribute approaches, and (c) Scaling approaches, allowing subjective data to be managed; among these, the models able (i) to handle subjective evaluations and judgments, expressed in explicit or implicit way, (ii) to obtain subjective [importance] weights at group level and at individual level, will be identified and described in the perspective of obtaining subjective weights. Pros and cons of these approaches in the perspective of subjective weighting will be discussed.

Keywords: weighting criteria, indicators construction, subjective weights.

1. Introduction

The methodology aimed at constructing indicators is very often presented in terms of "technology", by asserting the need to have specialist training in order to apply the procedure in a scientific and objective way. Actually the construction procedure, even though scientifically defined, is far from being objective and aseptic. As known, the consolidated methodology aimed at the construction of composite indicators (Nardo et al., 2005; Sharpe, 2004) defines different stages in order to develop the indicators. Each stage requires a decision / choice (methodological or not) to be taken:

1. **choosing the analytical approach** in order to verify the underlying dimensionality of selected elementary indicators (*dimensional analysis*)
2. **choosing and obtaining weights** in order to define the importance of each elementary indicator to be aggregated (*weighting criteria*)
3. **choosing and identifying the aggregating technique** in order to synthesize the elementary indicators values into composite indicators (*aggregating-over-indicators techniques*)
4. **choosing models and conceptual approaches** in order to assess (i) the robustness of the synthetic indicator in terms of capacity to produce correct and stable measures (*uncertainty analysis,*

sensitivity analysis), (ii) the discriminant capacity of the synthetic indicator (*ascertainment of selectivity* and *identification of cut-point or cut-off values*)

Even though some decisions are strictly technical, it is quite difficult to make these decisions objective since they may involve different kind of concerns. Generally they are taken through a process accepted and shared by the scientific community. However, in certain cases, the choice and decision may be shared by a larger community. One of the ways to obtain this is that to involving individuals in the process of social indicators construction. In other words, indicators construction is not simply a technical problem but should become part of a larger debate concerning how to construct indicators obtaining a larger *legitimacy*. Seen in this perspective, this topic can be placed in the ambit of an improvement of democratic participation to decisions ("res publica").

The weighting process. In indicator construction, weights aim at assigning differential *importance* weights to be aggregated. The usually discussion concerning the methodology applied in order to determine and assign weights to the indicators composing the synthesis always asserts that the choice of weights should be preferably derived from objective principle (Nardo et al., 2005; Ray, 2008; Sharpe, 2004). However, since developing and defining weights can be always interpreted in terms of **values judgment**, the procedure should include and involve individuals' contributions in attributing importance to different domains. In recent works (Hagerty & Land, 2007) further views were introduced about weighting in the context of composite indicators construction. In particular, creating composite indicator (describing social units at macro level) should take into account the agreement among citizens concerning the importance to be assigned to each indicator. The final composite should maximize this agreement. In their work, they provide a framework to jointly consider weights and social indicators as part of the research problem of constructing a composite indicator. This requires (i) a methodology allowing subjective weights to be collected (*subjective/individualized weighting* procedure) at individual-subjective level through subjective judgments, (ii) a methodology allowing subjective weights to be included in indicators by assigning the weights to the corresponding indicators. The issues to be faced in obtaining differential weights with particular reference to subjective measurement will be clarified in the following sections.

2. General concerns and principles underlying the weighting issue

In general terms, when we suppose that not necessarily all the measured indicators (sub-score) contribute with the same importance to the measurement and evaluation of the total variable (synthetic score), a weighting system needs to be defined in order to assign a weight to each indicator, before proceeding to the indicators aggregation. From the technical point of view, the weighting procedure consists in defining and assigning a weight to each sub-score. The weight will be used in the successive computation of the individual aggregate score; in particular, each weight is multiplied for the corresponding individual value of the sub-score.[1] A criterion should be adopted in order to define a weighting system, when it can not be implicitly identified. The weighting system should reproduce as accurately as possible the contribution of each sub-score to the construction of the synthetic score. In this perspective, defining a weighting system constitutes an improvement and refinement of the adopted model of measurement. In order to proceed to the difficult choice among the different weighting approaches, the researcher needs to take into account (Nardo et al., 2005; Ray, 2008) (a) the rationale and theoretical framework on which the measurement of the complex characteristics is founded and that will consequently regard the synthetic score, (b) the meaning and the contribution of each sub-score to the synthesis, and (c) the quality of data and the statistical adequacy of indicators.

---

[1] An alternative to the simple multiplication of weight and score is proposed by Hsieh (2003, 2004) – and discussed by Wu (2008) – by including the sum of importance scores as a denominator. This approach can be differentiated according to ranking and rating scores when directly used as weights. Hsieh (2003) identified different computational approaches in order to connect the weight with the score to be weighted. In particular, he proposed seven different weighting mechanisms of relative importance, three using discrete importance rating and four using ranking scores. The sum of weighted scores is divided with differently adequate denominators to obtained equivalent scales for the purpose of easy and intuitive comparison. Since in collecting subjective data to be directly used as weights, both scores can be adopted, rating and ranking scores data should be carefully assumed by considering that they can reflect different meanings in terms of weight and require different computational approaches, producing different results.

The identification of a system of weights should (i) consider in advance also some technical issues, related to the conditions for obtaining weights and concerning the level at which and the scale on which the weights should be determined (rescaling issue), (ii) make a decision in advance on:

- the proportional size of the weights (equal or differential weighting)
- the aggregation technique to be adopted (compensatory or non-compensatory)

2.1 Conditions for obtaining weights

The procedure for determining the weights has to take into account some basic conditions that can be technically formalized as follows:

$$CI_i = \sum_{j=1}^{K} x_{ij} w_{ij}$$

where

CI_i composite indicator for case i k number of indicators to be aggregated

x_{ij} indicator j to be aggregated for case i

w_{ij} weight j to be attribute to x for case i

Each weight w_{ij} should satisfy the following basic conditions

(i) the weights are non negative numbers: $0 \le w_{ij} \le 1$

(ii) the weights for each case i add up to unity: $\sum_{j=1}^{K} w_{ij} = 1$

(iii) the weights may require to be rescaled in order to have an identical range

(iv) the weights are relating in some way to the corresponding score (as we will see, this condition may require a decision to be taken)

Rescaling weights

Following their computation, weights may require to be rescaled. Re-scaling (i) normalises weights to have an identical range (0; 1), (ii) could distort the transformed indicator in presence of extreme values/or outliers, (iii) could widen the range of indicators lying within a small interval increasing the effect on the weights. The procedure can be performed as follows:

$$rw_{ij} = \frac{w_{ij}}{\max(w_j)} \quad \text{or} \quad rw_{ij} = \frac{w_{ij} - \min(w_j)}{\max(w_j) - \min(w_j)}$$

where

rw_{ij} rescaled value of the weight with reference to the object j for the respondent i

w_{ij} value of the weight with reference to the object j for the respondent i

The researcher has to carefully evaluate and make formally explicit not only the methodology to be adopted but also the results that would have been obtained with other methodologies, also reasonably applicable.

3. Statements in obtaining weights

3.1 Decisions to be taken

The first decision that needs to be made and that will be strongly influence the final results is between equal and different weighting. **Equal weighting** represents the preferred procedure, adopted in most of the applications. This happens mainly when (i) the theoretical structure attributes to each indicator the same adequacy in defining the variable to be measured, (ii) the theoretical structure does not allow hypotheses to be consistently derived on differential weightings, (iii) the statistical and empirical knowledge is not adequate for defining weights, (iv) the correct adoption and application of alternative procedures do not find any agreement. Although equal weighting, which does not necessarily imply unitary weighting, is certainly an explicit weighting scheme, the a priori decision to adopt the technique of equal weighting for methodological purposes makes the choice of weights apparently less subjective. A motivation for this approach is that it is objective in the sense that if adopted as a common technique of weighting, the subjective component would lie exclusively in the choice of indicators. There is an advantage of this approach: namely, that a debate over the inclusion of elementary indicators, that is, which indicators are important, can be conducted on a more basic level than a discussion that centers around the choice of numerical weights (Sharpe, 2004).

Differential weighting does not necessarily correspond to the identification of different weights but rather to the selection of the most appropriate approach in order to identify the weights among the identified ones (Nardo et al., 2005). Assigning differential weights can be just as doubtful, especially when the decision is not supported by (i) theoretical reflections that endow a meaning on each indicator or consider its impact on the synthesis, (ii) methodological concerns that helps to identify the proper techniques, consistently with the theoretical structure.Bobko et al. (2007) made a interesting review of the relevant literature across multiple disciplines and multiple decades on differential and unit weights. Their literature review indicates that unit weights have substantial predictive validity when compared with regression weights, but there is a lack of data on how other differential weighting strategies (e.g., weights generated by subject matter experts) compare to unit weights. Moreover, they provide a primary and a meta-analytic study by which they show how in their applications data and findings indicate that unit weights can be a highly appropriate approach for weighting under many circumstances. In subjective measurement, the effectiveness of weighted scores should be questioned with reference to (i) the theoretical issue of whether importance and satisfaction are distinct constructs, (ii) the psychometric properties of importance ratings (particularly, internal consistency and test-retest reliability), and (iii) the criteria used in assessing weighted scores. All these topics need more attention and care from the researchers (Russell & al., 2006). In order to avoid incoherencies between the theoretical meaning of weights and the way these weights are actually used, a consistent aggregating technique has to be chosen (Nardo et al., 2005). In particular, the choice of the weighting system must consider the compensability among the elementary indicators inside the synthetic score. In particular, this is allowed by the technique that will be used in aggregating the sub-scores. An aggregating technique is *compensatory* when it allows low values in some sub-scores to be compensated by high values in other sub-scores. If compensability is admitted, a unit showing a low value for one indicator will need higher values on the others in order to obtain a higher synthetic value.

3.2 Statistical approaches for obtaining weights

One of the ambit in which the issue of obtaining differential *importance* weights found consolidated applications is that of constructing composite indicators. As previously said, in this ambit, it is always asserted how the choice of weights would be preferably derived from objective principle (Ray, 2008). In this perspective, the statistical methods are traditionally considered and preferred (Nardo et al., 2005; Ray, 2008; Sharpe, 2004), above all (i) *Correlation Analysis*, (ii) *Principal Component Analysis*, (iii) *Data Envelopment Analysis*.

	Correlation Analysis	Principal Component Analysis	Data Envelopment Analysis
Assumption	Assigning equal weights to elementary indicators that are highly correlated.	The group of elementary indicators define a multidimensional variables.	The elementary indicators define particular dimensions, like capacity.
Goal	Less importance to be assigned to indicators that are highly correlated to the others (*double counting* effect).	Definition of a set of weights for each elementary indicator – components scores (*) – one for each dimension/component defining the latent variable. Weights allow one synthetic indicator for each component to be calculated. The resulting synthetic indicators will be consistent and independent from each other.	Identifying weights to be assigned to elementary indicators, with particular reference to concepts like "capacity". DEA estimates the *efficiency frontier* that can be used as a benchmark in order to measure and evaluate the relative performance of observed units.
Method	Since high correlation is considered as a sign of *double counting*, the procedure requires • averaging the correlation values registered between all the selected elementary indicators • defining the weight (inversely proportional to the correlation level)	• identification of the components explaining the greatest portion of total variance • for each elementary indicator, calculation of component weight by removing the part of the elementary indicator's contribution explained by its correlation with the other elementary indicators	The set of weights for each unit depends on its position defined in terms of its distance from • the efficiency frontier (corresponding to the best registered performance) or • the benchmark (corresponding to an ideal point) or different priority defined according to those aspects that turned out to be good performances (in some sense this requires the individual identification of a strategic- or priority-objective).
Benefit			It allows us to ✓ avoid an explicitly specification of a mathematical form describing the production function and the performance model ✓ uncover relationships that remain hidden for other methodologies ✓ handle many elementary indicators at the same time ✓ use any kind of input-output measurement ✓ identify, analyze, and quantified the sources of inefficiency, for every evaluated unit (in other words, it allows us to identify the elementary indicator showing the worse performance)
Problem	The **limit value** can not be defined at a statistical level because there is no statistical rule on this matter; in any case, such decision can not be made on a statistical base but in the ambit of the adopted conceptual framework.	The adoption of this approach has to consider that the **meaning of the weights** (component scores) is exclusive statistical.	The approach is **not always applicable**.

The adoption of statistical methods in weighting components of social indices has to be considered carefully since, by removing any control over the weighting procedure from the analysts, it gives a false appearance of mathematical objectivity that is actually difficult to achieve in social measurement (Sharpe, 2004).

4. Statements and approaches in obtaining subjective weights

In order to identify a subjective weighting system, a **model** should be chosen by consideringthe criterion of importance or preference to be adopted, (ii) the level at which weights are determined (*individual* or *group* weights), (iii) the techniques allowing subjective evaluations and judgments to be expressed by subjects in a directly or indirectly way, (iv) the approach allowing a subjective importance/preference continuum to be constructed in order to transform evaluations and judgments into data analyzable and interpretable in terms of importance/preference weights.

4.1 Decision to be taken

4.1.1 Obtaining subjective weights at individual or group level

In order to determine subjective weights,

- data should be collected at individual level
- weights can be defined at (i) individual level (individual data will be used in order to construct weights that could be different for each subject), and (ii) group level (individual data will be used in order to construct different weights for different group of individuals).[2]

The issue can be formally represented as follows:

	Subjective weighting at individual level
X	a matrix with *N* rows (*i*=1 … *N*, individuals) and *K* columns (*j*=1 … *K* object variables) in which x_{ij} score that individual *i* assigned to *j* object (e.g. satisfaction for family)
W	a matrix with *N* rows (*i*=1 … *N*, individuals) and *K* columns (*j*=1 … *K* object variables) in which w_{ij} importance that individual *i* assigned to *j* object (e.g. importance of family)
Z	a new matrix with *N* rows (*i*=1 … *N*, individuals) and *K* columns (*j*=1 … *K* weighted object variables) in which z_{ij} weighted score for individual *i* concerning *j* object

	Subjective weighting at group level
X	a matrix with *N* rows (for *i*=1 … *N*, individuals) and *K* columns (for *j*=1 … *K* object variables) in which x_{cij} score that individual *i* belonging to the *c* group assigned to *j* object (e.g. satisfaction for family) The group can be predefined or can be determined through clustering methods
W	a matrix with *G* rows (for *c*=1 … *G*, groups) and *K* columns (for *j*=1 … *K* object variables) in which w_{cj} importance that group *c* assigned to *j* object (e.g. importance of family)
Z	a new matrix with *N* rows (for *i*=1 … *N*, individuals) and *K* columns (for *j*=1 … *K* weighted object variables) z_{ij} weighted score for individual *i* concerning *j* object

The aim is (a) to determine the values of the W matrix (in the two versions, weights for individual and weights for groups), (b) to determine the interpretable values in Z matrix, and (c) to sum up the K weighted scores in a unique individual synthetic score. We need to identify methods supporting the two perspectives, individual and group weighting.

4.2 Multi-attribute approaches

In order to define importance of a group of elements (elementary indicators) to be identified at subjective level and consequently to identify subjective weights methods are required able to manage a certain number of combined comparisons. These comparisons can be managed by applying methods aimed at making decision among different available alternatives. These methods are encompassed among *Multi-Attribute Models* Usually. Weights obtained through these methods are considered more stable than those produced by direct evaluations. Among these models we can distinguish:

a. ***Multi-Attribute Decision Making*** (*MADM*): it represents a branch of the wider field of *Multiple Criteria Decision Making* (MCDM) and refers to making preference decisions (e.g., evaluation, prioritization, selection) over available alternatives that are characterized by multiple conflicting

[2] In both cases, the general basic conditions described above are equally valid in obtaining subjective weights.

attributes (Yoon, 1995). **Analytic Hierarchy Process (AHP)** (*pairwise comparison of attributes*) represents one of the techniques used in this ambit. *Applicability of AHP in order to obtain subjective weights:* in our perspective (obtaining subjective weights), the possibility to identify different hierarchies when applied to identical problems can turn out to be some kind of advantage, represented mainly by the possibility to obtaining subjective weights at individual level by a quite straightforward approach. However, the need to construct a hierarchy with many nodes might make this approach non-applicable in the context we are dealing with (large surveys).

b. ***Multi-Attribute Compositional Models***: these models are based upon a statistical de-compositional approach through which it is possible to manage subjective comparisons of attributes on different levels. Its goal is to determine which combination is preferred by the subject. Among these model, **Conjoint Analysis** (CA) is the most known. While AHP approach derives the "importance" of an alternative by summing up the scores of the elementary indicators, CA approach proceeds in the opposite direction, that is by disaggregating the preferences, expressed by the subject in combination (Edwards, 1982; Green and Rao, 1971; Hair et al., 1998; Louviere, 1991; Luce & Tukey, 1964; Malhotra, 1993; Yoon, 1995). *Applicability of conjoint model in order to obtain subjective weights*: the estimated part-worths allow the range of importance for each factor to be determined. By dividing each factor's range by the sum of all range values we can obtain the proportion, interpretable in terms of importance of each factor in the respondent's choice. The polarity is consistent to the response scale submitted to the respondents and is considered inside the analytical procedure. The approach (i) allows obtained proportions to be assigned to objects in terms of weights, (ii) does not require the rescaling procedure to be applied, (iii) does not allow a continuum of importance to be obtained, (iv) meets the requirement of the sum of weights (sum of the obtained proportions is equal one), (v) can be applied for obtaining subjective weights at both individual and group level. However, the approach should be applied with great caution since the obtained weights strongly depend upon the definition of the levels for each factor.

4.3 Scaling approaches

As known, the traditional approaches that enable to deal with subjective evaluations and judgments are the "**scaling models**". Let recall the features that can describe and characterize each scaling model (McIver & Carmines, 1979):

- **Dimensionality**, concerning the variable to which the combined individual score/s will be referred. Each dimension is related to different aspects of the defined variable. Two different dimensionalities can be distinguished, (a) uni-dimensionality (the definition of the considered variable assumes an unique and fundamental underlying dimension), (b) multidimensionality (the definition of the considered variable assumes several underlying aspects/dimensions).
- **Nature of data**, which depends on the researcher's interpretation, expressed in terms of appropriateness and consistency. Different interpretations lead to different scaling procedures. Let us examine the scaling models applicable according to the classical classification of subjective data, theorized by Coombs (Coombs, 1950, 1964; Flament, 1976; Jacoby, 1991; McIver & Carmines, 1979):[3] (a) *Single stimulus*. Many scaling models were conceived for this kind of data; they are very often applied, such as the *additive model* and the *cumulative models* (deterministic and probabilistic); (b) *Stimulus comparison*. The reference scaling models for this kind of data are the Thurstone model and the Q methodology; (c) *Similarities*. The reference scaling model for this kind of data is the *multidimensional scaling*; (d) *Preferential choice*. One of the reference scaling models is the *unfolding model* (Arcuri & Flores D'Arcais, 1974; Cox, 1994; Flament, 1976; Kruskal & Wish,

[3] Clyde Coombs developed his theory based on geometric interpretation of data (Jacoby, 1991). Synthetically, two entities in a single datum can vary in two different ways: (a) with regard to the set to which the entities belong to. The entities can belong to the same set (e.g., two individual who take the same test) or to two different sets (e.g., a stimulus and a response); (b) with regard to the relation in which the entities are involved that can be (i) a dominance relation (an individual answers a question by reporting a level exceeding a defined measure) or (ii) a proximity relation (two individual share an event). In Coombs's Data Theory, the combination of the two ways produces four types of data:

		pairs of points in observation	
		same set	**different sets**
relation between points in pair	**dominance**	*Stimulus comparison*	*Single stimulus*
	proximity	*Similarities*	*Preferential choice*

1978; McIver & Carmines, 1979; McKeown, 1988; Thurstone, 1927, 1959; Torgerson, 1958).

- **Scaling technique**, comparative or non-comparative (Maggino, 2007).
- **Criterion for testing the model**. It is finalized to check the fitting of the model to data and it is different from model to model. The rationale of the testing procedure is common to all the models but the criteria are different according to the chosen model (Maggino, 2007).
- **Standard of measurement**, concerning the treatment of the multiple measures and the assignment of the synthetic value (the final score can be assigned to individuals or to stimulus), according to the following pattern:

Standard of measurement		Multiple measures	With regard to the variable the objective of the measurement is to classify	Final score assigned to
The multiple measures allow to measure in more accurately	individual	**Stimulus (item)**	the individuals	**Individual**
	indicator	**Individual**	the elementary indicators	**Stimulus (item)**

- **Contribution to the measurement of each multiple measures**: the contribution can be *uniform* (that is, all the multiple measures contribute through the same evidence) or *differential* (that is, the multiple measures contribute through different evidence); in this perspective, a particular item characteristic can be considered, the *trace line*, that defines the relationship between the identified continuum and the frequency observed for each value of that continuum. This frequency can be interpreted in terms of "probability to obtain each value" (McIver & Carmines, 1979). In particular, two frequency distributions can be associated to each item, corresponding to two different probabilities respectively (i) *alpha*, probability relating to the expected value ("correct answer " or "agreement with the submitted sentence" o "answer that is in the direction of the measured variable"); (ii) *beta*, probability relating to the not-expected value ("incorrect answer" or "disagreement with the submitted sentence" or "answer that is in the opposite direction to the measured variable").

The following table (Maggino, 2007) summarizes the characteristics of the well-known scaling models:

				Scaling model's Characteristics				
				Dimensionality	**Nature of data**	**Scaling technique**	**Criterion for testing the model**	**Standard of measurement: final (synthetic) score assigned to**
Scaling models	**Additive**	Uni-dimensional		Uni	Single-stimulus	Not-comparative	Internal consistency	Cases
		Multidimensional		Multi	Single-stimulus	Not-comparative	Dimensionality of the items	Cases
	Cumulative	Thurstone model (differential scale)		Uni	Stimulus comparison	Comparative (pair comparison or rank-order)	Metrics between items	Items
		Q methodology		Uni	Stimulus comparison	Comparative (rank-order or comparative rating)		Items
		Deterministic	Guttman	Uni	Single-stimulus	Not-comparative	Scalogram analysis: reproducibility, scalability and ability to predict	Cases and items
			Multidimensional Scalogram Analysis (MSA)	Bi			Regionality and contiguity	Cases and items
			Partial Ordered Scalogram Analysis (POSA)	Bi			Correct representation	Cases and items
		Probabilistic	Monotone (one or more parameters)		Single-stimulus	Not-comparative	• parameters estimation (maximum likelihood) • goodness of fit (misfit and residuals analysis)	Cases and items (without condensation)
	Perceptual Mapping	Multidimensional scaling		Multi	Similarities	Comparative (pair comparison)	Goodness of fit of distances to proximities (stress, alienation)	Items
		Unfolding		Uni & Multi	Preferential choice	Comparative	Goodness of fit of distances to ordinal preferences	Cases and items
	Conjoint model			Multi	Preferential choice	Comparative (rank-order)	Goodness of fit of the model (part-worth) to the ranking	Items at individual level

4.3.1 Scaling models allowing subjective weights to be obtained

The observation of the characteristics of the models allows us to identify those that better can help us in pursuing our goal, the identification of subjective weights. In particular, since we are looking for a "subjective weight" that is able to give back the idea of "subjective importance" attributed to each element (item) in comparison with the other elements composing the set, we have to select those models that utilize data (i) whose nature is comparative or preferential (marked in yellow in the previous table),

(ii) produced by a comparative scaling technique (marked in pink in the previous table). At this point, the models that can be selected are (a) *Thurstone model* (*differential scale*) and *Q* methodology[4], comprised among the cumulative approaches, and (b) *unfolding model* and *conjoint model*, comprised among the "perceptual mapping" approaches.[5]
Since we need also to identify a procedure that can be applied in a survey context without particular efforts, the *Q* methodology will be excluded by our consideration. In our perspective, these models can be distinguished with reference to the possibility to define subjective weights at individual level or at group level (last column of the previous table), in particular (a) individual weighting: *conjoint model* (again), (b) group weighting: *Thurstone model* (*differential scale*), *unfolding model*.

Cumulative approach

The approach based upon the logic that can be defined as "cumulative" has the goal to "create" a continuum on which the elements (items) concerning a certain characteristic are positioned. In order to pursue this goal, the judgments expressed by a group of individuals are employed. The judgments can be expressed using the "paired comparison" scaling technique or the "rank ordering".
Historically, Louis Thurstone (1927, 1959) was the first researcher that was engaged in the creation a continuum with a increasing intensity concerning a certain characteristic by using the judgments expressed by a group of "judges" (Arcuri & Flores D'Arcais, 1974; McIver & Carmines; 1979; Torgerson, 1958). In particular, Thurstone was mainly concerned with the fundamental problem of how psychological stimuli could be measured and compared with one another. The proposed model is based upon a fundamental assumption, the *law of comparative judgments*. According to this law, each object (occupation) submitted to the individual judgment arises a response produced by a *discriminant process* referring to the considered attribute. This discriminant process is a theoretical construct and represents the evaluation expressed by an individual in comparing two objects with reference to the attribute. We can assume for each object/stimulus and each attribute the existence of several *discriminant processes*. This means that the value of the discriminant process as a result of repeated evaluations of each object can show variations related to the existence of the error of measurement. This variability assumes the existence of a distribution of the discriminant processes. The distribution of the discriminant processes is assumed to be normal, described by two parameters, mean and standard deviation. The most frequently occurring response represents the *modal discriminant process* that defines the scale value of the object by which each object can be located along the continuum. The basic assumption underlying the law of comparative judgment is that the degree to which any two objects can be discriminated is a direct function of the difference in their status as regards the attribute in question. If the great part of the respondents judges object A different from object B with reference to the continuum, the placement of objects on the continuum should reflect the degree to which respondents can discriminate among the perceived characteristic of the various objects. The greater the distance between object A and object B on the continuum, the greater the proportion of respondents that have agreed that object A differs from object B. On the contrary, the smaller the distance between object A and object B on the continuum, the more confusion exists about the relative difference between the two objects with reference to the considered characteristic (McIver & Carmines; 1979; Thurstone, 1927, 1959; Torgerson, 1958). Scales created by this method are called *Thurstone scales* or *differential scales*. Many analytical versions exist according to the experimental model adopted (assumptions) and on the number of cases and the number of objects involved. Values, calculated through the application of particular and simple analytical procedure, allow defined elements to be placed on the continuum and can be considered in terms of group subjective weights. The main problem shown by this approach concerns the theoretically possibility to meet its fundamental assumptions, e.g. uni-dimensionality of the psychological continuum (McIver & Carmines, 1979). The approach needs particular care from the applicative point of view, especially with reference to choice of (i) the objects that should be involved and that should share the

[4] The well-known method called *Budget Allocation* (BAL) can be assimilated to Q methodology: each respondent is asked to distribute a certain budget – constituted by an *X* scores – among the objects, by assigning higher scores to those objects that he/she considers more important. In some cases, the procedure can be extended in order to achieve weights through agreement among respondents (group-weights). This approach turns out to be practicable in case of low number of objectives (max 10-12) in order to save respondents a difficult and complicated task.
[5] *Perceptual mapping* represents an approach that attempts to visually display the perceptions of individuals. Typically the position of an element (item) is displayed relative to their competition. Perceptual maps can have any number of dimensions but the most common is two dimensions.

same continuum (ii) the technique by which the objects should be showed be shown and evaluated b the respondents objects'. With reference to this, it should be considered that the paired compariso technique should not be applied with a high number of objects that could make the respondents' task to heavy, in terms of both time and required attention (Arcuri & Flores D'Arcais, 1974). Some solution have been studied in order to make respondent's task lighter and easier. *Applicability of cumulativ model in order to obtain subjective weights*: the cumulative approach (i) allows a continuum c importance to be obtained, (ii) requires the continuum to be interpreted in terms of polarity, (iii) allow the objects to be positioned on this continuum according to a quantitative value interpretable in terms c weights, and (iv) produces weights that should be rescaled in order to meet the weights' condition presented above.

Unfolding approach

The unfolding approach is one of the models developed for the preferential choice data. It is aimed a representing subjects and objects (said stimuli) in a common space – usually unidimensional – such tha relative distances between them reflect psychological proximity between defined objects an individuals. The analytic approach, defined and introduced by Coombs (1950; McIver & Carmine 1979), allows one preference scale (or more scales) to be obtained by ranking the objects accomplishe by the subjects. The procedure requires the administration of a series of stimuli that have to be ordere by each subject according to a preference criterion. Each individual's preference ordering is called *scale*.[6] The basic assumption posed by the model states that one (or more) common latent attribu (referred to as joint scale or *J scale*) exists underlying the different observed preference orderings of group of individuals. The underlying dimensions can be determined as a result of the identification c the *ideal point* of the scale on which the subject is placed. The goal is to verify whether the differe individual *I* scales can be located in a single *J* scale. If so, then we can reasonably conclude that th subjects employ a common criterion in evaluating the various stimuli. In the opposite case, two differe possibilities exist: (a) subjects employ multiple criteria in the evaluation of the stimuli, (b) subjec respond to the stimuli in a personal way, in other words, a common underlying attribute does not exis Let us suppose that two subjects expressed their preferences with reference to five stimuli – *a, b, c, d, e* and that the preferences could be represented on a single dimension. The process of evaluating th consistency of the individual *I* scales to be represented on a common *J* scale is called unfolding the scales through two different perspectives, (a) *unfolding*: according to this perspective, individu preference orderings (*I* scales) can be used in order to determine the *J* scale (strength of preference), (*folding*: according to this perspective, the *J* scale can be used in order to draw the individual preferen orderings (*I* scales). Generally few individual scales (*I*) are employed given that the model applicati turns out to be more complex in presence of a great number of *I* scales (McIver & Carmines, 1979) *Applicability of unfolding model in order to obtain subjective weights*: the unfolding approach (i) allov a continuum of importance to be obtained, (ii) requires the continuum to be interpreted in terms polarity, (iii) allows the objects to be positioned on this continuum according to a quantitative val interpretable in terms of weights, (iv) produces weights that should be rescaled in order to meet th weights' conditions presented above.

References

Arcuri L., Flores D'Arcais G.B. (1974) *La misura degli atteggiamenti*, Martello – Giunti, Firenze.

Bobko P., P.L. Roth, M.A. Buster (2007) "The Usefulness of Unit Weights in Creating Composi Scores. A Literature Review, Application to Content Validity, and Meta-Analysis", Organization Research Methods, vol. 10, n. 4, 689-709.

[6] The *unfolding* input matrix is *two-mode two-ways*. The generic element a_{ij} represents the preference expressed by the *j*- individual with reference to the *i*-th object. The model allows the two modes of the matrix to be represented in a single spat representation: the *N* objects and the *m* individuals (*joint space analysis*).

[7] The *unfolding* approach is aimed to represent – on a single metric continuum – both stimuli and subjects from preferen expressed by a group of subjects by assuming that subjects employ a common criterion in expressing the preferences w reference to the stimuli. This model has been extended to higher dimensions, applicable when the preferences are supposed be expressed by respondents according to different criteria. The multi-dimensional approach should be carefully conside because the possibility exists to obtain degenerate solutions (local minimum).

Coombs C.H. (1950) "Psychological scaling without a unit of measurement", *Psychological Review*, 57, 148-158.
Coombs C.H. (1964) *A theory of Data*, Ann Arbor, MI: Mathesis.
Cox T.F., Cox M.A.A. (1994) *Multidimensional scaling*, Chapman & Hall, London.
Edward W., Newman J.R. (1982) *Multiattribute Evalutation*, Sage University Paper Series on Quantitative Applications in the Social Sciences, series no. 07-026, Newbury Park, CA:Sage.
Flament C. (1976) *L'analyse booléenne de questionnaire*, Mouton Éditeur, Paris, La Haye.
Green P.E., Rao V.R. (1971) "Conjoint measurement of quantifying judgmental data", *Journal of Marketing Research*, vol. 8, 355-363.
Jacoby W.G. (1991) *Data Theory and Dimensional Analysis*, Sage University Paper Series on Quantitative Applications in the Social Sciences, series no. 07-078, Newbury Park, CA:Sage.
Hagerty, M.R., Land, K.C. (2007) "Constructing Summary Indices of Quality of Life: A Model for the Effect of Heterogeneous Importance Weights", *Sociological Methods and Research*, Vol. 35, n. 4, 455-496.
Hair J.F. jr., Anderson R.E., Tatham R.L., Black W.C. (1998, 5th ed.) *Multivariate Data Analysis*, Prentice-Hall Inc., Upper Saddle River, New Jersey.
Hsieh, C.M. (2003) "Counting importance: The case of life satisfaction and relative domain importance", *Social Indicators Research*, n. 61, pp. 227–240.
Hsieh, C.M. (2004) "To weight or not to weight: The role of domain importance in quality of life measurement", *Social Indicators Research*, n. 68, pp. 163–174.
Kruskal J.B., Wish M. (1978) *Multidimensional Scaling*, Sage University Paper Series on Quantitative Applications in the Social Sciences, series no. 07-011, Newbury Park, CA:Sage.
Louviere J.J. (1991) "Experimental choice analysis: introduction and review", *Handbook of Marketing Research*, Oxford: Blackwell Publisher.
Luce R., Tukey J.W. (1964) "Simultaneous conjoint measurement: a new type of fundamental measurement", *Journal of Mathematics Psychology*, vol. 1, 1-27.
Maggino F. (2005) *L'analisi dei dati nell'indagine statistica*, Firenze University Press, Firenze.
Maggino F. (2007) *La rilevazione e l'analisi statistica del dato soggettivo*, Firenze University Press, Firenze.
Malhotra N.K. (1996) *Marketing research: an applied orientation*, Prentice-Hall International, Inc., Englewood Cliffs, New Jersey
McIver J.P., Carmines E.G. (1979) *Unidimensional Scaling*, Sage University Paper Series on Quantitative Applications in the Social Sciences, series no. 07-024, Newbury Park, CA:Sage.
McKeown B., Thomas D. (1988) *Q Methodology*, Sage University Paper Series on Quantitative Applications in the Social Sciences, series no. 07-066, Newbury Park, CA:Sage.
Nardo M., M. Saisana, A. Saltelli and S. Tarantola (EC/JRC), A. Hoffman and E. Giovannini (OECD) (2005) *Handbook on Constructing Composite Indicators: Methodology and Userguide*, OECD, Statistics Working Paper.Ray A.K. (2008) "Measurement of social development: an international comparison", *Social Indicators Research*, vol. 86, n.1, 1-46.
Russell L.B., A.M. Hubley, A. Palepu and B.D. Zumbo (2006) "Does Weighting Capture What's Important? Revisiting Subjective Importance Weighting with a Quality of Life Measure", *Social Indicators Research*, vol. 75, n.1.
Sharpe A., J. Salzman (2004) *Methodological Choices Encountered in the Construction of Composite Indices of Economic and Social Well-Being*, Center for the Study of Living Standards, Ottawa, CAN.
Thurstone, L.L. (1927). *A law of comparative judgement*. Psychological Review, 34, 278-286.
Thurstone, L.L. (1959) *The Measurement of Values*. Chicago, The University of Chicago Press.
Torgerson W.S. (1958) *Theory and Methods of Scaling*, John Wiley & Sons, Inc., New York, London, Sydney.
Yoon K.P., Hwang Ching-Lai (1995) *Multiple Attribute Decision Making*, Sage University Paper Series on Quantitative Applications in the Social Sciences, series no. 07-104, Thousand Oaks, CA:Sage.
Wu, C. H. (2008) "Can We Weight Satisfaction Score with Importance Ranks Across Life Domains?", *Social Indicators Research*, 86, 469–480.