Ranking farms with a composite indicator of sustainability

Ernest Reig-Martínez[*]

Universitat de València. Dpto. Economía Aplicada II. Campus de Tarongers, 46022 Valencia, Spain and Instituto Valenciano de Investigaciones Económicas (IVIE). C/ Guardia Civil, 22, 46020 Valencia, Spain. Email: Ernest.reig@uv.es

José A. Gómez-Limón

Instituto Andaluz de Investigación y Formación Agraria y Pesquera. Dpto. Economía Agraria. PO Box 3092, 14080 Córdoba, Spain. Email: josea.gomezlimon@juntadeandalucia.es

Andrés J. Picazo-Tadeo

Universitat de València. Dpto. Economía Aplicada II. Campus de Tarongers, 46022 Valencia, Spain. Email: Andres.j.picazo@uv.es

Agricultural Economics, 42(4), 2011

Abstract

The assessment of sustainability at farm level has been growing in popularity over the last few years. This paper contributes to this line of research by building up composite indicators for different facets –social, economic, environmental and global– of farm sustainability using a methodological approach that combines *Data Envelopment Analysis* and *Multi-Criteria Decision Making* methods, and assigns common weights to each individual sustainability indicator. This approach is applied to a database of 163 farms located in the Campos County, a region belonging to the dry lands of the Spanish Northern Plateau, using 12 individual indicators of sustainability. Our findings show that both economic and environmental composite sustainability indicators are positively correlated, but that this is not the case for the social indicator. We also check the influence of a set of variables on farm sustainability using bootstrapping statistical techniques, and showing that increasing farm size, membership on agricultural cooperatives and farmers' medium and upper agricultural-specific technical education all exert a significant positive influence on sustainability. These results provide clues for policy-makers that intend to design sustainability-increasing structural agricultural policies.

Keywords: Sustainability index, Agriculture, *Data Envelopment Analysis*, *Multi-Criteria Decision Making*, Spain.

JEL Classification: C61, D60, H41, Q10.

[*] *Corresponding author*. Phone: +34 963 828 339/49; Fax: +34 963 828 354.

1. Introduction

Sustainability is a widely used concept, which gained popularity after the famous Brundtland Report, which defined sustainable development as '*development that meets the needs of the present without compromising the ability of future generations to meet their own needs*' (WCED, 1987: 43). This definition reflects the simultaneous consideration of three important aspects: recognition of the limited availability of natural resources, acceptance of the need for intergenerational equity, and several social and spatial issues summarized as intra-generational equity (Becker, 1997). However, despite the impressive amount of scientific literature that followed, 'sustainability' remains a vague and elusive term when it comes to empirical research.

In the specific case of agriculture, considerable research efforts have been made to overcome the conceptual vagueness of sustainability by defining the appropriate scale of reference to assess sustainability and by developing an adequate theoretical framework to integrate the diverse aspects of sustainability (Hansen, 1996; Rao and Rogers, 2006; Raman, 2006). But there has been no agreement to date on a common analytic framework that could simultaneously meet the needs of policy-makers and researchers in this field.

Policy-making operates on a larger geographical scale than farming-system research and has pushed sustainability studies in the direction of developing analytical frameworks that comprise sets of indicators defined at national level and ready for international comparisons (OECD, 2001; EEA, 2005). However, researchers concerned with the interaction between agriculture and biodiversity or agriculture and the quality of the natural environment, tend to emphasize that economic and ecological interrelations are most pronounced at farm level, even if environmental information is frequently lacking (van Wenum et al., 1999). Consequently, defining agro-environmental indicators at national or regional level is not appropriate when indicators are intended to provide information regarding the environmental value of agricultural ecosystems. Therefore, the need to collect and use farm or local farming system-specific agro-environmental indicators has been widely recognized (Rigby et al., 2001; Van der Werf and Petit, 2002; Pacini et al., 2003). This approach, considering the farm as the basic unit for sustainability assessment, has increased in popularity in the last years, once it was also proved useful for policy-makers, as this is the unit that most public policies focus on. In fact, the ways in which public decision-makers attempt to address the 'governance' of the whole agricultural sector have their most direct impact on farms (Andersen et al., 2007; van Passel et al., 2007).

Parallel to the definition of an appropriate scale of analysis, a major stream of the literature has attempted to quantify sustainability. Both the need to lend scientific substance to the multi-faceted concept of sustainability, and a strong political demand for comprehensive assessments of the joint evolution of economic, social and environmental conditions have been instrumental in the pursuit of quantifying sustainability. This has proved to be a difficult task. However, even though establishing the sustainability of a local farming cultivation system or individual farm is plagued with all sort of difficulties, analysts and policy-makers *should* at least be able to rank systems or decisional units, according to *comparatively more* sustainable production practices (Reig et al., 2010). This will be the case if, for example, they aim to fix the amount of public support to be received for each individual farmer based on sustainability criteria. For this purpose, is required to have a single criterion based on the aggregation of the different dimensions (indicators) of sustainability.

A substantial amount of research has thus been conducted in order to build up composite indicators or indices of sustainability. Accordingly, the criteria for the selection of meaningful indicators and the requirements concerning normalization, aggregation and weighting of the underlying variables have focused the attention of researchers. In this sense it is worth quoting the works of Nardo et al. (2008) for a general methodological discussion on the issues arising from the construction of composite indicators and Böhringer and Jochem (2007) for a link to the literature and a critical survey of the main sustainability indices.

In this paper we want to draw attention to a problem that remains even after individual relevant indicators have been selected and normalized, which are preliminary stages necessary towards the construction of a composite indicator. This problem is the selection of the appropriate weights to be applied to these indicators, which are supposed to reflect their relative importance in a global assessment of sustainability. Researchers have traditionally opted for one of the following three main approaches for weighting individual variables or indicators. The simplest solution has been to arbitrarily impose specific weights, sometimes granting equal weights to the different components, i.e., the type of practice that has been followed in the construction of the *Human Development Index* (UNDP, 1990), which is currently calculated as the arithmetic average of three indicators, respectively related to life expectancy, education and *GDP* per capita. A second alternative has been to introduce social preferences (citizens' opinion) regarding the different dimensions of sustainability, an approach grounded on the consideration of sustainability as a social construction (Gómez-Limón and Sanchez-Fernandez, 2010), or weighting indicator relying on experts' opinion. While there are a variety of ways to summarize citizens' or experts' judgment, it is becoming increasingly popular to resort to multi-criteria methods such as the *Analytical Hierarchic Process* based on an ordinal pair-wise comparison of attributes (Saaty, 1980, 2001). Finally, a third approach is represented by those researchers that prefer to avoid subjectivity in the determination of weights and proceed to determine them endogenously. One way of doing so is by performing *Principal Components Analysis* (Jolliffe, 2002) on the original variables or indicators.

Our paper belongs to the research stream which aims to compute a weighting scheme for indicators by using an endogenous or positive approach. Weights are derived from the data without resorting to exogenous information or personal preferences. We are able to do so by making use of *Data Envelopment Analysis* (*DEA*), which is a widely used technique based on mathematical programming that allows to benchmark the performance of individual decision making units (*DMUs*) against frontiers of best practices based on the observed behavior of other units (Cooper *et al*, 2007). Benchmarking establishes the basis for computation of some measure of inefficiency for each *DMU*, also paving the way for some form of rank ordering of the *DMUs*. Notwithstanding its usefulness to build up composite indicators, *DEA* also has some shortcomings when employed to rank decisional units. For this reason, we have combined *DEA* and *Multi-Criteria Decision Making* (*MCDM*) methods as our preferred methodological approach in this particular case.

What we aim to demonstrate in this paper is the usefulness of a *DEA*-inspired methodology when determining relative levels of farm sustainability, aggregating data from a set of economic, social and environmental indicators into a single composite indicator. Afterwards we also show how some relevant socio-economic variables may influence farm sustainability, thus providing empirical evidence for the design of sustainability-enhancing agricultural policy measures. As far as we are aware, this is the first time that a *DEA-MCDM* model has been used to rank individual farms according to their sustainability, which may help policy-makers to more efficiently address the issue of 'greening' agricultural policies. We apply our methodology to a sample of farms belonging to an extensive rain-fed agricultural system on the Spanish Northern Plateau that has undergone important structural changes in recent times.

The rest of the article is organized as follows. Section 2 provides a review of the literature on *DEA* and the construction of composite indicators. Section 3 presents the sample and describes how partial indicators of sustainability have been computed. Section 4 presents the main findings for the estimates of sustainability and puts them into context. Section 5 is devoted to analyzing the determinants of sustainability. Finally, Section 6 summarizes and concludes.

2. *DEA* as a tool for the construction of composite indicators

2.1. Adopting DEA for sustainability assessment

The basic theoretical framework underlying *DEA* is a production function, in which it is assumed that a set of k = 1,…,K *DMUs* make use of a vector of inputs $x = (x_1,...,x_M) \in \Re_+^M$ to produce a vector of outputs $y = (y_1,..., y_R) \in \Re_+^R$. The *technology* that allows the transformation of inputs into outputs is represented by an output correspondence which is a mapping

$P:\Re_{+}^{M}\rightarrow P(x)\subseteq\Re_{+}^{R}$, where the *output set* $P(x)$ represents the set of all feasible vectors of outputs y given a vector of inputs x. Furthermore, it is commonly assumed that the technology satisfies the properties initially suggested by Shephard (1970). In a basic *DEA* model, the efficiency of DMU_o is defined by the maximum of a ratio that transforms inputs into outputs:

$$
\begin{aligned}
& Max_{u_{ro},v_{mo}} \quad \frac{\sum_{r=1}^{R} u_{ro} y_{ro}}{\sum_{m=1}^{M} v_{mo} x_{mo}} \\
& Subject\ to: \\
& \qquad \frac{\sum_{r=1}^{R} u_{ro} y_{rk}}{\sum_{m=1}^{M} v_{mo} x_{mk}} \leq 1 \qquad k=1,\ldots,K \\
& \qquad u_{ro} \geq 0 \qquad\qquad r=1,\ldots,R \\
& \qquad v_{mo} \geq 0 \qquad\qquad m=1,\ldots,M
\end{aligned}
\tag{1}
$$

The weights u_{ro} and v_{mo} are chosen in order to place DMU_o under the most favorable light, meaning that they are computed by maximizing its efficiency ratio. Therefore, weights are peculiar for each unit and are selected according to the convenience of the *DMU* under evaluation (DMU_o in this case), subject to the constraints that the efficiency ratios of all *DMUs* in the sample computed with those weights have an upper bound of one. Accordingly, the dominance of a DMU_o over any other DMU_k requires finding positive weights u_{ro} and v_{mo} such that:

$$
\sum_{r=1}^{R} u_{ro} y_{ro} - \sum_{m=1}^{M} v_{mo} x_{mo} \geq \sum_{r=1}^{R} u_{ro} y_{rk} - \sum_{m=1}^{M} v_{mo} x_{mk} \text{ for all other } DMU_k \tag{2}
$$

Expression (1) can be also used to assess the relative performance of a decisional unit, i.e. concerning sustainability, after a suitable transformation to a linear form has been undertaken (see Cooper et al., 2007). The problem can be further simplified by assuming a single input for each unit and making it equal to unity. In the context of production economics this assumption implies that one input provides varying amounts of several services (Lovell et al., 1995), or, as in this case, gives rise to different intensities in several features that are relevant for the assessment of sustainability. Thus, for DMU_o the following model can be computed:

$$
\begin{aligned}
& Max_{\mu_{ro}} \quad h_o = \sum_{r=1}^{R} \mu_{ro} I_{ro} \\
& Subject\ to: \\
& \qquad \sum_{r=1}^{R} \mu_{ro} I_{rk} \leq 1 \qquad k=1,\ldots,K \\
& \qquad \mu_{ro} \geq 0 \qquad\qquad r=1,\ldots,R
\end{aligned}
\tag{3}
$$

where I_{rk} stands for the value of indicator r for DMU_k, and μ_{ro} is the weight attached to indicator r in the assessment of sustainability of DMU_o. Notice that we use I as a substitute for y as a reminder that we are no longer using outputs in our objective function, but measurable attributes of the decisional units. Instead of measuring decisional unit efficiency in the input-output transformation, the objective function now entails the achievement of the maximum value for a composite indicator obtained from a set of indicators corresponding to different facets of sustainability.

2.2. Weighting issues in DEA on constructing composite indicators

The *DEA* approach boasts important advantages as a basis for the measurement of sustainability. First, it can deal with a variety of value and physical data. This is important because of the multifaceted nature of the concept of sustainability, which involves monetary-valued economic indicators, but also social data and biophysical environmental variables employing other measurement units. Second, it provides a built-in method of data standardization, as decisional units are ranked from zero to one, according to their level of efficiency. And finally, *a priori* exogenous information is not required to calculate weights, which are obtained by solving individual linear program optimization. The ability of *DEA*-based models to build up composite sustainability indicators that reflect a variety of economic, social, and environmental factors has been shown by Callens and Tyteca (1999), who translate the *efficient* or *inefficient* characterization of a decisional unit into *sustainable* and *unsustainable*. They strike, however, a note of caution, warning that efficient by no means implies sustainable, the former being only a necessary condition for the latter. Best practice is only a relative concept while sustainability has to do with absolute magnitudes concerning the absorption capacity of ecosystems. *DEA* remains, thus, an acceptable method to assess *relative*, not absolute, sustainability levels and it is in this vein that we build up our sustainability indicator.

The strength of *DEA* lies in that it provides a reasonable method to detect inferiority defined in broad terms, i.e., with regards to sustainability. But, let us maintain the conventional use of the term inefficiency instead of inferiority: a decisional unit is judged inefficient when it performs worse than its peers, even if contemplated under the most favorable idiosyncratic set of weights. This *benefit-of-the-doubt* approach (Cherchye et al., 2007) fits international multidimensional comparisons particularly well when the use of a fixed common set of weights may prevent acceptance of the entire exercise by the national political authorities, some of which can feel disadvantaged when comparing the performance of the concerned member state within an international organization (Cherchye et al., 2008). But *DEA* is not so convincing when trying to discriminate within the sub-sets of efficient and inefficient units. Each inefficient unit score is calculated in relation to a different set of efficient units acting as a reference. It renders any direct comparison among inefficiency levels meaningless (Sinuany-Stern and Friedman, 1998, Kao and Hung, 2005), unless the units being compared share the same reference set, which seems highly improbable.

As regards the units that are classified as efficient, the problem is that a decisional unit could be deemed efficient because of a variety of spurious reasons. Maybe an excessive number of inputs and outputs have been observed in relation to the number of decisional units, or an unrealistic extreme set of highly idiosyncratic and unbalanced weights has been used. In the first case it has been recommended to use *Principal Component Analysis* in combination with *DEA* to reduce the problem of dimensionality (Adler and Golany, 2002), but it is the second problem that has attracted more attention, leading to restrictions on weights when using the multiplier version of the *DEA* model, and to the development of non-radial models that can incorporate all the sources of inefficiency when using the envelopment version (Silva-Portela and Thanassoulis, 2006).

The shortcomings of *DEA* when ranking decisional units have been recognized in the literature, which has also proposed a variety of ways to deal with this problem (Sarkis, 2000; Adler et al., 2002; Angulo-Meza and Estellita-Lins, 2002; Kao and Hung, 2005; Podinovski and Thanassoulis, 2007). Restricting flexibility in the selection of weights is an option that has been repeatedly mentioned. The original *DEA* model (Charnes et al., 1978) excluded zero weights through the incorporation of a non-Archimedean small number which assures that all inputs and outputs are used in the computation of the efficiency measures. But there is no way to determine what value this parameter should exactly take, because different values would give rise to different results. Instead, other methods, referred to as the *Assurance Region* and the *Cone Ratio*, have been employed to restrict the domain of the weights or *multipliers* (see Cooper et al., 2007 for an explanation). Value judgments have sometimes been used to provide prior information and set up restrictions on weights, as in Shang and Sueyhoshi (1995), which combine the use of *Analytical Hierarchic Process* –for the selection of the upper and lower bounds restricting the weights–, and the cross-efficiency matrix method.

Cross-efficiency was originally proposed by Sexton (1986) and Sexton et al. (1986) and has become a relatively popular method to rank efficient *DMUs*. In essence, the cross-efficiency method enlarges the number of different sets of multipliers that determine the efficiency ranking of a *DMU*. The *DMU* being ranked is not only contemplated under its own most favorable set of weights, but also according to the sets of multipliers that have been chosen in the first place as the most favorable for the *others DMUs*. Then, a matrix is constructed to show the ratings achieved by each *DMU* under the optima input and output multipliers of the others and an average efficiency figure is computed. Recently, new proposals have been made to improve the selection of suitable weight sets to be used in computing cross-evaluation (Lam, 2010).

More simple methods have also been suggested to compensate the use of potentially unbalanced weights. One of these seeks to combine two different indexes to form a composite indicator (Zhou et al., 2007). The first index is obtained by adopting a model that helps each *DMU* to select the *best* set of weights to assess its performance *vis-á-vis* the others, while the other selects the *worst* set of weights. Then both are combined using a parameter that changes according to the preferences of the analyst. Benchmarking, originating in Torgersen et al. (1996), has also been a popular approach. It consists in ranking an efficient unit according to its frequency of appearance in the reference sets of inefficient *DMUs*.

Andersen and Petersen (1993) introduced a technique called *Super-Efficiency Analysis*, which relaxed the upper bound of one for efficient units in the multiplier version through the exclusion of the unit being evaluated from the constraint set. Recently Nahra et al. (2009) claimed that this method has significant advantages over traditional *DEA* scores. Finally, other authors, starting with Friedman and Sinuany-Stern (1997), have applied multivariate statistical tools to achieve a complete ranking of *DMUs*.

2.3. Linking DEA and MCDM for constructing composite indicators

One stream of the literature pursuing better discrimination among decisional units and solving the problems created for the selection of unrealistic weights has profited from the links between *Data Envelopment Analysis* (*DEA*) and *Multi-Criteria Decision Making* (*MCDM*) techniques. The literature has recognized for a long time now that *MCDM* and *DEA* formulations are deeply related (Joro et al., 1998) and may even coincide if the input and output vectors employed in the *DEA* context are contemplated as vectors of attributes or criteria for evaluating *DMUs*, and output maximization or input minimization are viewed as the associated objectives. As an example, *DEA* and *MCDM* parallelism has been used to estimate the weights of different objectives for decision-makers (André et al., 2010). The multiplier formulation of *DEA*, based on the ratio between a weighted sum of outputs and a weighted sum of inputs for the *DMU* that is being analyzed, may therefore be seen as a *value function* under the *MCDM* perspective. This value function aggregates inputs and outputs into a single measure of worth (Stewart, 1996).

The utilization of multi-criteria decisional methods has created a new path to solving the problems posed by conventional *DEA* models in ranking *DMUs*. They have improved the discriminating power of *DEA* through the introduction of multiple criteria concerning efficiency measurement, while preserving the same set of constraints. Golany (1988) seems to be the first researcher to have proposed the integration of *DEA* and multiobjective programming, a specific branch of *MCDM*, to find common weights. Later on, Li and Reeves (1999) proposed a *Multiple Criteria Data Envelopment Analysis* (*MCDEA*) model that makes use of *minimax* and *minsum*-type objective functions, together with the classical objective function in *DEA*:

$$
\begin{aligned}
& Min_{u_{ro},v_{mo}} \;\; d_o \;\; or \;\; Max \;\; h_o = \sum\nolimits_{r=1}^{R} u_{ro} y_{ro} \\
& Min \;\; D \\
& Min \;\; \sum\nolimits_{k=1}^{K} d_k \\
& Subject \;\; to: \\
& \qquad \sum\nolimits_{m=1}^{M} v_{mo} x_{mo} = 1 \\
& \qquad \sum\nolimits_{r=1}^{R} u_{ro} y_{rk} - \sum\nolimits_{m=1}^{M} v_{mo} x_{mk} + d_k = 0 \qquad k = 1,\dots,K \\
& \qquad D - d_k \geq 0 \qquad k = 1,\dots,K \\
& \qquad u_{ro} \geq 0 \qquad r = 1,\dots,R \\
& \qquad v_{mo} \geq 0 \qquad m = 1,\dots,M \\
& \qquad d_k \geq 0 \qquad k = 1,\dots,K
\end{aligned}
\tag{4}
$$

In this model u_{ro} and v_{mo} are decisional variables. The first objective is the classical objective appearing in the *DEA* multiplier model, with d_o representing deviation between the weighted sum of outputs and the weighted sum of inputs of DMU_o and taking values between zero and the unity. Accordingly, DMU_o is deemed efficient if $d_o = 0$, or, equivalently, if $h_o = 1$. Variable D, in the second objective, represents the maximum value that can be reached across all d_k deviation variables. The third objective function accounts for summation over all d_k. Efficiencies defined under the second and third objective are far more demanding than under classical *DEA* criterion. Achieving $d_o = 0$ under *DEA*-efficient criterion does not imply efficiency under the *minimax* or *minsum* criteria but, by definition, *minimax* or *minsum* efficiency requires $d_o = 0$.

The utilization of *minimax* or *minsum* criterion yields a fewer number of efficient units and allows less flexibility in the allocation of input and output weights than conventional *DEA*. As a result, the discriminatory power of the model improves in comparison with the conventional *DEA* model in (1).

While several methods have been developed to discriminate between efficient units, a problem remains concerning the ranking of inefficient units. A solution based on obtaining a structure of common weights has been proposed to achieve a complete ranking of all *DMUs*. Again, multi-objective programming has been put to use in combination with *DEA*. More specifically, compromise programming has been employed (Kao and Hung, 2005; Despotis, 2002, 2005; Chen et al., 2009; Zohrehbandian et al., 2009), with a two-step procedure. The first step is to select a reference point, represented by a vector of efficiency scores. Some researchers chose that point by allowing each *DMU* to select the weights that maximize its efficiency score, using model (3). The second step consists of computing a set of common weights in such a way as to minimize the distance between the corresponding efficiency scores and the *ideal* or more favorable ones that have been chosen in step one.

In compromise programming an ideal solution is defined as that which would simultaneously optimize each objective individually, and is generally not feasible (Cohon, 1978). In order to determine the degree of proximity between both sets of efficiency scores, (i.e. *ideal* and *common weights-based*) a generalized family of distance measures (Zeleny, 1974) is applied:

$$d_p = \left[\sum_{k=1}^{K} \left(h_k^* - h_k \right)^p \right]^{\frac{1}{p}} \qquad p \geq 1 \tag{5}$$

where p represents the distance parameter, h_k^* stands for the *ideal* or *reference* efficiency score, and h_k for the score calculated under common weights. Changing the value of the p parameter, different distance metrics come into play and also different objectives can be formulated. Within the *MCDM* literature, some researchers have made use of this family of distance norms to seek a balance between environmental and economic objectives (Romero, 1996) and to explore compromises between alternative definitions of sustainability for natural systems (Díaz-Balteiro and Romero, 2004).

Following this methodological stream, Despotis (2002, 2005) introduced the *Global Efficiency* approach, which seeks to find a common set of weights, assessed across all the *DMUs*, to achieve global efficiency scores. In the process, the minimax and minsum objective functions, formerly described in model (4), are now combined additively. The author's original model was not linear, but after suitable transformations, Despotis (2005) proposed the following linear program:

$$\begin{array}{lll}
Min_{d_k,\mu_r,z} \; t\dfrac{1}{K}\sum_{k=1}^{K} d_k + (1-t)z & & \\
Subject\ to: & & \\
\qquad \sum_{r=1}^{R} \mu_r I_{rk} + d_k = h_k^* & k = 1,\dots,K & \\
\qquad (d_k - z) \leq 0 & k = 1,\dots,K & \qquad (6) \\
\qquad d_k \geq 0 & k = 1,\dots,K & \\
\qquad \mu_r \geq \varepsilon & & \\
\qquad z \geq 0 & &
\end{array}$$

ε being a non-Archimedean small number which assures that all attributes I_r are used in the computation of the sustainability scores.

The first term of the objective function, when considered on its own, represents the mean deviation between the *DEA*-efficiency scores, namely h_k^* (computed with the most self-favorable weights for each *DMU*), and the global efficiency scores for all units, whereas the second term represents, through the non-negative variable z, the maximal deviation between the aforementioned efficiency scores. Different sets of common weights are generated by varying the parameter t between 0 and 1, thus granting more or less relative importance to the norms respectively implied by the first and second terms of the objective function. Each value of the parameter t may produce a different set of common weights for inputs and outputs, thus generating a different global efficiency pattern.

When t takes the value 1 the parameter p takes the value 1 in the aforementioned family of distance functions (5). It does correspond to the norm L_1 (*city-block* or *Manhattan* concept of distance). The objective function now solely consists of the minsum objective. When t takes the value 0, then we are operating with norm L_∞ (*Tchebychev* distance) because parameter p equals infinite in formula (5). It means that minimax becomes the sole objective function. A series of alternative ranks can thus be obtained, according to different t values, and all *DMUs* can then be ranked with regards to their average global efficiency score.

Finally, and prior to the calculation of the sustainability scores as proposed by expression (6), the *benefit of the doubt* scores of sustainability need to be obtained using a *DEA* model which allows for idiosyncratic weights. In order to do so, in this paper we propose using the additive output-oriented *Slacks-Based Measure* (*SBM*) model introduced by Tone (2002), which integrates radial efficiencies as well as slacks into a single scalar measure of efficiency. Formally, the *DEA*-based score of performance for DMU_o, namely variable h_o^*, computed with the most self-favorable weights for this unit comes from the following program:

$$
\begin{aligned}
& h_o^* = Min_{\lambda_k, s_r^+} \frac{1}{1+\frac{1}{R}\sum_{r=1}^{R} s_r^+ / I_{ro}} \\
& \textit{Subject to}: \\
& \qquad x_o \geq \sum_{k=1}^{K} \lambda_k x_k \\
& \qquad I_{ro} = \sum_{k=1}^{K} \lambda_k I_{rk} - s_r^+ \qquad r = 1,\dots,R \\
& \qquad \lambda_k \geq 0 \qquad k = 1,\dots,K \\
& \qquad s_r^+ \geq 0 \qquad r = 1,\dots,R
\end{aligned}
\tag{7}
$$

where s_r^+ is the slack in the indicator of sustainability r, and λ_k measures the intensity with which DMU_k enters in the composition of the efficient reference set to which DMU_o is being compared. Furthermore, the parameter h_o^* is upper bounded to one, with a unity score indicating the best performance.

Resorting to the calculation of common weights to increase discrimination between decision units, within a *DEA-MCDM* aggregation framework, as in Despotis (2002, 2005), has introduced a new perspective in the calculation of sustainability composite indices. As far as we are aware this procedure has not been previously applied to the assessment of farm-level sustainability. Also, the employment of a slacks-inclusive measure of efficiency in the first step of our *DEA-MCDM* analysis avoids the shortcomings of radial efficiency measures in the computation of *benefit of the doubt* sustainability indicators, contributing to an improved definition of the aforementioned *ideal* or *reference* efficiency score.

3. Data and sample: Computation of indicators of sustainability

3.1. Rain-fed agriculture in Campos County (Spain)

The empirical application of the methodology proposed has been implemented on a representative sample of 163 farms belonging to the rain-fed agricultural system of Campos County[1], in the province of Palencia, located in the central part of the Spanish Northern Plateau (about 800 m.a.s.l.). Characterized by a continental climate, production in the region is mainly based on annual extensive crops, particularly winter cereals.

This county has a surface area of 304,483 hectares, of which 86% (261,505 hectares) is considered as utilized agricultural area (*UAA*). Rain-fed lands account for 83% (254,992 hectares) of county's territory, the major crops being barley (52%), wheat (26%), alfalfa (5%), sunflower (4%), oats (4%) and pulses (3%). We chose this agricultural system for our case study firstly due to practical interest, bearing in mind that it can be treated as a representative case of extensive (low-input-low-output) agriculture, where environmental functions play a relevant role (Kallas et al., 2007). The second reason is that the farms are relatively uniform (all of them operate under the same edafo-climatic, market and legal frameworks) and that data are readily available, both factors being advantageous for sustainability analysis.

3.2. The selection and computation of indicators of sustainability

In order to develop the sustainability assessment proposed for the aforementioned case study, the first stage was to select a set of indicators that cover the three components of sustainability: economic, social and environmental. For this purpose, a general catalogue of indicators of agricultural sustainability for each of the dimensions of sustainability was built. This catalogue was based on an extensive review of the literature, from which we gathered the indicators utilized in previous studies carried out by both institutions and individual researchers. On the basis of this catalogue, the selection of the most appropriate indicators was undertaken by a panel of 16 experts (seven university researchers, five agricultural public servants and four agricultural extension officers) chosen on the basis of their broad knowledge of the agricultural system analyzed and the different dimensions of sustainability. In order to select the indicators, the ex-

[1] The case study considered in our research is based on the same dataset as that in Gómez-Limón and Sanchez-Fernandez (2010). However, these authors follow a rather different methodological stream, using *Principal Component Analysis* and *Analytic Hierarchy Process* as weighting techniques to compute composite indicators of sustainability. Furthermore, they employ a fairly different approach to determine the factors influencing farm-level sustainability.

perts followed the criteria of analytical soundness, measurability and policy relevance suggested by Sauvenier et al. (2006). Furthermore, experts were also asked to follow the pragmatic criterion of 'economic valuation' proposed by Pannell and Glenn (2000). Thus, the only indicators to be chosen were those capable of being calculated easily and cheaply through the information provided by agricultural producers.

After employing this procedure, we finally obtained a set of 12 indicators[2]: three indicators for the economic dimension (income of agricultural producers –*INCOME*–, contribution of agriculture to GDP –*CONTGDP*– and insured area –*INSUAREA*), four for the social component (agricultural employment –*AGRILABO*–, work-force stability –*LABOSTAB*–, risk of abandoning agricultural activity –*RISKABAN*– and economic dependence on agricultural activity –*ECODEPEN*) and five environmental indicators (soil cover –*SOILCOV*–, nitrogen balance –*NITROBAL*–, pesticide risk –*PESTRISK*–, energy balance –*ENERGBAL*– and environmental subsidy areas –*AGROENV*). A brief explanation of the definitions and significance of each of the indicators chosen can be found in Table 1. Further technical details regarding why each indicator was chosen and how they were calculated can be found in Sanchez-Fernandez (2009).

Table 1. Sustainability indicators. About here

The abovementioned indicators have different measurement units, so we decided to use a normalization procedure. Selecting a suitable method for normalization (see Nardo et al., 2008 for a review) is no trivial affair and warrants special attention, as different normalization methods can yield different results (Ebert and Welsh, 2004). Among the various normalization techniques available, we employ *re-scaling* or *ranging* normalization. The formulas used are the following (Nardo et al., 2008):

$$I_{rk} = \frac{\left(C_{rk} - C_{r\min}\right)}{\left(C_{r\max} - C_{r\min}\right)} \tag{8}$$

$$I_{rk} = \frac{\left(C_{r\max} - C_{rk}\right)}{\left(C_{r\max} - C_{r\min}\right)} \tag{9}$$

where I_{rk} and C_{rk} are, respectively, the normalized and crude values of indicator r for farm k, while $C_{r\min}$ and $C_{r\max}$ are, respectively, the minimum and maximum crude values of indicator k found in the sample of farms considered.

[2] While Gómez-Limón and Sanchez-Fernandez (2010) make use of 16 base indicators to compute their composite sustainability indicator, we only use 12 of these base indicators. The reason is twofold. First, one indicator in Gómez-Limón and Sanchez-Fernandez (2010) that refers to the sustainable use of irrigation water has been dropped, as we only deal with rain-fed farms. Second, we have excluded another three less relevant indicators in order to increase the discriminating power of our *DEA*-based models. Even after having reduced the number of base indicators, the problem of lack of discrimination still remains, as mentioned later on. However, this problem can be now overcome by combining *DEA* and *MCDM* techniques.

Expression (8) is used in the cases where indicators mean '*the greater the crude value is, the more sustainable the farm*' (*INCOME*, *CONTGDP*, *INSUAREA*, *AGRILABO*, *RISKABAN*, *ECODEPEN*, *SOILCOV*, *ENERGBAL* and *AGROENV*). On the contrary, expression (9) is used when indicators mean '*the lower the crude value is, the more sustainable the farm*' (*LABOSTAB*, *NITROBAL* and *PESTRISK*). Thus, following this procedure, the values of the normalized indicators (I_{rk}) vary within a dimensionless range [0,1], whereby 0 corresponds to the worst possible value of the indicator (the least sustainable) and 1 to the best (the most sustainable).

3.3. Data gathering

We have relied on the data provided by an *ad hoc* survey as the main source of primary information to calculate the variables mentioned above. With this purpose, a specific questionnaire was designed including relevant information concerning productive processes (technology and inputs used and output obtained for every crop grown), as well as farm and farmer features as explanatory variables of sustainability used in a second stage analysis.

The universe of this survey was the 3,960 farms that, according to the last *Agricultural Census*, operate in Campos County. Given the difficulty of implementing random sampling, stratified sampling was chosen based on the affiliation of these producers to farmers' unions (*ASAJA*, *UPA* and *COAG*). The survey was completed by personal interviews during the period where farmers went to union offices in order to fill in forms to obtain Common Agricultural Policy subsidies and payments (March-April, 2008). Following this procedure, 163 valid questionnaires (farms) were finally obtained.

The primary information supplied by the survey has been complemented with secondary additional information in order to calculate individual farm values for each indicator. This information has been mainly collected from scientific literature (technical coefficients), although official statistics and legal documents have also been used. Tables 2a and 2b depict a summary descriptive analysis of these variables.

Tables 2a and 2b. Sample description. About here

4. Results: Ranking farms in Campos County with a *DEA-MCDM* approach

This section presents and discusses our estimates of sustainability. Composite indicators of sustainability have been computed at the farm level aggregating the indicators described above according to model (6). Previously, the *benefit of the doubt* scores of sustainability allowing for idiosyncratic weights have been computed using model (7).

Following the methodology commented above, first, *dimensional* composite indicators have been built for economic, social and environmental sustainability, considering in each case only the indicators chosen for each of these dimensions. In a second step, a *global* composite indicator of sustainability has been also assessed, considering all the indicators selected for integrated sustainability assessment.

Before commenting on the results of these composite indicators, it is worth remembering that an infinite number of solutions (scores for the composite indicator) for model (6) is possible depending on the metric considered, i.e., the value given to parameter t. In order to illustrate this point, Fig. 1 shows how the score of the global composite indicator varies for two particular farms when t is parameterized.

Fig. 1. Scores of the global composite indicator depending on the value of t. About here

As explained above, different values of t involve different proprieties of the composite indicator being calculated, all of which are equally relevant from a technical point of view (no single value of this parameter can be considered as the 'right' one). In any case, it is worth obtaining a single score for the composite indicator so as to simplify the analysis. With the purpose of aggregating all these alternative composite scores depending on t into one sole score of sustainability, Despotis (2002) proposes obtaining an equal number of distinct composite scores for each *DMU* and then to take the average value as the unique sustainability score. However, this approach could lead to different results depending on the specific composite scores (values of t) considered for assessing the average. In an attempt to avoid this source of subjectivity, in this paper we propose considering the value of the definite integral of the composite indicator from t=0 to t=1 as the sole aggregated index. The resulting value of this integral corresponds with the area below the line representing the value of the composite indicator, as plotted in Fig. 1.

Such integration can be calculated numerically, considering numerous enough scores for the composite indicator along the range [0,1] of t. For our case study, we have computed composite sustainability scores allowing the parameter t to vary between 0 and 1 at intervals of 0.01. As a result, we have obtained up to 101 sets of estimates of sustainability for each farm and composite indicator. The values obtained following this calculus are reported next.

Summarized descriptive statistics for the composite indicators assessed can be observed in Table 3. Each composite indicator for a specific dimension of sustainability has been computed independently, using the corresponding set of variables, as displayed in the Table. The global sustainability indicator makes use of all 12 variables. It is important to note that after applying model (7), a total of 94 farms out of 163 are tied with scores equal to unity for the global sustainability indicator. Solving this lack of discrimination is one of the advantages of using the *DEA*-

MCDM model for the computation of common weights. In fact, not a single tie appears in the ranking performed with model (6).

Table 3. Descriptive statistics for dimensional and global sustainability. About here

Regarding the values achieved for the global sustainability composite indicator, it is worth highlighting that these scores are distributed with a mean of 0.561 and a standard deviation of 0.098. In this sense, it is noteworthy that although all farms operating within a particular agricultural system share the same edafo-climatic (crop alternatives), technological (productive options), market or legal frameworks, they can be relatively heterogeneous in terms of sustainability performance (see Fig. 2 for our case study). This reflects the crucial role of farmers' productive decisions, which finally determine the level of sustainability of each individual farm. Thus, there is room to incentivize agricultural producers to modify the way they manage their holdings through appropriate policy instruments in order to upgrade their sustainability performance.

Fig. 2. Cumulative probability of the global sustainability composite indicator. About here

Once the dimensional composite indicators have been calculated, the next question posed is whether all these indices rank the farms analyzed in a similar way. For this purpose, we have used the Spearman's rank correlation coefficient, verifying the statistical dependence of the different dimensions of sustainability by pairs. By doing so, we are able to confirm which dimensions are complementary or, in contrast, which are in conflict. The results achieved in this sense are shown in Table 4.

Table 4. Spearman's rank correlation coefficients. About here

Results show that economic and environmental composite indicators are correlated positively and significantly. This could suggest that most farms in our sample are inefficient from both economic and ecological points of view, that is, they are able to decrease the current use of inputs and environmental pressure (an improvement in environmental sustainability) while still maintaining or improving their economic performance (economic sustainability). That is, steps towards extensification would be worth taking as a measure to increase both economic and environmental sustainability. As will be commented next, farm size is probably the variable that links farms' economic and environmental performance in this agricultural system (smaller farms do not achieve the economies of scale required to be efficient from an economic and environmental perspective). However, further analysis should be performed in order to confirm these hypotheses, such research being far beyond the scope of this paper.

Furthermore, it is also worth indicating that the social composite indicator is negatively correlated with the economic and environmental dimensions of sustainability. In order to explain this fact, it is essential to take into account the particular features of this agricultural system. As in other similar low-input-low-output agricultural systems worldwide, the analyzed area is undergoing a profound process of structural adjustment characterized by rising capital/labor ratios and the enlargement of agricultural holdings (Moreno-Pérez and Ortiz-Miranda, 2008). Both features, larger and more mechanized farms, are prerequisites to maintain farms in the agricultural business. Farms failing to successfully adapt to this process have disappeared[3]. Within this framework, it is easy to understand how most of the more labor-intensive (more socially sustainable) farms perform disappointingly from an economic point of view (economic sustainability), taking into account that labor productivity falls in these cases below current wages (with labor provided by salaried workers) or below opportunity costs (with own-family labor). Furthermore, these more socially sustainable (labor-intensive) farms normally grow more intensive crops, thus generating higher environmental pressure (lower environmental sustainability).

Finally, as is logically assumed, we can see that all single-dimensional composite indicators are positively correlated with the global sustainability composite indicator. However this correlation is not significant for the social dimension. This is probably because of the existing trade-offs between economic and environmental sustainability on the one hand and social sustainability on the other.

5. Explaining farms' sustainability

In the empirical literature on performance assessment, it has been common practice to perform analyses aimed at investigating the determinants of performance. Our interest here is also to explore the factors that might explain sustainability in the agricultural system of Campos County. In doing so, we make use of the procedure proposed by Simar and Wilson (2007), based on truncated regression and bootstrapping techniques, to explain the composite indicator of global sustainability. With the purpose of accommodating our analysis to the left-truncated distribution functions developed in Simar and Wilson's paper, the variable to be explained is the inverse of the scores of global sustainability[4]. Thus this transformed variable ranges from one to infinity, the higher the value the less sustainable or, in other words, the more unsustainable.

That said, explaining farmers' sustainability requires taking the following steps. In the first place, maximum likelihood needs to be used to obtain estimates of $\boldsymbol{\beta}$ and σ_η in the truncated regression of the inverse of the sustainability scores estimated with *DEA-MCDM* techniques on a set of covariates z, using the subset $i = 1,\ldots,N$ of inefficient observations, i.e., observations

[3] In fact, the number of holdings in the region of Castilla y León has declined dramatically in recent years. According to official data (National Statistics Institute (*INE*): http://www.ine.es/), in 1999 there were 175,454 farms in the region, while in 2007 there were only 93,142. This implies a decrease of 47% in the number of agricultural holdings over this eight-year period.

[4] Picazo-Tadeo and García-Reche (2007) and Picazo-Tadeo *et al.* (2009) use a similar transformation.

with a score of sustainability smaller than one or, in other words, an inverse of this score greater than one. Formally:

$$\left(\frac{1}{SUSTAINABILITY}\right)_i = z_i\boldsymbol{\beta} + \eta_i \qquad (10)$$

Secondly, the following three steps need to be repeated L times to obtain a set of bootstrap estimates of $\boldsymbol{\beta}$ and σ_η. Step one consists of drawing η_i for each i = 1,…, N from the following normal distribution:

$$N\left(0,\hat{\sigma}_\eta^2\right) \ \textit{left truncated at point} \ \left(1\text{-}z_i\hat{\boldsymbol{\beta}}\right) \qquad (11)$$

Step two requires computing, yet again for each i = 1,…, N, the following expression:

$$\left(\frac{1}{SUSTAINABILITY}\right)_i^* = z_i\hat{\boldsymbol{\beta}} + \eta_i \qquad (12)$$

In step three, maximum likelihood is used to estimate the truncated regression:

$$\left(\frac{1}{SUSTAINABILITY}\right)_i^* = z_i\boldsymbol{\beta} + \eta_i \qquad (13)$$

Jointly, steps one to three yield the following set of bootstrap estimates of $\boldsymbol{\beta}$ and σ_η:

$$B\left[\left(\hat{\boldsymbol{\beta}}^*,\hat{\sigma}_\eta^*\right)_b\right]_{b=1}^{L} \qquad (14)$$

Finally, values in B and the original estimates are used to construct estimated confidence intervals for $\boldsymbol{\beta}$ and σ_η.

Concerning the environmental features capable of influencing sustainability, we include frequently used structural variables such as age, farm size, level of education, specific professional training of farmers and membership of agricultural cooperatives.[5] Besides, the number of replications in the bootstrap procedure has been settled equal to 1,000. Table 5 displays the estimated parameters (column 1) and their confidence intervals (columns 2 to 5).

Table 5. Truncated regression. Bootstrapped confidence intervals. About here

The results obtained show different significant relationships. Before commenting on these relationships, and in order to interpret the signs of the estimated parameters in the truncated re-

[5] Other variables representing the percentage of income derived from farming or the percentage of farm surface area subjected to agri-environmental programs had to be omitted from bootstrapping because of their highly positive correlation with farm size.

gression correctly, let us recall here that what we are explaining is the inverse of the scores of global sustainability. That said, in the first place it can be verified that global sustainability rises as *farm size* increases. The fact that larger farms are more sustainable can be explained by: a) the existence of economies of scale in agricultural production, which makes for more efficient production (Álvarez and Arias, 2004; Karagiannis and Sarris, 2005) and, therefore, greater economic sustainability, b) the development of a more diversified range of crop-mixes (more stable work-load over the year) and the generation of sufficient income to permit the continuity of agricultural activity, allowing for greater social sustainability (Moreno-Pérez and Ortiz-Miranda, 2008), and c) higher generation of environmental benefits because of the more extensive productive techniques implemented (Burton and Walford, 2005; Cahill and Hill, 2005) and the higher degree of participation in agro-environmental programs (Atance and Barreiro, 2006; Vanslembrouck et al., 2002). Campos County is an area where agricultural censuses have shown a trend in the last twenty years towards increasing farm size, through farm amalgamation. Larger farms have become more diversified than smaller ones and have adopted some environmentally friendly cultivation techniques, such as conservation tillage and the abandonment of the moldboard plow (Moreno-Pérez and Muñoz, 2007). These techniques have improved the ability of the Campos agricultural system to protect the habitat of threatened bird species, such as the great bustard or *Otis tarda* (Palacín et al., 2003), with more than half of the world population living in Spain, and particularly in the cereal steppes of the Northern Plateau.

Second, global sustainability improves when the *level of specialized training in agriculture* is intermediate or advanced. The results show that farms of producers who have taken professional training courses or have got a university degree in agriculture are significantly more sustainable than those belonging to farmers who have not taken any specialist courses. Hence, the level of technical knowledge acquired should be analyzed as an indicator of the professionalism and specialization of farmers. We may therefore assume that as the level of technical knowledge of producers rises, they will be able to run their farms better, making them more profitable and eco-compatible (Phillips, 1994; Atance et al., 2006).

Third, the global sustainability composite indicator also increases when the farm-owner is a *member of an agricultural cooperative*. Here we can note the important role played by cooperatives in farm management, such as a) supplying inputs, b) commercializing harvests and c) providing technical advice. We may therefore assume that cooperative membership enables the farmer to improve both economic returns and environmental sustainability (optimization of the use of resources through applying external technical expertise), as shown in Valentinov (2007).

Fourth, it is also worth mentioning that sustainability decreases when the farmer has a higher *level of education*. This may seem to be counter-intuitive, although an explanation can be given. More educated people in Spanish rural areas, particularly those with university degrees in non-agricultural subjects, tend to be engaged in non-farming sectors, where revenues and social prestige are higher (Gómez-Benito et al., 1999). Particularly in the case of young rural women in Spain, it has been ascertained that access to higher education has directly led to abandonment of rural professional careers (Díaz, 2006). Within this general trend, there are still some educated people managing farms. However, we should not expect them to manage their own farms particularly well. There may be many reasons for this, of which the most important may be a lack of vocation or interest (they failed when trying to get a better job outside agricultural sector), or absenteeism (i.e., living at some distance from their farms in order to practice their main profession).

6. Conclusions and policy implications

Defining farming sustainability indicators at national or regional level is no longer appropriate when indicators are intended to yield information regarding the economic, social and environmental value of particular agricultural ecosystems. Therefore, the need to collect and use farm or local farming system-specific sustainability indicators has been widely recognized in the literature. *Data Envelopment Analysis* (*DEA*) is particularly well suited to determining relative levels of sustainability, because it can easily deal with multi-faceted attributes of decision making units (*DMUs*), operating an endogenous procedure to the computation of weights for each of these attributes. This technique has been widely used to aggregate indicators under a *benefit of the doubt* approach (Cherchye et al., 2007). Nevertheless, a conventional *DEA* model only establishes a dichotomy in the classification of *DMUs*, as belonging to the efficient or inefficient subsets, and does not allow for a complete ranking of them, because it fails to discriminate among units within each of these two sets. The main contribution of this paper is to show how a composite sustainability indicator can be built up to obtain a complete ranking of individual farms, by using endogenous common weights for a set of economic, social and environmental indicators. This type of aggregation is achieved through a two-step model that combines *DEA* and *Multi-Criteria Decision Making* (*MCDM*). Unlike other researchers (Despotis, 2002, 2005), we compute a non-radial *DEA* model in the first step of our analysis.

This *DEA-MCDM* model is applied to a sample of 163 farms in a rain-fed farming system on the Spanish Northern Plateau (Campos County), using 12 economic, social and environmental indicators. Our model perfectly ranks all farms, with no ties, and establishes a positive rank correlation between the economic and environmental facets of sustainability at farm level. In contrast, the social composite indicator is negatively correlated with the economic and environmental indicators. Moreover, we have aimed to determine, using bootstrapping techniques, some determinants of farm-level sustainability. Sustainability rises as farm size increases, as farmers are endowed with intermediate or advanced agricultural-specific levels of education and with membership in agricultural cooperatives. Surprisingly, sustainability decreases with a rise in farmers' general level of education, but we believe in this case that education acts as a proxy for part-time farming and absenteeism, given that farmers with non-agricultural university degrees are prone to pursuing non-agricultural careers, managing their farms with less professional care.

We believe that our findings offer substantial insight into the empirical assessment of sustainability at farm level, which is a topic full of practical consequences for policy-making. Policy-makers may need farm-level indicators for a variety of reasons, including the evaluation of competing green claims from the farming community and detecting the need for new policies to mitigate the environmental impacts of agricultural policies (Russillo and Pintér, 2009). Ideally, public policies intended to improve farmer's management of rural resources should adopt payments directly linked to the level of countryside services provided by landowners, which will vary according to different land uses. Therefore, a need arises '*to identify indicators that can proxy for bundles of countryside services and to link payments to them… [and] policy mechanisms will often be linked to agricultural input use, processes and activity levels rather than directly to the output of countryside services*' (Hodge, 2000: 269; 271). More specifically, a scoring system based on reliable sustainability composite indicators may reduce transaction costs in the implementation of governmental contracts with farmers. This is a topic of paramount importance as institutionalization of environmental contracting, or 'paid stewardship' for farmers, is one of the main directions taken by the European agri-environmental policy (Latacz-Lohmann and Hodge, 2003).

Furthermore, sustainability indicators can also help to monitor the environmental and social consequences of structural changes in farming and to design policy packages to boost sustainable rural development. This is particularly important in the Spanish cereal pseudo-steppes, like Campos County, where both intensification pressure, through irrigation, and abandonment of marginal lands would result in a loss of biological diversity (Suárez et al., 1997). Our empirical findings support the idea that farm size, farmers' membership of cooperatives and intermediate and advanced specialized agricultural education and training all exert a positive influence on farm sustainability. Therefore, policy-makers may wish to make use of structural agricultural policies to raise the sustainability of the analyzed dry land agricultural system by promoting farm-amalgamation and partnerships between farmers, and by expanding intermediate and advanced agricultural technical education and training. Larger farms are clearly more diversified than smaller ones in Campos County and apply more environmentally friendly cultivation techniques, membership in farm cooperatives positively influences management quality, and intermediate and advanced specialized agricultural education also help farmers to run more sustainable farms. Our paper therefore shows that computing composite sustainability indicators constitutes a first step in the design of policy measures aimed at boosting sustainability at farm and local-farming system level.

Acknowledgments

The authors would like to express their gratitude to the anonymous reviewers. The research was co-financed by the Spanish Ministry of Science and Innovation and FEDER through projects AGL2010-17560-C02-01 and 02/AGRI. Financial support from projects ECO2008-05908-C02-02/ECON (Spanish Ministry of Science and Innovation and FEDER) and PROMETEO 2009-098 (Generalidad Valenciana) is also acknowledged.

References

Adler, N., Golany, B., 2002. Including principal components weights to improve discrimination in data envelopment analysis. *J. Oper. Res. Soc.* 53(9), 985-991.

Adler, N., Friedman, L., Sinuany-Stern, Z., 2002. Review of ranking methods in the data envelopment analysis context. *Eur. J. Oper. Res.* 140(2), 49-265.

Álvarez, A., Arias, C., 2004. Technical efficiency and farm size: a conditional analysis. *Agr. Econ.* 30(3), 241-250.

Andersen, P., Petersen, N.C., 1993. A procedure for ranking efficient units in data envelopment analysis. *Manage. Sci.* 39(10), 1261-1264.

Andersen, E., Elbersen, B., Godeschalk, F., Verhoog, D., 2007. Farm management indicators and farm typologies as a basis for assessments in a changing policy environment. *J. Environ. Manag.* 82(3), 353-362.

André, F.J., Herrero, I., Riesgo, L., 2010. A modified DEA model to estimate the importance of objectives with an application to agricultural economics. *Omega* 38, 371-382.

Angulo-Meza, L., Estellita-Lins, M.P., 2002. Review of methods for increasing discrimination in data envelopment analysis. *Ann. Oper. Res.* 116(1-4), 225-242.

Atance, I., Barreiro, J., 2006. CAP MTR versus environmentally targeted agricultural policy in marginal arable areas: impact analysis combining simulation and survey data. *Agr. Econ.* 34(3), 303-313.

Atance, I., Gómez-Limón, J.A., Barreiro, J., 2006. El reto de la multifuncionalidad agraria: Oferta de bienes privados y públicos en el Sur de Palencia. *Revista Española de Estudios Agrosociales y Pesqueros* 210, 155-200.

Becker, B., 1997. *Sustainability Assessment: A Review of Values, Concepts, and Methodological Approaches*. Issues in Agriculture n. 10. Consultative Group on International Agricultural Research (CGIAR) – World Bank, Washington.

Böhringer, C., Jochem, P.E.P., 2007. Measuring the immeasurable – A survey of sustainability indices. *Ecol. Econ.* 63(1), 1-8.

Burton, R.J.F., Walford, N.S., 2005. Multiple succession on family farms in the South East of England: a counterbalance to agricultural concentration? *J. Rural Stud.* 21(3), 335-347.

Cahill, C., Hill, B., 2005. Policies affecting or facilitating resource adjustment in agriculture in the European Union, in: Blandford, D., Hill, B., eds., *Policy reform and agricultural adjust-*

ment in the agricultural sectors of developed countries. CABI Publishing, Wallingford, Oxon (UK).

Callens, I., Tyteca, D., 1999. Towards indicators of sustainable development for firms. A productive efficiency perspective. *Ecol*. *Econ*. 28(1), 41-53.

Charnes, A., Cooper, W.W., Rhodes, E., 1978. Measuring the efficiency of decision making units. *Eur*. *J*. *Oper*. *Res*. 2(6), 429-449.

Chen, Y-W., Larbani, M., Chang, Y-P., 2009. Multiobjective data envelopment analysis. *J*. *Oper*. *Res*. *Soc*. 60(11), 1556-1566.

Cherchye, L., Moesen, W., Rogge, N., van Puyenbroek, T., 2007. An introduction to 'benefit of the doubt' composite indicators. *Soc*. *Indic*. *Res*. 82(1), 111-145.

Cherchye, L., Moesen, W. Rogge, N., van Puyenbroeck, T., Saisana, M., Saltelli, A., Liska, R., Tarantola, S. 2008. Creating composite indicators with DEA and robustness analysis: the case of the Technology Achievement Index. *J*. *Oper*. *Res*. *Soc*., 59(2), 239-251.

Cohon, J.L., 1978. *Multiobjective programming and planning*. Academic Press, New York.

Cooper, W.W., Seiford, L.M., Tone, K., 2007. *Data envelopment analysis. A comprehensive text with models, applications, references and DEA-Solver software.* Springer-Bussiness Media – LCC, New York.

Despotis, D.K., 2002. Improving the discriminating power of DEA: focus on globally efficient units. *J*. *Oper*. *Res*. *Soc*. 53(3), 314-323.

Despotis, D.K., 2005. A reassessment of the human development index via data envelopment analysis. *J*. *Oper*. *Res*. *Soc*. 56(8), 969-980.

Díaz, C., 2006. Cambios generacionales en las estrategias de inserción sociolaboral de las jóvenes rurales. *Revista Española de Estudios Agrosociales y Pesqueros* 211, 307-338.

Diaz-Balteiro, L., Romero, C., 2004. In search of a natural systems sustainability index. *Ecol*. *Econ*. 49(3), 401-405.

Ebert, U., Welsch, H., 2004. Meaningful environmental indices: a social choice approach. *J. Environ. Econ. Manag.* 47(2), 270-283.

EEA (European Environment Agency), 2005. *Agriculture and environment in EU-15 – the IRENA indicator report*. European Environment Agency, Copenhagen.

Friedman, L., Sinuany-Stern, Z., 1997. Scaling units via the canonical correlation analysis and the data envelopment analysis. *Eur*. *J*. *Oper*. *Res*. 100(3), 629-637.

Golany, B., 1988. An interactive MOLP procedure for the extension of DEA to effectiveness analysis. *J*. *Oper*. *Res*. *Soc*. 39(8), 725-734.

Gómez-Benito, C., González, J.J., Sancho-Hazak, R., 1999. *Identidad y profesión en la agricultura familiar*. Centro de Investigaciones Sociológicas, Madrid.

Gómez-Limón, J.A., Sanchez-Fernandez, G., 2010. Empirical evaluation of agricultural sustainability using composite indicators. *Ecol*. *Econ*. 69(5), 1062-1075.

Hansen, J.W., 1996. Is agricultural sustainability a useful concept? *Agr. Sys*. 50(1), 117–143.

Hodge, I. 2000. Agri-environmental Relationships and the Choice of Policy Mechanism. *World Econ*. 23(2), 257-273.

Kao, C., Hung, H-T., 2005. Data envelopment analysis with common weights: the compromise solution approach. *J. Oper. Res. Soc.* 56(10), 1196-1203.

Jolliffe, I.T., 2002. *Principal components analysis*. Springer – Verlag, New York.

Joro, T., Korhonen, P., Wallenius, J., 1998. Structural comparison of data envelopment analysis and multiple objective linear programming. *Manage. Sci.* 44(7), 962-970.

Kallas, Z., Gómez-Limón, J.A., Barreiro-Hurlé, J., 2007. Decomposing the value of agricultural multifunctionality: Combining contingent valuation and the analytical hierarchy process. *J. Agr. Econ*. 58(2), 218-241.

Karagiannis, G., Sarris, A., 2005. Measuring and explaining scale efficiency with the parametric approach: the case of Greek tobacco growers. *Agr. Econ*. 33(3), 441-451.

Lam, K.F., 2010. In the determination of weight sets to compute cross-efficiency ratios in DEA. *J. Oper. Res. Soc*. 61(1), 134-143.

Latacz-Lohmann, U., Hodge, I., 2003. European agri-environmental policy for the 21st century. *Aust. J. Agr. Resour. Ec.*, 47(1), 123-139.

Li, X-B., Reeves, G.R., 1999. A multiple criteria approach to data envelopment analysis. *Eur. J. Oper. Res.* 115(3), 507-517.

Lovell, C.A.K., Pastor, J.T., Turner, J.A., 1995. Measuring macroeconomic performance in the OECD: A comparison of European and non-European countries. *Eur. J. Oper. Res.* 87(3), 507-518.

Moreno-Pérez, O.M., Muñoz, C., 2007. Aspectos territoriales del proceso de ajuste en las llanuras cerealistas castellano-leonesas, in: Arnalte, E., ed., *Políticas agrarias y ajuste estructural en la agricultura española.* Serie Estudios. Ministerio de Agricultura, Pesca y Alimentación, Madrid.

Moreno-Pérez, O.M., Ortiz-Miranda, D., 2008. Understanding structural adjustment in Spanish arable crop farms: policies, technology and multifunctionality. *Span. J. Agric. Res.* 6(2), 153-165.

Nahra, T.A., Mendez, D., Alexander, J.A., 2009. Employing super-efficiency analysis as an alternative to DEA: An application in outpatient substance abuse treatment. *Eur. J. Oper. Res.* 196(3), 1097-1106.

Nardo, M., Saisana, M., Saltelli, A., Tarantola, S., Hoffman, A., Giovannini, E., 2008. *Handbook on Constructing Composite Indicators: Methodology and User Guide*. Joint Research Center - OECD, Paris.

OECD (Organisation for Economic Cooperation and Development), 2001. *Environmental indicators for agriculture. Volume 3: Methods and results*. Organisation for Economic Cooperation and Development, Paris.

Pacini, C., Wossink, A., Giesen, G., Vazzana, C., Huirne, R., 2003. Evaluation of sustainability of organic, integrated and conventional farming systems: a farm and field-scale analysis". *Agr. Ecosys. Environ.* 95(1), 273-288.

Palacín, C., Alonso, J.C., Martín, C.A., et al., eds., *Atlas de las aves reproductoras de España*. D.G. Conservación de la Naturaleza – SEO, Madrid.

Pannell, D.J., Glenn, N.A., 2000. A framework for the economic evaluation and selection of sustainability indicators in agriculture. *Ecol. Econ*. 33(1), 135-149.

Picazo-Tadeo, A.J., García-Reche, A., 2007. What makes environmental performance differ between firms? Empirical evidence from the Spanish tile industry. *Environ. Plan. A* 39(9), 2232-2247.

Picazo-Tadeo, A.J., Reig-Martínez, E., Estruch, V., 2009, Farming efficiency and the survival of valuable agro-ecosystems: A case study of rice farming in European Mediterranean wetlands. *Open Environ. Sci*. 3, 69-78.

Phillips, J.M., 1994. Farmer education and farmer efficiency: A meta-analysis. *Econ. Devel. Cult. Change* 43(1), 149-165.

Podinovski, V.V., Thanassoulis, E., 2007. Improving discrimination in data envelopment analysis: some practical suggestions. *J. Prod. Anal.* 28(1), 117-126.

Raman, S., 2006. *Agricultural sustainability. Principles, processes and prospects*. Haworth Press, Binghamton, NY (USA).

Rao, N.H., Rogers, P.P., 2006. Assessment of agricultural sustainability. *Current Sci.* 91(4), 439-448.

Reig, E., Aznar, J., Estruch, V., 2010. A comparative analysis of the sustainability of rice cultivation technologies using the analytic network process. *Span. J. Agric. Res.* 8(2), 273-284.

Rigby, D., Woodhouse, P.H., Young, T., Burton, M., 2001. Constructing a farm level indicator of sustainable agricultural practice. *Ecol. Econ.* 39(3), 463-478.

Romero, C., 1996. Multicriteria decision analysis and environmental economics: An approximation. *Eur. J. Oper. Res.* 96(1), 81-89.

Russillo, A., Pintér, L., 2009. *Linking farm-level measurement systems to environmental sustainability outcomes: Challenges and ways forward*. International Institute for Sustainable Development, Winnipeg, Manitoba (Canada).

Saaty, T.L., 1980. *The analytic hierarchy process. Planning, priority setting, resource allocation.* McGraw-Hill, New York.

Saaty, T.L., 2001. *Decision making for leaders. The analytic hierarchy process for decisions in a complex world*. RWS Publications, Pittsburgh.

Sanchez-Fernandez, G., 2009. Análisis de sostenibilidad agraria mediante indicadores sintéticos: aplicación empírica para sistemas agrarios de Castilla y León. Ph.D. dissertation. Polytechnic University of Madrid, Madrid.

Sarkis, J., 2000. A comparative analysis of DEA as a discrete alternative multiple criteria decision tool. *Eur. J. Oper. Res.*, 123(3), 543-557.

Sauvenier, X., Valckz, J., van Cauwenbergh, N. et al., 2006. *Framework for assessing sustainability levels in Belgian agricultural systems – SAFE*. Final Report – SPSD II CP 28. Belgian Science Policy, Brussels.

Sexton, T.R., 1986. The methodology of data envelopment analysis, in: R.H. Silkman ed., *Measuring efficiency: An assessment of data envelopment analysis, new directions for program evaluation*. Jossey Bass, San Francisco, CA.

Sexton, T.T., Silkman, R.H., Hogan, A., 1986. Data envelopment analysis: Critique and extensions, in: Silkman, R.H., ed., *Measuring efficiency: An assessment of data envelopment analysis, new directions for program evaluation*. Jossey Bass, San Francisco, CA.

Shang, J., Sueyoshi, T., 1995. A unified framework for the selection of a flexible manufacturing system. *Eur. J. Oper. Res.* 85(2), 297-315.

Shephard, R.W., 1970. *Theory of cost and production functions*. Princeton University Press, Princeton.

Silva-Portela, M.C., Thanassoulis, E., 2006. Zero weights and non-zero slacks: Different solutions to the same problem. *Ann. Oper. Res.* 145(1), 129-147.

Simar, L., Wilson, P.W., 2007. Estimation and inference in two-stage, semi-parametric models of production processes. *J. Econometrics* 136(1), 31-64.

Sinuany-Stern, Z., Friedman, L., 1998. DEA and the discriminant analysis of ratios for ranking units. *Eur. J. Oper. Res.* 111(3), 470-478.

Stewart, T.J., 1996. Relationships between Data Envelopment Analysis and Multicriteria Decision Analysis. *J. Oper. Res. Soc.* 47(5), 654-665.

Suárez, F., Naveso, M.A., De Juana, E. ,1997. Farming in the drylands of Spain: birds of the pseudosteppes, in: Pain, D.J., Pienkowski, M.W., eds., *Farming and birds in Europe: The Common Agricultural Policy and its implications for bird conservation*. Academic Press, San Diego (USA).

Tone, K., 2002. A slack-based measure of super-efficiency in data envelopment analysis. *Eur. J. Oper. Res.* 143(1), 32-41.

Torgersen, A.M., Forsund, F.R., Kittelsen, S.A.C., 1996. Slack adjusted efficiency measures and ranking of efficient units. *J. Prod. Anal.* 7(4), 379-398.

UNDP (United Nations Development Program), 1990. *Human Development Report 1990*. Oxford University Press, New York.

Van der Werf, H.M.G., Petit, J., 2002. Evaluation of the environmental impact of agriculture at the farm level: a comparison and analysis of 12 indicator-based methods. *Agr. Ecosyst. Environ.* 93(1), 131-145.

Van Passel, S., Nevens, F., Mathijs, E., van Huylenbroeck, G., 2007. Measuring farm sustainability and explaining differences in sustainable efficiency. *Ecol. Econ*. 62(1), 149-161.

Van Wenum, J., Buys, J., Wossink, A., 1999. Nature quality indicators in agriculture, in: Brouwer, F., Crabtree, B., eds., *Environmental indicators and agricultural policy*, CABI Publishing, Wallingford.

Vanslembrouck, I., van Huylenbroeck, G., Verbeke, W., 2002. Determinants of the willingness of Belgian farmers to participate in agri-environmental measures. *J. Agric. Econ.* 51(3), 489-511.

Valentinov, V., 2007. Why are cooperatives important in agriculture? An organizational economics perspective. *J. Instit. Econ.* 3(1), 55-69.

WCED (World Commission on Environment and Development), 1987. *Our common future. The Brundtland Report.* Oxford University Press, Oxford.

Zeleny, M., 1974. A concept of compromise solutions and the method of the displaced ideal. *Comput. Oper. Res.*1, 479-496.

Zhou, P., Ang, B.W., Poh, K.L., 2007. A mathematical programming approach to constructing composite indicators. *Ecol. Econ.* 62(2), 291-297.

Zohrehbandian, M., Makui, A., Alinezhad, A., 2009. A compromise solution approach for finding common weights in DEA: an improvement to Kao and Hung's approach. *J. Oper. Res. Soc.* 61(4), 604-610.

Table 1

Sustainability indicators

DIMENSION	INDICATOR	DEFINITION
Economic	*Income of agricultural producers* (***INCOME***)	A farmer's net income may be estimated via a short-term analysis using the gross profit made by the farm (i.e. the difference between revenue -sales plus subsidies- and variable costs). This indicator is measured in €/ha.
	Contribution of agriculture to GDP (***CONTGDP***)	This indicator measures the wealth generated by agricultural activity for society as a whole and is equivalent to agricultural income minus subsidies: also measured in €/ha.
	Insured area (***INSUAREA***)	This indicator reflects the stability of farmers' profitability in the face of hypothetical losses due to unanticipated external events. It is quantified as the percentage of the farm area covered by a yield or income insurance.
Social	*Agricultural employment* (***AGRILABO***)	Agricultural employment is an indicator of the social implications of agriculture in the provision and distribution of income. This indicator is quantified in terms of hours of work per ha.
	Stability of work-force (***LABOSTAB***)	The stability of the workforce has been quantified as the percentage of labor demanded during critical periods (the three months with highest labor requirements) over the total use of labor. The higher the value of this indicator, the less stable the population in rural areas.
	Risk of abandonment of agricultural activity (***RISKABAN***)	The continuity of agricultural activity depends on a) the farmer's age and b) the profitability of the farm. This index was therefore constructed to range from a maximum of 1 (farmer less than 55 years old and above-average income) to a minimum of zero (farmer more than seventy years old and below-average income).
	Economic dependence on agricultural activity (***ECODEPEN***)	This indicator quantifies the total percentage of the farmer's income derived from agriculture. As his financial dependence on this activity increases, so does the stability of the rural population.
Environmental	*Soil cover* (***SOILCOV***)	The soil cover indicator represents the number of days of the year on which vegetation covers the soil. Taking into account that rain is the most relevant erosion agent in this region (rainfall concentrated both in autumn and throughout the year as scattered heavy rains), it is assumed that the higher the value of this indicator, the greater the protection it affords against erosion.
	Nitrogen balance (***NITROBAL***)	The difference between the nitrogen contained in the inputs (fertilizers) and outputs (crops) is measured in kg N/ha. The difference between these two values gives us the amount of nitrogen released to the environment every year.
	Pesticide risk (***PESTRISK***)	This indicator provides information about the overall toxicity released into the environment through the pesticides used in agricultural production. This toxicity has been estimated by adding the potential lethality of live organisms (measured through the Lethal Dose 50 or LD50 by oral administration in rats) of the different active matters used, measured in kg of rat per hectare. Thus, as the value of this indicator rises, the biocide effects of farming operations also increase.
	Energy balance (***ENERGBAL***)	The energy balance of a farm measured in terms of kcal/ha can be calculated employing the input/output focus. The higher the value of this indicator, the more sustainable the farm is from an environmental perspective.
	Agro-environmental subsidy areas (***AGROENV***)	Four different agro-environmental schemes are available for rain-fed lands: 1) environmental fallow, 2) ecological agriculture, 3) extensification scheme to protect flora and fauna and 4) scheme for sunflower. Farmers subscribing these contracts are subjected to the strictest environmental requirements of the agricultural code of practices. Hence, the percentage of the area of the farm devoted to programs of this type is a good indicator of the maintenance of biodiversity.

Table 2a

Sample description (continuous variables)

	CRUDE VALUES		NORMALIZED VALUES	
	Mean	*Standard deviation*	*Mean*	*Standard deviation*
INDICATORS OF SUSTAINABILITY				
Economic indicators				
Income of agricultural producers (*INCOME* in €/ha)	306.2	(110.1)	0.534	(0.156)
Contribution of agriculture to GDP (*CONTGDP* in €/ha)	126.3	(102.3)	0.466	(0.166)
Insured area (*INSUAREA* in %)	64.9	(42.7)	0.649	(0.427)
Social indicators				
Agricultural employment (*AGRILABO* in hours/ha)	7.6	(1.5)	0.280	(0.142)
Stability of work-force (*LABOSTAB* in %)	43.4	(10.4)	0.480	(0.213)
Risk of abandonment (*RISKABAN* in 0-1 scale)	0.137	(0.214)	0.195	(0.306)
Economic dependence (*ECODEPEN* in %)	86.4	(25.0)	0.864	(0.250)
Environmental indicators				
Soil cover (*SOILCOV* in %)	70.4	(7.7)	0.417	(0.152)
Nitrogen balance (*NITROBAL* in kg N/ha)	20.6	(27.5)	0.703	(0.110)
Pesticide risk (*PESTRISK* in kg rat/ha)	0.635	(0.573)	0.867	(0.120)
Energy balance (*ENERGBAL* in 10^3 kcal/ha)	6,116	(1,535)	0.350	(0.118)
Agro-environmental subsidy areas (*AGROENV* in %)	5.8	(9.2)	0.216	(0.341)
SOCIO-ECONOMIC VARIABLES				
Age (years)	44.4	(9.4)	--	
Land (hectares)	124.5	(119.0)	--	
Time dedicated to agricultural training (hours/year)	36.4	(30.8)	--	
NUMBER OF FARMS		163		

Table 2b

Sample description (categorical variables)

	Frequency
SOCIO-ECONOMIC VARIABLES	
Member of an agrarian organization	58.9%
Education	
School-leaving certificate	19.0%
Primary school	39.3%
Secondary school	34.4%
University	7.3%
Agricultural training	
None	12.3%
Basic: agricultural extension	74.2%
Medium: professional training	9.2%
Upper: university	4.3%
NUMBER OF FARMS	163

Table 3

Descriptive statistic for dimensional and global sustainability composite indicators

	Economic sustainability	*Social sustainability*	*Environmental sustainability*	*Global sustainability*
Average	0.359	0.250	0.377	0.561
St. deviation	0.119	0.071	0.102	0.098
Maximum	0.634	0.466	0.660	0.973
Minimum	0.000	0.119	0.075	0.373

Table 4

Spearman's rank correlation coefficients (significance levels between brackets[a])

	Economic sustainability	*Social sustainability*	*Environmental sustainability*	*Global sustainability*
Economic sustainability	1			
Social sustainability	-0.1997 (0.0106)	1		
Environmental sustainability	0.2498 (0.0013)	-0.1842 (0.0186)	1	
Global sustainability	0.5556 (0.0000)	0.1125 (0.1528)	0.7417 (0.0000)	1

[a] The null hypothesis is the hypothesis of independence.

Table 5

Truncated regression. Bootstrapped confidence intervals. The dependent variable is the inverse of global sustainability

	Estimated parameter	*99% confidence*		*95% confidence*	
		Lower bound	*Upper bound*	*Lower bound*	*Upper bound*
Age (years)	0.00238	-0.00387	0.00898	-0.00280	0.00797
Farm size (hectares)	-0.00039	-0.00084	0.00010	**-0.00079**	**-0.00002**
Education: primary school [a]	0.11575	-0.01970	0.27212	-0.00306	0.23912
Education: secondary school and university [a]	0.23444	**0.08344**	**0.39950**	**0.10496**	**0.37324**
Agricultural training: basic [b]	-0.02705	-0.21212	0.15883	-0.17969	0.12693
Agricultural training: medium and upper [b]	-0.27611	**-0.51382**	**-0.02571**	**-0.46749**	**-0.06160**
Time dedicated to agricultural training (hours/year)	-0.00110	-0.00273	0.00065	-0.00252	0.00044
Member of an agricultural cooperative	-0.11990	**-0.22707**	**-0.01185**	**-0.21202**	**-0.03285**
Constant	1.80040	1.38952	2.19430	1.47340	2.12490
Sigma	0.28741	0.25479	0.33715	0.26238	0.32780
Number of observations		163			

[a] The category omitted is school-leaving certificate.

[b] The category omitted is no specialized agricultural training.

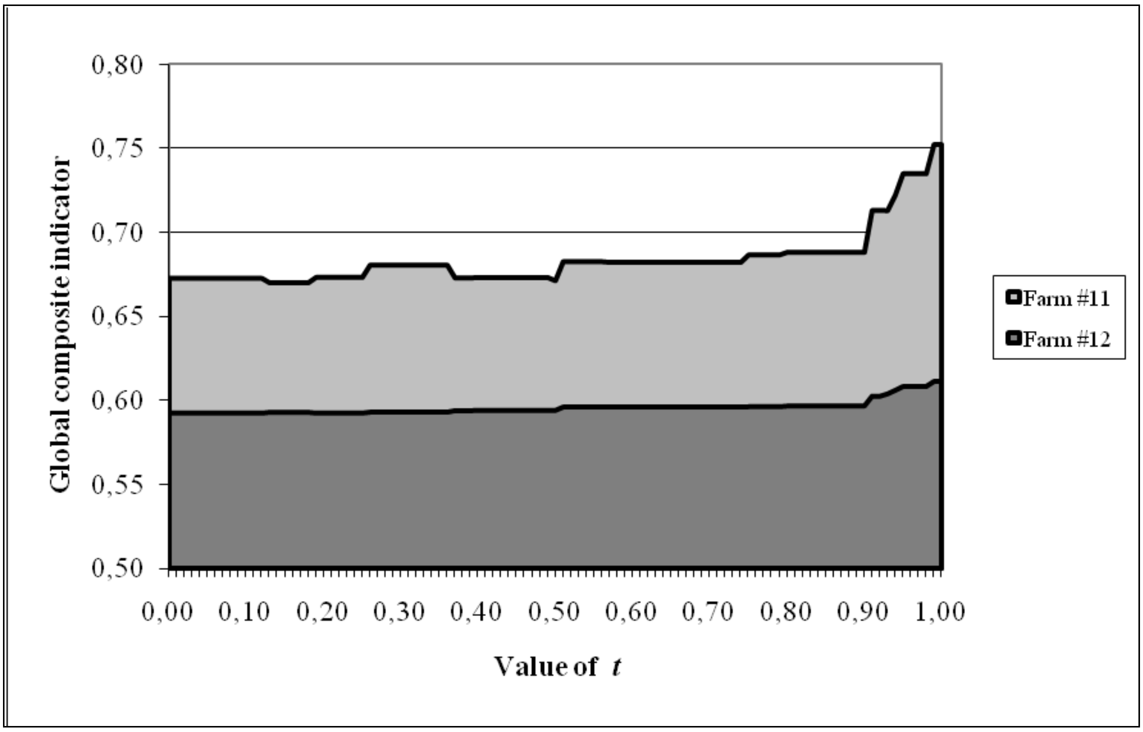

Fig. 1. Scores of the global sustainability composite indicator depending on the value of *t*.

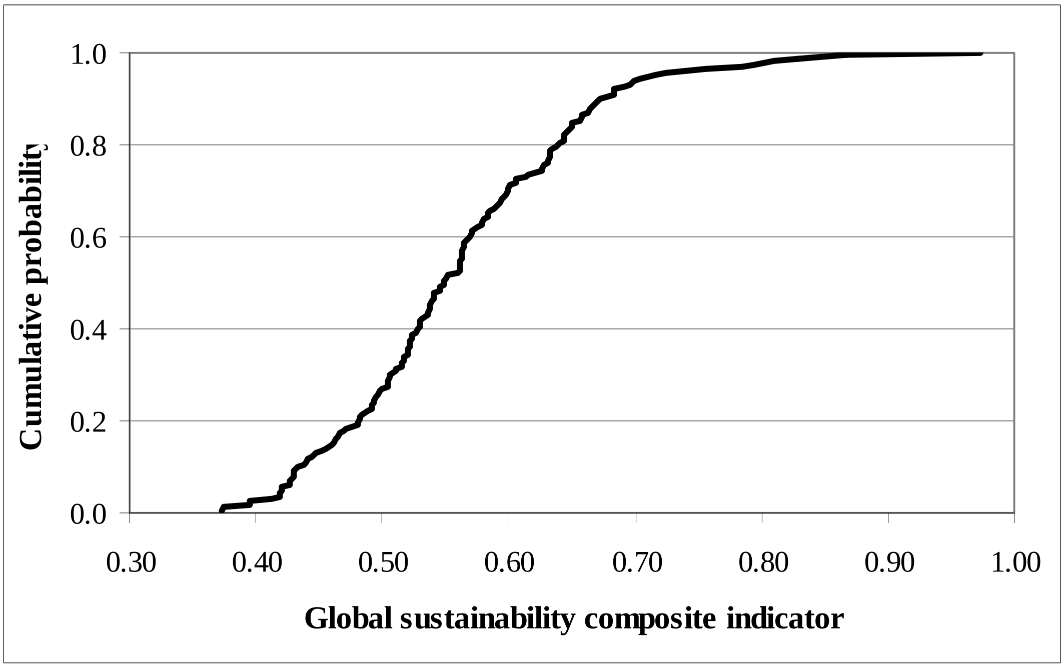

Fig. 2. Cumulative probability of the global sustainability composite indicator.