Technical and scale efficiency in the Italian Citrus Farming: A comparison between Stochastic Frontier Analysis (SFA) and Data Envelopment Analysis(DEA) Models

Madau, Fabio A.

University of Sassari (Italy)

September 2012

TECHNICAL AND SCALE EFFICIENCY IN THE ITALIAN CITRUS FARMING: A COMPARISON BETWEEN STOCHASTIC FRONTIER ANALYSIS (SFA) AND DATA ENVELOPMENT ANALYSIS (DEA) MODELS

Fabio A. Madau
University of Sassari, 07100 Sassari, Italy

ABSTRACT
This paper aims to estimate technical and scale efficiency in the Italian citrus farming. Estimation was carried out from two different approach: a non parametric and a parametric approach using a Data Envelopment Analysis (DEA) model and a Stochastic Frontier Analysis (SFA) model, respectively. Several studies have compared technical efficiency estimates derived from parametric and non parametric approaches, while a very small number of studies have aimed to compare scale efficiency obtained from different methodological approaches. This is one of the first attempts that aims to put on evidence possible difference in scale efficiency estimations in farming due to methods used. Empirical findings suggest that the greater portion of overall inefficiency in the sample might depend on producing below the production frontier than on operating under an inefficient scale. Furthermore, we found that the estimated technical efficiency from the SFA model is substantially at the same level of this estimated from DEA model, while the scale efficiency arisen from SFA is larger than this obtained from DEA analysis.

Keywords: Technical efficiency, Scale efficiency, Data Envelopment Analysis, Stochastic Frontier Analysis, Citrus farming,

J.E.L.: C13, Q12

---

1. INTRODUCTION

Since the Farrell (1957) seminar paper on efficiency estimation several parametric and non parametric procedures have been proposed in order to calculate efficiency and productivity of firms. However two main approaches have been mainly proposed in literature: the Stochastic Frontier Analysis (SFA) and the Data Enveopment Analysis (DEA). The former is a parametric technique originally and independently proposed by Aigner *et al.* (1977) and Meeusen and van der Broeck (1977) and the latter is a non parametric approach originally proposed by Charnes *et al*. (1978).
Both approaches have their advantages and disadvantages and the suitability of method to the data depends on the industry to be examined (Ruggiero, 2007). The main advantage of SFA is that it takes into account stochastic variation of the output due to ability to handle random noise that can affect output. On the other hand, SFA model is limited to the production of single output and some distributional assumptions need to be made on functional forms of production frontier and in order to separate the stochastic component from the inefficiency term. It is a fact that assumptions about the random error component, fixity of parameters, and production frontier specification can affect empirical results. *Vice versa*, DEA can take into account multi-output productions and no assumptions need to be made on functional forms for the production or cost frontiers and on distributions of the errors. On the other hand, it is a deterministic

method and it cannot separate inefficiency component from noise. This approach also produces biased estimates in presence of measurement error and other statistical noise[1].

Several studies on comparing the two approaches have been proposed in literature (see *e.g.* Gong and Sickles, 1992; Hjalmarsson *et al.* 1996; Sickles, 2005), also focusing attention on agriculture (Kalaitzandonakes and Dunn 1995; Sharma *et al.* 1997; Wadud and White, 2000; Minh and Long 2009). They have mostly investigated on differences between estimated technical efficiency scores and their distribution on the observed sample. Sharma *et al.* (1997) underlined that it is expected that the DEA efficiency scores would be less than those obtained under the specifications of stochastic frontier due to DEA method attributes any deviation of the data points from the frontier to inefficiency. However, empirical findings obtained from these studies confirm that the opposite may occur if the DEA frontier is fitted tightly to the sample data. Generally, the differences in the estimated results from two approaches could be mainly attributed to the different characteristics of the data, the choice of input and output variables, measurement and specification errors, as well as estimation procedures (Ruggiero 2007; Minh and Long 2009).

On the contrary, poor relevance has been given on comparison in measuring scale efficiency despite its important role in conditioning economic efficiency. Indeed, scale efficiency is a measure inherently relating to the returns to scale of a technology at any specific point of the production process (Førsund and Hjalmarsson 1979). It measures how close an observed plant is to the optimal scale, *i.e.* it describes the maximally attainable output for that input mix[2] (Frisch, 1965).

In our opinion, more effort should be produced in comparing estimated scale efficiencies calculated from SFA and DEA approaches[3]. More attempts need to be done especially in agriculture due to the fact that a great number of papers have estimated scale efficiency in this sector but in most of these studies, the measure is calculated using a DEA model (Karagiannis and Sarris 2005; Bravo-Ureta *et al.*, 2007). This is a relevant issue because differences in scale efficiencies interpretation and scale properties might derive form inherent differences between parametric and non parametric models. Therefore, it is expected to obtain some differences according to the methodology applied for estimating scale efficiencies in terms of scores and their distribution on the sample (Banker *et al.*, 1986; Førsund. 1992).

Efficiencies measures arisen from DEA are technology invariant due to the DEA method of enveloping the data for construction the frontier and observed units at both ends of the size distribution may be identified as efficient simply for lack of other comparable units. It means that real economies of scale at large (or small) units will be difficult to detect and it may be an identification problem whether scale inefficiency of technically efficient units is real or due to the specification on variable returns to scale and the method of enveloping the data (Førsund. 1992). On the other hand, since the statistical theory is well developed for the parametric approaches, SFA allows us to make statistical inferences about estimated scale efficiency (Kumbhakar and Tsionas,

[1] However, some authors have proposed models in which properties of SFA and DEA are integrated in order to overcame disadvantages of both methods (Ruggiero, 2004; Simar and Zeleneyuk, 2011).

[2] This definition substantially corresponds to Banker's (1984) concept of *most productive scale size* (MPSS) in the DEA context.

[3] Among the others, Banker *et al.* (1986); Ferrier and Lovell (1990), Bjurek *et al.* (1990) Førsund (1992) have compared scale efficiencies and scale properties obtained from parametric and non parametric approaches.

2008). DEA generally does not permit to make it because any statements regarding the statistical properties of estimated efficiency measures including scale efficiencies can be formulated[4]. Furthermore, DEA and other deterministic models attributes any deviation of each observation from the frontier to inefficiency, while SFA models allow us to separate inefficiency component from noise.
According to Orea (2002) and Karagiannis and Sarris (2005), the approach followed for the DEA is hardly transferable using a SFA approach with flexible functional forms for the production frontier. DEA calculates scale efficiency by dividing the technical efficiency estimated under the hypothesis of constant return to scale and technical efficiency estimated imposing variable return to scale technology. In case of parametric approach, hypothesis that variable returns to scale technology is enveloped from the constant returns to scale technology is weak by a theoretical point of view. Indeed, there is nothing to guarantee that the variable returns to scale technology is enveloped from the constant returns to scale technology in the parametric context.
A model for estimating scale efficiency within a parametric flexible stochastic approach was proposed by Ray (1998). Following this methodology, a scale efficiency measure is obtained from the estimated parameters of the production frontier function under the variable returns to scale hypothesis and from the estimated scale elasticity. Ray's (1998) model has the advantage of being easily tractable from the econometric point of view and being particularly suitable for a translog frontier function. In spite of these operational advantages, the model proposed by Ray (1998) has been scarcely adopted for estimating scale efficiency in agricultural studies.
The objective of this paper is to contribute in the existing literature providing a comparison between SFA and DEA approaches for estimating technical and scale efficiency. In particular both parametric and non parametric approaches were applied to estimate technical and scale efficiencies exhibited by the Italian citrus farming[5]. While several studies have compared technical efficiency estimates derived from parametric and non parametric approaches, this is one of the first attempts that aims to put on evidence possible difference in scale efficiency estimations in farming due to methods used, especially considering a stochastic specification of the production frontier in the parametric model. Regarding the parametric approach, a non-neutral production function model and the Ray (1998) model were applied to estimate technical and scale efficiency in the Italian citrus farming, respectively.

2. THE ITALIAN CITRUS FRUIT-GROWING SECTOR

Citrus fruit growing is one of the largest categories in the Italian vegetable and fruit sector. Since 2006, the value of production has amounted to more than 1 billion euro, accounting for about 10% of the total value of vegetables and fruits produced (Giuca 2008). Oranges comprise about 54% of citrus fruit production, whereas the contribution of lemons and tangerines to overall production (in terms of value) is equal to 17% and 19%, respectively.

[4] As reported by Kumbhakar and Tsionas (2008), however some progress have been made into the DEA context in terms of bootstrapping and statistical properties of DEA findings.
[5] More in depth discussions of empirical findings from DEA and SFA are reported in Madau (2010) and Madau (2011), respectively. Obviously, findings from both methods are showed and discussed also in this paper for better supporting discussion on comparison between DEA and SFA results.

The land area cultivated to citrus fruits corresponds to about 122,000 ha, while the number of farms is about 85,000 (Ismea 2008). Substantially, the farms are situated in the southern regions of Italy and, specifically, more than 70% of the farms and about 80% of cultivated land are located in only two regions: Sicily and Calabry. Since the early 1990s, however, land area covered by citrus fruits has decreased by about 30% (in 1990, it amounted to 184,000 ha) and the number of citrus farmers decreased by about 45% (about 170,000 in 1990). In this period, exports have slightly increased, while imports has grown sixfold (Giuca 2008).

Several reasons for this deterioration can be explored. First, the increasing competition in the world citrus fruit market has penalised Italian farmers because of structural and organisational problems that historically characterised the Italian citrus fruit sector. Specifically, Italian farms appear significantly small (on average, the area is 1.44 ha) and most of the citrus farms are located in less favourable areas where economic and productive alternatives are limited. Furthermore, despite the small size, many farms are fragmented in more plots of land, with evident implications on the ability to operate under efficient conditions.

These and other factors have contributed in the last few years to Italy's declining competitiveness and efficiency in the world citrus fruit market. Structural constraints seem to negatively affect the performance of the Italian sector and inhibit economic development of citrus farming. The detection of technical and scale efficiencies can offer us more information about the nature of these problems. If significant technical and/or scale inefficiency were found, this would indicate that structural problems prevent farm expansion and the rational use of technical inputs. An analysis of the relationship between technical and scale (in)efficiency would allow us to determine direction priorities - technical efficiency or scale efficiency oriented measures - in order to improve overall efficiency in the farms.

3. METHODOLOGICAL BACKGROUND

Both for non parametric and parametric calculation of scale efficiency a preliminary step is estimating the frontier function and the correspondent measures of technical efficiency. As well-known, technical efficiency is defined as the measure of the ability of a firm to obtain the best production from a given set of inputs (*output-increasing oriented)*, or as a measure of the ability to use the minimum feasible amount of inputs given a level of output (*input-saving oriented*) (Greene 1980; Atkinson and Cornwell 1994)[6]. This section illustrates how technical and scale efficiency output-oriented measures can be obtained from the DEA and the SFA models[7].

3.1 Non parametric estimation: Data Envelopment Analysis (DEA)

Data Envelopment Analysis (DEA) is a *non parametric* approach to estimate efficiency originally proposed by Charnes *et al.* (1978) and based on the Farrell's model (1957). DEA consents the estimation of efficiency in multi-output situations and without

[6] When firm operates in a constant returns to scale area the input and output-oriented measures coincide (Fare and Lovell 1978).

[7] For a more detailed description of the methodologies and the impact of such a choice on the empirical findings see Wadud and White (2000) and Bravo-Ureta *et al.* (2007).

assuming *a priori* functional form for frontier production (Roland and Vassdal, 2000). Therefore DEA assumes that the production function is unknown and solving a linear programming problem it calculates efficiency by comparing each production unit against all other units. The best practice frontier is represented by a piecewise linear envelopment surface. Therefore, TE scores arisen from DEA are invariant to technology, because obtained trough comparisons among an observation and each others and not with respect to an estimated frontier.
The discussion on DEA presented here is brief and concerns the *output-oriented* Constant Return to Scale (CRS) DEA and Variable Return to Scale (VRS) DEA. The *output-oriented* CRS DEA model for a single output is described below. TE is derived solving the following linear programming model (Ali and Seiford 1993):

$$\max_{\theta,\lambda} \quad \theta_i$$
$$\text{subject to} \quad \sum_{j=1}^{n} \lambda_j y_j - \theta_i y_i - s = 0$$
$$\sum_{j=1}^{n} \lambda_j x_{kj} + e_k = x_{ki}$$
$$\lambda_j \geq 0; \quad s \geq 0; \quad e_k \geq 0 \tag{1}$$

where θ_i is the proportional increase in output possible for the i-th DMU (Decision Making Unit that in this study is a farm), λ_j is an $N\times 1$ vector of weights relative to efficient DMUs, s is the output slack; and e_k is the k-th input slack. Banker *et al.* (1984) suggest to adapting the CRS DEA model in order to account for a variable returns to scale situation. Adding the convexity constraint $N1'\lambda = 1$, the model can be modified into VRS DEA[8].
The proportional increase in output which is possible is accomplished when output slack, s, becomes zero. A DMU results efficient when the values of θ_i and λ_i are equal to 1; and $\lambda_j = 0$. On the contrary, a DMU is inefficient when $\theta_i > 1$, $\lambda_i = 0$; and $\lambda_j \neq 0$. Solving (1) we can obtain a measure of TE that reflects "distance" between the observed and optimal output production for a certain inputs bundle:

$$TE_i = \frac{Y_i}{Y_i^*} = \frac{1}{\theta_i} \qquad 0 \leq TE_i \leq 1 \tag{2}$$

where Y_i and Y_i^* are the observed and maximum possible (optimal) output, respectively.
A measure of scale efficiency (SE) can be obtained by comparing TE^{CRS} and TE^{VRS} scores. Any difference between the two TE scores indicates there is scale inefficiency that limits achievement of an optimal (constant) scale:

$$TE_i^{CRS} = TE_i^{VRS} * SE_i \tag{3}$$

Therefore, it can be calculated as (Coelli 1996a):

[8] Banker *et al.* (1984) proposed this modification for the input-oriented model, but this constraint can be added also for output-oriented VRS models.

$$\mathrm{SE}_i = \frac{\mathrm{TE}_i^{CRS}}{\mathrm{TE}_i^{VRS}} \qquad 0 \leq \mathrm{SE}_i \leq 1 \tag{4}$$

where $\mathrm{SE}_i = 1$ indicates full scale efficiency and $\mathrm{SE}_i < 1$ indicates presence of scale inefficiency.

However, a shortcoming of the SE score is that it does not indicate if a farm is operating under increasing or decreasing return to scale. This is resolvable by simply imposing a *non-increasing return of scale* (NIRS) condition in the DEA model, *i.e.* changing the convexity constraint $N1'\lambda = 1$ of the DEA VRS model into $N1'\lambda \leq 1$. If $\mathrm{TE}^{\mathrm{NIRS}}$ and $\mathrm{TE}^{\mathrm{VRS}}$ are unequal, then farms operate under increasing return to scale (IRS); if they are equal a decreasing return to scale (DRS) exists.

3.2 Parametric Estimation: Stochastic Frontier Analysis (SFA)

SFA was originally and independently proposed by Aigner *et al.* (1977) and Meeusen and van der Broeck (1977). In these models, the production frontier is specified which defines output as a stochastic function of a given set of inputs. The presence of stochastic elements makes the models less vulnerable to the influence of outliers than with deterministic frontier models. It concerns that the error term ε may be separated in two terms: a random error and a random variable explanatory of inefficiency effects:

$$y_{it} = f(x_{it}, t; \beta) \bullet \exp \varepsilon \tag{5a}$$

$$\varepsilon = (v_{it} - u_{it}) \qquad i = 1,2,\ldots N \qquad t = 1,2,\ldots T \tag{5b}$$

where y_{it} denotes the level of output for the i-th observation at year t; x_{it} is the row vector of inputs; t is the time index, β is the vector of parameters to be estimated; $f(\cdot)$ is a suitable functional form for the frontier (generally Translog or Cobb-Douglas); v_{it} is a symmetric random error assumed to account for measurement error and other factors not under the control of the firm; and u_{it} is an asymmetric non-negative error term assumed to account for technical inefficiency in production.

The v_i's are usually assumed to be independent and identically distributed $N(0, \sigma_v^2)$ random errors, independent of the u_{it}'s that are assumed to be independent and identically distributed and with truncation (at zero) of the normal distribution $|N(0, \sigma_u^2)|$. The Maximum Likelihood Estimation (MLE) of (5) allows us to estimate the vector β and the variance parameters $\sigma^2 = \sigma_u^2 + \sigma_v^2$ and $\gamma = \sigma_u / \sigma_v$; where $0 \leq \gamma \leq 1$. The TE measure is obtained by the ratio of y_{it} to the maximum achievable level of output:

$$\mathrm{TE} = \frac{y_{it}}{y^*} = \exp(-u_{it}) \tag{6}$$

where y* is the output that lies on the frontier. Furthermore, assuming a semi-normal distribution for u_{it} and according to Jondrow *et al.* (1982), the degree of technical efficiency of each firm could be estimated.

In order to estimate inefficiency effects, some authors proposed a *two-stage* method, in which the first stage consists in technical efficiency estimation using a SFA approach, and the second stage involves the specification of a regression model that relaxes technical efficiency with some explanatory variables (Pitt and Lee 1981; Kalirajan 1982; Parikh and Shah 1994).

One-stage SFA models in which the inefficiency effects (u_i) are expressed as a function of a vector of observable explanatory variables were proposed by Kumbhakar *et al.* (1991), Reisfschneider and Stevenson (1991), Huang and Liu (1994). In this model, all parameters ま frontier production and inefficiency effects ま are estimated simultaneously. This approach was adapted by Battese and Coelli (1995) to account for panel data. They proposed an *one-stage* approach where the functional relationship between inefficiency effects and the firm-specific factors is directly incorporated into the MLE. The inefficiency term u_{it} has a truncated (at zero) normal distribution with mean m_{it}:

$$u_{it} = m_{it} + W_{it} \tag{7a}$$

where W_{it} is a random error term which is assumed to be independently distributed, with a truncated (at -m_{it}) normal distribution with mean zero and variance 《 (*i.e.* W_{it} ち- z_{it} such that u_{it} is non-negative).

The mean m_{it} is defined as:

$$m_{it} = Z\,(z_{it}, \delta) \qquad i = 1,2,\wedge.N \qquad t = 1,2,\wedge.T \tag{7b}$$

where Z is the vector (Mx1) of the z_{it} firm-specific inefficiency variables of inefficiency; and δ is the (1xM) vector of unknown coefficients associated with z_{it}. So we are able to estimate inefficiency effects arisen from the z_{it} explanatory variables[9].

Orea (2002) argues that the *non parametric* approach difficultly can be directly transferred into a parametric approach in order to calculate scale efficiency. Indeed when parametric approach is used, hypothesis that VRS technology is enveloped from CRS technology is weak by a theoretical point of view.

As mentioned above, Ray (1998) proposed a model in which scale efficiency can be calculated from the estimated parameters of the production frontier and from scale elasticity estimations. For a translog frontier function:

$$\ln y_{it} = \text{伲} + \sum_{j=1}^{n} \text{伲}_j \ln x_{jit} + \frac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{l} \text{伲}_{jk} \ln x_{jit} \cdot \ln x_{kit} + (v_{it} - u_{it}) \tag{8}$$

and assuming an output-oriented approach for the technical efficiency estimation, scale elasticity at farm-specific input bundle is equal to:

$$E_{it} = \sum_{j=1}^{n}\left(\beta_j + \sum_{k=1}^{l}\beta_{jk}x_{kit} + \beta_{ji}t\right) \tag{9}$$

Remanding to Ray (1998) for a more detailed description of the methodology, it follows that the output-oriented scale efficiency (SE^O) corresponds to:

---

[9] According to Battese and Corra (1977), Battese and Coelli (1995) suggest to replacing the parameter 似 with $\gamma = \sigma_u^2 / \left(\sigma_u^2 + \sigma_v^2\right)$ because of it can be searched between zero and one and this property allows us to obtain a suitable starting value for an iterative maximisation process.

$$\mathrm{SE}_{it}^{O} = \exp\left[\frac{(1-\mathrm{E}_{it})^2}{2\beta}\right] \tag{10}$$

where:

$$\beta = \sum_{j=1}^{n}\sum_{k=1}^{l}\beta_{jk} \tag{11}$$

with β that is assumed to be negative definite as to guarantee that $0 < \mathrm{SE}_{it}^{O} \leq 1$[10].

This output-oriented scale efficiency measures the role of scale in conditioning technical efficiency. Scale efficiency reflects the relative output expansion by producing at optimal scale on the frontier for the observed factor proportions of a firm whose technical inefficiency has been eliminated (Karagiannis and Sarris 2005). In other terms, following the Frisch's definition, scale efficiency measures the distance to full efficient scale after moving a production unit to the frontier in the vertical direction.

As reported by Ray (1998), scale efficiency (10) and scale elasticity (9) are both equal to one only at an MPSS, *i.e.* where constant returns to scale prevails. Elsewhere they differ and SE is <1 irrespective of whether E_{it} is greater than or less than unity. It means that the magnitude of scale elasticity reveals nothing about the level of SE at the points different by the MPSS.

On the basis of the definition of scale efficiency measured by (10), the sub-optimal scale is associated with increasing returns to scale. When $\mathrm{E}_{it} > 0$ (increasing returns to scale) then SE increases with an increase in output and the optimal scale should be reached expanding the observed output level. *Vice versa*, output should be contracted to reach the optimal scale when a plan operates in a decreasing returns to scale (supra-optimal) area ($\mathrm{E}_{it} < 0$)[11].

In order to explain scale efficiency differentials among plans, Karagiannis and Sarris (2005) used a *two-stage* approach. At the first stage, SEs are estimated using the formula (10) and successively, at the second stage, the SE scores are regressed against a set of explanatory variables. Following the procedure proposed by Reinhard *et al.* (2002), these authors in the second stage used a MLE technique to estimate this stochastic frontier regression model:

$$\ln \mathrm{SE}_{it}^{O} = m_{it} + \varepsilon_{it} \qquad \text{with} \tag{12a}$$

$$m_{it} = \mathrm{Z}(z_{it}, \rho) \qquad \text{and} \tag{12b}$$

$$\varepsilon_{it} = (v_{it}^{*} - u_{it}^{*}) \qquad i = 1,2,\ldots N \qquad t = 1,2,\ldots T \tag{12c}$$

where z_{it} represents the same set of variables used in the inefficiency model (9), ρ are the parameters to be estimated, ε_{it} is the error term composed by v_{it}^{*} that represents the

[10] Negative definiteness of β is a sufficient but not necessary condition (Ray, 1998).

[11] Among the advantages of this measure, Ray (1998) argues that "*this (scale efficiency measure) should make findings from econometric models more directly comparable with the evidence from nonparametric DEA models, where scale efficiency measures are routinely reported* (p. 193)".

statistical noise (independently and identically distributed with $N\ (0, \sigma^2_{v*})$ random variable truncated at $-m_{it}$) and by u^*_{it} that represents the conditional scale inefficiency remaining even after variation in the z_{it} has been accounted for ($u^*_{it} \sim N\ (-m_{it},\ \sigma^2_{u*})$).
The *two-stage* approach in the SFA models has been criticized by several authors because it is inconsistent in it's assumption regarding independence of the inefficiency effects (Battese and Coelli 1995; Kumbhakar and Lovell 2000). With specific reference to technical efficiency, the rationale underlying is that the specification of the regression of the second stage - in which the estimated technical efficiency scores are assumed to have a functional relationship with the explanatory variables - conflicts with the assumption that u_i's are independently and identically distributed (TE is the dependent variable in the second stage procedure).
However, as underlined by Reinhard *et al.* (2002), a *two-stage* procedure can consistently be used as long as the efficiency scores are calculated from the first-stage parameter estimates, instead of being estimated econometrically at the first stage. In the case of the procedure illustrated above for computing scale efficiency effects, no such assumption is made with respect to the dependent variable SE because SE scores are obtained from the parameter estimates and the estimated values of scale elasticity. Thus, Reinhard *et al.* (2002) recommended application of the *two-stage* procedure for estimating scale efficiency effects.

4. DATA AND THE EMPIRICAL MODELS

Data were collected on a balanced panel data of 107 Italian citrus farms. All the selected farms participated in the official Farm Accountancy Data Network (FADN) during the period 2003-2005 and they are specialized in citrus fruit-growing (more then 2/3 of farm gross revenue arises from citrus production). Farms with less than two European Size Units (ESU) were excluded from the sample[12]. Therefore, both non parametric and parametric analyses are based on a total of 321 observations (see Table 1 for summary statistics about farms).

4.1 DEA model

We applied both CRS and VRS DEA models in order to calculate scale efficiency. Estimation of technical and scale efficiency was carried out performing separated analysis for each considered year.
The dependent variable (Y) represents the output and it is measured in terms of gross revenue from the *i*-th farm. The aggregate inputs, included as variables of the production function, are 1) X_1 the total *land* area (hectares) devoted to citrus fruit-growing by each farm; 2) X_2 the expenditure (euro) for seeds, fertilizers, water and other *variable inputs* used in the citrus fruits-growing; 3) X_3 the value (euro) of *machineries* used in the farm; 4) X_4 the value (euro) of *capital* (amount of fixed inputs such as buildings and irrigation plant, except for machineries); 5) X_5 the expenditure (euro) for *other inputs*, consisting in fuel, electric power, interest payments, taxes, etc.; 6) X_6 the total amount (annual working hours) of *labour* (including family and hired workers);

[12] In FADN, ESU indicates the farm economic size. ESU is defined on the basis of farm potential gross value added (Total Standard Gross Margin).

Regarding the machineries and capital variables, they were measured in terms of annual depreciation rate so to have a measure of annual utilization, on average, of the capital stock[13]. All variables measured in monetary terms were converted into 2003 constant euro value.

Table 1 - Summary statistics for citrus farms in the sample (mean values)

Variable	**2003**	**2004**	**2005**
Gross revenue (*euro*)	54,508	53,861	56,542
Land area (*hectares*)	13.21	13.26	13.41
Expenditure for seeds, fertilizers, etc. (*euro*)	3,878	4,866	5,066
Machineries (*annual depreciation rate, euro*)	2,395	2,489	2,962
Capital (*annual depreciation rate, euro*)	5,050	5,182	5,052
Other expenditures (*euro*)	1,240	939	1,322
Labour (annual working *hours*)	2,785	2,814	2,772
Age of farm owner	59.1	59.7	60.7
Size (*ESU**)	4.7	4.7	4.7
Altitude (*metres*)	104	104	104
Number of plots of land	1.6	1.7	1.7

** ESU = European Size Units*

Furthermore, a set of explanatory variables of efficiency were selected in order to evaluate their effect on technical and scale efficiency. More precisely, individual estimated technical and scale efficiency were regressed to: 1) Z_1 the *age* of the farm owner; 2) Z_2 a *dummy* variable that reflects the *size* of the farm measured in terms of ESU that can assume a value involved from 3 to 7[14]; 3) Z_3 the variable *altitude* that reflects the average altitude (in metres) by each farm; 4) Z_4 the *number of plots* of land in which farm is fragmentized; Z_5 a *dummy* variable that reflects the placement (or not) of each farm in a *Less-favoured area* such as defined by the EEC Directive 75/268 (0 = Less-Favoured zone; 1 = non Less-favoured zone); Z_6-Z_{11} that represent a set of *dummy* variables indicating the regional location of farms (R_{cam} = Campany; R_{cal} = Calabria; R_{apu} = Apulia; R_{bas} = Basilicata; R_{sic} = Sicily; R_{sar} = Sardinia).

Variables such as age of farmers, farm size, and regional location have been widely used in the efficiency analyses applied to agriculture. The first is generally used as a proxy of farmer skills, experience, and learning-by-doing (the rationale is that the expected level of efficiency increases with experience). The second was implemented to evaluate the role of farm economic size in conditioning efficiency (a positive sign is expected, *i.e.* efficiency tends to increase in larger farms). The third serves to estimate the presence of territorial and geographic variability that may affect efficiency.

Altitude and location in a less-favoured area are variables used in some efficiency analysis to account for geoclimatic and socioeconomic heterogeneities (Karagiannis and Sarris 2005; Madau 2007). On the other hand, the number of lots has not been a variable generally employed in the efficiency analyses in agriculture. But, in our opinion and as highlighted above, it could be significant in conditioning both farm technical and scale efficiencies in the Italian citrus farming. Indeed, the subdivision of the farm land area

[13] As underlined by Madau (2008), value of capital goods is estimated in different ways into the efficiency analyses. Some authors have considered the total amount of value, whereas other authors have expressed capital in terms of annual capacity utilization. In this case, the capital measure depends on the adopted criteria for calculate capacity utilization.

[14] Any observed farm exhibits an ESU Class 8 or 9.

into more plots of land could be an obstacle toward achieving full (technical and scale) efficiency on the part of farmers.

4.2 SFA model

We assumed a Translog functional form as frontier technology specification for the citrus farms. The adopted model corresponds to the Huang and Liu (1994) non-neutral production function model applied on panel data, which assumes that technical efficiency depends on both the method of application of inputs and the intensity of input use (Karagiannis and Tzouvelekas 2005)[15]. It means that the inefficiency term u_{it} explained by (7) is equal to:

$$u_{it} = \delta_0 + \sum_{i=1}^{N} \delta_{it}\, z_i + \sum_{m=1}^{M} \delta_m \ln x_{mit} + W_{it} \qquad i = 1,2,\wedge.N \qquad t = 1,2,\wedge.T \qquad (13)$$

The Translog stochastic function production model is specified as formula (8) and involves seven variables: the variables X_1-X_6 correspond to the same bundle of inputs selected for the DEA model and X_7 is a variable that represents the *time* (year) and it can assume value equal to 1 (2003), 2 (2004) or 3 (2005).
In the inefficiency model (13) we found the same set of explanatory variables used for the DEA model. In addition, according to the non-neutral model proposed by Huang and Liu (1994), (in)efficiency is expected to depend by the inputs used in the production. Therefore, the same pool of variables (included time) used to describe the frontier function production (x_{it}) were included in the inefficiency model.
Finally, applying the second-stage regression (12), scale efficiency effects were calculated using the same bundle of variables used for the technical efficiency effects model, with the exception of inputs that describe the frontier production.

5. ANALYTICAL FINDINGS

5.1 Estimated results from non parametric approach

Technical efficiency scores arisen by application of the DEA model were estimated using the DEAP 2.1 program created by Coelli (1996a).
Results indicate that output-efficiency technical efficiency obtained for the CRS and VRS frontiers are, on average, equal to 0.623 and 0.711, respectively (Table 2). These measures were calculated as averages on the triennial period of observation (2003-2005). Considering the latter measure まso called むpure efficiencyめbecause devoid of scale efficiency effects まand since technical efficiency scores are calculated as an output-oriented measure, the results imply that citrus fruits-growing farmers would be able to increase output by about 30% using their disposable resources more effectively (at the present state of technology).
Scale efficiency is calculated applying formula (4). The mean scale efficiency for the Italian citrus fruits producers in Italy is equal to 0.894. It means that adjusting the scale of the operation, citrus farms could improve their efficiency by 10.6%.

---

[15] Substantially, this model corresponds to the Battese and Coelli (1995) model with a non-neutral specification for the production frontier function.

Imposing the NIRS condition, we found that the most of the farms exhibit an increasing returns to scale (Table 3). Of the 107 farms, 71 (66.3%) show increasing (sub-optimal) returns to scale, 22 (20.6%) show constant (optimal) returns to scale, and 14 (13.1) show decreasing (supra-optimal) returns to scale. Therefore, this implies that scale inefficiency is mainly due to the farms operating under a sub-optimal scale, - *i.e.* farms where their output levels are lower than optimal levels and they should be expanded to reach the optimal scale. It was found that typology of returns to scale do not vary in each farms during the time of observation.

Tab.2 Estimated technical efficiency and scale efficiency using DEA

Efficiency	**TE^{CRS}**	**TE^{VRS}**	**SE**
Mean*	**0.623**	**0.711**	**0.894**
s.d.	*0.242*	*0.256*	*0.163*
Min	0.226	0.257	0.287
Max	1.000	1.000	1.000

* calculated on the basis of a triennial period

In the most of sub-optimal scale farms, scale efficiency is sensitively low (the average SE in this group is less than 0.700). On the contrary, supra-optimal scale farms appear more efficient, in terms of ability to operate under an adequate scale (mean SE equal to 0.934). Both TE and SE scores vary substantially across farms. To explain some of these variations, the efficiency scores were regressed on the farm-level characteristics. A Tobit regression model was used, since the efficiencies vary from zero to unity.

Table 3 Scale efficiency and returns to scale from DEA

	Observations		**Scale Efficiency**
	n.	**%**	
Total sample (mean)	321	100	0.894
Supra-optimal scale	42	13.1	0.934
Optimal scale	66	20.6	1.000
Sub-optimal scale	213	66.3	0.692

Age of farmer is negatively related to technical efficiency, but the estimated coefficient is not statistically significant (Table 4). Farm size is positively related to efficiency level. The results indicate that improvement of technical efficiency depends, among the others, on citrus farms attaining an adequate size (magnitude is equal to 0.036). Altitude slightly affects technical efficiency, while, as expected, the number of lots is negatively correlated to technical efficiency.

The findings imply that technical efficiency tends to decrease in the case of partitioning farms in more plots. The magnitude of this effect is 0.056, indicating that the presence of a plurality of lots affects sensitively efficiency from a technical point of view. Furthermore, farms situated in less-favoured areas tend to be more inefficient than those located in normal zones (magnitude is equal to -0.018). Finally, the fact that all the *dummy* variables reflecting geographical location of citrus farms are not significant by a statistical point of view suggests that this factor should not be considered an explanatory factor of technical efficiency variability.

Table 5 shows results on relationship between scale efficiency and possible sources of inefficiency. Age of farmers is negatively related to scale efficiency, even if magnitude is not sizeable (-0.003). It implies that citrus farms managed by younger farmers should be more scale efficient than farms managed by older farmers. Farm size might positively affect scale efficiency. It is the factor that contributes the most to conditioning scale efficiency (magnitude is equal to 0.042). This suggests that large-sized farms tend to have higher scale efficiency than small-scale farms. Altitude and location in a less-favoured area are not significant variables by a statistical point of view. On the contrary, the number of plots of land represents the second most important factor in the order of importance that affects scale efficiency (-0.040). The consistent negative sign of the estimated coefficient indicates that in-farm land fragmentation might be a relevant structural constraint to achieving an adequate scale efficiency by part of citrus farmers.

Tab.4 耒Technical efficiency effects from DEA

Variables		***Coefficient***	**s.e.**
Constant	δ_0	0.607 ***	0.226
Age	δ_1	-0.001	0.002
Size	δ_2	0.036 *	0.023
Altitude	δ_3	-0.001	0.003
N. of plots of lands	δ_4	-0.056 *	0.032
Less-favoured zones	δ_5	-0.018 *	0.011
Campany	δ_6	0.110	0.169
Calabry	δ_7	0.082	0.131
Apulia	δ_8	0.225	0.174
Basilicata	δ_9	0.049	0.143
Sicily	δ_{10}	0.090	0.127
Sardinia	δ_{11}	redundant	

*** = significance at 1% level ** = significance at 5% level * = significance at 10% level

Tab.5 耒Scale efficiency effects from DEA

Variables		***Coefficient***	**s.e.**
Constant	δ_0	0.756 ***	0.145
Age	δ_1	-0.003 **	0.001
Size	δ_2	0.042 **	0.020
Altitude	δ_3	-0.001	0.002
N. of plots of lands	δ_4	-0.040 **	0.018
Less-favoured zones	δ_5	0.005	0 012
Campany	δ_6	0.269 *	0.167
Calabry	δ_7	-0.002	0.089
Apulia	δ_8	-0.045	0.102
Basilicata	δ_9	-0.093 *	0.049
Sicily	δ_{10}	-0.179 **	0.082
Sardinia	δ_{11}	redundant	

*** = significance at 1% level ** = significance at 5% level * = significance at 10% level

Finally, the findings show that there are statistically significant differences in scale efficiency between farms located in different geographical regions of Italy, implying that location sensitively influences scale efficiency.

Specifically, farms that operate in Campany should tend to be more scale efficient than the others (the variable that reflect location in Sardinia is redundant).

5.2 *Estimated results from parametric approach*

Parameters for the function and inefficiency model were estimated simultaneously. ML estimation was obtained using the computer program FRONTIER 4.1, created by Coelli (1996b). ML estimates for the preferred frontier model were obtained after testing various null hypotheses in order to evaluate suitability and significance of the adopted model.As testing procedure we adopted the *Generalised likelihood-ratio test*, which allows us to evaluate a restricted model with respect to the adopted model (Bohrnstedt and Knoke, 1994)[16].

The test was applied in order to estimate the more suitable to the data functional form o the frontier (Transolg or Cobb-Douglas specification; non-neutral or neutral specification), presence of inefficiency effects, nature of inefficiency effects, presence of an intercept in the inefficiency model, presence of farm-specific factors, presence of regional effects and, finally, presence of Age and Altitude effects (because of poor estimated statistical significance).

Table 6 reports the results of these t-tests and in the light of these the model was estimated to obtain the preferred form. MLE for the more appropriate model are shown, as reported above, in the Table 7.

Since the Translog function takes into account also interaction among involved inputs, the production elasticities were computed using the traditional formula for the estimation of the elasticity of the mean output with respect to the *k*-th input (except for the time variable):

$$\frac{\partial \ln E(Y)}{\partial \ln(x_k)} = \beta_k + 2\beta_{kk}\, x_{ki} + \sum_{j \neq k} \beta_{kj}\, x_{ji} \qquad (14)$$

Application of (14) indicates that, at the point of approximation, the estimated function satisfies the *monotonicity* (all parameters show a positive sign) and *diminishing marginal productivities* (magnitude is lower than unity for each parameter) properties (Table 8).

The estimated production elasticities suggest that land is the foremost important input followed by expenditure for seeds and other technical inputs, labour, and machineries. It means that enlargement of the land area would affect significantly farm productivity.

Specifically holding all other inputs constant, an increase of 1% in land area would result in a 0.47% increase in output. According to other research findings, the high elasticity of the land area is not surprising in presence of small size farms because this factor could be considered a quasi-fixed input (Alvarez and Arias 2004; Madau 2007).

Except for land area, these findings suggest that production of Italian citrus farms is sensitively elastic with respect to these factors, which should allow farmers to easily vary their own use level in the short run - elasticity of seeds (and other technical inputs) and labour is equal to 0.27 and 0.18, respectively - while the other quasi-fixed inputs (capital and machinery) affect productivity less (elasticity equal to 0.04 and 0.11,

[16] The statistic test λ has approximately a chi-square (or a mixed-square) distribution with a number of degrees of freedom equal to the number of parameters (restrictions), assumed to be zero in the null-hypothesis. When λ is lower than the correspondent critical value (for a given significance level), we cannot reject the null-hypothesis.

respectively). The time variable shows a negative sign, but the magnitude is not relevant, implying that time does not significantly affect production.

Table 6 まHypothesis testing for the adopted model for SFA

Restrictions	**Model**	**L(H_0)**	λ	***d.f.***	$\chi^2_{0.95}$	**Decision**
None	Translog, non neutral	-89.81				
$H_0 : \bar{\ }_{m} = 0$	Neutral	-100.18	20.74	6	12.59	Rejected
$H_0 : \beta_{ij} = 0$	Cobb-Douglas	-168.34	157.06	21	32.67	Rejected
$H_0 : \gamma = \delta_0; \delta_r; \delta_m = 0$	No inefficiency effects	-120.45	61.28	6	11.91*	Rejected
$H_0 : \gamma = \delta_0; \delta_m = 0$	No stochastic effects	-101.80	23.98	9	19.92*	Rejected
$H_0 : \delta_0 = 0$	No intercept	-90.44	1.26	1	3.84	Not rejected
$H_0 : \delta_r; \delta_m = 0$	No firm-specific factors	-118.76	57.90	11	19.68	Rejected
$H_0 : \delta_{6 \wedge .} \delta_{11} = 0$	No *Regional* effects	-94.88	10.14	6	12.59	Not rejected
$H_0 : \delta_{1;}\ \delta_3 = 0$	No *Age* and *Altitude* effects	-92.65	5.68	2	5.99	Not rejected

* Critical values with asterisk are taken from Kodde and Palm (1986). For these variables the statistic λ is distributed following a mixed χ^2 distribution.

Returns to scale were found to be clearly increasing (1.144). Therefore, the hypothesis of constant returns to scale is rejected. It means that citrus farmers should enlarge the production scale by about 14%, on average, in order to adequately expand productivity, given their disposable resources.

As to the estimated technical efficiencies, the analysis reveals that, on average, citrus farms are 71% efficient in using their technology (Table 7). Since technical efficiency scores are calculated as an output-oriented measure, the results imply that farmers would be able to increase output by about 30% using their disposable resources more effectively (at the present state of technology).

The estimated ratio-parameter γ is significant (for 去= 0.01) and it indicates that differences in technical efficiency among farms is relevant in explaining output variability in citrus fruits-growing (1/3 of the variability on the whole). Estimation of this parameter 伢 suggests that about 58% of the general differential between observed and best-practice output is due to the existing difference in efficiency among farmers. Therefore, technical efficiency might play a crucial role into the factors affecting productivity in the citrus farming.

Empirical findings concerning the sources of efficiency differentials among farms are presented in Table 7. Farm size is positively related to efficiency level. The results indicate that improvement of technical efficiency strongly depends on citrus farms attaining an adequate size (magnitude is equal to 0.495). Specifically, farm size increase should affect positively both productivity (returns to scale more than unity) and efficiency (negative sign of *Size* variable).

As expected, the number of lots is negatively correlated to technical efficiency, implying that technical efficiency tends to decrease in the case of partitioning farms in more plots, also if the magnitude of this effect is low (0.014) Finally, farms situated in less-favoured areas tend to be more inefficient than those located in normal zones (0.012)[17].

[17] Similar results were found by Madau (2007)

Tab. 7a – ML Estimates for SFA parameters and for TE (preferred model) - continue

Variables	**Parameter**	**Coefficient**	**s.e**
	FRONTIER MODEL		
Constant	β_0	-0.608	*0.139*
Land Area	β_1	-1.827	*0.422*
Expenditure for seeds, fertilizers, etc.	β_2	1.515	*0.463*
Machineries	β_3	1.662	*0.378*
Capital	β_4	-0.526	*0.453*
Other expenditures	β_5	0.193	*0.287*
Labour	β_6	0.576	*0.569*
Year	β_T	-1.697	*0.696*
(Land Area) x (Land Area)	β_{11}	0.052	*0.037*
(Land Area) x (V. expenditure)	β_{12}	0.102	*0.038*
(Land Area) x (Machineries)	β_{13}	0.085	*0.028*
(Land Area) x (Capital)	β_{14}	0.021	*0.036*
(Land Area) x (O. expenditures)	β_{15}	0.010	*0.035*
(Land Area) x (Labour)	β_{16}	0.060	*0.060*
(Land Area) x (Year)	β_{1T}	-0.111	*0.053*
(V. expenditure) x (V. expenditure)	β_{22}	0.032	*0.031*
(V. expenditure) x (Machineries)	β_{23}	0.016	*0.026*
(V. expenditure) x (Capital)	β_{24}	-0.053	*0.037*
(V. expenditure) x (O. expenditures)	β_{25}	0.099	*0.033*
(V. expenditure) x (Labour)	β_{26}	-0.328	*0.069*
(V. expenditure) x (Year)	β_{2T}	-0.011	*0.055*
(Machineries) x (Machineries)	β_{33}	0.074	*0.017*
(Machineries) x (Capital)	β_{34}	-0.088	*0.033*
(Machineries) x (O. expenditures)	β_{35}	-0.125	*0.035*
(Machineries) x (Labour)	β_{36}	-0.198	*0.051*
(Machineries) x (Year)	β_{3T}	-0.020	*0.035*
(Capital) x (Capital)	β_{44}	0.030	*0.024*
(Capital) x (O. expenditures)	β_{45}	0.085	*0.024*
(Capital) x (Labour)	β_{46}	0.046	*0.058*
(Capital) x (Year)	β_{4T}	0.076	*0.051*
(O. expenditures) x (O. expenditures)	β_{55}	0.029	*0.021*
(O. expenditures) x (Labour)	β_{56}	-0.137	*0.072*
(O. expenditures) x (Year)	β_{5T}	0.089	*0.048*
(Labour) x (Labour)	β_{66}	0.228	*0.076*
(Labour) x (Year)	β_{6T}	0.141	*0.065*
(Year) x (Year)	β_{TT}	0.026	*0.076*
	INEFFICIENCY MODEL		
Constant	δ_0	-	-
Age	δ_1	-	
Size	δ_2	-0.495	*0.087*
Altitude	δ_3	-	-
Number of plots of land	δ_4	0.014	*0.031*
Less-favoured zones	δ_5	0.012	*0.010*
Campany	δ_6	-	-
Calabria	δ_7	-	-
Apulia	δ_8	-	-
Basilicata	δ_9	-	-
Sicily	δ_{10}	-	-
Sardinia	δ_{11}	-	-
Land Area	δ_{SUP}	-0.679	*0.147*
Expenditure for seeds, fertilizers, etc.	δ_{SV}	0.359	*0.105*
Machineries	δ_{QM}	-0.043	*0.062*
Capital	δ_{QC}	0.068	*0.110*
Other expenditures	δ_{AS}	0.319	*0.149*
Labour	δ_{LAV}	-0.740	*0.214*
Year	δ_T	0.091	*0.135*

Tab. 7b 夫ML Estimates for SFA parameters and for TE (preferred model)

Variables	Parameter	Coefficient	s.d.
	VARIANCE PARAMETERS		
σ^2	σ^2	0.127	*0.016*
γ	γ	0.333	*0.131*
γ*	γ*	0.579	
Log-likelihood function		-92.66	
	TECHNICAL EFFICIENCY		
Mean			**0.710**
s.d			*0.266*
Maximum			1.000
Minimum			0.060

Regarding the relationship between technical efficiency and technical inputs, ML estimation shows that all inputs have a significant part to play in determining efficiency (Table 7). Land area, labour, and machinery carry a negative sign, implying that an increase in each variable positively affects technical efficiency.
Finally, the empirical findings suggest that farmers tend to become less efficient over time even if the magnitude is really low (0.091)[18].

Tab. 8 夫Estimated elasticities and returns to scale from SFA

Input	Elasticity	s.d.
Land area	0.466	*0.219*
Expenditure for seeds, fertilizers, etc	0.265	*0.146*
Machineries	0.112	*0.101*
Capital	0.037	*0.050*
Other expenditures	0.080	*0.102*
Labour	0.182	*0.073*
Returns to scale	**1.144**	***0.372***
Time	-0.001	*0.145*

Scale elasticities and scale efficiencies were estimated applying formulas (9) and (10). Table 9 shows that the average scale efficiency is 81.8%. It implies that observed farms could have further increased their output by about 18% if they had adopted an optimal scale. Results also indicate that about 80% of the observations exhibit increasing returns to scale. They operate under a suboptimal scale, *i.e.,* their output levels are lower than optimal levels and they should be expanded to reach the optimal scale. In these farms, scale efficiency is sensitively lower than the average (77.5%) and the average scale elasticity is abundantly upper than unity (1.237).
On the other hand, only about 6% of the observations are characterised by operating under an optimal scale, while about 15% of the panel reveals decreasing returns to scale. The relationship between scale efficiency and farm size seems to be confirmed by analytical results on the scale efficiency effects (see Table 11 below). These were obtained from application of (12) to the estimated data. The original proposed model 夫 the second-stage regression of the scale efficiency scores to the variables described

18 We also calculated efficiency scores for each year separately in order to assess if technical change exists over the observed period. However findings arisen from application of SFA to cross-sectional data year-by-year suggest that no technical progress exists because of not significant differences between the estimated annual technical efficiency scores(on average).

above - was tested using the *Generalised likelihood-ratio test* procedure in order to evaluate if a restricted model is preferable. Specifically three tests were applied concerning hypotheses on presence of intercept in the inefficiency effects, role of the regional areas in conditioning the farm scale inefficiency and presence of the Less-favoured area parameter, respectively. On the basis of the *t*-test results (reported in Table 10), we estimated the preferred model that is different from the proposed one for the absence of the intercept and the Less-favoured area variable. Estimated findings of scale inefficiency effects are reported in Table 11.

Tab. 9 表Estimated scale efficiency and scale elasticity from SFA

	Observations		Scale efficiency	Scale elasticity
	n.	%		
Total sample (mean)	321	100	**0.818**	**1.144**
s.d			0.213	0.416
Maximum			1.000	1.588
Minimum			0.012	0.662
Supra-optimal scale	47	14.7	0.978	0.897
Optimal scale	19	5.9	1.000	1.000
Sub-optimal scale	225	79.4	0.775	1.237

Table 10 表Hypothesis testing for the scale efficiency effects model from SFA

Restrictions	Model	L(H_0)	λ	*d.f.*	$\chi^2_{0.95}$	Decision
None	Translog, non neutral	123.92				
$H_0 : \delta_0 = 0$	No intercept	123.92	0.01	1	3.84	Not rejected
$H_0 : \delta_6 \ldots \delta_{11} = 0$	No *Regional* effects	114.71	18.42	6	12.59	Rejected
$H_0 : \delta_5 = 0$	No *Less-favoured area* effects	112.88	2.08	1	3.84	Not rejected

Farm size is the factor that contributes the most to conditioning positively scale efficiency (magnitude is equal to 0.040). This suggests that large-sized farms tend to have, as expected, higher scale efficiency than small-scale farms.

Table 11 表Scale efficiency effects (preferred model) from SFA

Variables	Parameter	Coefficent	s.e
Constant	δ_0	-	-
Age	δ_1	0.006	*0.001*
Size	δ_2	0.040	*0.017*
Altitude	δ_3	0.019	*0.024*
Number of plots of land	δ_4	-0.030	*0.014*
Less-favoured zones	δ_5	-	-
Campany	δ_6	0.044	*0.013*
Calabry	δ_7	0.002	*0.009*
Apulia	δ_8	-0.016	*0.050*
Basilicata	δ_9	-0.011	*0.057*
Sicily	δ_{10}	0.055	*0.061*
Sardinia	δ_{11}	0.051	*0.053*
Year	δ_T	-0.066	*0.099*

The number of plots of land represents the second most important factor in the order of importance that affects scale efficiency (-0.030). The consistent negative sign of the

estimated coefficient indicates that in-farm land fragmentation might be a relevant structural constraint to achieving an adequate scale efficiency by part of citrus farmers.The low magnitude (0.006) of the farmers' age parameter suggests that this variable has little influence on the observed efficiency differentials. In other words, older and more experienced farmers tend to be more scale efficient than younger farmers, but even though significant, this is not a sensitive cause of inefficiency. Also, altitude has positive and significant effects on scale efficiency (0.019). Most likely, this is probably linked to citrus fruit varieties grown by many farmers in Sardinia, which are more suited for cultivation in hilly areas. Similar to technical efficiency effect estimation, the relationship between time and scale efficiency is negative (-0.066).
This lends support to the assertion that (technical and scale) efficiency tends to decrease over time. Finally, the findings show that there are statistically significant differences in scale efficiency between farms located in different geographical regions of Italy. Farms located in Apulia and Basilicata tend to be less scale-efficient than those located in the other southern regions. Specifically, farms situated in the two insular regions (Sicily and Sardinia) report a higher magnitude (0.055 and 0.051, respectively), implying that location in these regions positively and sensitively influences scale efficiency.

6. A COMPARISON BETWEEN SFA AND DEA ESTIMATES AND DISCUSSION

In this paper we applied two approaches to estimate technical and scale efficiencies on a sample of citrus fruit farms, in which the non parametric approach is based on DEA technique, while the parametric is based on a SFA model.
We found that technical efficiency estimated from DEA model under variable returns to scale hypothesis and from SFA show not significant differences (averages equal to 0.711 and 0.710, respectively). *Vice versa*, significant difference (for $\alpha = 0.05$) is revealed between DEA CRS and SFA model (0.623 and 0.710, respectively).
However, as reported above, while DEA attributes any deviation to the frontier to estimated inefficiency component, technical efficiency computed from SFA corresponds to real inefficiency devoid of noise effects. Therefore, distribution of scores on the sample should give us more information about differences between estimated technical efficiencies calculated from SFA and DEA. As reported in Table 12, findings arisen form DEA (under variable returns to scale) suggest that the main share of farms reveals an optimal degree of efficiency (more than 20%), while a full efficiency is achieved by less than 2% of the sample in case of estimation trough SFA model. On the contrary, the share of farms that report an efficiency score close to the frontier ($0.900 < TE < 1.000$) amounts to 34.9% and 14.9% for SFA and DEA models, respectively. It might depend on the DEA method of constructing the frontier and its inherent difficulty under variable returns to scale hypothesis to detect the real efficiency due to possibility of overestimating number of full efficient units (Førsund, 1992; Kumbhakar and Tsionas, 2008).
In the light of differences in distribution of the scores on the sample, we computed the Spearman rank correlations between efficiency ranking of the observed sample (Table 13). All the correlations coefficients are positive and highly significant. The strongest correlation is obtained between the rankings from the SFA and the DEA VRS model. It confirms that hypothesis of constant returns to scale should be rejected, as reported above, in the SFA model implying that under the same set of data and assuming

variable returns to scale for the DEA frontier, the SFA model holds no real advantage over DEA in estimating technical efficiency scores and efficiency variability.

Table 12 – Frequency distributions of technical and scale efficiency estimates from the SFA and from DEA VRS models

Efficiency score	**TECHNICAL EFFICIENCY**			
	SFA		**DEA ^{VRS}**	
	Observations	**%**	**Observations**	**%**
< 0.200	13	4.0%	-	-
0.201 – 0.300	27	8.4%	12	3.7%
0.301 – 0.400	17	5.3%	27	8.5%
0.401 – 0.500	26	8.1%	51	15.9%
0.501 – 0.600	22	6.9%	42	13.1%
0.601 – 0.700	23	7.1%	39	12.1%
0.701 – 0.800	19	5.9%	18	5.6%
0.801 – 0.900	57	17.8%	18	5.6%
0.901 – 0.999	112	34.9%	48	14.9%
1.000	5	1.6%	66	20.6%
Total	321	100.0%	321	100.0%

Efficiency score	**SCALE EFFICIENCY**			
	SFA		**DEA ^{VRS}**	
	Observations	**%**	**Observations**	**%**
< 0.200	3	0.9%	-	-
0.201 – 0.300	8	2.5%	3	0.9%
0.301 – 0.400	9	2.8%	9	2.8%
0.401 – 0.500	17	5.4%	6	1.9%
0.501 – 0.600	18	5.6%	9	2.8%
0.601 – 0.700	18	5.6%	6	1.9%
0.701 – 0.800	34	10.6%	24	7.4%
0.801 – 0.900	54	16.8%	57	17.8%
0.901 – 0.999	141	43.9%	147	45.8%
1.000	19	5.9%	60	18.7%
Total	321	100.0%	321	100.0%

Concerning scale efficiency estimates, there are significant differences from the two methods (for $\alpha = 0.05$). The mean scale efficiency relative to SFA model (0.818) is lower than that estimated from the DEA model (0.894). Table 12 shows that distribution of scale efficiency scores on the sample is similar between DEA and SFA measures, except to share of farms that reveal full efficiency. Using SFA model, 5.9% of the sample reports an optimal degree of scale efficiency, while this percentage amounts to 18.7% in case of application of DEA model. According to Førsund (1992), it could depend on identification problem of full efficient observations by part of DEA model because units located at the end of size distribution may be identified as efficient simply for lack of other comparable units. *Vice versa*, since the mean DEA scale efficiency score is higher than the correspondent SFA measure, a larger number of full efficient farms computed trough DEA might be attributed to real differences due to the empirical methodologies adopted to estimate the frontier and efficiency.
Computation of the Spearman rank correlations suggests that correlation between scale efficiency ranking from the SFA and the DEA models is positive and significant but magnitude is not sensitively high (Table 13). It implies that choice of the

methodological approach might influence estimation of scale efficiency. This is a relevant point arisen from this study and it confirms how scale efficiency (and generally efficiency measures) can vary according to the model adopted for estimating frontier function on a given sample of farms, as underlined or found by several authors (Banker *et al.*, 1986; Førsund. 1992; Sharma *et al.*, 1997; Wadud and White, 2000; Ruggiero, 2007).

Table 13 表Spearman rank correlation matrix of TE and SE rankings obtained from different models

TE	**Estimated average**	**Spearman rank correlation (p)**		
		TE^{CRS}	TE^{VRS}	TE^{SFA}
TE^{CRS}	0.623	1.000		
TE^{VRS}	0.711	0.715	1.000	
TE^{SFA}	0.710	0.610	0.922	1.000

SE	**Estimated average**	**Spearman rank correlation (p)**	
		TE^{DEA}	TE^{SFA}
TE^{DEA}	0.894	1.000	
TE^{SFA}	0.818	0.547	1.000

However, scale efficiency is found to be high, on average, from application of both methods. Since the technical efficiency score is, on average, lower than the scale efficiency score this implies that the greater portion of overall inefficiency in the sample might depend on producing below the production frontier than on operating under an inefficient scale. It means that the search for an optimal scale would not become a priority for citrus farmers, while it would be a priority increasing ability in using disposable technical inputs because of technical efficiency is higher than scale efficiency. In other terms, it means that farm size issue is much less important relative to the amount of technical efficiency

Furthermore, both DEA and SFA analyses suggest that scale inefficiency is mainly due to the farms operating under a sub-optimal scale. Indeed, we found that the most of the observed farms operate under increasing returns to scale for both methods also if the incidence of sub-optimal scale farms on the total citrus farms is higher if scale efficiency is measured trough SFA (66.3% vs. 79.4% for DEA and SFA, respectively). In addition, both analyses suggest that these sub-optimal-scale farms must have adjusted their output levels to a greater extent than the supra-optimal-scale ones. In these latter farms, the margin that separate them from the optimal scale seem to be really narrow, as suggested by the estimated scale efficiency that is, on average, close to unity (SE equal to 0.934 and 0.978 for DEA and SFA, respectively), while in the sub-optimal scale farms this margin is large (scale efficiency equal to 0.692 and 0.775 for DEA and SFA, respectively). Therefore it implies that scale inefficiency is mainly due to the farms operating under a suboptimal scale.

These findings are not surprising, considering that recent studies have focused on realities characterised by the presence of small-sized farms and have found similar results about diffusion of suboptimal-scale-efficient farms (Coelli *et al.* 2002; Karagiannis and Sarris 2005; Latruffe *et al.* 2005; Cisilino and Madau 2007). The underlying rationale is that these realities are often characterised by a large number of small-sized farms that generally face capital, structural, and infrastructural constraints (*e.g.*, vast land fragmentation, huge number of single-household farms, insignificant presence of land market). They usually do not have adequate farming implements or up-to-date technologies or they are not allowed to reach their optimum size under their particular circumstances. Thiele and Brodersen (1999) argue that these market and structural constraints are among the main factors that usually impede achievement of efficient scales by part of farmers. Regarding the Italian citrus farms, Idda (2006) and Carillo *et al.* (2008) found that, often, the input mix is unbalanced (with respect to the rational and efficient composition of the input bundle) in favour of a high ratio of capital to land area and labour to land area. This should be mainly caused by a scarce flexibility in the land market, which forces farmers to expand the use of other inputs (except for land), especially labour and capital, with practical implications on the scale efficiency. Therefore, the presence of a quasi-fixed factor such as land should negatively affect scale efficiency and should favour exhibition of increasing returns to scale.

Estimation of the (technical and scale) inefficiency effects show that it is slightly sensitive to the method used. Computation of DEA reveals that technical efficiency should significant depend on farm size (positive effect), on number of plots of land and (negative effect) and on location in a less-favoured area (negative effect). Application of SFA approach seem to confirm these findings because the only three factors appeared significant by a statistical point of view are those mentioned above and the sign of the effect is the same estimated from DEA.

It must be underlined that the fact that farm size affect technical efficiency is an empirical finding that is often found in the literature, even if studies show controversial results about the relationship between technical efficiency and farm size (Sen 1962; Kalaitzandonakes *et al.* 1992; Ahmad and Bravo-Ureta 1995; Alvarez and Arias 2004).

On the other hand, estimation of scale efficiency effects show similar results in DEA and SFA application. In both analysis farm size (positive effect), number of plots of land (negative effect) and geographical location of farm should be the main factors that affect scale efficiency in the Italian citrus farming.

6. Conclusions

This paper aimed to evaluate technical and scale efficiencies on a sample of citrus farms located in Italy. Using two different approach (parametric and non parametric) we found that some margins exist to increase efficiency, both using better disposable inputs and operating on a more appropriate scale. Empirical findings arisen from the two methods used suggest that the overall inefficiency should depend on producing below the production frontier and on operating under a rational scale.

The former reason might be more important since technical inefficiency appears greater than scale inefficiency.

However, the estimated technical efficiency from the SFA model is substantially at the same level of this estimated from DEA model, while the scale efficiency arisen from SFA is larger than this obtained from DEA analysis.

Computation of both DEA and SFA analyses suggest that most of the scale-inefficient farms operate under increasing returns to scale, *i.e.,* under a sub-optimal scale. Regarding factors that affect inefficiency, the results indicate that farm size and the number of plots significantly and sensitively influence both technical and scale efficiencies. More specifically, the larger and less fragmented farms tend to show higher technical and scale efficiencies.

Finally, the correlation between the efficiency rankings of the two approaches is positive and significant both for technical and scale efficiency ranking, also if magnitude of the Spearman rank correlation coefficient is higher for technical efficiency than for scale efficiency rankings.

REFERENCES

Ahmad M, Bravo-Ureta B (1995) An econometric analysis of dairy output growth. American *Journal of Agricultural Economics* 77: 914-921.

Aigner DJ, Lovell CAK, Schmidt, PJ (1977) Formulation and Estimation of Stochastic Frontier Production Function Models. *Journal of Econometrics* 6: 21-37.

Alvarez A, Arias C (2004) Technical Efficiency and Farm Size: A Conditional Analysis. *Agricultural Economics* 30: 241-250.

Atkinson SE, Cornwell C (1994) Estimation of Output and Input Technical Efficiency Using a Flexible Functional Form and Panel Data. *International Economic Review* 35: 245-255.

Balk BM (2001) Scale Efficiency and Productivity Change. *Journal of Productivity Analysis* 15: 159-183.

Banker RD (1984) Estimating the Most Productive Scale Size using Data Envelopment Analysis. *European Journal of Operational Research* 17: 35-44

Banker RD, Charnes A, Cooper WW (1984) Some Models for Estimating Technical and Scale Inefficiency in Data Envelopment Analysis. *Management Science* 30: 1078-1092.

Banker RD., Conrad RF, Strauss RP (1986) A Comparative Application of Data Envelopment Analysis and Translog Methods: An Illustrative Study of Hospital Production. *Management Science* 32: 30-44.

Battese GE, Coelli TJ. (1995) A Model for Technical Inefficiency Effects in a Stochastic Frontier Production Function for Panel Data. *Empirical Economics* 20: 325-232.

Battese GE, Corra GS (1977) Estimation of a Production Frontier Model: With a Generalized Frontier Production Function and Panel Data. *Australian Journal of Agricultural Economics* 21: 169-179.

Bjurek H, Hjalmarsson L, Forsund FR (1990) Deterministic Parametric and Non parametric Estimation of Efficiency in Service Production. A Comparison. *Journal of Econometrics* 46: 213-227.

Bohrnstedt GW, Knoke D (1994) *Statistics for Social Data Analysis*. F.E. Peacock Publishers Inc., Itasca.

Bravo-Ureta BE, Solis D, Moreira Lopez V, Maripani JF, Thiam A, Rivas T (2007) Technical Efficiency in farming: a Meta-Regression Analysis. *Journal of Productivity Analysis* 27: 57-72.

Carillo F, Doria P, Madau FA (2008) *L'analisi della redditività delle colture agrumicole attraverso l'utilizzo dei dati RICA*. Stilgrafica, Rome.

Charnes A, Cooper WW, Seiford LM (1994) *Data Envelopment Analysis: Theory, Methodology and Application*. Kluwer Academics, Dordrecht, Boston and London.

Charnes A, Cooper WW, Rhodes E (1978) Measuring the Efficiency of Decision Making Units. *European Journal of Operational Research* 2: 429-444.

Cisilino F, Madau FA (2007) *Organic and Conventional Farming: a Comparison Analysis through the Italian FADN*. Poster Paper discussed at 103rd EAAE Seminar, Barcelona, April 23-25, 2007.

Coelli T (1996a) *A Guide to DEAP Version 2.1: A Data Envelopment Analysis (Computer) Program*. CEPA Working Paper 8/96, Department of Econometrics, University of New England, Armidale.

Coelli TJ (1996b) *A Guide to Frontier Version 4.1: A Computer Program for Stochastic Frontier Production and Cost Function Estimation*. CEPA Working Papers 7/96, Department of Econometrics, University of New England, Armidale.

Coelli T, Rahman S, Thirtle C (2002) Technical, Allocative, Cost and Scale Efficiencies in Bangladesh Rice Cultivation: A Non parametric Approach. *Journal of Agricultural Economics*. 53: 607-626.

Färe R, Lovell CAK (1978) Measuring the Technical Efficiency of Production. *Journal of Economic Theory* 19: 150-162.

Farrell MJ (1957) The Measurement of Productive Efficiency. *Journal of the Royal Statistical Society*, A 120 part 3: 253-290.

Ferrier GD, Lovell CAK (1990) Measuring Cost Efficiency in Banking: Econometric and Linear Programming Evidence. *Journal of Economettics* 46: 229-245.

Førsund FR (1992) A Comparison of Parametric and Non parametric Efficiency Measures: The Case of Norwegian Ferries. *Journal of Productivity Analysis* 3: 25-43.

Førsund FR, Hjalmarsson L (1979) Generalized Farrell Measures of Efficiency: An Application to Milk Processing in Swedish Dairy Plants. *The Economic Journal* 89: 294-315.

Frisch R (1965) *Theory of Production*. D. Reidel, Dordrecht.

Giuca S (ed.) (2008) *Le politiche agricole regionali a sostegno dell'agrumicoltura it*aliana. Stilgrafica, Rome.

Gong B., Sickles R. (1992) Finite Sample Evidence on the performance of Stochastic Frontiers and Data Envelopment Analysis using Panel Data. *Journal of Econometrics* 51: 259-284.

Greene WH (1980) On the Estimation of a Flexible Frontier Production Model. *Journal of Econometrics* 13: 101-115.

Herrero I (2000) DEA: *A Review of Some of the Main Papers*. in Pascoe S, Fousakis P, Herrero I, Juliussen V, Mardle S, Roland BE, Vassdal T (eds.) *Technical Efficiency in EU Fisheries: Methodological Report*. TEMEC Working Paper I, University of Portsmouth, pp. 56-73

Hjalmarsson L, Kumbhakar SC, Hesmati A (1996) DEA, DFA and SFA: A Comparison. *Journal of Productivity Analysis* 7: 303-327.

Huang CJ, Liu JT (1994) Estimation of a Non-Neutral Stochastic Frontier Production Function. *Journal of Productivity Analysis* 5: 171-180.

Idda L (ed.) *Aspetti economici e prospettive dell'agrumicoltura in Italia: il caso delle regioni Puglia, Basilicata e Sardegna*. Tipografia Editrice Gallizzi, Sassari (Italy).

Ismea (2008) *Agrumi, report economico e finanziario*. Imago Media, Dragoni (Italy).

Jondrow J, Lovell CAK, Materov IS, Schmidt P (1982) On the Estimation of *Technical Inefficiency in the Stochastic Frontier Production Functions Model. Journal* of Econometrics 19: 233-238.

Kalaitzandonakes NG, Wu S, Ma J (1992) The Relationship between Technical Efficiency and Size Revisited. *Canadian Journal of Agricultural Economics* 40: 427-442.

Kalaitzandonakes NG, Dunn EG (1995) Technical Efficiency, Managerial Ability and Farmer Education in Guatemalan Corn Production: A Latent Variable Analysis. *Agricultural and Resource Economic Review* 24: 36-46.

Kalirajan K (1982) On Mesauring Yield Potential of the High Yielding Varieties Technology at Farm Level. *Journal of Agricultural Economics* 33: 227-236.

Karagiannis G, Sarris A (2005) Measuring and Explaining Scale Efficiency with the Parametric Approach: the Case of Greek Tobacco Growers, *Agricultural Economics* 33: 441-451.

Karagiannis G, Tzouvelekas V (2005) Explaining Output Growth with a Heteroscedastic Non-neutral Production Frontier: The Case of Sheep Farms in Greece. *European Review of Agricultural Economics* 32: 51-74.

Kodde DA, Palm FC (1986) Wald Criteria for Jointly Testing Equality and Inequality. *Econometrica* 54: 1243-1248.

Kumbhakar SC, Ghosh S, McGuckin T (1991) A Generalized Production Frontier Approach for Estimating Determinants of Inefficiency in U.S. Diary Farms. *Journal of Business Economics Statistics* 9: 279-286.

Kumbhakar SC, Tsionas EG (2008) Scale and Efれiency Measurement using a Semiparametric Stochastic Frontier Model: Evidence from the U.S. Commercial Banks. *Empirical Economics* 34: 585-602.

Kumbhakar SC, Lovell CAK (2000) *Stochastic Frontier Analysis*. Cambridge University Press, New York.

Latruffe L, Balcombe K, Davidova S, Zawalinska K (2005) Technical and Scale Efficiency of Crop and Livestock Farms in Poland: Does Specialization Matter?. *Agricultural Economics* 32: 281走96.

Madau FA (2007) Technical Efficiency in Organic and Conventional Farming: Evidence from Italian Cereal Farms. *Agricultural Economics Review* 8: 5-21.

Madau FA (2008) *Analisi della distanza economica tra aziende biologiche e convenzionali: un approccio di frontiera*. in Abitabile C., Doria P (eds.) *La produzione agricola alimentare tra biologico e convenzionale*. Working Paper SABIO n. 5, Stilgrafica, Rome, pp. 111-46.

Madau FA (2010) Technical and Scale Efficiencies in the Italian Citrus Farming. *The Empirical Economics Letters* 9: 609-618.

Madau FA (2011) Parametric Estimation of Technical and Scale Efficiencies in the Italian Citrus Farming. *Agricultural Economics Review* 12: 91-111.

Meeusen W, Broeck J. van den (1977) Efficiency Estimation from Cobb-Douglas Production Function with Composed Error. *International Economic Review* 18: 435-444.

Minh NK, Long GT (2009) Efficiency Estimates for Agricultural Production in Vietnam: A Comparison of Parametric and Non parametric Approaches. *Agricultural Economics Review* 10: 62-78.

Orea L (2002) Parametric Decomposition of a Generalized Malmquist Production Index. *Journal of Productivity Analysis* 18: 5-22.

Parikh A, Shah K (1994) Measurement of Technical Efficiency in the North-West Frontier Province of Pakistan. *Journal of Agricultural Economics* 45: 132-138.

Pitt MM, Lee LF (1981) Measurement and Sources of Technical Inefficiency in the Indonesian Weaving Industry. *Journal of Development Economics* 9: 43-64.

Ray S (1998) Measuring and Explaining Scale Efficiency from a Translog Production Function. *Journal of Productivity Analysis* 11: 183-194.

Reifschneider D, Stevenson R (1991) Systematic Departures from the Frontier: A Framework for the Analysis of Firm Inefficiency. *International Economic Review* 32: 715-723.

Reinhard S, Lovell CAK, Thijssen GJ (2002) Analysis of Environmental Efficiency Variation. American *Journal of Agricultural Economics* 84: 1054-1065.

Ruggiero J. (2004) Data Envelopment Analysis with Stochastic Data. *Journal of the OR Society* 55: 1008-1012.

Ruggiero J. (2007) A Comparison of DEA and the Stochastic Frontier Model using Panel Data. *International Transactions in Operational Research* 14: 259-266.

Seiford LM (1996) Data Envelopment Analysis: The Evolution of the State of the Art (1978-1995). *Journal of Productivity Analysis* 7: 99-138.

Seiford LM, Thrall RM (1990) Recent Developments in DEA: The Mathematical Programming Approach to Frontier Analysis. *Journal of Econometrics* 46: 99-138.

Sen AK (1962) An Aspect of Indian Agriculture. Economic Weekly 14: 243-266.

Sickles R. (2005) Panel Estimators and the Identification of Firm-specific Efficiency levels in Parametric, Semiparametric and Nonparametric Settings. *Journal of Econometrics* 126: 305-334.

Thiele H, Brodersen CM (1999) Differences in Farm Efficiency in Market and Transition Economies: Empirical Evidence from West and East Germany. European *Review of Agricultural Economics* 26: 331-347.

Sharma KR, Leung P, Zaleski HM (1997) Productive Efficiency of the Swine Industry in Hawaii: Stochastic Frontier vs. Data Envelopment Analysis. *Journal of Productivity Analysis* 8: 447-459

Simar L., Zelenyuk V. (2011) Stochastic FDH/DEA Estimators for Frontier Analysis. *Journal of Productivity Analysis* 36: 1-20

Wadud A, White B. (2000) Farm Household Efficiency in Bangladesh: A Comparison of Stochastic Frontier and DEA Methods. *Applied Economics* 32: 1665-1673.