Modern Economy, 2011, 2, 614-624
doi:10.4236/me.2011.24069 Published Online September 2011 (http://www.SciRP.org/journal/me)

Who Meets the Standrads: A Multidimensional Approach

Antonio Villar
Department of Economics, *Pablo de Olavide University & Ivie*, *Seville*, *Spain*
E-mail: *avillar@upo.es*
Received June 9, 2011; *revised July* 15, 2011; *accepted July* 26, 2011

Abstract

We consider here the evaluation of the performance of a society with respect to a given set of targets. We provide a characterization of an intuitive evaluation formula that consists of the mean of the shares of the achievements in the targets. The criterion so obtained permits one not only to endogenously determine who meets the standards and who does not, but also to quantify the degree of fulfilment. Two empirical illustrations are provided: the compliance of the European Union Stability and Growth Pact, on the one hand, and the evaluation of research excellence in the Spanish universities, on the other hand.

Keywords: Meeting the Standards, Bonus/Malus Criterion, Multidimensional Targets, Additive Monotonicity

1. Introduction

Consider an organization consisting of several units whose performance is to be evaluated with respect to a vector of targets or reference values previously set. Depending on the problem under consideration, those targets may represent absolute values, relative performance thresholds, or a mixture of them. We can think that the purpose of the evaluation is the allocation of some resources among those who qualify and/or prestige or recognition. The evaluation procedure itself may be conceived as a simple dichotomous criterion concerning the achievement of the targets, it may attempt at providing quantitative estimates of the overall degree of fulfilment, or something in between (e.g. classification in different categories).

We shall refer to the organization as a *society* and to the incumbent units as *agents*. The key feature of the problem is the existence of a society with many agents whose performance is to be evaluated with respect to a given set of multidimensional targets, to be called *standards*. Note that in some cases meeting the standards may imply getting values below the thresholds.

Deciding who meets the standards in a multidimensional scenario is not immediate. Two extreme positions can be considered. On the one hand, there is the most demanding interpretation by which meeting the standards means achieving all target values simultaneously. On the other hand, there is the other extreme interpretation according to which achieving some target is a sufficient criterion. Each of those polar views makes the decision on who meets the standards rather trivial. The drawback is that in both cases we may find very unfair outcomes, as we can be treating equally highly different performances. The difficult problem is, of course, how to handle the intermediate cases. That is, when agents in society exceed some of the prescribed targets but fail to reach some others (a relevant case in practice and a usual source of conflicts). The bottom line is whether we admit or not compensations among achievements, both across dimensions and across agents, and what kind of compensations should be considered (we shall refer to this feature as the *substitutability* problem).

Let us consider two cases that illustrate well the key features of this type of evaluation problem.

Example 1: ***The European Stability and Growth Pact (SGP)***. *The SGP is an agreement among the* 16 *members of the European Union that take part in the Eurozone*, *to facilitate and maintain the stability of the Economic and Monetary Union*. *It involves setting reference values for some key public finance variables and aims at enforcing fiscal discipline after the monetary union (member states adopting the euro have to meet the Maastricht convergence criteria, and the SGP ensures that they continue to observe them). The basic reference values are two*: *(a) An annual budget deficit no higher than* 3% *of GDP*; *(b) A national debt lower than* 60% *of GDP*. *The question is: Are the countries in the Eurozone complying with the SGP*?

Example 2: ***Research excellence in the Spanish Universities.*** *It is well known that Spanish universities are not subject to regular evaluation processes*, *contrary* *to what happens to research groups or Faculty. As a consequence, society tends to assume that all universities are similar and the market does not discriminate graduates from different universities. Yet there are some data that would allow evaluating the research performance of the Spanish universities. The question is: Can we identify the set of universities that excel in research, out of the distribution of the results in the different research dimensions*?

Those examples illustrate two specific cases of the evaluation problem under consideration. In both examples the evaluation may require not only identifying those who meet the standards, but also to estimate their degree of success. In Example 1 the standards are fixed externally whereas in Example 2 the standards are relative to the actual performance. Therefore, we can also consider the question of whether some specific objectives have been reached in Example 1, whereas this type of question is meaningless in Example 2. Also observe that meeting the standards in Example 1 means having values of the index *below* the thresholds, whereas in Example 2 it means values *above* the thresholds.

This type of problem can be regarded as a case of multicriterion decision making (e.g. [1] and [2]). The proposed solutions may be interpreted as a class of *compromise solutions* on specific domains that evaluate the achievements in terms of some distance function (see, for instance, [3] and [4]). Our approach, however, stems from the principles that are applied for the analysis of development, inequality and poverty. Roughly speaking development measures allow to estimate the achievements, the targets play a similar role to the poverty thresholds, and inequality enters the picture as measuring the degree of substitutability among the achievements. See [5-10].

The paper is organized as follows. Section 2 contains the basic model. We present there the key assumptions and the essential ideas of this contribution by means of a simple and intuitive evaluation function: an arithmetic mean of the shares of the achievements in the targets. The axioms we use for that are rather standard: weighted anonymity (any two agents with the same weight and the same realizations are indistinguishable), weighted neutrality (all dimensions that enter with the same weight are equally important), a normalization property, and additive monotonicity (an increase in the realizations entails an increase in the evaluation function that depends positively on the size of that increment). Section 3 introduces a more flexible evaluation model, allowing for different degrees of substitutability between agents and dimensions, by characterizing the uniparametric family of generalized means. Section 4 contains an empirical illustration of this approach by analyzing the two examples presented above: the performance of the countries in the Eurozone, regarding the EU Stability and Growth Pact, and the selection of the set of excellent Spanish universities from a research viewpoint. A few final comments are gathered in Section 5.

2. The Basic Model

2.1. Measuring the Achievements

Let $N = \{1, 2, \cdots, n\}$ denote a *society* with n *agents* and let $K = \{1, 2, \cdots, k\}$ be a set of *characteristics*, with $k \geq 2$. A *realization* is a matrix $Y = \{y_{ij}\}$ with n rows, one for each agent, and k columns, one for each dimension. The entry $y_{ij} \in \mathbb{R}$ describes the value of variable j for agent i. Therefore, $\mathbb{R}^{nk}$ is the space of realization matrices and we assume implicitly that all dimensions can be approximated quantitatively by real numbers.

There is a parameter vector of reference values $z \in \mathbb{R}^k_{++}$ that describes the standards fixed for the different dimensions. We shall not discuss here how those thresholds are set, even though the importance of that choice is more than evident.

In order to deal with agents of different size or importance (e.g. families, firms, regions, countries), there is a vector $\rho \in \mathbb{R}^n_{++}$ that tells us the weights with which the different agents enter into the evaluation. Similarly, in order to allow for the presence of targets of different merit, we introduce a vector $\beta \in \mathbb{R}^k_{++}$ that puts weights on the different dimensions.

An evaluation problem, or simply a problem, is a point $P = (Y, z, \rho, \beta)$ in the space $\Omega = \mathbb{R}^{nk} \times \mathbb{R}^k_{++} \times \mathbb{R}^n_{++} \times \mathbb{R}^k_{++}$. We denote by $M(P) \subset N$ the set of agents who meet the standards in problem P.

In order to evaluate the overall achievements of the society with a realization matrix Y, relative to the reference vector z, and weighting vectors ρ, β, we look for a continuous function $\varphi : \Omega \to \mathbb{R}$ that associates to each problem $P \in \Omega$ a real value $\varphi(P)$ that provides a measure of its performance. This function is determined by a set of intuitive and reasonable properties that we introduce next.

The first property we consider, *weighted anonymity*, establishes that all weighted agents are treated alike. That is, if we permute agents' realization vectors together with their associated weights, the evaluation does not change.

Weighted Anonymity: Let $(Y, z, \rho, \beta) \in \Omega$ and let $\pi(Y)$, $\pi(\rho)$ denote a permutation of the indices of the rows of matrix Y and the corresponding entries of vector ρ. Then,

$$\varphi(Y, z, \rho, \beta) = \varphi(\pi(Y),\ z,\ \pi(\rho),\ \beta).$$

The second property, *weighted neutrality*, says that all weighted dimensions are equally important. That can be expressed, in line with the definition above, as follows:

Weighted Neutrality: Let $(Y, z, \rho, \beta) \in \Omega$ and let $\mu(Y),\ \mu(\beta)$ denote a permutation applied to the indices of the columns of matrix Y and the corresponding entries of vectors z and β. Then,

$$\varphi(Y, z, \rho, \beta) = \varphi(\mu(Y),\ \mu(z),\ \rho,\ \mu(\beta)).$$

Next property, *normalization*, makes the value of the index equal to zero when $Y = \mathbf{0}$ (the null matrix) and equal to $\sum_{i \in N} \rho_i \times \sum_{j \in K} \beta_j$ when $Y = Z$ (where $Z = (z, z, \cdots, z)$ is the matrix whose columns repeat the target vector $\mathbf{z}$ for each agent).[1] Formally:

Normalization:

$$\varphi(\mathbf{0}, z, \rho, \beta) = 0,\ \ \varphi(Z, z, \rho, \beta) = \sum_{i \in N} \rho_i \times \sum_{j \in K} \beta_j .$$

Our last property, *additive monotonicity,* establishes conditions on the behaviour of the evaluation function when the matrix of the agents' achievements changes from Y to $Y' = Y + \Delta Y$, for some $\Delta Y \in \mathbb{R}^{nk}$. The property requires the change of the index to be a monotone function g of the change ΔY in the realization matrix. This is a very natural property that is most useful when the data on the agents' performance is collected from several sources, or across different time periods, or when there are mistakes to be corrected. The new data can be integrated by simply computing the value of that function g regarding those new data and adding up the result to the original value of the index. Formally:

Additive Monotonicity: Let $(Y, z, \rho, \beta) \in \Omega$ and let $\Delta Y \in \mathbb{R}^{nk}$. Then,

$$\varphi((Y + \Delta Y), z, \rho, \beta) - \varphi(Y, z, \rho, \beta) = g(\Delta Y, z, \rho, \beta)$$

for some increasing function $g : \Omega \to \mathbb{R}$.

Note that this requirement is cardinal in nature and involves a separability feature of the overall index. Indeed, it implies that increasing the achievement of an agent in a given dimension by one unit will have the same impact on the index, no matter the level at which this happens (perfect substitutability of weighted agents and weighted dimensions).[2]

Remark *It is easy to see that additive monotonicity and normalization together imply additivity, that is,*

$$\varphi((Y + Y'), z, \rho, \beta) = \varphi(Y, z, \rho, \beta) + \varphi(Y', z, \rho, \beta).$$

The following result shows that all those requirements yield an evaluation function that corresponds to the arithmetic mean of the weighted shares of the achievements in the targets. Formally:

Theorem 1: *A continuos function* $\varphi : \Omega \to \mathbb{R}$ *satisfies weighted anonymity, weighted neutrality, normalization, and additive monotonicity, if and only if it takes the form:*

$$\varphi(Y, z, \rho, \beta) = \sum_{i \in N} \sum_{j \in K} \rho_i \beta_j \frac{y_{ij}}{z_j} \quad (1)$$

Moreover, those properties are independent.

Proof

1) The function in (1) satisfies all those properties. We prove now the converse.

Let $P = (Y, z, \rho, \beta) \in \Omega$ and let $\Delta_{ij}(a) \in \mathbb{R}^{nk}$ be a matrix with all elements other than (i, j) equal to zero and the (i, j) entry equal to a.

By applying repeatedly additive monotonicity we can write:

$$\varphi(P) = \sum_{i \in N} \sum_{j \in K} g(\Delta_{ij}(y_{ij}), z, \rho, \beta)$$

Let now $[\mathbf{1}, \mathbf{1}, \cdots, \mathbf{1}]a$ denote a uniform matrix whose generic element is a and take $z = \mathbf{1}s$, $\rho = \mathbf{1}p$, $\beta = \mathbf{1}d$, for some positive scalars s, p, d where $\mathbf{1}$ is the unit vector in the corresponding space. Note that, in this special case and in view of the weighted anonymity and weighted neutrality properties, we have:

$$g[\Delta_{ij}(a), \mathbf{1}s, \mathbf{1}p, \mathbf{1}d] = g[\Delta_{ht}(a), \mathbf{1}s, \mathbf{1}p, \mathbf{1}d], \quad \forall\, i, j \in N,\ \forall\, h, t \in K$$

Therefore, we can write:

$$\begin{aligned} &\varphi([\mathbf{1}, \cdots, \mathbf{1}]a,\ \mathbf{1}s, \mathbf{1}p, \mathbf{1}d) \\ &= kn \times g[\Delta_{ij}(a), \mathbf{1}s, \mathbf{1}p, \mathbf{1}d] \\ &\Rightarrow g[\Delta_{ij}(a), \mathbf{1}s, \mathbf{1}p, \mathbf{1}d] \\ &= \frac{\varphi([\mathbf{1}, \cdots, \mathbf{1}]a,\ \mathbf{1}s, \mathbf{1}p, \mathbf{1}d)}{nk} \end{aligned}$$

From that it follows:

$$\begin{aligned} &\varphi(Y, z, \rho, \beta) \\ &= \frac{1}{nk} \sum_{i \in N} \sum_{j \in K} \varphi([\mathbf{1}, \cdots, \mathbf{1}]y_{ij},\ \mathbf{1}z_j,\ \mathbf{1}\rho_i, \mathbf{1}\beta_j) \end{aligned} \quad (2)$$

Now observe that our assumptions imply that φ is linearly homogeneous, that is,

$$\varphi(\lambda Y, z, \rho, \beta) = \lambda \varphi(Y, z, \rho, \beta),$$

for all $\lambda > 0$. Let now $f : \mathbb{R}^4_{++} \to \mathbb{R}_{++}$ be given by: $f(y_{ij}, z_j, \rho_i, \beta_j) := \varphi([\mathbf{1}, \cdots, \mathbf{1}]y_{ij},\ \mathbf{1}z_j,\ \mathbf{1}\rho_i,\ \mathbf{1}\beta_j)$. As

[1] This simply extends the idea that the index is equal to one when Y = Z and all agents and all targets are equally important, *i.e.* $\rho_i = 1/n$ for all $i \in N$, $\beta_j = 1/k$, for all $k \in K$.

[2] This property may have an ethical content when agents are made of several individuals (e.g. the branches or the divisions of a firm) and the evaluation involves some rewards. It ensures the neutrality of the rule with respect to the order in which data are computed.

this function inherits the linear homogeneity property and satisfies normalization, by taking $y_{ij} = z_j$ and $\lambda = \frac{y_{ij}}{Z_j}$, we have:

$$f\left(\frac{y_{ij}}{z_j} z_j, z_j, \rho_i\right) = nk\rho_i\beta_j \frac{y_{ij}}{z_j}$$

Therefore, plugging those values into Equation (2), for all i, j, we get:

$$\varphi(Y, z, \rho, \beta) = \sum_{i\in N}\sum_{j\in K}\rho_i\beta_j\left(\frac{y_{ij}}{z_j}\right)$$

2) To separate the properties let us consider the following indices, for $\rho_i = 1/n$ for all i (anonymity), $\beta_j = 1/k$ for all j (neutrality):

2,a) $\varphi^A\left(Y, z, \mathbf{1}\frac{1}{n}, \mathbf{1}\frac{1}{k}\right) = \sum_{i\in N}\sum_{j\in K}\frac{y_{ij}}{z_j}$. It satisfies anonymity, neutrality, and additive monotonicity but not normalization.

2,b) $\varphi^B\left(Y, z, \mathbf{1}\frac{1}{n}, \mathbf{1}\frac{1}{k}\right) = \min_{i\in N}\left\{\frac{y_{ij}}{z_j}\right\}$. It satisfies anonymity, neutrality, and normalization but not additive monotonicity.

2,c) $\varphi^C\left(Y, z, \mathbf{1}\frac{1}{n}, \mathbf{1}\frac{1}{k}\right) = \frac{1}{k}\sum_{i\in N}\sum_{j\in K}\omega_i\left(\frac{y_{ij}}{z_j}\right)$, with $\sum_{i\in N}\omega_i = 1$ and $\omega_i \neq 1/n$ for some i. It satisfies neutrality, normalization, and additive monotonicity but not anonymity.

2,d) $\varphi^D\left(Y, z, \mathbf{1}\frac{1}{n}, \mathbf{1}\frac{1}{k}\right) = \frac{1}{n}\sum_{i\in N}\sum_{j\in K}\delta_j\left(\frac{y_{ij}}{z_j}\right)$, with $\sum_{j\in K}\delta_j = 1$ and $\delta_j \neq 1/k$ for some j. It satisfies anonymity, normalization, and additive monotonicity but not neutrality. **Q.e.d.**

This theorem tells us that assuming weighted anonymity, weighted neutrality, normalization, and additive monotonicity amounts to measuring social performance as the (weighted) arithmetic mean of the agents' relative achievements.

It is interesting to observe that equation (2) provides an implicit estimation of the performance of agent i with respect to dimension j, $e_{ij}(Y, z, \rho, \beta)$, that is given by the evaluation of a fictitious society with a uniform realization matrix $[\mathbf{1}, \cdots, \mathbf{1}]y_{ij}$, a uniform reference vector $\mathbf{1}z_j$, and a uniform weighting system $\mathbf{1}\rho_i$, $\mathbf{1}\beta_j$. That is,

$$e_{ij}(y_i, z, \rho, \beta) = \varphi\left([\mathbf{1}, \cdots, \mathbf{1}]y_{ij},\ \mathbf{1}z_j,\ \mathbf{1}\rho_i,\ \mathbf{1}\beta_j\right) \quad (3)$$

This allows us to estimate the *overall contribution of an agent*, by simply computing:

$$\begin{aligned} C_i(Y, z, \rho, \beta) &= \frac{1}{k}\sum_{j\in K}\varphi\left([\mathbf{1}, \cdots, \mathbf{1}]y_{ij},\ \mathbf{1}z_j,\ \mathbf{1}\rho_i\right) \\ &= n\rho_i\sum_{j\in K}\beta_j\frac{y_{ij}}{z_j} \end{aligned} \quad (4)$$

that is, as $n\rho_i$ times the weighted sum of all her relative achievements. Trivially, when $\rho_i = 1/n$ we have the weighted sum of the y_{ij}/z_j values.

Similarly, we can have a measure of the overall success of society in a given dimension, as:[3]

$$\begin{aligned} S_j(Y, z, \rho, \beta) &= \sum_{i\in N}\varphi\left([\mathbf{1}, \cdots, \mathbf{1}]y_{ij},\ \mathbf{1}z_j,\ \mathbf{1}\beta_j\right) \\ &= k\beta_j\sum_{i\in N}\rho_i\frac{y_{ij}}{z_j} \end{aligned} \quad (5)$$

2.2. The Agents Who Meet the Standards and the Targets that Have Been Reached

Let us consider now the question of who meets the standards and whether we can consider that a given target has been collectively achieved. In our model those problems are solved endogenously by the very formula that measures the overall performance. In order to facilitate the exposition, we focus on the case in which meeting the standards means achieving values above the established thresholds. In that case, an agent with $y_{ij} > z_j$, for all j, certainly meets the standards.

Consider now an agent h in the limit case in which $y_{hj} = z_j$, for all $j \in K$. According to equation (3), the overall performance of this agent is given by:

$$C_h\left((Y_{-h}, z), z, \rho, \beta\right) = n\rho_h\sum_{j\in K}\beta_j$$

(where (Y_{-h}, z) describes a matrix whose hth row is precisely z). Therefore, the set $M(P)$ of agents who meet the standards in problem P is given by:

$$M(P) = \left\{ i \in N \Big/ \sum_{j\in K}\beta_j\frac{y_{ij}}{z_j} \geq \sum_{j\in K}\beta_j \right\} \quad (6)$$

(note that we allow for the existence of agents in $M(P)$ whose achievements are below the target in some dimension, provided they are compensated with over compliance in other dimensions).

Equation (6) permits one to directly identify the set of those who meet the standards in the k-dimensional space in which we plot on $\mathbb{R}^k$ all agents' vectors of

[3]Note that computing the success in a given dimension makes sense when the thresholds are externally given and may not be meaningful when they correspond to functions of the actual values of the realization matrix.

relative achievements,

$$y_i(z,\beta)=(\beta_1 y_{i1}/z_1,\ \beta_2 y_{i2}/z_2,\cdots,\ \beta_k y_{ik}/z_k),$$

for all $i \in N$. Indeed, the set $M(P)$ is given by all those agents whose vectors of relative achievements are above the hyperplane defined by $\sum_{j\in K} y_{ij}(z)=\sum_{j\in K}\beta_j$.

When the reference values $z\in\mathbb{R}^k_{++}$ are externally given (*i.e.* they correspond absolute thresholds), we can also consider whether a specific objective has been reached by society. According to equation [5], objective j is achieved provided:

$$S_j(Y,z,\rho,\beta)\geq\sum_{i\in N}\rho_i=S_j\left(\left(Y^{-j},1z_j\right),z,\rho,\beta\right),$$

where $\left(Y^{-j},\mathbf{1}z_j\right)$ describes a matrix whose jth column is equal to z_j in all entries. Therefore, the set of objectives that have been collectively achieved are those that satisfy the following condition:

$$\sum_{i\in N}\rho_i\frac{y_{ij}}{z_j}\geq\sum_{i\in N}\rho_i,\quad j\in K \qquad (7)$$

3. A More Flexible Formlation

The additive structure of the evaluation function φ in Theorem 1 implies a particular trade-off between the different achievements, as the evaluation only depends on the sum of the agent's relative realizations but not on their distribution. So each agent can substitute any relative realization for another one at a constant rate (equal to $-\beta_j/\beta_t$, for all $j,t\in K$) no matter the level at which this happens. Similarly, the relative achievements of one agent in a given dimension can be substituted by those of another one, once more at a constant rate (here we find a marginal rate of substitution equal to $-\rho_i/\rho_h$ for all $i,h\in N$).

One might be willing to consider evaluation criteria that incorporate variable degrees of substitutability (e.g. decreasing marginal rates of substitution which implies penalizing the inequality of realizations across agents and/or dimensions, which may actually be a reason to introduce such a criterion). The simplest way of allowing for variable substitutability across agents and dimensions is by looking for a uniparametric extension of the formula in Theorem 1, so that controlling a single number permits one to regulate the degree of substitutability. To arrive at such a formula, let us start by performing the following exercise. Let $P=(Y,z,\rho,\beta)\in\Omega$, be a problem with Y strictly positive (*i.e.* $y_{ij}>0$ for all i,j) and consider the transformation $Y(\alpha)$ of Y given by $y_{ij}(\alpha)=(y_{ij})^\alpha$, for all i,j, and the transformation $z(\alpha)$ of vector z given by $z_j(\alpha)=(z_j)^\alpha$ for all j, some scalar α. Call $P(\alpha)$ to this transformed problem. Applying Theorem 1 to $P(\alpha)$ we get:

$$\varphi(P(\alpha))=\sum_{i\in N}\sum_{j\in K}\rho_i\beta_j\left(\frac{y_{ij}}{z_j}\right)^\alpha$$

The parameter α controls the impact of the individual deviations of the targets on the evaluation index. The larger the value of α the larger the impact of values above the reference level and viceversa. In particular, the parameter α controls the degree of concavity (for $\alpha<1$) or convexity (for $\alpha>1$) of the function.

Note that we require $y_{ij}>0$ for all entries of matrix Y, in order to avoid inconsistencies. We therefore, set $\widehat{\Omega}=\mathbb{R}^{nk}_{++}\times\mathbb{R}^k_{++}\times\mathbb{R}^n_{++}\times\mathbb{R}^k_{++}$ as our reference space from now on.

What should be the relationship between the evaluation of problems $P(\alpha)$ and P? The following property answers that question:

α -**Power**: Let $P=(Y,z,\rho,\beta)\in\Omega$ and let $P(\alpha)=(Y(\alpha),z(\alpha),\rho,\beta)\in\Omega$ denote a problem derived from the previous as follows. Each y_{ij} in Y is substituted by $(y_{ij})^\alpha$ and each z_j in z is substituted by $(z_j)^\alpha$, for $\alpha\in\mathbb{R}$. Then,

$$\varphi(P)=\left[\varphi(P(\alpha))\right]^{1/\alpha}$$

This property mimics the principle applied by the variance to the measurement of differences to the mean. If we take the power α of all relevant parameters of the problem, then we re-scale the resulting formula by taking the inverse power.

The following result is trivially obtained:[4]

Theorem 2: *An index $\varphi:\widehat{\Omega}\to\mathbb{R}_+$ satisfies weighted anonymity, weighted neutrality, "normalization", additivity and α -power, if and only if it takes the form:*

$$\varphi(Y,z)=\begin{cases}\left[\sum_{i\in N}\sum_{j\in K}\rho_i\beta_j\left(\frac{y_{ij}}{z_j}\right)^\alpha\right]^{1/\alpha}, & \alpha\neq0\\ \prod_{i\in N}\prod_{j\in K}\left(\frac{y_{ij}}{z_j}\right)^{\rho_i\beta_j}, & \alpha=0\end{cases} \qquad (8)$$

Moreover, those properties are independent.

Theorem 2 identifies the generalized mean of order α as the right formula to evaluate the performance of the society, where α is the parameter that incorporates our concern for equality across agents and dimensions (or the degree of substitutability).

Remark *Theorem 1 is not a particular case of Theorem 2 because the domain on which the evaluation function is defined is different.*

The set of those who meet the standards is now given

[4]The first part of the normalization property has to be adjusted to the new domain, by letting $\lim_{Y\to0}\varphi(Y,\cdots)=0$. We call "normalization" (with inverted commas) to this modified property.

by all agents whose vectors of weighted relative realizations, $y_i(z,\beta) \in \mathbb{R}^k_{++}$, are above (resp. below) the hyper-surface defined by:

$$\sum_{j \in K} \beta_j \left(y_{ij} / z_j \right)^{\alpha} = \sum_{j \in K} \beta_j .$$

Therefore, choosing α (the elasticity of substitution) amounts to fix the bonus/malus frontier. In particular, $\alpha \to -\infty$ (resp. $\alpha \to +\infty$) corresponds to the extreme case in which an agent meets the standards when she is above the targets in *all* dimensions simultaneously (resp. above *some* target); that is, the *max* (resp. the *min*) function. As for the intermediate cases, we find two of special relevance: the *arithmetic mean*, associated to the value $\alpha = 1$, discussed in the former section, and the *geometric mean*, associated to the value $\alpha = 0$. A similar reasoning applies to the case of achieving some target, with respect to the hyper-surface

$$\sum_{i \in N} \rho_i \left(\frac{y_{ij}}{z_j} \right)^{\alpha} = \sum_{i \in N} \rho_i .$$

From a different viewpoint the parameter α may be regarded as an equality coefficient in the following sense: the smaller the value of α the more weight we attach to a more egalitarian distribution of the agents' achievements, both among themselves and with respect to the different dimensions. The case $\alpha = 1$ shows no concern for the distribution, as only the sum of the achievements matters (*inequality neutrality*). Values of α smaller than one correspond to inequality aversion. The geometric mean, in particular, penalizes moderately the unequal distribution of the achievements, whereas the extreme case $\alpha \to -\infty$ (resp. $\alpha \to +\infty$) implies caring only about the smallest (resp. the highest) achievement of each agent.

This can be illustrated as follows. Take the evaluation function of a given agent,

$$C_i^{\alpha}(Y,z,\rho,\beta) = \left(n\rho_i \sum_{j \in K} \beta_j \left(\frac{y_{ij}}{z_j} \right)^{\alpha} \right)^{1/\alpha} \quad (9)$$

The parameter α controls de degree of substitutability among the different dimensions on an indifference curve, $C_i^{\alpha}(Y,z,\rho,\beta) = q$. The smaller the value of α the more difficult to substitute the achievement in one dimension by that in another. In the limit, no substitution is allowed so that meeting the standards implies surpassing all target levels.

Similarly, assuming that the reference values correspond to absolute thresholds externally given, the evaluation of the global performance with respect to a given target, $j \in K$, is given by:

$$S_j(Y,z) = \left[k\beta_j \sum_{i \in N} \rho_i \left(\frac{y_{ij}}{z_j} \right)^{\alpha} \right]^{1/\alpha} \quad (10)$$

The parameter α tells us now about the substitutability between individuals within a given dimension. The higher the value of α the easier to substitute the achievement of one individual by the achievement of another and viceversa.

4. Empirical Illustrations

Let us consider the application of our model to the evaluation of the two problems presented in Examples 1 and 2 in Section 1.

4.1. The European Stability and Growth Pact (SGP)

The SGP establishes that all member states of the Eurozone have to satisfy the following two requirements: (a) An annual budget deficit no higher than 3% of the GDP; (b) A national debt lower than 60% of the GDP. Let us take those values as the thresholds applicable to evaluate the performance of the states in the Eurozone, ignoring all implementation issues and the re-interpretations and refinements introduced later. **Table 1** provides the data on budget deficit and national debt for the 16 countries in the Eurozone, between 2006 and 2009. The question is to determine which countries do satisfy those criteria and which do not (note that here meeting the standards means producing outcomes which are *below* the thresholds).

Table 1 suggests several ways of interpreting the evaluation problem. On the one hand, we may consider that satisfying the performance criteria means meeting the standards every single year. In that case we would have four separate evaluation problems. On the other hand, one may also consider the evaluation for the whole period, as the performance of the countries is affected by the economic cycle. In that case we treat deficits and debt data corresponding to different years as if they were different variables.[5]

Table 2 provides the summary data of the countries' performance under the two evaluation approaches. The set of agents meeting the standards is given by:

$$M(P) = \left\{ i \in N \Big/ \frac{1}{k} \sum_{j \in K} \frac{y_{ij}}{z_j} \leq 1 \right\} \quad (6')$$

Therefore, we present the data in **Table 2** by showing in each cell the value $\frac{1}{k} \sum_{j \in K} \frac{y_{ij}}{z_j}$, so that we can easily identify those who meet the standards. **Table 2** includes data for t = 2006, 2007, 2008 and 2009, for deficit and debt together (first four columns), the data on deficit and debt on the whole period (last two columns).

[5]Here we assume that the two dimensions are equally important and also that all years are equally weighted. Note, however, that our model would easily accommodate differences in those respects.

Table 1. Public Debt and deficit in the Eurozone (2006-2009).

Country	2006		2007		2008		2009	
	Deficit	Debt	Deficit	Debt	Deficit	Debt	Deficit	Debt
Belgium	–0.3	88.1	0.2	84.2	1.2	89.8	6	96.7
Germany	1.6	67.6	–0.2	65	0	66	3.3	73.2
Greece	3.6	97.8	5.1	95.7	7.7	99.2	13.6	115.1
Spain	–2	39.6	–1.9	36.2	4.1	39.7	11.2	53.2
France	2.3	63.7	2.7	63.8	3.3	67.5	7.5	77.6
Ireland	–3	24.9	–0.1	25	7.3	43.9	14.3	64
Italy	3.3	106.5	1.5	103.5	2.7	106.1	5.3	115.8
Cyprus	1.2	64.6	–3.4	58.3	–0.9	48.4	6.1	56.2
Luxembourg	–1.4	6.5	–3.6	6.7	–2.9	13.7	0.7	14.5
Malta	2.6	63.7	2.2	61.9	4.5	63.7	3.8	69.1
Netherlands	–0.5	47.4	–0.2	45.5	–0.7	58.2	5.3	60.9
Austria	1.5	62.2	0.4	59.5	0.4	62.6	3.4	66.5
Portugal	3.9	64.7	2.6	63.6	2.8	66.3	9.4	76.8
Slovenia	1.3	26.7	0	23.4	1.7	22.6	5.5	35.9
Slovakia	3.5	30.5	1.9	29.3	2.3	27.7	6.8	35.7
Finland	–4	39.7	–5.2	35.2	–4.2	34.2	2.2	44
Average	1.3	68.3	0.6	66	2	69.4	6.3	78.7

Source: Eurostat (Euroindicators 2010).

Table 2. Performance of the Eurozone.

Country\Year	**Deficit and Debt together**				**Global**	**All years**	
	2006	2007	2008	2009		*Deficit*	*Debt*
Belgium	0.68	0.74	0.95	1.81	1.04	0.59	1.50
Germany	0.83	0.51	0.55	1.16	0.76	0.39	1.13
Greece	1.42	1.65	2.11	3.23	2.10	2.50	1.70
Spain	0.00	–0.02	1.01	2.31	0.83	0.95	0.70
France	0.91	0.98	1.11	1.90	1.23	1.32	1.14
Ireland	–0.29	0.19	1.58	2.92	1.10	1.54	0.66
Italy	1.44	1.11	1.33	1.85	1.43	1.07	1.80
Cyprus	0.74	–0.08	0.25	1.49	0.60	0.25	0.95
Luxembourg	–0.18	–0.54	–0.37	0.24	–0.21	–0.60	0.17
Malta	0.96	0.88	1.28	1.21	1.08	1.09	1.08
Netherlands	0.31	0.35	0.37	1.39	0.60	0.33	0.88
Austria	0.77	0.56	0.59	1.12	0.76	0.48	1.05
Portugal	1.19	0.96	1.02	2.21	1.34	1.56	1.13
Slovenia	0.44	0.20	0.47	1.22	0.58	0.71	0.45
Slovakia	0.84	0.56	0.61	1.43	0.86	1.21	0.51
Finland	–0.34	–0.57	–0.42	0.73	–0.15	–0.93	0.64
Eurozone	**0.79**	**0.65**	**0.91**	**1.71**	**1.01**	**0.85**	**1.18**

The data show that, according to the criterion in [6'] there are only two countries that meet the standards year by year between 2006 and 2009: Luxembourg and Finland. There are 7 more countries that satisfy the criteria when considering the whole period: Germany, Spain, Cyprus, Netherlands, Austria, Slovenia and Slovakia. And there are two countries that do not meet the standards in any of the years considered: Greece and Italy.

Let us now consider whether the Stability and Growth Pact has been fulfilled collectively along the years analyzed in **Tables 1** and **2**. To do so we let the weight ρ_i of each country be given by its relative GDP. We observe that, taking the two objectives together there is only one year in which the Eurozone did not satisfy the criteria of the SGP (last row of **Table 2**). Yet the deviation was bad enough as to conclude that for the whole Eurozone and the whole period, the pact has not been fulfilled (as $\varphi(.) = 1.01$). Looking at each objective individually, we observe that the Eurozone has collectively reached the deficit target (nine countries did it individually) but has failed to satisfy the debt target (even though eight countries met that objective). All together the Eurozone has failed to meet the standards, even though nine of the countries have succeeded in doing it.

4.2. Research Excelence in the Spanish Universities

We now consider the evaluation of research excellence in the Spanish public universities, out of the data reported in [11]. This paper analyzes the performance of the Spanish universities and provides an overall ranking using a set of variables whose relative weights are determined by the opinion of researchers obtained by a specific survey. Values are relative to the size of the permanent faculty in each university and are normalized so that the top university in each dimension gets a mark of 100.[6]

Here we take three out of the six variables computed by those authors, as we understand they are the most relevant ones. These variables are: publications (in terms of ISI papers), individual research productivity achievements, IRPA for short,[7] and success in getting research funds competitively. In order to define "excellence" we take a relative vector of reference values given by: $z_1 = 75$ for ISI publications, $z_2 = 85$ for individual productivity achievements, and $z_3 = 50$ for research funds. Those values correspond, approximately, to percentile 85 within each category. As for the weights of the variables we re-scale those in the study that imply the following: $\beta_1 = 0,348$ (papers), $\beta_2 = 0,328$ (IRPA), and $\beta_3 = 0,324$ (funds). **Table 3** provides the data corresponding to the 48 Spanish universities analyzed.

The object of this exercise is to determine the set of universities that are "excellent" from the point of view of their research realizations in 2009.[8] If we consider the extreme value $\alpha = -\infty$, that is, those universities that are above the thresholds in all dimensions, we find that there are only three universities that meet those standards of excellence: *Universitat Atònoma de Barcelona*, *Universidad Pablo de Olavide,* and *Universitat Pompeu Fabra*. If we take the case $\alpha = 0$ (the geometric mean), we find five additional universities entering the bonus set: *Universidad Autónoma de Madrid, Universitat de Barcelona, Universidad Carlos III, Universidad Miguel Hernandez,* and *Universitat Rovira i Virgili*. Reducing the level of exigency to $\alpha = 1$ (the arithmetic mean) does not add new universities to that set. Finally, for the other extreme value, $\alpha = +\infty$ (namely, the set of universities that satisfy at least one of those criteria), we find that the set of excellent universities includes five more: *Alcalá*, *Girona*, *Lleida*, *Rey Juan Carlos*, and *Valencia*.

Table 4 gives the data of the 8 universities that meet the excellence standards using the geometric and/or the arithmetic mean. The table contains their relative arithmetic mean scores, information about the region in which those universities are placed, and whe- ther they are *new* (created in the last twenty years, say), *modern* (created in the 60's) or *traditional* (with a history of hundreds of years). Even though discussing those data is not the purpose of this exercise, it is quite noticeable the success of the Catalan universities and the dominance of new and modern universities over the traditional ones.

5. Final Comments

We have provided here a criterion to evaluate the performance of a society with respect to a collection of targets. This criterion materializes in a simple an intuitive formula, a mean of order α of the shares of the realizations in the targets, which has been characterized by means of standard requirements. The order of the mean is a parameter that determines the substitutability between the achievements and therefore the admissible degree of compensation among the various dimensions and the different agents. From this perspective the model can be regarded as producing endogenously a system of *shadow prices* that permits one to aggregate the different

[6]By "permanent Faculty" is understood here those people who are civil servants (*funcionarios*) within the categories that require a doctoral degree. That should be taken into account in order to interpret the results.

[7]The "tramos de investigación", a voluntary individual research evaluation carried out every six years by a central agency, that results in a small salary increase.

[8]The results presented here correspond to the original figures after rounding them up to integer numbers plus at most two digits.

Table 3. Research Performance of the Spanish Universities.

Universities	ISI Articles	Research Bonuses	Research Funds
A Coruña	19.95	65.08	25.44
Alcalá	43.72	85.71	36.65
Alicante	48.95	80.95	31.2
Almería	38.82	65.08	23.69
Autónoma Barcelona	91.88	90.47	51.12
Autónoma Madrid	72.61	95.24	45.79
Barcelona	84.16	80.95	46.18
Burgos	34.99	63.49	28.68
Cádiz	30.16	66.67	19.03
Cantabria	51.53	80.95	34.6
Carlos III	62.01	100	55.53
Castilla-La Mancha	57.63	77.78	36.6
Complutense Madrid	27.65	77.78	32.13
Córdoba	60.51	77.78	19.53
Extremadura	39.38	77.78	19
Girona	64.91	66.67	60.68
Granada	42.92	77.78	26.94
Huelva	42.66	63.49	22.3
Islas Baleares	40.68	82.54	48.85
Jaén	56.33	66.67	38.46
Jaume I	40.5	79.36	33.06
La Laguna	31.28	58.73	14.48
La Rioja	35.56	69.84	25.45
Las Palmas de G.C.	19.82	50.79	17.34
León	29.86	73.01	21.99
Lleida	51.15	69.84	49.94
Málaga	30.27	69.84	20.47
Miguel Hernández	97.28	90.47	49.78
Murcia	41.51	77.78	25.67
Oviedo	37.55	76.19	23.57
Pablo de Olavide	80.58	92.06	62.22
País Vasco	19.23	68.25	31.44
Politécnica Cartagena	53.5	69.84	26.67
Politécnica Cataluña	46.93	74.6	42.55
Politécnica Madrid	30.04	50.79	26.89
Politécnica Valencia	62.32	63.49	34.95
Pompeu Fabra	100	87.3	100
Pública Navarra	44.22	74.6	27.28
Rey Juan Carlos	51.48	71.43	52.43
Rovira i Virgili	90.63	84.12	46.01
Salamanca	36.58	79.36	33.77
Santiago Compostela	49.51	82.54	34.39
Sevilla	36.29	76.19	25.71
UNED	20.88	66.67	20.67
Valencia	55.91	87.3	30.03
Valladolid	31.52	69.84	23.22
Vigo	56.65	66.67	31.11
Zaragoza	46.4	79.36	29.86

Source: Buela-Casal *et al*. (2010).

Table 4. Evaluation of the Spanish universities that meet the research standards.

Universities	*Score*	*Region*	*Type*
Pompeu Fabra	100	Catalonia	New
Pablo de Olavide	78.15	Andalucía	New
Miguel Hernández	77.51	Valencia	New
Autònoma de Barcelona	76.38	Catalonia	Modern
Rovira i Virgili	72.00	Catalonia	New
Carlos III	71.33	Madrid	New
Barcelona	69.16	Catalonia	Traditional
Autónoma de Madrid	69.09	Madrid	Modern

dimensions.

We have discussed in some detail the linear case, corresponding to the value $\alpha = 1$. There are good reasons to singularize this special case:

(a) It entails a principle very easy to understand: the arithmetic mean. This aspect may be important when the evaluation involves incentives, because understanding properly the incentives scheme is usually a necessary condition for its effectiveness.

(b) It permits one to perform the evaluation in the context of poor data. There are many situations in which we only have average values of realizations across agents but not individual data. Since the arithmetic mean of the original data coincides with the mean of the average values, we can apply this procedure even in the absence of rich data.

(c) It allows handling both positive and negative values of the variables.

(d) It fits well in those cases in which it is not clear whether one should penalize or foster diversity. Recall that values of α smaller than 1 penalize progressively the dispersion of the achievements whereas values of α greater than 1 do the contrary. So choosing α above or below unity amounts to promoting the differentiation of the agents' performance (specialization) or the homogeneous behaviour (uniformity). The linear case represents preference neutrality regarding pooling or separating behaviour.

Needless to say there are contexts in which values $\alpha \neq 1$ will be more suitable (e.g. when meeting the standards involves safety issues or when similar behaviour is preferable).

We have introduced the notions of weighted anonymity in order to deal with agents of different size or importance, and with targets of different relevance. The "size" of the agents will typically be related to the number of units within each agent (or the absolute value of their realizations, as in the Stability and Growth Pact, discussed above). We can also think of a more complex determination of those weights, in particular when individual outcomes may be partially interdependent. A case in point is that in which agents in society constitute a network (think for instance of the evaluation of research teams). In that case the weights may be associated to some measure of centrality, as in [12] and [13].

The presence of targets of different relevance is also common in many problems (e.g. weighting progressively less the past realizations when evaluating the outcomes along a given period of time). A different problem is that of handling targets with different *degrees of priority*, that is, targets that admit different degrees of substitutability (e.g. a group of targets have to be fulfilled before any other group is taken into account). The analysis of that case is left for future research.

6. Acknowledgements

Thanks are due to Ohiana Aristondo, Carmen Herrero, and Francisco André for their comments. I would like also to thank the hospitality of the Oxford Poverty and Human Development Iniciative (OPHI) while writing this paper. Financial support from Ministerio de Educación y Ciencia, under project ECO2010-21706 and Junta de Andalucía, under project SEJ-6882, is gratefully acknowledged.

7. References

[1] R. L. Kenney and H. Raiffa, "Decisions with Multiple Objectives: Preferences and Value Trade-offs," John Wiley, New York, 1976.

[2] P. L. Yu, "Multiple Criteria Decision Making: Concepts, Techniques and Extensions," Plenum, New York, 1985.

[3] C. Romero, "A Note on Distributive Equity and Social Efficiency," *Journal of Agricultural Economics*, Vol. 52 No. 2, 2001, pp. 110-112. doi:org/10.1111/j.1477-9552.2001.tb00928.x

[4] F. André, M. A. Cardenete and C. Romero, "Designing Public Policies," Springer Verlag, Berlin, 2010.

[5] F. Bourguignon and S. R. Chakravarty, "The Measurement of Multidimensional Poverty", *Journal of Economic Inequality*, Vol. 1, No. 1, 2003, pp. 25-49.

doi:org/10.1023/A:1023913831342

[6] S. R. Chakravarty, "A Generalized Human Development Index", *Review of Development Economics*, Vol. 7, No. 1, 2003, pp. 99-114. doi:org/10.1111/1467-9361.00178

[7] C. Herrero, R. Martinez and A. Villar, "Multidimensional Social Evaluation. An Application to the Measurement of Human Development", *The Review of Income and Wealth*, Vol. 56, No. 3, 2010, pp. 483-497. doi:org/10.1111/j.1475-4991.2009.00375.x

[8] S. Seth, "A Class of Sensitive Multidimensional Welfare Indices," Vanderbilt University, Nashville, Mimeo 2010.

[9] K. Tsui, "Multidimensional Poverty Indices," *Social Choice and Welfare*, Vol. 19, No. 1, 2002, pp. 69-93. doi:org/10.1007/s355-002-8326-3

[10] A. Villar, "A New Approach to the Measurement of Multidimensional Poverty," Pablo de Olavide University Working Paper, Seville, 2010.

[11] G. Buela-Casal, M. P. Bermúdez and J. C. Sierra, "Quevedo-Blasco, R. & Castro, A. (2010), Ranking de 2009 en Investigación de las Universidades Públicas Españolas," *Psicothema*, Vol. 22, pp. 171-179.

[12] C. Ballester, A. Calvó-Armengol and Y. Zenou, "Who's Who in Networks. Wanted: the Key Player," *Econometrica*, Vol. 74, No. 5, 2006, pp. 1403-1417. doi:org/10.1111/j.1468-0262.2006.00709.x

[13] B. Ruhnau, "Eigenvector-centrality: A Node Centrality?" *Social Networks*, Vol. 22, No. 4, 2000, pp. 357-365. doi:org/10.1016/S0378-8733(00)00031-9

