FACEPA

Farm Accountancy Cost Estimation and

Policy Analysis of European Agriculture

Methodology for analysing competitiveness, efficiency and economy of scale.

Use and applications of DEA

FACEPA Deliverable No. D5.1.3 – April 2009

INEA – Italy

Luca Cesaro

Sonia Marongiu

Filippo Arfini

Michele Donati

Maria Giacinta Capelli

The research leading to these results has received funding from the European Community's Seventh Framework Program (FP7/2007-2013) under *grant agreement* n° 212292.

Executive Summary

Efficiency and performance are a core concept of economics research. The relationship between them has been analyzed under many points of view, using different techniques and investigating the main determinants of efficiency. One of these is the non-parametric method of DEA, based on a finite sample of observed production units, which uses a linear programming method and does not need to estimate a pre-established functional form. It follows the Farrell approach (1957) and was proposed in 1978 by Charnes, Cooper and Rhodes. DEA constructs an efficient frontier using the best performing farm business of the sample. DEA being a deterministic tool, some problems can arise in the case of outliers and sampling variations. To solve them, the recent literature has shown that a statistical model can be defined to determine the statistical properties of non-parametric frontier estimators. Bootstrapping is one way to analyze the sensitivity of efficiency scores relative to the sampling variations of the estimated frontier.

The advantage of DEA is its flexibility and the possibility of using it for different farm types and scenario analysis. The possibility to calculate technical efficiency over time makes the analysis of farm efficiency changes over a period possible. This permits additional information to be acquired, for example, about the effect on efficiency from the application of an agricultural policy, the introduction of taxations or subsidies, transition reforms, a change in support scheme etc. Moreover, specific techniques allow comparisons to be made between different countries.

Another instrument that permits productivity change to be assessed is the Malmquist Index, based on the ratio of two distance functions. It is a measure of total factor productivity used in situations where prices do not exist or where they have little economic meaning.

One of the most important objectives of the FACEPA project is to implement the analytical approaches using the existing FADN database. It is a very important source of information for DEA because this method requires data at a farm level and these are harmonized for all European countries. The lack of price information makes an empirical analysis of price efficiency difficult, but the rich literature presented shows that the FADN database has been used in a wide range of cases, including the analysis of transition economies and the access of New Member States.

Contents

Abbreviations and Acronyms

AE	Allocative Efficiency
AWU	Annual Work Unit
CAP	Common Agricultural Policy
CEECs	Central and Eastern European Countries
COP	Cereals, Oilseeds and Proteinseeds
CRS	Constant Returns to Scale
DEA	Data Envelopment Analysis
DFA	Deterministic Frontier Analysis
EE	Economic Efficiency
EU	European Union
FACEPA	Farm Accountancy Cost Estimation and Policy Analysis of European Agriculture
FADN	Farm Accountancy Data Network
FDH	Free Disposal Hull
HC	Health Check
INEA	Istituto Nazionale di Economia Agraria (National Institute of Agricultural Economics)
MPI	Malmquist Productivity Index
SE	Scale Efficiency
SFA	Stochastic Frontier Analysis
SPS	Single Payment Scheme
TE	Technical Efficiency
TF	Type of Farming
VRS	Variable Returns to Scale

List of Figures and Tables

Introduction

This paper is a part of FACEPA WP5 whose aim is to analyze the applications and extensions of the Cost of Production model to assess farm performance. More specifically, this part of the project will analyze

- the relationship between production cost structure and farm performance
- the efficiency and competitiveness of the crop and dairy sectors in the participant countries
- the differences between economic performance of individual and corporate farms in the new Member States
- the impact of CAP reform on farm economic performance
- agriculture's competitiveness in the European countries under different support schemes

All these subjects will be discussed in 4 deliverables, the contents of which will be as follows:

- 5.1 Methodology for analyzing competitiveness, efficiency and economy of scale
- 5.2 Allocative and technical efficiency change and technological progress in farm economic performance
- 5.3 Performance of crop and dairy farms
- 5.4 Assessment of the impact of CAP reforms on farm economic performance

Deliverable 5.1 focuses on the methodologies to analyze farm competitiveness and efficiency and constitutes the introduction for the other deliverables, which will apply the concept to case studies and illustrate some scenario analyses in the CAP Reform context.

Production frontier analysis is one of the tools available for this kind of study and to compare different farm production processes. The idea to define an efficient frontier function against which to measure the current performance of productive units was introduced by Farrell in 1957. Since then, different techniques have been used to calculate or estimate the efficient frontiers.

These techniques are classified into two groups:

- Parametric methods: based on the econometric estimation of the cost of production function. This requires the definition *a priori* of the functional form of the efficient frontier (for example, using Deterministic Frontier Analysis – DFA or Stochastic Frontier Analysis – SFA)
- Non-parametric methods: based on mathematical programming techniques. In this case the relative efficiency of similar farms is determined using linear programming techniques (for example Data Envelopment Analysis – DEA or Free Disposal Hull –

FDH). No functional form is pre-established but one is calculated from the sample observations in an empirical way.

In this study the non-parametric method of Data Envelopment Analysis (DEA) is analyzed, paying attention to its origins, developments and application. A rich literature on the application of DEA using the FADN database allows the most common use of this methodology to be understood.

The first chapter explains the different concepts of efficiency, as presented by Farrell in the first application of the frontier production function approach. The possibility of decomposing the overall efficiency into its components (technical and allocative) permits a better understanding of the necessary improvement to reach the optimal allocation on the frontier. The analysis can be made following two schemes (input and output).

The second chapter describes the non-parametric method of DEA that, starting from Farrell's approach, also permits the scale effects on efficiency to be taken into account. In fact, in the case of variable returns to scale, the optimal behaviour could depend on increasing or decreasing returns to scale. For constant return to scale, the DEA model assumes that the production units are operating at an optimal scale. Another important task in the efficiency analysis is to identify the common features in the most efficient farms in order to estimate the impact of variables under (or not) the farmer's control. The chapter explains this important aspect together with an illustration of further developments of DEA methodologies that use statistical models to determine statistical properties of non-parametric frontier estimators (the bootstrapping method is the most important).

In the third chapter, the Malmquist Productivity Index approach is illustrated. It is a quantity approach methodology used to measure productivity changes in a farm's economic performance over time. It also has useful properties for empirical work because it can be applied in situations where prices do not exist.

The last chapter reviews the literature on DEA application. One of the most important purposes of the FACEPA research project is to implement the different analytical approaches using existing FADN databases at national and European level. So, all the work presented started from FADN information to analyze efficiency. The chapter is divided into three sections to take into account the different objectives of the papers (determinants of efficiency, comparison among different farm types and bootstrapping method).

1. Different concepts of efficiency

In the immediate post-war years there was a general interest in growth, efficiency and productivity. The first theories and papers appeared in 1957, with the Solow (macroeconomic approach) and Farrell (microeconomic approach) studies. In particular, Farrell (following the works of Debreu and Koopmans, 1951) involves new insights into two important issues: how to define efficiency and productivity and how to calculate the benchmark technology and efficiency measures.

In Farrell's approach, the measurement of economic efficiency is linked to the use of a *frontier production function*, in opposition to the notion of *average performance* underlying most of the econometric literature on the production function up to the time of Farrell contribution. Farrell's efficiency measures are completely data-based, so no specific functional form needs to be predefined. His work focused on the following points:

- efficiency measures are based on radial uniform contractions or expansions from inefficient observations (observed) to the frontier (unobserved)
- the production frontier is specified as the most pessimistic piecewise linear envelopment data (the function being as close as possible to the observations)
- the frontier is calculated solving a system of linear equations, obeying the two conditions on the unit isoquant (slope not positive and no observed point lies between it and the origin; input-oriented approach).

Efficiency and productivity are core concepts of economics and Farrell introduced a method to decompose the economic (overall) efficiency of a production unit into its technical and allocative components.

Technical Efficiency (TE): refers to the achievement of the maximum potential output from given amounts of inputs, taking into account physical production relationships. It can be measured within two main frameworks: input and output-oriented. In an input-oriented framework, technical efficiency gives the potential input reduction that a farm could apply without reducing its output level. In an output-oriented framework, technical efficiency gives information about the potential output increase that a farm could implement without increasing its use of inputs. In the case of constant returns to scale, both orientations give close results. On the contrary, in the case of variable returns to scale (increasing or decreasing) an additional component, scale efficiency, must be take into account in the calculation of technical efficiency.

Allocative Efficiency (AE): measures the distance between the farm and the point of maximum profitability, given market prices of inputs and outputs. In other words, the allocative efficiency shows whether the use of different proportions of production factors guarantees the attainment of maximum production with a particular market price.

Economic Efficiency (EE): is the product of technical and allocative efficiency (overall efficiency). It can be interpreted as the potential reduction in production costs (cost efficiency) or the potential increase in revenue (revenue efficiency) that a farm could apply in order to operate at the point of technical and allocative efficiency. Economic efficiency enables conclusions to be drawn on whether the farm operates at optimal or suboptimal size.

As previously stated, technical efficiency (and, consequently, overall efficiency) can be measured in different ways, depending on the approach used. In particular, the choice depends on whether farms may be constrained in their input reduction or output expansions. In both cases, the measurement of the adjustment necessary to attain technical efficiency must be limited to economically efficient points and this means that it is necessary to take a given price structure into account.

Farrell's efficiency measures followed an input-oriented scheme, whereas an analysis of output-oriented efficiency was introduced in other works (*Färe, Grosskopf, Lovell, 1985*). As previously stated, under constant returns to scale the two measures are equivalent.

The following figure shows the concept of efficiency in both frameworks.

Figure 1.A illustrates the input-oriented scheme:

- ii' is the production frontier represented by an isoquant ii' that captures the minimum combination of inputs needed to produce a unit of output
- X_1 and X_2 are the two inputs used to obtain one output
- $p_x p'_x$ is the isocost line, whose slope is the ratio of input prices ($-p_{x1}/p_{x2}$)

Every combination of inputs along the isoquant is considered technically efficient while any point above and to the right of it is technically inefficient. In other words, units that are technically efficient will be located at the frontier, while those that are not appear below the frontier since they obtain less output than technically possible. The technical efficiency measure can be estimated as the relationship between the obtained output and what would be attained if the unit were located at the frontier.

Point P is the observed farm that uses two input quantities to obtain the output. The isoquant ii' represents the various combinations of the two factors that a perfectly efficient farm might use to produce the same output. So, the farm Q represents an efficient farm that produces the same output as P but using a fraction OQ/OP of each factor. This is the technical efficiency (TE) of the farm P.

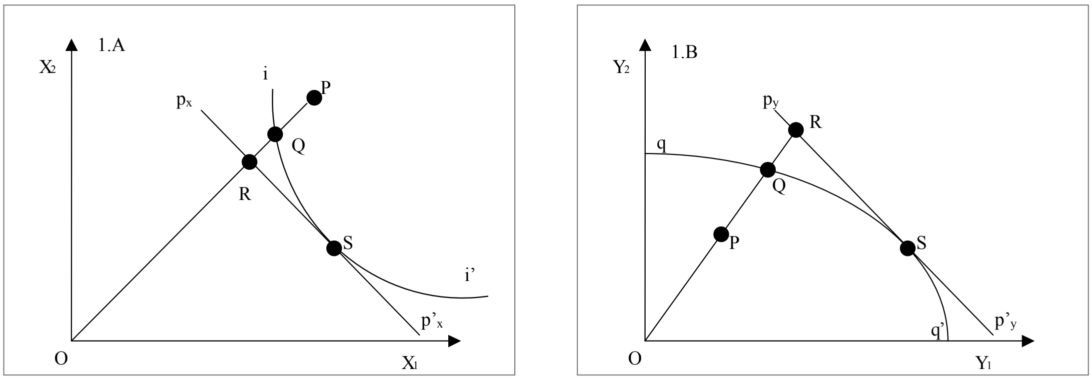
If information on market price is known and cost minimization is assumed in such a way that the input price ratio is reflected by the slope of the isocost line, allocative efficiency can be derived from the unit isoquant. S is the optimal method of production and, as Q, this point represent 100% of technical efficiency. However, the cost of production at Q will be only a fraction, OR/OQ of those at S. This is the measure of allocative efficiency (AE) .

The full economic efficiency (EE) is achieved for the farm operating at the tangency point between the isoquant and the price line (farm S).

Farm economic efficiency score for P is given by the ratio OR/OP, so by the product for technical and allocative efficiency.

$$EE = TE*AE = OQ/OP*OR/OQ = OR/OP$$

***Figure 1*: Economic efficiency in the input-oriented (1.A) and output-oriented (1.B) framework**



The same representation can be made using the output-oriented approach (figure 1.B).

- qq' is a production possibility curve for a farm
- Y_1 and Y_2 are the two outputs produced with one input
- p_y p'_y is the output price line, whose slope is the ratio of output prices ($-p_{y1}/p_{y2}$)

Farms that lie on the curve (such as Q and S) are fully efficient, while farm P is inefficient.

Considering farm P, the technical efficiency is found with the radial output ratio line, so that P is projected on the frontier, exactly in Q. The distance PQ represents the proportional amount by which both outputs could be increased and the ratio OP/OQ gives the technical efficiency score of farm P.

Allocative efficiency can be calculated with the help of the output price line (pp'), projecting the technically efficient farms on this line: the allocative efficiency for farm P (and also Q) is OQ/OR.

Economic efficiency is achieved for the farm operating at the tangency point between the production possibility curve and the price line (farm S).

As in the previous case, the farm economic efficiency score for P is given by the ratio OP/OR, so by the product for technical and allocative efficiency.

All the efficiency scores are included between 0 and 1: 1 indicates full efficiency, with lower scores indicating lower efficiency.
It is important to highlight that the correspondence between radial technological adjustment and economic efficiency disappears when technology is non-homothetic. In this case economically efficient adjustment would follow a non-linear expansion path and a radial adjustment can provide a biased estimate of the economically efficient adjustment. A similar situation appears, for example, when the set of technical possibilities is constrained by the presence of quasi-fixed production factors. The implications of this problem were recognized by Farrell in his work, though he maintained a focus on radial technical adjustments.
An important advantage of the efficiency analysis is that it takes a comprehensive approach to the farms, taking all inputs and outputs into account simultaneously. Considering farm performance in terms of economic, technical and allocative input and output efficiency, it is possible to study a part of the profit maximization process.

2. Methodologies to measure farm performance and efficiency: the non-parametric methods of DEA

Starting from Farrell's seminal measures for technical and allocative efficiency, different techniques have gradually been developed to calculate and estimate the efficient frontier against which to measure farm performance and efficiency. These techniques can be classified in different ways. One of them distinguishes between parametric and non-parametric approaches. In the parametric approach the functional form of the efficient frontier is imposed a priori while in the non-parametric scheme the frontier is calculated on the basis of the sample observations.

Among the parametric methods, the most important one is the Stochastic Frontier Analysis (SFA) proposed by *Aigner et al. (1977)* and *Meeusen and Van den Broeck (1977)* that estimates the parameters of a production function using econometric methods and, in particular, introducing two error terms, one for noise and one for inefficiency:

$$lnY = f(X) + \nu - \mu$$

where Y is the observed output, f(X) is the production function using factors X, ν is a standard error term and μ is a non-negative random term, with *exp(-μ)* representing the technical efficiency. In this case, the determinants of technical efficiency are estimated simultaneously with the production function.

As concerns the non-parametric methods, the most important one is the Data Envelopment Analysis (DEA), proposed by *Charnes et al. (1978)*. DEA uses linear programming to construct an efficient frontier using the best performing farm of the sample. The technique identifies efficient production units, which belong to the frontier, and inefficient ones, which remain below it. All the deviations from the frontier are attributed to inefficiency and not to measurement errors as in the parametric method. As previously illustrated, a farm on the frontier has an efficiency score of 1, while far from the frontier the efficiency score is lowest. The method does not require any distributional or specification assumption. A linear programming is needed to solve a maximization program under constraints. Technical efficiency is calculated under the assumption of constant returns to scale while scale efficiency can be obtained by the residual between efficiency under constant returns and efficiency under variable returns to scale. A key inspiration for the application of this method has been the theory of Farrell (1957).

So, generally speaking, the main differences between these methodologies are that the former assumes a functional form for the frontier (parametric), it is stochastic and uses econometric methods, while DEA does not assume a functional form for the frontier (non-parametric), it is deterministic and involves mathematical programming. Its deterministic nature is a disadvantage because DEA does not distinguish between technical inefficiency and statistical noise effects. Deterministic models, in fact, assume that all the deviations from the frontier are under the control of the agent, but there are some circumstances that cannot be controlled by the agent but can determine sub-optimal performances. For example: uncertainty, socio-economic and demographic factors, environment, etc. These factors should not be considered as technical efficiency but the deterministic model does so, contrary to the stochastic procedures that introduce a random error in the specification of the frontier model.

Moreover DEA results are very sensitive to outliers and to the sampling variation.

Another difference is that, with DEA, multiple outputs and inputs can be considered simultaneously and they can thus be quantified using different units of measurement.

There are different contributions in the literature that compare the stochastic frontier and DEA approach. Some studies obtained similar conclusions from both methodologies while in others the absolute level and relative ranking of farm efficiency were influenced by the method used. *Bauer et al. (1998)* proposed a set of conditions to measure how far the different approaches are mutually consistent:

1. the efficiency scores obtained by the different methods should have comparable means, standard deviations and other distributional properties
2. the methods should rank the farms in approximately the same order
3. the methods should identify largely the same farms as "best practice" and as "worst practice" (i.e. the efficient and inefficient farms).

2.1 The development of DEA analysis

The Farrell's approach to frontier estimation was reformulated by *Charnes, Cooper and Rhodes (1978)* as a mathematical programming model using the Data Envelopment Analysis (DEA) defined as

> *a mathematical programming model applied to the observational data that provides a new way of obtaining empirical estimates of extremal relations – such as the production functions and/or efficient production possibility surfaces that are a cornerstone of modern economics.*

This methodology was used initially to measuring the relative efficiency of a set of Decision-Making Units (DMUs). The aim of their model, known as the CCR model, was to define a frontier envelopment surface for all sample observations. This surface is determined by the efficient DMUs that lie on it, while the others can be considered as inefficient and an individual inefficiency score will be calculated for each one.

The CCR model has an input orientation and assumed constant returns to scale (CRS). Subsequent papers have considered alternative sets of assumptions: *Banker, Charnes and Cooper (1984)* proposed a model (BCC model) considering variable returns to scale (VRS).

In the CRS model the measure of global technical efficiency may be obtained by comparing large scale units with small scale units and vice versa. The VRS model allows for variations in returns to scale: in this case an additional constraint is necessary that ensures the evaluation of pure technical efficiency regardless of issues of scale.

The analysis presented in the next section is developed following an input oriented scheme, but the DEA model can also be output oriented or un-oriented. As previously stated, input oriented models try to maximize the proportional decrease in input variables, while output oriented ones will maximize the proportional increase in the output vector. The choice of one model or the other is based on the characteristics of the dataset analyzed.

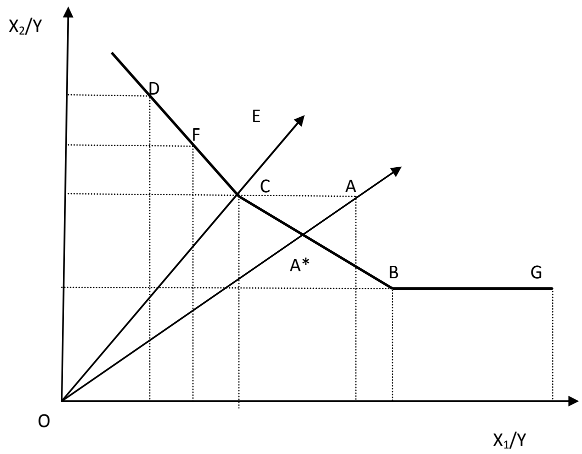
After the first important applications of DEA, there have been a large number of studies and analyses that have extended and applied the DEA methodology in different areas of agricultural economics. Moreover, DEA has also been used in other sectors, especially in the management science literature and in the service industries where there are multiple outputs (such as banking, health, telecommunications and electricity distribution).

The following sections give details and describe differences between the CRS and VRS models.

2.1.1 The CCR model (Constant Returns to Scale)

As previously stated, the efficiency measure of Farrell has been applied by *Charnes, Cooper and Rhodes (1978)* using DEA methodology. Farrell's method consisted of projecting each observed unit onto an efficient unit isoquant. The CCR model generalizes Farrell's approach, considering multiple outputs and reformulating the calculation of individual efficiency measures, solving a linear programming problem for each production unit. In this model, Constant Return to Scale and input orientation are assumed: this means that all the units under analysis are performing at an optimal scale. The CCR model can be interpreted through a simple example (Figure 2).

Figure 2: **The CCR model**



A, B, C, D, E, G are 6 DMUs that used different amounts of the two inputs (X_1 and X_2) to produce various amounts of a single output Y. The line DG represents the frontier unit isoquant derived by DEA techniques from data on the population of the DMUs. DEA method calculates the efficient frontier by finding the segments DC, CB and BG that envelope all the DMUs performances. So, the frontier is not a proper isoquant, but a linear approximation in which the observations at vertices (D, C, B, G) represent real DMUs, while the units between them (F, A*) are hypothetical units calculated as weighted averages of inputs. They are combinations of the real units.

How is the level of inefficiency of each unit determined?

The technical efficiency of A would be represented by the ratio OA*/OA, where A* is a linear combination of referents B and C that utilizes the inputs in the same proportions as A, since A and A* lie on the same ray.

The technical efficiency of E could be directly measured by comparison with C, situated on the efficient isoquant. The ratio OC/OE determines the technical efficiency of E.

As concerns G, although it is situated on the efficient frontier, it cannot be considered technically efficient because it uses the same amount of input X_2 as B but more input X_1 to produce the same level of output.

Unlike stochastic frontier techniques, the individual technical efficiency scores are calculated and not estimated.

Given a set of N farms, the data relate to K inputs and M outputs on each of them. For the *i-th* farm, these are represented by the vectors x_j and y_j. The data of all N farms are represented by $KX \times N$ input matrix (X) and by $MY \times N$ output matrix (Y). The purpose of DEA is to construct a non-parametric envelopment frontier over the data points such that all observed points lie on or below the production frontier.

The DEA model involves optimizing a scoring function (H), defined as the ratio of the weighted sum of outputs and the weighted sum of inputs subject to the constraints that the similar ratios for every farm be less or equal to one, implying that efficient units will have a score of one.

For each i-th farm, the linear problem is the following:

$$\max_{u,v} H = (\mathrm{u}'y_i/\mathrm{v}'x_i)$$

$$st \qquad (\mathrm{u}'y_j/\mathrm{v}'x_j) \leq 1, \qquad j = 1,2,\dots,N$$

$$\mathrm{u,v} \geq 0$$

where $\mathrm{u}'y_i/\mathrm{v}'x_i$, is the scoring function (u is an $M \times 1$ vector of output weights and v is an $K \times 1$ vector of input weights). The goal is to find values for u and v that maximise the efficiency score of the i-th farm subject to the constraint that all the efficiency measures must be less than or equal to one. This ratio formulation ensures that 0<Max H<1: a unit will be efficient if and only if this ratio equals unity otherwise it is considered as relatively inefficient.

The ratio formulation of the model has an infinite number of solutions (if u and v are solutions, then αu and αv are solutions), so to avoid this problem it is necessary to impose the constraint:

$$vx_i = 1$$

The maximization then becomes

$$\max_{\mu,v} (\mu' y_i)$$

$$st \qquad v'x_i = 1$$

$$\mu' y_j - v'x_j \leq 1, \qquad j = 1,2,\dots,N$$

$$\mu, v \geq 0$$

This transformation of u and v into μ and v, is identified with a *multiplier* form of the DEA linear programming problem.

Introducing the duality in linear programming, one can derive an equivalent *envelopment* form of this problem, the final objective of which is to determine the linear combination of referents that for each farm minimizes the value of θ:

$$\mathrm{TE_{crs}} = \min_{\theta,\lambda} \theta$$

$$st \qquad -y_i + Y\lambda \geq 0 \quad (1)$$

$$\theta x_i - X\lambda \geq 0 \quad (2)$$

$$\lambda \geq 0$$

where θ is a scalar that represents the minimum level to which the use of inputs can be reduced without altering the output level. So, the scalar θ provides the value of the global technical efficiency score for the i-th farm. The solution of this linear problem reports the group that for each farm analyzed yields at least the same level of output (first constraints) but consuming just a proportion (θ) of each of the inputs used by the farm (second constraints).

Indeed, as in Farrell's definition (1957), θ will satisfy the condition of less than or equal to 1: if it is equal to one, the farm is considered technically efficient (a point on the frontier). It also means that the use of all inputs cannot be reduced at the same time, although a variation in the use of one of them may improve efficiency. If the index is less than one there is some degree of technical inefficiency.

λ is a $N \times 1$ vector of constants that represents the weights to be used as multipliers for the input levels of each unit not located at the frontier. It indicates the input levels that an inefficient unit should aim at in order to achieve efficiency.

It is important to underline that a value of θ must be found for each farm, because the linear programming problem must be solved N times, one for each farm in the sample.

2.1.2 The BCC model (Variable Returns to Scale)

As previously stated, the CRS DEA model assumes that the DMUs are operating at an optimal scale. This model permits a measure of global technical efficiency to be obtained without variations in returns to scale. In the real world, however, this optimal behaviour is often precluded by some factors such as imperfect competition, constraints, finance, etc. To take this circumstance into account, *Banker, Charnes and Cooper (1984)* have extended DEA to the case of variable returns to scale (VRS).

This model distinguishes between pure technical efficiency and scale efficiency (SE), identifying if increasing, decreasing or constant returns to scale are present.

As a consequence, the assumptions of CRS linear have to change by adding further convexity constraints

$$N1'\lambda = 1$$

Hence, the envelopment form of the input oriented VRS DEA model is specified as

$$\mathrm{TE}_{\mathrm{vrs}} = \min_{\theta,\lambda} \theta$$

$$st \qquad -y_i + Y\lambda \geq 0$$

$$\theta x_i - X\lambda \geq 0$$

$$N1'\lambda = 1$$

$$\lambda \geq 0$$

where $N1$ is a $N \times 1$ vector of one.

θ is the input technical efficiency score under VRS, having a value $0 \leq \theta \leq 1$. As in the previous case, if the θ value is equal to one, the farm is on the frontier; the vector λ is an $N \times 1$ vector of weights that defines the linear combinations of the peers of the i-th farm.

Because the VRS DEA model is more flexible and envelops the data in a tighter way than the CRS DEA model, the VRS technical efficiency score is equal to or greater than the CRS or overall technical efficiency score. This relationship can be used to measure the scale efficiency of the farms:

$$SE = \frac{TE_{crs}}{TE_{vrs}}$$

SE = 1 implies scale efficiency or CRS while SE < 1 indicates scale inefficiency that can be due to the existence of either increasing or decreasing returns to scale. As a consequence, some units that are VRS efficient can be inefficient under the CRS scheme because their size deviates from the optimal scale.

A weakness in this procedure is that it cannot provide an indicator of whether the DMU is operating under increasing or decreasing returns to scale. This may be determined by calculating an additional DEA equation with non-increasing returns to scale (NIRS) imposed. The previous VRS DEA model may be changed replacing the $N1'\lambda = 1$ restrictions with $N1'\lambda \leq 1$ and another envelopment surface will permit to distinguish between different scales in the production structure. In particular:

- if $TE_{nrs} = TE_{vrs} \neq TE_{crs}$ then the units are producing at decreasing returns to scale
- if $TE_{nrs} \neq TE_{vrs} = TE_{crs}$ then the units are producing at increasing returns to scale
- if $TE_{nrs} = TE_{vrs} = TE_{crs}$ production is characterized by constant returns to scale

2.1.3 Further extensions to CCR and BCC models

The models described so far are based on a continuous definition of input and output variables. *Charnes, Cooper and Rhodes (1981)* developed further strategies to apply the DEA method to discrete level variables. In these cases, DEA is solved separately for each category of observations, assessing any difference in the mean efficiency of all the sub-samples considered.

DEA is also evolved to treat variables that are not under the DMU's direct control (non-discretionary variables). In this case it is possible to isolate these variables in order to analyze their effect on the DMU performance[1].

As previously stated, DEA is a very flexible methodology because it does not impose the weights of inputs and outputs: they are calculated solving a mathematical programming problem and it is not necessary to know them a priori. This can generate high or low multipliers in relation to the economic theory that support the specification of the model. To avoid this, several approaches have been developed. The most relevant ones are the Assurance Region method (AR) and the Cone-Ratio method (CR). The AR approach (developed by *Thompson, Singleton, Thrall and Smith, 1986*) is used in the case of existence of large differences in input

[1] *Banker and Morey (1986)* adapted DEA to allow an analysis of efficiency on the basis of exogenous and non-exogenous fixed inputs and outputs.

and output weights among the different DMUs. It imposes a constraint on the relative magnitude for some inputs and outputs, defining the region of weights to some special area consistent with the underlying economic theory. The CR approach (developed by *Charnes, Cooper, Wei and Huang, 1989*) is an extension of CCR model that constrains the multipliers to belong to closed cones.

Other extensions to the CCR model are the Multiplicative models that introduce multiplicative measures of relative efficiency through the use of multiplicative envelopment surfaces (*Charnes, Cooper and Seiford, 1983*). There are also models that take into account the measurement of allocative efficiency on the basis of price information, assuming a behavioural objective (cost minimization or revenue/profit maximization), using the Malmquist index approach.

DEA is a non-parametric technique that has a deterministic nature: the literature describes it as a non-statistical method. Nevertheless, recent literature has shown that it is possible to define a statistical model to determine the statistical properties of non-parametric frontier estimators. *Grosskopf (1996)* provided a good survey of statistical inference in non-parametric linear programming frontier models, also analyzing the asymptotic properties of the estimators. This type of result shows some limitations when used in conjunction with small samples. Moreover, extra noise is introduced when estimates of the unknown parameters of the limiting distributions are used to construct estimates of confidence intervals. To solve these problems, bootstrapping techniques have been introduced. The bootstrap technique permits the sensitivity of efficiency scores relative to the sampling variation of the frontier to be analyzed, avoiding the problems of asymptotic sampling distributions. DEA results, in fact, may be affected by sampling variation in the sense that distances to the frontier are underestimated if the best performers in the population are not included in the sample. To account for this *Simar and Wilson (1998, 2000)* proposed a bootstrapping method allowing the construction of confidence intervals for DEA efficiency scores which relies on smoothing the empirical distribution. This technique consists of a simulation of a true sampling distribution by mimicking a data generating process, using in this case the outputs from DEA. In this way, a new dataset is created and DEA is re-estimated using this dataset. Repeating the process many times allows a good approximation to be achieved of the true distribution of the sampling.

Generally speaking, statistical inference based on a non-parametric frontier approach to the measurement of economic efficiency is available either by using asymptotic method or by using bootstrap. A problem still remains concerning the high sensitivity of DEA to extreme values and outliers and also the way for allowing stochastic noise to be considered in a non-parametric frontier framework.

2.2 The effects of exogenous variables in the analysis of technical efficiency

As stated in the previous section, the non-parametric method of DEA permits the technical efficiency score to be estimated for a sample of farms. A second, important task in this kind of analysis is the identification of the common features of the most efficient farms in order to estimate the impact on efficiency of variables under (or not) the farmer's control.

The literature refers to two different approaches to account for the effects of these variables. The first is a one-stage procedure that directly includes the exogenous variables in the estimation of the efficiency score. The second is a two-stage procedure that first estimates the efficiency using inputs and outputs and then analyzes the effect of exogenous variables by means of regression analysis or analysis of variance. This second approach is more useful for political analyses because it provides an easy and clear evaluation of the relations between efficiency and its determinants. Testing which of them are significant and which are positive or negative, it is possible to highlight which aspects of the farms should be targeted in order to improve farm efficiency.

In agricultural studies, the variables usually take into account to explain the farm efficiency level are the location and size, the farmer's age, qualifications, experience and specialization, the use of services, the combination of inputs. So, together with the structural and socio-economic variables, other variables representing the quality of the production factors are often included in this analysis. The agricultural literature refers to some difficulties in collecting this kind of information and, in many cases, additional surveys may be necessary (for example questionnaires and specific surveys).

To regress the single variables (or group of variables) in the two-stage process, the Ordinary Least Squares (OLS) method is commonly used.

In addition to the previously described variables, another group of factors that can affect farm performance are the long-term aspects connected with the farm strategic management. Many authors have found relationships between strategy and farm performances. Although the literature on strategic management is about large organizations, it is also applicable to farms. Generally speaking, there are different levels of farm environment influencing strategy:

1. external environment: macro-economic conditions over which the farmer has no control
2. operational environment: situation on the market over which the farmer has some control
3. internal environment: the resources under direct control of the farmer (for example long-term decisions about resource allocation and use or long-term decisions about fixed costs, quality of buildings and machinery, farm layout, etc.)

4. micro-social environment: the social environment in which the farm operates (the social situation of the farmer or the family situation, support from the family, etc.)

The first two factors are generally considered in terms of geographical location: apart from differences like soil quality and climate, different geographical regions may be characterized by differences in business culture.

3. The Malmquist Productivity Index and its use to analyze agricultural efficiency and productivity changes

The Malmquist Productivity Index was originally introduced by *Caves, Christensen and Diewert (1982)* and is a total factor productivity index based on the ratio of two distance functions. This approach measures productivity change comparing observed change in output with the imputed change in output that would be possible from the observed input changes. This imputation is based on the production possibilities set for either the current or the subsequent period. For any time interval two Malmquist ratios are available, depending on whether the reference technology is that of the initial period or subsequent period. Following the work of Färe et al. (1994) the geometric mean of the two ratios measures the Malmquist Productivity Index. Moreover, in case of variables returns of scale, the Index can be output or input oriented.

The Malmquist Index has some properties that are useful in empirical work because it can be used in situations where either prices do not exist or where existing prices have little economic meaning and it can be decomposed into economically relevant sources of productivity changes. So, it requires only quantity data. Moreover this approach does not require the assumption of efficient production but it identifies the "best practice" countries in every period, which gives an efficient production frontier and measures each country's output relative to the frontier.

To understand the structure and the means of Malmquist Productivity Index the output oriented version is illustrated.

Let's suppose having for each time period t=1,…, T a certain production technology S^t that transform the inputs x^t into outputs y^t

$$S^t = \{(x^t, y^t)\}$$

This means that the technology at t consists of the set of all possible input/output pairs. At time t there will be an output distance function $D^t(x^t, y^t)$ for every set of input/output belonging to S^t.

$D^t(x^t, y^t)=1$ only if (x^t, y^t) is on the boundary of frontier technology, that is only if the production is technically efficient. To define the Malmquist index we need to define distance functions comparing output at one period with the technology of another period such as $D^t(x^{t+1}, y^{t+1})$. This distance function measures the maximum proportional change in outputs required to make (x^{t+1}, y^{t+1}) feasible in relation to the technology at time t. If this point is outside the set of feasible production in period t technical change has occurred.

The Malmquist Productivity Index (MPI) is equal to

$$MPI^t = \frac{D^t(x^{t+1}, y^{t+1})}{D^t(x^t, y^t)}$$

The relative technology is S^t and, relative to that technology MPI^t is the ratio of the efficiency of (x^{t+1}, y^{t+1}) to the efficiency (x^t, y^t). If $MPI^t > 1$ productivity has increased between t and t+1. Alternatively it is possible to define another MPI using S^{t+1} as the reference technology:

$$MPI^{t+1} = \frac{D^{t+1}(x^{t+1}, y^{t+1})}{D^{t+1}(x^t, y^t)}$$

As previously stated, the Malmquist Productivity Index is calculated as geometric means of the two indexes above

$$MPI = \left[\frac{D^t(x^{t+1}, y^{t+1})}{D^t(x^t, y^t)} \cdot \frac{D^{t+1}(x^{t+1}, y^{t+1})}{D^{t+1}(x^t, y^t)}\right]^{1/2}$$

This expression can be factored as

$$MPI = \frac{D^{t+1}(x^{t+1}, y^{t+1})}{D^t(x^t, y^t)} \left[\frac{D^t(x^{t+1}, y^{t+1})}{D^{t+1}(x^{t+1}, y^{t+1})} \cdot \frac{D^t(x^t, y^t)}{D^{t+1}(x^t, y^t)}\right]^{1/2}$$

Where it is possible to separate the change in relative efficiency and the shift in technology between the two periods t and t+1

$$Efficiency\ change = \frac{D^{t+1}(x^{t+1}, y^{t+1})}{D^t(x^t, y^t)}$$

$$Technical\ change = \left[\frac{D^t(x^{t+1}, y^{t+1})}{D^{t+1}(x^{t+1}, y^{t+1})} \cdot \frac{D^t(x^t, y^t)}{D^{t+1}(x^t, y^t)}\right]^{1/2}$$

The change in relative efficiency is the change in the distance of observed production from maximum feasible production between t and t+1.

A Malmquist Index with value greater than unity reveals improved productivity. Efficiency and technical changes indices exceeding unity reflect gains in those components.

It is important to note that if $x^t = x^{t+1}$ and $y^t = y^{t+1}$ there has been no change in inputs and outputs between the periods and the productivity change index signals no change, so MPI = 1. In this case the component measures of efficiency change and technical change are reciprocals but not necessarily equal to one, because a change in efficiency might exactly offset a technological change.

3.1 Application of the Malmquist Index: a literature review

The Malmquist Index approach has been used in the literature for different kinds of analyses and one of these is the comparison of intercountry agricultural productivity and efficiency.

These kinds of studies have proliferated in recent years to take advantage of database improvements, to test alternative methodologies and to explore various research questions.
Many studies focused on the situation of African countries, former socialist countries or on the differences between developed countries and a less-developed one. Some of them have investigated changes in productivity either at farm or sector level in agriculture.

One interesting study was by *Trueblood and Coggins (2001)* for 115 countries over the period 1961-1991. These authors first compared the overall average technical efficiency in each period under VRS and CRS assumptions and then examined productivity growth using the Malmquist index approach. When analysis of this type occurs (with many and different countries) one important issue concerns whether it is more appropriate to compare countries' technical efficiency relative to a global frontier or regional frontiers. Another issue in the intercountry comparison regards the opportunity to compare the efficiency and productivity of farms in less developed countries to those located in more developed countries. In this case, it may happen that a farm could be technically efficient given its economic environment (which perhaps lacks infrastructure, research, services etc.) and technically inefficient by international standards. When the comparison is made among very different countries, it is usual to find technical inefficiencies and, relative to the agricultural sector, the possible causes are lags in technological development, diffusion and adoption of new technologies, pervasive market, trade distortions etc. To compare such countries, the FAO database has been used and in particular the aggregate output expressed in monetary terms (international dollars, a measure obtained calculating weighted world prices for each commodity and multiplying it for the quantities of the commodities). With regard to inputs, the study considers land (arable and permanent cropland), labour (males and females economically active in the agricultural sector, not on an hourly flow basis), fertilizers (the sum of nitrogen, potash and phosphates expressed in nutrient-equivalent terms), livestock, physical capital (n° of tractors in use). The results of the analysis show an increasing in the agricultural productivity of developing countries over the time and a difference in agricultural productivity between developed and developing countries. Moreover, the study highlights the differences in productivity growth among regions.

4. Some applications of DEA to the FADN database: a literature review

As previously stated, the estimation of technical efficiency based on DEA is an important tool that permits specific adjustment to be individuated in the production activities. The application of DEA in the agricultural sector makes it necessary to pay attention to the characteristics of some inputs, the application requirements of which need specific considerations.

First of all, agriculture is characterized by production factors that are fixed or quasi-fixed so very slow to adjust. For example, land rental markets operate over long periods of time and do not allow instantaneous adjustments. As concerns family labour, transaction costs for labour market participation are often too high and this prevents the complete integration between farms and rural labour markets. Thus, in the case of agriculture the quasi-fixity of inputs may create a persistent technical inefficiency and, sometimes, an inefficiency of agricultural policy measures that would encourage adjustment in the use of production factors. Moreover, the Farrell measure of technical efficiency provides insights for total factor employment and proposes an equiproportional reduction for all the factors necessary to attain such efficiency. But in the case of agriculture there are important questions to take into account, such as the different impact of agricultural inputs on the crop or environment. An equiproportional input reduction or increase could not have a proportional effect on the output.

This section presents some analyses concerning the use of FADN database in the application of DEA methodology. DEA requires data at farm level, so FADN is an important source of information, not only because it gathers observations at micro-economic level, but also because it provides additional information about farm structure. The availability of a relevant quantity of data, makes the FADN database very useful to develop DEA methodology.

One problem of FADN is that the database does not contain reliable information about input prices and this is sometimes necessary for an efficiency analysis. Price efficiency assumes the existence of technical efficiency and hence this is a necessary condition for economic efficiency. The lack of accurate farm-level price data and the adjustment lags to current prices make it difficult to capture price efficiency empirically.

The application of DEA to the FADN database permits farm performance and efficiency to be analyzed between different farm types and European regions. Moreover, using time series, it is possible to obtain information on changes in farm performance and efficiency over time, in order to have some indications about the impact of different policies or support schemes etc. In particular, for the new EU Member States, this kind of analysis could be a useful tool to better understand farm performances and economic efficiencies after market deregulation and adjustments made for EU membership. In particular, many studies have been done with the aim of analyzing the consequences of the transition process from the previous protected economic system to a market economy and the consequences of the application of the Common Agricultural Policies in the traditional farming sector.

The chapter is divided into three sections in order to have a clearer framework of the economic literature about DEA.

4.1 Application to the investigation of the determinants of efficiency

Bojnec and Latruffe (2009) in their analysis of technical efficiency of Slovenian farms, evidenced the performance of Slovenian agriculture for different sectors and over an extended period. In particular, they investigated the determinants of technical efficiency of individual farms in the period 1994-2003. This work compares the results from the SFA method and DEA method, using the output-oriented approach that estimates the potential output increases without further input increases.

The technical efficiency of individual Slovenian farms is analyzed using two methodologies. First of all, the score is calculated using the DEA method: the output is the value of total revenue while the inputs are utilized land (UAA, ha), labour (AWU), the value of total assets as a capital variable and the value of variable inputs. Then, based on the literature of farm technical efficiency in transition economies, three groups of determinants of technical efficiency have been identified. Their influence has been investigated regressing the technical efficiency over this set of explanatory variables that can be defined using FADN data:

1. *Family character of the farms*: includes the share of hired labour in total labour used, the share of rented land in the total utilized area and the share of marketed output in total output. The first two determinants capture the effects of the reliance on external factors on technical efficiency while the last is a proxy for farm commercialization and its non-subsistence character.
2. *Farm specialization*: calculated using the Herfindahl index (sum of the squared shares of each activity in the farm total revenue)
3. *Reliance on subsidies*: ratio of subsidies for production to total revenue.

The calculation of technical efficiency score over time shows the trend of farm performance over the studied period and, in this specific case, gives additional information on the changes during the transition process.

Technical efficiency regression on the determinants of efficiency gives further indications about the positive or negative influence of all the selected variables. In this case, there are interesting political issues resulting from the analysis. The most important one is the reliance on subsidies. The transition process has aimed to support individual farm development and most of this support has been delivered in the form of market-price supports. The result is that subsidies have had a negative effect on technical efficiency, reducing farmers efforts, and this conforms to the theory and previous studies on other countries. But further research is necessary to analyze productivity and efficiency after the accession of new Member States because a reduction in technical efficiency does not contradict a possible technology improvement with the help of subsidies.

The Slovenian FADN dataset has been used in another analysis by *Bojnec and Latruffe (2008)* to compare efficiency scores applying DEA and SFA methodology. The period is the same (1994-2003) and the aim of the analysis is the investigation of 13 farm sector[2] performances. The farm sectors have been classified with a cluster analysis based on the log-likelihood distance between observations. The farms are classified considering socio-economic type, areas and production types (according to the shares of revenue).

The agricultural inputs used in this analysis are: utilized agricultural area (UAA, in hectares), total labour (AWU, hours per year), the value of assets and the value of intermediate consumption. The SFA considers the total output in value while in DEA the total output is divided into crops, livestock and other outputs, permitting also a representation of different technologies of farm sectors. Unlike the previous case, this analysis calculates economic and allocative efficiencies using total cost of the farm business and input prices[3]. The interesting results of this analysis can be summarized in three points:

1. input market deregulation and liberalization increase the economic efficiency of Slovenian farms that are shifting from traditional small-scale subsistence towards family owned and entrepreneurial activities, with a different structure of managerial and management practice
2. there are differences in farm performance and efficiency among the farm sectors: whereas they use similar technological, management and business practices their performances and economic efficiency differ by their output market specialization. Macroeconomic and institutional environment and industry structure can explain these differences

---

[2] Crop, dairy using own feed, cattle using own feed, pigs using own feed, sheep using own feed, other livestock using own feed, livestock using purchased feed, fruit, grapes and wine (share of the specific farm production > 50% of total revenue), mixed (combination of different revenues where the single product contribution is < 50% of total revenue), vegetables, forestry (share of the specific farm production > 75% of total revenue) and combined (incomes from crop and livestock using own feed where the single production contribution is < 75% of total revenue).

[3] Land prices are farm land rentals, labour prices are hired farm labourer's wages and capital prices are farm depreciation costs.

3. there are differences in technical, scale, allocative and economic efficiencies over time. In particular, the scores for some farms are increasing as a consequence of a better competitiveness in the EU and other international markets.

Fousekis, Spathis and Tsimboukas (2001) use the 1997 data set of FADN to measure and decompose the overall efficiency of sheep farming in some disadvantaged and mountainous areas of Greece (Epirus, Sterea and Thessaly). More specifically, the objective is to decompose the overall efficiency in technical and scale efficiency. To do this, the CRS and VRS models have been used on a sample of 101 farms.
In order to specify the production technology, the model considered FADN data for outputs (milk and meat) and inputs (labour in AWU, capital, purchased feedstuff, farm-produced feedstuff and animal stock).
The objective has been pursued by applying both the VRS and CRS DEA models to a sample of 101 FADN farms. If the assumption of CRS is true, the application of the CRS DEA model will lead to correct estimates of technical efficiency. However, if returns to scale are increasing or decreasing, the CRS DEA model will estimate an overall efficiency (technical and scale efficiency). In other words, if there is a scale inefficiency, the CRS DEA model will underestimate the true pure technical efficiency and, at the same time, will provide no information on scale inefficiency and its causes. So the VRS and CRS DEA efficiency results are quite different.
The different determinants of technical efficiency have also been identified in this case. FADN does not provide any information on personal characteristics of farmers (age, education, experience, etc.), so farm characteristics are taken into account. More specifically: typical gross margin, % of subsidies on total family farm income, % of value of milk to gross sales revenue and % value of farm-produced feedstuff to value of purchased feedstuff.
The typical gross margin can be considered as a proxy of farm size and, as reported in other analyses, the negative sign of the coefficient would mean that smaller farms tend to attain a higher efficiency level than larger farms. Also in this case, the contribution of subsidies to the family farm income has a negative correlation coefficient with technical efficiency. It is important to highlight that correlation coefficients are partial measures of association so they must be interpreted carefully, especially because the economic theory does not provide any guide on the factors affecting pure technical efficiency.

Iraizoz, Rapun and Zabaleta (2003) use the stochastic parametric frontier and DEA methodology to estimate technical efficiency in the horticultural sector in Navarra (Spain), of two important crops in particular: asparagus and tomato. They do the analysis considering the changes in agricultural policies after the CAP Reform, which has made agricultural sector more competitive and more market-oriented. In this context, an improvement in farm efficiency is fundamental and the measurement of existing inefficiencies in agricultural production becomes much more useful. An interesting part of this work analyzes the relationships between the technical efficiency measure and other relevant variables such as size, combinations of inputs and traditional measures of farm performances. The identification of the source of technical efficiency is essential to design private or public policies to improve farm performances.

The data used in the analysis are figures coming from the Spanish FADN, which refer to 46 horticultural farms where asparagus is cultivated extensively as a crop on non-irrigated land, while tomato is an intensively cultivated crop on irrigated land. In this work:

- output is measured by sales of asparagus and tomato gross production[4] (both in euros)
- inputs are labour (AWU, number of hours worked per year), land (UAA, hectares), capital (average annual inventory of machinery and buildings, euros) and cultivation costs (cost of seeds, fertilizers and other specific cultivation costs, euros)

Both the CRS and VRS DEA models have been estimated for the same number of farms and output and input variables, obtaining different efficiency scores: global technical efficiency, pure technical efficiency and scale efficiency. The study also explains the efficiency, focusing on the second stage of the analysis of technical efficiency, that is the relationship between the scores and different exogenous variables. Interesting conclusions arise from the regression. First of all, in the case of asparagus production there are not significant relationships between efficiency and size, so to be more efficient it is not necessary to be bigger. On the other hand, for both commodities the correlation between efficiency and land productivity index (value of output per hectare) has been significant: farms that perform well are able to use their inputs efficiently. There is the same positive result for the labour productivity index (value of output per labour unit). In the case of tomato production, a negative relationship exists between capital per hectare of land and technical efficiency and this result could suggest an overcapitalization of the farms. In this case, farmers should match their farm size more closely to their real conditions: the agricultural policies could include measures to improve the capacity of farmers to apply technologies, to gain access to services, improve educational level and so on.

This study also makes an analysis of the most important determinants of technical efficiency, making a regression of efficiency scores against the output (expressed in physical units), land productivity (monetary output per hectare), labour productivity (monetary output per labour unit), costs and capital per hectare. The result is a positive correlation between output and efficiency scores and this could indicate the presence of scale economies in the sample. Moreover, farms with higher costs and more capital should be more intensive, with much higher

[4] Sales + farm use + farmhouse consumption + (closing valuation – opening valuation)

output and, so, with a positive correlation with the efficiency score. A negative correlation in this case can indicate farm overcapitalization.

Piot-Le Petit, Vermersch, Weaver (1997) analyze agriculture's environmental externalities applying DEA to French agriculture. This work is very interesting because it estimates technical efficiency under three hypotheses of the quasi-fixity of farm inputs. One question that may arise in the calculation of a technical efficiency score concerns the specification of a variation range before setting a standard of efficiency. One example is the distinction between short and long-term problems. In this work, inputs are divided into two groups to take into account the possibilities of changing them according to the efficiency objectives.

The data are those of the French FADN, the cereal sector in particular. The production vector includes two outputs expressed in monetary terms: cereal output (wheat and maize) and other outputs (mainly oilseed). The inputs include:

- quasi-fixed inputs: area cultivated under cereals (ha), under other cereals (ha) and family labour (AWU)
- variable inputs (expressed in monetary terms): equipment, fertilizers, pesticides, seeds, others

The tested hypotheses are the following:

- short-term: land and labour are maintained quasi-fixed
- long-term: all inputs are variable so the farm can adjust all the variable levels
- intermediate scenario where the level of a quasi-fixed input is modified after an opportunity of adjustment (family labour may be adjusted through off-farm employment)

The result shows that in general the technical efficiency measure increases when the number of fixed factors decreases. So, the estimated score in the short-term is lower than in the intermediate and long-term. When factors are quasi-fixed, both the level of efficiency and distribution of efficiency among farms may be overestimated when based on the long-term and all factors are variable. So, also in non-parametric approaches it is important to specific the input variability and fixity.

Moreover the work has another important conclusion: it is evidence to support the notion that adjustment of technical efficiency of farm use of fertilizers and pesticides could be an important opportunity for reducing the associated environmental impacts. For example: the introduction of a tax on mineral fertilizer should increase the opportunity cost of organic fertilizer and this works well when farmers are initially inefficient. In this case a tax could help to improve farmers' efficiency and would be consistent with environmental issues. As a consequence, knowledge on the technical efficiency of farmers would be useful to a public authority in charge of agricultural and environmental policies to implement the best policy.

Hansson and Ohlmér (2008) investigated how managerial practices can contribute to improve dairy farm efficiency. In their work, DEA efficiency scores were based on 507 Swedish dairy farms and these scores have been used to make a regression with some operational managerial practices.

These practices are defined as aspects describing animal health, breeding and feeding practices that can be changed on a short-term basis. The analysis also considers economic efficiency that measures the overall efficiency including both technical and allocative efficiencies. Moreover, the study is conducted at the whole farm level, considering all the inputs and outputs of dairy farms and not just the milk production.

Farm level accounting data and prices are obtained from Statistics Sweden, stratified according to geographical location and size. All data were collected over several years (1998-2002) for 507 individual farms, but each farm is represented by its own yearly average. The first-stage efficiency analysis is based on this information.

Data on the operational managerial practices are obtained by combining a mail questionnaire and data from a dairy cow recording scheme conducted by the Swedish Dairy Association. A total of 360 farmers replied to the questionnaire, but only 169 filled it in completely, so, for the second-stage regressions there are missing values.

In this case, an alternative approach would have been to base the first-stage analysis on the same farms available for the second-stage but this might have led to biased efficiency scores.

The inputs taken into account are fodder, fertilizer, seed (measured in kg), labour (measured in hours, family and employees), capital (measured in SEK) and energy (measured in units), while the outputs are milk, livestock, crops, forage (measured in kg) and other (the remaining outputs of the farm, mostly allowances measured in euros). Quantities are obtained by dividing the revenue or expenses by its price.

With regard to the managerial practice variables, three groups are take into account:

- animal health practices: age at first calving, calving interval, length of dry period
- breeding practices (breeding percentage, breeds percentage)
- feeding practices (analysis of forage, analysis of fodder grain, feed ration, mix of forage)

The managerial practice variables were regressed on the farm level efficiency scores. Six equations were estimated: long and short-term economic, technical and allocative efficiencies. The regression results show that animal health practices do not affect long and short-term efficiency. Regarding the breeding percentage, if it equals the replacement rate, this has a

significantly negative effect on the long-term technical efficiency: a reason is that farms that breed more heifers than needed can evaluate them before letting them become dairy cows. So, the farmers can choose to keep only the most promising heifers. If the breeding equals the replacement need this evaluation is impossible. The breed percentage (Swedish Red and White Breed) does not affect the efficiency scores, so there are no differences in quality between the heifers. Feeding practices seem to affect the economic and allocative efficiency scores more and an explanation is that analyses of feeding facilitate more optimal feeding use, so lead to higher efficiency.
In general, although it is possible to improve the operational managerial practices on a day-to-day basis, the significant effects are more likely to have effects in the long-term, when the farmer labour and capital can also be adjusted to optimal levels. This study is interesting because it makes the distinction among different kinds of efficiency and differentiates between short and long-term. Furthermore, it investigates empirical results and not a simulation model.

4.2 Application to the comparison of groups of farms to find the most productive technology

Cisilino and Madau (2007) use Italian FADN data to compare organic and conventional farming. These two systems are characterized by major differences in production techniques and, as consequence, in the technical-productive paradigm and agronomic techniques. These differences create some problems in the comparison analysis. The approach defines conventional farms as an approximation, which means *how an organic farm should be if it were conventional*. Considering this approach the similarity between organic and conventional farms operating in the same context is founded on the same levels of potential production and on the same level of available resources. So the hypothesis is that there is technological homogeneity between the two productive systems. There are many problems under this hypothesis and different studies and researches have tried to solve them. Notwithstanding the fact that the debate on farm’s selection process for comparison is still open, there are specific requirements that organic and conventional farms have to achieve to permit this analysis (similar environmental conditions, same localization, same equipment of productive factors, same business type).

A very interesting study has been done by *Arfini and Donati (2008bis)* who used an integrated model (PMP and DEA) to analyze the impact of the Health Check (HC) measures in four different European farming regions (Emilia Romagna, Anatoliki-Makedonia-Thraki, Kassel and

Ostra Mellansverige). They used FADN data for 2004 to assess the performances, the capacity to respond to new scenarios and the level of technical efficiency of a sample of different farms, specialized in cereal and livestock productions. More specifically, the PMP approach is used to represent the farm characteristics and to simulate the effects of the new agricultural policy measures, while DEA permits the estimation of a farm technical efficiency index, before and after reform. The aim of the paper is to assess the effects of the HC on the competitiveness of farms, considering the role of decoupled aid, the farm reaction to market changes and the maintenance of farm environmental functions. In particular, the HC reform introduces the possibility to apply the regionalization of aid and to increase the modulation up to a maximum of 22%. Both measures have involved a process of redistribution not only among farms but also among sectors. Moreover, they have a direct impact on the level of technical efficiency at a farm level, together with other important aspects of farming (productive, economic and structural characteristics).

As concerns the DEA analysis, the output is the saleable production (expressed in monetary terms, euros), while inputs are represented by land (hectares), labour (family and extra-family work, AWU) and total variable costs (expressed in monetary terms, as a measure of inputs available on the market as a whole).

The interesting result of this work is that the transition from coupled to decoupled payments (Single Payment Scheme, SPS) does not significantly change the levels of efficiency (mean level per Region and Farm Type). The transition from historical SPS to the regionalized SPS introduced with the HC without variations in price does not change the farming system, while in the presence of these variations the farms reorganize their production and improve their level of technical efficiency.

In this context, considering technical efficiency as a proxy for level of competition, the evidence is that an important influencing factor is represented by the price evolution and, as a consequence, by the capacity of farmers to allocate the production factors in a technically efficient manner depending on the changes in market conditions.

Latruffe, Fogarasi (2009) makes an interesting contribution to the analysis of farms' technical efficiency in Central and Eastern European Countries (CEECs) during the transition period. Unlike many analyses, this work evaluates farm performances in comparison with Western Countries. This kind of comparison is very important because since enlargement, all farms in all countries compete on the same market, respecting almost the same policies. The expectation on the relative performance of CEECs compared to the EU-15 is ambiguous. On the one hand, the low technical efficiency might be due to the difficulties in making investments because of a lack of financing. On the other, most of the land is farmed by corporate farms, whose large production scale might give an advantage in the overall efficiency in these countries. This paper makes a comparison between France and Hungary in two important sectors: milk (TF41) and COP (TF13). The objective is to investigate how French and Hungarian farms differ in terms of performance and which country leads the technology. Change of farm performance over the whole period is studied using the Malmquist indices that can be decomposed into the part of change due to a change in farm position with respect to the frontier (technical efficiency change) and the part due to a shift of the frontier over time (technological change). Indices equal to 1 indicate no change while indices greater (less) than 1 shows progress (or regress).

The role of subsidies in the farm performance is compared using econometric regression. Public subsidies are among the main factors explaining technical efficiency or inefficiency and, following the empirical evidence, they have a negative impact on farm technical efficiency. Regressions are carried out in the merged sample and the variables used are in turn:

- the technical efficiency scores calculated under each country's frontier
- the technology ratios
- the TFP change indices calculated under each country's frontier

Thus, three regressions are made for the dairy sector and three for COP in the merged sample. This is the first study that do this investigation considering the metafrontier.

The study is done in more than one stage. Firstly, technical efficiency and Total Factor Productivity (TFP) are calculated for the countries separately. Secondly, the same measures are calculated with a common technology frontier, i.e. with a DEA metafrontier constructed on the merged sample of both countries in order to understand which country is lagging behind in terms of technology. In other terms, the metafrontier permits the technology gap to be assessed between two groups of observations: it is constructed enveloping several groups of observations whose technology differs and comparing it with the respective frontier of each group in order to identify technology gaps between each group's frontier and the enveloping metafrontier.

So, two kinds of DEA frontier and technical efficiency scores are calculated:

- one frontier for each country (TE_C)
- one frontier merging in a unique sample (metafrontier, TE_M)

A technology ratio is computed for each farm dividing the two efficiency scores:

$$Techology\ ratio = \frac{TE_C}{TE_M}$$

By construction this ratio is equal to or less than 1. Average technology ratios for French and Hungarian farms are compared: if they are different, it indicates a gap between both countries' technology, with the higher average revealing the more productive technology.

For dairy farms, the DEA model includes two outputs (the amount of milk produced in litres and the value of other outputs in euros) and five inputs (the UAA in hectares, the total labour used in AWU, the value of total assets in euros, the value of intermediate consumption in euros and the number of livestock units calculated with the standard European definition). For COP farms there is a single output (the value of total output in euros) and the same inputs as dairy except the livestock units.

FADN data for the period 2001-2004 are used.

The results of this study reveal that in both sectors, Hungarian farms are strong leaders in terms of technology despite a technological deterioration (technological change indices less than 1). Hungary has technological advantages for large scale over the small scale of French farms, especially in the dairy sector.

4.3 Appli cations of the bootst rapping method to overcome the sampling problem

Notwithstanding the flexibility of DEA, its results may be affected by sampling variations: this means that distances to the frontier are likely to be underestimated if the best performers in the population are not included in the sample.

An interesting application of bootstrapping procedure has been made by *Balcombe, Davidova, Latruffe (2006)*. The objective of the study is to assess the extent to which the Malmquist productivity changes derived used DEA are affected by sampling variation, in the first stage calculating productivity indices and in the second stage investigating the farm-specific change in productivity.

The sample is composed of 250 Polish farms over a 5-year period (1996-2000) and the data were collected by the Institute of Agricultural and Food Economics in Warsaw.

In the model three outputs are included in value terms: crop, livestock and non-agricultural output. Four inputs are used: land (UAA, hectares), labour (AWU), capital (proxied by the value of depreciation of fixed assets plus interest paid on loans) and intermediate consumption (aggregate value of seeds, fertilizers, chemicals, feed and fuels).

The study follows the method introduced by *Simar and Wilson (1999)* which adapted the bootstrap methodology to the case of Malmquist index, using a bivariate density estimate that accounts for the temporal correlation via the covariance matrix of data from adjacent years. The set of bootstrap Malmquist indices provided by this procedure allows to account for the bias and to construct confidence intervals. In this study 2000 bootstrap interactions were performed and the 95% of confidence intervals were constructed.

A second-stage regression is performed using the information provided by the bootstrap procedure to investigate factors determining farm productivity change. A heteroschedastic panel regression (Maximum Likelihood) is employed[5]: the idea is to assume that Malmquist indices are observed with noise:

$$y_{i,t} = y*_{i,t} + u_{i,t}$$

Where $y*_{i,t} = \gamma_t \beta_0 + x'_{i,t}\beta_1 + e_{i,t}$

- $y_{i,t}$ is the Malmquist index for the i-th farm at the t-th period, calculated with DEA and used as dependent variable in the regression
- $y*_{i,t}$ is the true Malmquist index that is unobserved and that may be estimated
- $u_{i,t}$ $e_{i,t}$ are error terms (assumed to be normal, independent of the explanatory variables and independent of each other)
- γ_t is a fixed time effect
- $x'_{i,t}$ is a vector of explanatory variables
- β_0, β_1 are parameters to be estimated

About the error terms, $e_{i,t}$ is homoschedastic while $u_{i,t}$has a variance different:

$$Var(u_{i,t}) = \eta_{i,t}\lambda$$

$\eta_{i,t}$ is the variance of the i-th farm's DEA Malmquist as given by the bootstrap distribution while λ is a parameter reflecting the degree to which the bootstrap standard errors contribute to the variance of the errors in equations.

The explanatory variables are land, capital to labour ratio (proxy of farm technology), share of hired labour (represents the farm integration into the labour market), share of marketed output in total output (availability of financial resources), share of other income in total income (concentration of farming), farmer's age and agricultural education.

The interesting conclusion of this work is that the comparison between the point estimates of the Malmquist index and the analysis based on the confidence intervals shows different results: productivity might have been static rather than decreasing. This result contrasts with previous studies that highlighted a productivity regress for Polish farms. Therefore, this study put on evidence the uncertainty surrounding the findings regarding productivity change measured through the Malmquist DEA method using point estimates only.

As concerns the results of the heteroschedastic panel regression, the results are coherent with the theory and with previous researchers. In particular, this study underlines a negative influence of the share of hired labour on productivity growth: family labourers have more incentives than hired labour to act efficiently. With regard to the share of other income, often in transition economies the diversification is included in a survival strategy but this could imply poor management practices and therefore low technical efficiency.

---

[5] Usually Ordinary Least Squares regressions are used in this kind of analysis. This is a novelty.

The bootstrapping procedure and Tobit model have been used in the work of *Davidova, Latruffe (2007)* that for the first time makes an analysis of the relationship between farm financial structure and technical efficiency in Central and Eastern European farms during the transition to a market economy. In particular, Czech farms have been analyzed. The study tests three financial theories:

- Free cash flow theory: suggests that farmers who are in debt need to meet their repayment obligations and are therefore motivated to improve their efficiency (positive relationship)
- Agency costs theory: says that highly indebted farmers might incur higher costs and thus be less technically efficient (negative relationship)
- Credit evaluation theory: suggests that banks prefer borrowers who are low risk, thus more technically efficient because banks evaluate loan applications according to the applicant's probability of repayment. The land being the main asset in agriculture, land poor farmers often have rationed or are denied access to credit (negative relationship)

Under a theoretical point of view, the relationship between financial structure and farm efficiency appears ambiguous.

In this paper, the analysis shows the difference of this relationship between individual farms and corporate farms, the main farm management types in Czech agriculture. While the former use bank credit extensively, the corporate farms have liabilities that stem from the reform process, which are not debts to the banks but the so-called transformation debts (debts to the State or from State farms). Considering that the reform debts impede further access of corporate farms to commercial loans, it is expected that the relationship between financial structure and technical efficiency is different for the individual and commercial farms.

The study employs an output-oriented model using FADN data. The specialization of livestock and crop farms has been defined for farms where at least 65% of the value of total agricultural output comes from crops or livestock. The livestock sub-sample contains 88 farms (53 individual and 35 corporate) and the crop sub-sample contains 256 farms (221 individual and 35 corporate).

The output variables (expressed in value terms) are crops, livestock and other outputs. The inputs are land (UAA, ha), labour (AWU), capital (depreciation and interest) and intermediate consumption (value terms). The value units are expressed in Czech Koruna (CZK). Four frontiers were estimated: one for each specialization (livestock and crop) for each management form (individual and corporate farms).

The second-stage regression has been made using a Tobit model, defined by the following equations:

$$y = \begin{cases} y* \; if \; y*>1 \\ 1 \; otherwise \end{cases}$$

$$y* = Y\beta + X_e\gamma + u = Z\delta + u$$

where Y is the indebtedness variable, X_e are the other explanatory variables of the models, β and γ are the parameters to be estimated, u is an error term. The dependent variable y is the observed output-oriented efficiency score (greater than 1). For inefficient farms whose efficiency score is strictly more than 1, the latent variable $y*$ is the observed efficiency.

The dependent variable used in each regression for the four sub-samples is the total technical efficiency score given by the output-oriented DEA model, while the following independent variables are taken into account:

- Size variables (livestock units for livestock farms and two land area dummies for individual and corporate crop farms)
- Technology proxies (the ratio of capital to labour and the ratio of land to labour)
- Quality of labour (percentage of hired labour) and institutions for land tenure (percentage of rented land)
- Indebtedness variables: debt to asset ratio (long-term), current debt to current asset ratio (liquidity), ratio of bank debt to total assets

Four regional dummies are used to take into account the five Czech agri-environmental regions (maize, sugar beet, cereal, potato and mountain-forage used as reference).

Application of the bootstrapping technique leads to the conclusion that the point estimates alone overstate farm technical efficiency and this is another demonstration of the usefulness of estimating confidence intervals of the DEA efficiency scores. In general, overall corporate farms in the Czech Republic are more clustered towards the technically efficient frontier than the individual farms. This could mean that the management experience in the post-reform of individual farms is more deficient than those of former state and collective farms. But it is true that individual farms are more heterogeneous and there are farms lagging behind the best practice. This aspect is highlighted by the bootstrapping method.

As concerns the relation between financial structure and technical efficiency, the results detect differences between the two management types. Credit evaluation seems to be a problem of individual farms and both are influenced by the agency costs theory.

References

Arfini, Donati, *Health Check and Farm Efficiency: a comparative assessment of our European Agricultural Regions,* 109th EAAE Seminar, Viterbo (Italy), November 2008

Arfini, Donati, *Health Check e efficienza delle aziende agricole: una valutazione comparative su quattro Regioni agricole europee,* XLV Convegno di Studi della Sidea, Portici (Italy), September 2008bis

Balcombe, Davidova, Latruffe, *The use of bootstrapped Malmquist indices to reassess productivity change findings: an application to a sample of Polish farms,* Paper presented at the International Association of Agricultural Economists Conference, Gold Coast, Australia, August 12-18 2006

Banker, Charnes, Cooper, *Some models for estimating technical and scale inefficiencies in Data Envelopment Analysis,* Management Science, 30 (9), pp.1078-1092, 1984

Bauer, Berger, Ferrier, Humphrey, *Consistency conditions for regulatory analysis of financial institutions: a comparison of frontier efficiency methods,* Journal of Economics and Business, 50, pp.85-114, 1998

Bojnec S., Latruffe L., *Determinants of technical efficiency of Slovenian farms,* Paper presented at the I Mediterranean Conference of Agro-Food Social Scientists, 103rd EAAE Seminar, Barcelona, Barcelona (Spain), 23rd - 25th 2007

Bojnec S., Latruffe L., *Measures of farm business efficiency*, Industrial Management Data Systems, 108, (2), pp. 258-270, 2008

Caves, Christensen, Diewert, *The economic theory of Index Numbers and the measurement of input, output and productivity,* Econometrica 50, pp.1393-1411, 1982

Charnes, Cooper, Golany, Seiford, Stutz, *Foundations of Data Envelopment Analysis for Pareto-Koopmans efficient empirical production functions*, Journal of Econometrics 30 (1–2), pp. 91–107, 1985

Charnes, Cooper, Rhodes, *Evaluating program and managerial efficiency: an application of DEA to program follow through,* Management Science, 27(6), pp.668-697, 1981

Charnes, Cooper, Rhodes, *Measuring the efficiency of decision making units,* European Journal of Operational Research, 2, pp. 429-444, 1978

Charnes, Cooper, Rousseau, Semple, *Data Envelopment Analysis and axiomatic notions of efficiency and reference set* in *Research Report CCS 558*, Center For Cybernetic Studies, The University of Texas, Austin, 1987

Charnes, Cooper, Seiford, *Data Envelopment Analysis: Theory, Methodology, and Application*, Springer, 1994

Charnes, Cooper, Seiford, *Invariant multiplicative efficiency and piecewise Cobb-Douglas envelopments*. Operations Research Letters 23, pp. 101–103, 1983

Charnes, Cooper, Wei, Huang, *Cone-Ratio Data Envelopment Analysis and Multi-Objective Programming,* International Journal of Systems Science, 20(7), pp. 1099-1118, 1989

Cisilino, Madau, *Organic and Conventional Farming: a Comparison Analysis through the Italian FADN*, 103rd EAAE Seminar, Barcelona, Spain, April 2007

Coelli, Prasada Rao, Battese, *An Introduction to Efficiency and Productivity Analysis*, Springer, 1998

Coelli, *Recent Developments in Frontier Modelling and Efficiency Measurement*, Australian Journal of Agricultural Economics, 39, (3), pp 219-245, 1995

Cooper, Seidorf, Tone, *Data Envelopment Analysis*, Kluwer Academic Publishers, Boston, 2002

Davidova, Latruffe, *Relationships between Technical Efficiency and Financial Management for Czech Republic Farms,* Journal of Agricultural Economics, 58 (2), pp. 269-288, 2007

Debreu, *The coefficient of Resource Utilisation,* Econometrica, 19, pp. 273-292, 1951

Färe, Grosskopf, Lovell, *The Measurement of Efficiency of Production*, Kluwer-Nijhoff Publishing, Boston, 1985

Farrell M.J., *The measurement of productive efficiency,* Journal of the Royal Statistical Society, 120, (3), pp. 253-290, 1957

Forsund, Sarafoglou, *On the origins of Data Envelopment Analysis*, Department of Economics, University of Oslo, Memorandum 24/2000,

Førsund, Sarafoglou, *The tale of two research communities: the diffusion of research on productive efficiency*, Department of Economics, University of Oslo, Memorandum 08/2003

Fousekis, Spathis, Tsimboukas, *Assessing the Efficiency of Sheep Farming in Mountainous Areas of Greece. A Non Parametric Approach*, Agricultural Economics Review, 2, (2), August 2001

Fulginiti, Perrin, *Productivity in LDC Agriculture: non parametric Malmquist Measures,* GATT Research Paper 94, March 1994

Greene, *Frontier Production Functions*, EC-93-20, 1993, Stern School of Business, New York University

Grosskopf, *Statistical Inference and non parametric efficiency: a selective survey,* Journal of Productivity Analysis, 7, pp. 161-176, 1996

Hansson, Ohlmér, *The effect of operational managerial practices on economic, technical and allocative efficiency at Swedish dairy farms,* Livestock Science, Elsevier, 2008

Iraizoz, Rapun, Zabaleta, *Assessing the technical efficiency of horticultural production in Navarra, Spain,* Agricultural Systems, 78, pp. 387-403, 2003

Kalvelagen, *Efficiently Solving DEA Models with GAMS*, GAMS Documentation, Washington, November 2004

Koopmans, (conference paper, publ. 1951), *Optimum Utilization of the Transportation System*, in Proceedings of the International Statistical Conferences, 5, pp. 136-145, 1949

Koopmans, *An analysis of production as an efficient combination of activities*, T.C. Koopmans Ed., Activity Analysis of Production and Allocation, Cowles Commission for Research in Economics, Mon. n° 13, Wiley, New York, 1951

Latruffe, Fogarasi, *Farm performance and support in Central and Western Europe: a comparison of Hungary and France,* Working Paper SMART, LERECO 09-07, March 2009

Ortega, Ward, Andrew, *Measuring Technical Efficiency in Venezuela: The Dual-Purpose Cattle System (DPCS)*, Department of Food and Resource Economics, University of Florida, 2004

Piot-Lepetit, Vermersch, Weaver, *Agriculture's environmental externalities: DEA evidence for French Agriculture,* Applied Economics 29, pp. 331-338, 1997

Seiford, Thrall, *Recent developments in DEA, the mathematical programming approach to frontier analysis*, Journal of Econometrics, 46, pp. 7-38, 1990

Simar, Wilson, *Sensitivity analysis of efficiency scores: how to bootstrap in non-parametric frontier models,* Management Science, 44, pp. 49-61, 1988

Simar, Wilson, *Statistical Inference in Nonparametric Frontier Models: The State of the Art*, Journal of Productivity Analysis, 13, pp. 49-78, 2000

Theodoridis, Psychoudakis, Christofi, *Data Envelopment Analysis as a Complement to Marginal Analysis,* Agricultural Economics Review, 7 (2), pp.55-65, 2006

Trueblood, Coggins, *Intercountry agricultural efficiency and productivity: a Malmquist index approach,* Washington DC, US Department of Agriculture, Economic Research Service, June 2001

FACEPA

Farm Accountancy Cost Estimation and Policy Analysis of European Agriculture

www.ekon.slu.se/facepa