Empirical Investigation of the Real Input-Output Relation in Agricultural Production

Georgia KOUTOUZIDOU* and Alexandros THEODORIDIS**

Jel Classification: C14, Q12, C67

Abstract

This study describes the real nature of the input-output relation by combining parametric and non-parametric methods. Data Envelopment Analysis is applied as complement to regression analysis in order to overcome the drawback of the parametric approach, which assumes that all units operate efficiently. By using data from a sample of 365 farms, technical efficiency is estimated; farms are then divided into production function models, which correspond to different efficiency levels. The results show that the parameters of the whole data set and those of the efficient group differ substantially, indicating that the existence of inefficiency leads to inconsistent estimations.

Keywords: *Data Envelopment Analysis, regression analysis, input-output relation, production function*

Rèsumé

Dans cet article, nous allons explorer la nature de la relation intrants-extrants, en combinant des méthodes paramétriques et non-paramétriques. Outre l'analyse de régression, nous allons appliquer l'analyse d'enveloppement des données (DEA) pour surmonter les limites de l'approche paramétrique, qui suppose des unités de production efficaces. En utilisant les données relatives à un échantillon de 365 exploitations, nous allons évaluer l'efficacité technique. Ensuite, les exploitations seront divisées selon des modèles de fonction de production, correspondant aux différents niveaux d'efficacité. Les résultats vont démontrent que les paramètres de l'ensemble des données retenues et les paramètres du groupe efficace sont substantiellement différents, indiquant ainsi que l'inefficacité conduit à des erreurs d'estimations.

Mots-clés: Analyse d'Enveloppent des Données, analyse de régression, relation intrants-extrants, fonction de production.

1. Introduction

An empirical problem that arises from the estimation of production functions, in order to elicit their real-world implications, is the assumption that producers fully employ the existing technology, utilizing the factors of production efficiently (Heady and Dillon, 1961; Yotopoulos and Nugent, 1976; Bernini et al., 2004). Prevailing methods for estimating production functions, like conventional econometric analysis, do rest on the assumption that producers operate on their production functions, maximizing output obtainable from the inputs they use; that is producers are assumed to satisfy the first-order condition for output maximization allocating inputs efficiently and ending up on their production function (Kumbhakar and Lovell, 2000). Nevertheless, this is far from reality, since not all producers succeed in utilizing the minimum inputs required to produce the outputs they choose to produce, given the technology at their disposal, meaning that all producers are not technically efficient (Kumbhakar and Lovell, 2000; Cooper et al., 2005).

The conventional estimation techniques neglect the level of technical efficiency of the units, yielding parameter estimates that are bound to be biased, untrustworthy and meaningless (Yotopoulos and Nugent, 1976). By ignoring inefficiency, the conventional framework is mispecified and estimates of productivity are contaminated with variation in efficiency (Lovell, 1993). Consequently, if one ignores differences in technical efficiency among farms, one biases the parameter values obtained from the estimation of the production function. Several parametric methods that allow for inefficiency have been developed, among these Corrected OLS (Winsten, 1957; Gabrielsen, 1975), Modified OLS (Richmond, 1974) and Stochastic Frontier Analysis (Aigner et al., 1977; Meeusen and van den Broeck, 1977). These parametric methods provide estimations of efficiency levels and of economic parameters but the efficiency criterion is mixed up with the selection of the appropriate functional form (Thiry and Tulkens, 1992). On the contrary, non-parametric methods applied for efficiency assessment, like Data Envelopment Analysis (DEA), provide more flexibility, since they do not require an explicit specification of the functional form of the production frontier.

The purpose of this study is to describe the parametric production frontier and to approximate the real nature of the input-output relation and of the factors' returns, in the sense that the estimated production characteristics will be more consistent with the economic theory through the application of two instruments; one non-parametric for efficiency measurement and one parametric for parameters estimation. This technological relationship is approximated with the contribution of the Data Envelopment Analysis (DEA), which is an established non-parametric technique for the

* Department of Agricultural Economics, School of Agriculture, Aristotle University of Thessaloniki.
** Lab. of Animal Production Economics, Faculty of Veterinary Medicine, Aristotle University of Thessaloniki.

assessment of the performance of production units like crop farms. Hence, this study deals with the combination of two approaches – DEA and regression analysis – in order to reveal the efficient, "true" input-output relation in an agricultural production process and to provide estimates with clear economic content that can be used as guides to production decisions.

The performance of the farms is evaluated by implementing DEA and then the obtained subsets of efficient and inefficient observations are estimated by means of the appropriate production function. Usually, DEA and regression analysis are applied competitively for the evaluation of the performance of production units (Thanassoulis, 1993; Cubbin and Tzanidakis, 1998). However, in this study, DEA is applied complementary to the conventional regression analysis aiming to overcome and eliminate the drawbacks of the parametric approach, which assumes that all production units are technically efficient and operate on their production function, that is, on the boundary of their production set.

The concept of eliminating the inefficient production units with the implementation of non-parametric approaches before fitting a parametric model (semi-parametric method) has been used by Simar (1992), Thiry and Tulkens (1992) and Bernini et al. (2004), However, they did not use DEA for the filtering of the data, but they applied free disposal hull (FDH) method. Tofallis (2001) proposed a hybrid approach for the estimation of the parametric production frontier with the use of DEA, but he implemented only an illustrative application. This article's contribution is that it attempts to establish a methodological link between the two approaches using farm accounting data from crop farms in Greece for the estimation of the frontier production function.

The paper is organized as follows: Section 2 describes the data and the theoretical models and sets out the methodological approach. Section 3 presents the empirical results and Section 4 concludes the paper.

2. Data, Models and Methods

The farm accounting data for this application were collected through a farm management survey, of 365 crop farms, carried out during the 2005-2006 period in the prefectures of Serres, Karditsa and Pella (Greece). The eligible study area is characterized by high homogeneity of land and climate conditions which entails that the existing cropping patterns do not differentiate among the three regions. The sample farms present similar structural and economic characteristics and apply the same production technology, allowing, therefore, the empirical application of the DEA approach. The farms cultivate annual crops of primary importance, namely wheat, lucerne, maize and cotton. The size of the sample and the fact that it includes farms of the same or similar farm type fit the requirements for this empirical application without imposing any restrictions. The number and the structure of the sample farms enable a detailed and in-depth empirical analysis. It should be mentioned that this two-step procedure for the estimation of the efficient production characteristics requires a quite large data set, because if the number of the inefficient observations is large, the number of degrees of freedom might be decreased excessively and may prevent the estimation of the efficient production frontier.

The required technical and economic indices for each sample farm that were collected through the farm management survey included:

- The cultivated land, expressed in hectares, and its corresponding cost in €.
- The labor requirements, expressed in hours, and its cost in €.
- The composition and the value of the different types of capital, expressed in €.
- The crop yield and the producer prices, as well as the provided subsidies.

Based on the above indices, the following were calculated for each sample farm:

- The annual expenses of buildings and machinery, expressed in €.
- The variable cost (seeds, fertilizers, pesticides and herbicides, hired mechanic labor, fuel, etc.), expressed in €.
- The gross output: the value of production plus the provided subsidies, expressed in €.

The approach applied consists mainly of two steps. In the first step, the level of technical efficiency of the sample farms is estimated with the implementation of an output-oriented DEA model. This estimated level of technical efficiency is used as a classification criterion for the 365 farms, which are divided into sub-samples (efficiency groups) for the estimation of separate production functions that correspond to different levels of efficiency. Hence, the data set is partitioned according to the estimated level of TE. Gross output is introduced in the specified DEA model as output, and the cultivated land, human labor used, variable cost and annual expenses of fixed capital are used as inputs.

DEA has received considerable attention from economists and has become increasingly popular in the measurement of the technical efficiency of production units. It is a non-parametric method which has been developed based on mathematical programming and input-output models (Charnes et al., 1978; Banker et al., 1984). DEA defines a non-parametric, best practice frontier, which is approximated by piecewise linear facets and measures the efficiency of each unit relative to that frontier. Efficient units lie by definition on that frontier and act as benchmarks for the rest of the units, the inefficiency of which is indicated in proportion to their distance from the frontier.

The original DEA formulation of Banker et al. (1984) is expressed as follows:

$$Max\ \theta = \sum_{r=1}^{s} y_{rj} u_r - \sum_{i=1}^{m} x_{ij} v_i \qquad (1)$$

subject to:

$$\sum_{r=1}^{s} y_{rj}u_r - \sum_{i=1}^{m} x_{ij}v_i \leq 1,\ j\ =\ 1,...,n$$

$$\sum_{r=1}^{s} u_r + \sum_{i=1}^{m} v_i = 1 \qquad (2)$$

$$u_r \geq 1,\ r = 1,...,s$$

$$v_i \geq 1,\ i = 1,...,m$$

where, n is the number of units, m the number of inputs, s the number of outputs, x_{ij} the level of input i used from unit j (i = 1, 2,…, m), y_{rj} the level of output r produced from unit j (r = 1, 2,…, s), u, the weight of output r, v, the weight of input i and TE=1/θ is the technical efficiency, which express the increase in outputs, given the level of inputs (output-oriented). A production unit is efficient if TE = 1, while when TE < 1 the unit is considered inefficient. Comprehensive reviews and extensions of the model can be found in Seiford and Thrall (1990), Cooper et al. (2000), Coelli et al. (2005).

In the second step of this application the Cobb-Douglas and the Translog production functions are estimated using the method of Least Squares for the entire data set, as well as for each of the data sets formulated according to the estimated level of technical efficiency. The variables incorporated in the specified production function are the same used in the DEA model. Specifically, gross output is the dependent variable and cultivated land, human labor used, variable cost and annual expenses of fixed capital are the independent variables.

The Cobb-Douglas has been widely applied in agricultural economics and its use has been justified by Griliches (1957), Mundlak and Hoch (1965) and Zellner et al. (1966). The main characteristic of the Cobb-Douglas function is that it is relatively easy to estimate, although, it is unduly restrictive. In order to limit the restrictive properties imposed on the production process by the Cobb-Douglas function, the Translog production function (Christensen et al., 1973) is very often employed and tested against the restricted Cobb-Douglas functional form.

The selection of the appropriate functional form that represents adequately the data is based on the performance of the Likelihood Ratio test (LR-test), which is:

$$\lambda = -2[\text{likelihood}(H_0) - \text{likelihood}(H_1)] \qquad (3)$$

where H_0 and H_1 are the values of the log-likelihood function under the null and alternative hypothesis (namely, Cobb-Douglas and Translog function), respectively. The λ statistic has an asymptotic chi-squared distribution, with degrees of freedom equal to the parameters assumed to be zero in the null hypothesis.

The description of the production frontier in parametric terms is obtained from the estimation of the model that includes the farms that utilize efficiently the existing technology, according to the results of DEA (level of technical efficiency equal to one). The estimated regression coefficients are compared with those estimated from the production function that includes as observations the technical inefficient farms by carrying out a Chow test, in an attempt to find if there are differences between the estimated function for the entire sample and for the efficient sub-sample. This way, the initial assumption is investigated, which gave the incentive for this application; ignoring the differences of technical efficiency among farms it will lead to misspecification of the production technology and will provide inconsistent estimates from the economic point of view.

The Chow test has been designed for testing the equality of regression coefficients of two equations (Kmenta, 1990). The 365 farms are categorized on the basis of the presence of technical inefficiency into two separate groups of observations: the first group N_1 includes the efficient farms, while the second group N_2 includes the inefficient ones.

$$y_i = \begin{cases} x_i b_{i1} + \varepsilon_i & i \in N_1 \\ x_i b_{i2} + \varepsilon_i & i \in N_2 \end{cases} \qquad (4)$$

The least square method is applied to each group, as well as to the entire sample, thus allowing for different coefficient estimates.

The equality of the values of the parameters estimated for each group of observations is tested with the use of the Chow-test. The appropriate test statistic becomes:

$$F = \frac{\left[S - (S_1 + S_2)\right] / k}{(S_1 + S_2) / (N_1 + N_2 - 2k)} \sim F_{k, N_1 + N_2 - k} \qquad (5)$$

where S, S_1 and S_2 are the residual sum of squares of the entire sample and of the efficient and inefficient sample, respectively, k is the number of the estimated parameters, and $N_1$and N_2 is the number of observations in the efficient and inefficient group, respectively.

The measurement of efficiency is a practical decision tool for adopting management strategies that can induce farmers to improve their productivity and, hence, their competitiveness. In this context, a descriptive analysis of the sample farms based on the presence of technical inefficiency is carried out. Thus, a technical and economic description of the efficient farm structure is provided and the sources of inefficiency of the sample farms are partly determined.

3. Results

The technical efficiency of the 365 farms is estimated with an output-oriented DEA model. The frequency distribution of the estimated technical efficiency and its mean for each sub-sample (efficiency groups) is presented in Table 1.

The mean technical efficiency is 0.777, which reveals that there are substantial production inefficiencies among sample farms, and justifies the motivation of the study that not

Table 1. *Frequency distribution of efficiency estimates.*

TE	DEA		
	Number of farms	%	Mean TE
< 0.6	44	12.05	0.523
0.6 – 0.7	61	16.71	0.647
0.7 – 0.8	82	22.47	0.752
0.8 – 0.9	105	28.77	0.847
0.9 – 1.0	40	10.96	0.940
= 1.0	33	9.04	1.000
Total	365	100.00	0.777

all farms are utilizing their inputs efficiently. This means that the average farm produces 77.7% of maximum feasible output and could increase its production by 22.3% if it were operating at the efficient level.

There are wide variations in the level of technical efficiency among sample farms. Of the 365 sample farms, 33 farms, 9.04% of the total number of farms, were found to be fully technical efficient (Table 1). The technical inefficient farms are 332, that is the 90.96% of the sample. The frequency distribution of the estimated level of technical efficiency indicates that 44 farms, i.e. 12.05%, are placed in the low efficiency group (0% - 60%), with mean value of technical efficiency 52,3%, 61 farms (16,71% of the sample) lie in the 60% to 70% range of technical efficiency,with mean value 64,7%, 82 farms (22,47%) lie in the 70% to 80% range, with mean value 75,2%, 105 farms (28,77%) lie in the 80% to 90% range, with mean value 84,7% and 40 farms, which consist the 10,96% of the total number of sample farms are placed in the high efficiency group (90%-100%), with mean value of technical efficiency 94,0%.

Table 2 describes the mean value of the variables used in the estimation of the production functions for each of the six efficiency groups, as well as for the whole sample.

The gross output achieved by the 365 sample farms is, on average, 39577 €, the cultivated land is 15.4 hectares per farm, the human labor employed is 1016 hours per farm, while variable costs and annual expenses of fixed capital are 13068 and 12266 € per farm, respectively. The average gross output of the farms which have been categorized according to the estimated level of technical efficiency ranges between 28747 and 54654 €. The average cultivated land ranges between 13.4 and 18.8 hectares, while the average human labor ranges from 891 to 1326 hours. The variable cost and the annual expenses of fixed cost results range from 12295 to 18624 € and 10982 to 14153 €, respectively. Based on the economic data of the survey the composition of the total production cost of the average sample farm is calculated. The results reveal that the shares of land and labor expenses in total production cost are 18.4% and 10.0%, respectively, while the shares of the variable and fixed capital cost are 37.0% and 34.6%, respectively, revealing that the sample farms are capital intensive and depend heavily on the use of purchased variable inputs and on high investments on buildings and machinery.

The parameters that describe the "true", efficient, technological nature of the input-output relationship are approximated through the estimation of the production function in which all farms lay on the efficiency frontier. Therefore, separate Cobb-Douglas and Translog production functions are estimated for the 365 sample farms and for each sub-set composed with the contribution of DEA. During this empirical analysis, in order to narrow down the choice of models and define the appropriate one, a number of different specifications of production functions concerning the variables involved in the model were applied. The selection criteria included the degree to which the results were satisfactory, reasonable and consistent with economic theory, the level of significance and the sign of the estimated coefficients and the level of the adjusted goodness of fit measure (Hendry and Richard, 1982; Kmenta, 1990).

A General Likelihood Ratio test (LR-test) was performed to test which of the Cobb-Douglas and Translog production functions is the appropriate form to be estimated in this study. The results of the LR-test for the six efficiency groups and the whole sample (Table 3) indicate that in all cases the null hypothesis, which states that the Cobb-Douglas production function is preferred to Translog, is rejected and the alternative is adopted. Hence, Cobb-Douglas is an inadequate presentation of the data under examination and it is rejected confidently in favor of Translog production function. The results presented from now on refer solely to the Translog production function.

The estimated Translog production functions and the calculated elasticities of production for the whole sample and for the six subsets which correspond to different efficiency levels are presented in Table 4. The estimated coefficients of the variables in the Translog production function do not have any direct interpretation; thus, the elasticity of output for each input has to be

Table 2 - *Summary statistics of the sample data.*

Level of TE	<0.6	0.6-0.7	0.7-0.8	0.8-0.9	0.9-1.0	1.0	Total
Number of farms	44	61	82	105	40	33	365
Gross Output (€)	28747	32335	37849	44344	41127	54654	39577
Land (ha)	15.5	15.6	14.5	15.6	13.4	18.8	15.4
Labor (hours)	1256	1120	919	891	864	1326	1016
Variable Capital (€)	13614	12397	12267	12403	12295	18624	13068
Fixed Capital (€)	13946	12954	11016	12033	10982	14153	12266

Table 3 - H*ypothesis testing of Cobb-Douglas και Translog parameter values using LR-test.*

TE	λ	Critical value $x^2_{10,5\%}$	Hypothesis testing
< 0.6	48.88	18.31	H_0: rejected
0.6 – 0.7	76.75	18.31	H_0: rejected
0.7 – 0.8	76.82	18.31	H_0: rejected
0.8 – 0.9	116.18	18.31	H_0: rejected
0.9 – 1.0	62.08	18.31	H_0: rejected
= 1.0	40.97	18.31	H_0: rejected
Total	45.35	18.31	H_0: rejected

calculated as the first derivative of the output with respect to each input (Debertin, 1986), using formula (3):

$$\varepsilon_j = \frac{\partial \ln y}{\partial \ln y_j} = \beta_j + 2\beta_{jj}\bar{x}_j \sum_{j \neq k} \beta_{ik}\bar{x}_k \qquad (6)$$

The elasticity, ε_j, measures the responsiveness of output to a 1% change in the *j*th input. The estimations provide a good fit of the data according to the value of adjusted R^2, which is increased in the case of the efficient subset (adjusted $R^2 = 0.967$) compared to that obtained from the full data set (adjusted $R^2 = 0.886$). For the entire data set output elasticity of land is 0.51, elasticity of labor is 0.19, elasticities of variable and fixed capital are 0.23 and 0.04, respectively, while for the frontier production function that includes as observations the technical efficient farms elasticity of land is 0.69, elasticity of labor is 0.16 and elasticity of variable and fixed capital is 0.18 and 0.07, respectively, although the output elasticity of fixed capital is insignificant. The estimated output elasticities for all inputs were positive indicating that the estimated Translog production function is a well-behaved production technology. The shares of factors are relatively consistent with *a priori* expectations, showing that the elasticity of output with respect to land is the highest among all inputs, followed by the elasticity of output with respect to variable capital. This finding is common in all efficiency groups indicating that land and variable capital have major influence on output. This result is consistent with the fact that the sample farms cultivate annual crops. The sum of the output elasticities for the whole sample is estimated to be 0.97, implying slightly decreasing returns to scale in the production which shows that the farms operate at a non-optimal scale.

The production function that approximates the efficient frontier technological relationship and its estimation leads to parameters with clearer economic content is that which includes as observations the farms that utilize efficient the factors of production. Hence, it is ascertained that the production function that describes the frontier production function is the Translog production function estimated for the 33 technically efficient sample farms.

However, the question that emerges is whether the estimated parameters of the frontier production function differ from the parameters obtained from the estimated production function for the full data set. The stability of these estimated regression coefficients are tested using the Chow-test.

The implementation of the Chow-test requires the sample data to be divided into two separate groups using as a criterion the presence or not of technical inefficiency and then the hypothesis $H_0: b_{i1} = b_{i2}$, $i = 1,2,3,4$, is tested, where b_{i1} are the regression coefficients estimated from the sample of the 33 efficient farms and b_{i2} are the coefficient parameters estimated from the sample of the 332 inefficient farms. Thus, the null hypothesis states that estimated regression coefficients of the two equations are equal, and that the value of the parameters of the two groups is stable. The acceptance of the null hypothesis suggests that the estimation of the input-output relation from the group of efficient observations does not differ from that obtained from the total sample farms which are characterized by the presence of inefficiency.

The calculated value F of the stability test is 7.58, while the critical value $F_{14,337}$ at 1% significance level is 2.135, implying that null hypothesis is rejected and the alternative is adopted, where $b_{i1} \neq b_{i2}$.

From the results it is suggested that the coefficients obtained from the production function which contains the entire data set and which is conventionally estimated in regression analysis, neglecting the presence of inefficiency, differ substantially from the coefficients of the production function which includes only the technically efficient farms and which is considered to approximate the true input-output relation. Therefore, it is indicated that DEA can be used complementary to regression analysis and can lead to a clearer economic result and to a more realistic and reliable description of the input-output functional relationship. The supposition that if the existing technical inefficiency of the farms is ignored will lead to misinterpretation of the structure of production technology was fully confirmed by the aforementioned findings.

Table 5 describes the main technical and economic characteristics of the mean efficient, inefficient and entire sample farm, respectively. These characteristics have been reduced to 1000 € of gross output in order to partly determine the sources of inefficiency and the magnitude of the factors that contribute to the improvement of technical efficiency.

Results imply that for the achievement of 1000 € of gross output the mean efficient farm cultivates 0.06 and 0.05 hectares less than the inefficient and total farm, saving, re-

Table 4 - *Parameters of Translog production functions.*

Variables	Level of TE						Total
	< 0.6	0.6-0.7	0.7-0.8	0.8-0.9	0.9-1.0	=1.0	
Intercept	-12.08	-6.87	-11.31***	3.06	-2.02	7.38	0.78
ln *land*	-7.36***	0.14	-2.12**	2.97***	0.42	3.86**	-0.39
ln *labor*	-0.80	0.01	0.34	-0.16	-0.19	-0.81**	-0.46*
ln *vc*	8.88***	1.58	3.22***	-1.60**	2.43	-1.87	0.95
ln *fc*	-0.49	1.26*	1.62***	1.12***	-0.51	0.48	0.99**
$\ln^2$ *land*	-0.90***	-0.16***	-0.21**	0.13*	-0.06	0.25**	-0.09***
$\ln^2$ *labor*	-0.16***	-0.001	-0.001	-0.08***	-0.01	0.02	0.01
$\ln^2$ *vc*	-0.94**	-0.16	-0.31***	0.11	-0.41**	0.17	-0.09
$\ln^2$ *fc*	0.05	-0.01**	-0.14***	-0.08***	-0.05	-0.04	-0.06***
ln *land* ln *labor*	-0.38*	-0.08**	-0.04	-0.06	-0.19***	-0.18***	-0.19***
ln *land* ln *vc*	1.75***	0.18	0.37**	-0.44***	0.38	-0.41**	0.23**
ln *land* ln *fc*	0.30	0.07	0.16*	0.09	-0.18	-0.07	0.10
ln *labor* ln *vc*	0.42**	-0.08*	0.06	0.23***	0.17**	0.09	0.09*
ln *labor* ln *fc*	0.11	-0.01	0.03	-0.05**	-0.01	0.08**	0.07***
ln *vc* ln *fc*	-0.26	0.04	0.06	0.04	0.27**	0.02	-0.09
e_{land}	0.46***	0.45***	0.46***	0.60***	0.46***	0.69***	0.51***
	(0.107)	(0.046)	(0.0320	(0.039)	(0.070)	(0.117)	(0.045)
e_{labor}	0.10**	0.22***	0.25***	0.16***	0.21***	0.16***	0.19***
	(0.045)	(0.0130	(0.014)	(0.015)	(0.033)	(0.035)	(0.019)
e_{vc}	0.29***	0.29***	0.23***	0.14***	0.25***	0.18*	0.23***
	(0.106)	(0.047)	(0.031)	(0.039)	(0.059)	(0.119)	(0.046)
e_{fc}	0.20***	0.10***	0.10***	0.10***	0.13***	0.07	0.04
	(0.061)	(0.025)	(0.019)	(0.013)	(0.039)	(0.057)	(0.027)
Adj R^2	0.936	0.987	0.991	0.994	0.991	0.967	0.886
No of obs	44	61	82	105	40	33	365
df	29	46	67	90	25	18	350

Note: *, **, *** indicates 10%, 5% and 1%, respectively. The figures in parentheses are standard errors.

spectively, 27 € and 24 €. The divergence in human labor employed is small, with the mean efficient farm using only 1 hour more, which costs 1 €. The variable capital cost for the mean efficient farm is 341 €, 12 € and 11 € more than in the inefficient and sample farm. Fixed capital cost is 259 €, 58 € and 51 € less than the inefficient and sample farm, while regarding the total production cost it emerges that the mean efficient farm spends 73 and 64 € less, respectively. These results indicate that a reduction in the cultivated land, which, however, according to its cost appears to be of high productivity, and of the annual expenses of fixed capital appears to favor a higher level of efficiency, in combination with the level use of the other factors of production. This result is confirmed by the calculation of land and capital cost per unit of labor input. The land and fixed capital cost per unit of labor is 5.8 and 10.7 € per hour of labor used for the mean efficient farm 0.7 and 1.6 € per hour of labor less than in the mean inefficient farm, verifying, therefore, the finding that a reduction in land and fixed capital ameliorates technical efficiency.

Table 5 - *Main technical and economic characteristics per output of the mean efficient, inefficient and total sample farm.*

TE	< 1.0	= 1.0	Total
Number of farms	332	33	365
Gross output (€)	1000	1000	1000
Land (ha)	0.40	0.34	0.39
Labor (hours)	25	26	26
Rent (€)	168	141	165
Labor cost (€)	90	89	90
Variable capital cost (€)	329	341	330
Fixed capital cost (€)	317	259	310
Total cost (€)	904	831	895

4. Conclusions

The paper deals with the estimation of production functions by taking into account the possibility of technically inefficient performance of production units. The estimation of the production functions is carried out in a two-stage manner by combining two approaches; the non-parametric DEA and the conventional regression analysis. Through this procedure, we can not only exploit the advantages of each approach, but also partially control some of their limitations and drawbacks. The empirical application of this two-step procedure, which did not receive much attention in the past, is based on farm accounting data of 365 crop farms of the same or similar farm type.

The results of the DEA indicated, as expected, the presence of inefficiency among farms and 332 farms of the sample were relatively technical inefficient. The sample data set was partioned into six subsets of efficient and inefficient observations which were estimated using the Cobb-Douglas and the Translog production functions. The LR-test showed that the Translog production function provides an adequate representation of the data, while the Chow-test indicated that the parameter coefficients obtained from the full and the efficient data set, differ substantially. This finding verified the initial incentive of this study that filtering the data set using a performance management tool like DEA can improve the economic quality of the estimations; therefore, it can prove to be meaningful and consistent with the economic theory production characteristics.

In conclusion, this study shows that the description of the true technological nature of inputs and outputs in a production process is approximated with the use of DEA which can be applied complementary to the conventional regression analysis, eliminating the limitation of the latter that all units allocate their sources efficiently and bridges the conceptual gap between the two methods.

References

Aigner D.J., Lovell C.A.K., Schmidt P., 1977. Formulation and estimation of stochastic frontier production function models, Journal of Econometrics, 6, 21-37.

Banker R.D. Charnes A., Cooper W.W., 1984. Some models for estimating technical and scale inefficiency in Data Envelopment Analysis, Management Science, 30, 1078-1092.

Bernini C., Freo M., Gardini A., 2004. Quantile estimation of frontier production function, Empirical Economics 29, 373-381.

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

Christensen L., Jorgenson D., Lan L., 1973. Transcendental logarithmic production frontiers, The Review of Economics and Statistics, 55, 28-45.

Coelli T.J., Prasada Rao D.S., O'Donnell C.J., Battese G.E., 2005. An Introduction to Efficiency and Productivity Analysis, Springer, New York.

Cooper W.W., Seiford M.L., Tone K., 2000. Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver software, Kluwer Academic Publishers, Boston.

Cubin J., Tzanidakis G., 1998. Regression versus data envelopment analysis for efficiency measurement: an application to the England and Wales regulated water industry. Utilities Policy, 7, 75-85.

Debertin D. (1986). Agricultural production economics. Macmillan Publishing Company, New York.

Gabrielsen A., 1975. On estimating efficient production functions. Working paper No A-35, Chr. Michelsen Institute, Department of Humanities and Social Sciences, Bergen, Norway.

Griliches Z., 1957. Specification bias in estimates of production function. Journal of Farm Economics, 39, 8-20.

Heady E.O., Dillon J.L., 1961. Agricultural production functions. Iowa State University Press, Iowa.

Henry D.F., Richard J.F., 1982. On the formulation of empirical models in dynamic econometrics. Journal of Econometrics, 20, 3-33.

Kmenta J., 1990. Elements of econometrics, Macmillan Publishing Company, Iowa.

Kumbhakar S.C., Lovell C., 2000. Stochastic frontier analysis, Cambridge University Press.

Lovell C.A.K., 1993. Production frontiers and productive efficiency, in: Fried H.O., Knox Lovell C.A., Schmidt, S.S. (Eds.), The Measurement of Productive Efficiency: Techniques and Applications, Oxford University Press, New York, 3-67.

Meeusen W., van den Broeck J., 1977. Efficiency estimation from Cobb-Douglas production functions with composite error, International Economic Review, 18, 435-444.

Mundlak Y., Hoch I., 1965. Consequences of alternative specifications in estimation of Cobb-Douglas production functions, Econometrica, 33, 814-828.

Richmond J. 1974. Estimating the efficiency of production, International Economic Review, 15, 515-521.

Seiford M.L., Thrall R.M., 1990. Recent developments in DEA: The mathematical programming approach to frontier analysis, Journal of Econometrics 46, 7-38.

Simar L., 1992. Estimating efficiencies from frontier models with panel data: a comparison of parametric, non-parametric and semi-parametric methods with bootstrapping, Journal of Productivity Analysis, 3, 171-203.

Thanassoulis E., 1993. A comparison of regression analysis and data envelopment analysis as alternative methods for performance assessments, Journal of Operational Research Society, 44, 1129-1144.

Thiry B., Tulkens H., 1992. Allowing for inefficiency in parametric estimation of production functions for urban transit firms, Journal of Productivity Analysis, 3, 45-65.

Tofallis C., 2001. Combining two approaches to efficiency assessment, Journal of Operational Research Society, 52, 1225-1231.

Winsten C.B., 1957. Discussion on Mr. Farrell's paper, Journal of the Royal Statistical Society, A 120, 282-284.

Yotopoulos P., Nugent J., 1976. Economics of development, Harper-Row Publ., New York.

Zellner A., Kmenta, J., Drèze J., 1966. Specification and estimation of Cobb-Douglas production function models, Econometrica 34, 784-795.