Economics Bulletin

Volume 30, Issue 1

A stochastic production frontier model with a translog specification using the generalized maximum entropy estimator

Pedro Macedo
Department of Mathematics, University of Aveiro, Portugal

Elvira Silva
Faculty of Economics, University of Porto, Portugal

Abstract

In this paper, an empirical application of the generalized maximum entropy estimator in a stochastic production frontier model with a translog specification is discussed to investigate technical efficiency in a wine region of Portugal. The empirical results indicate technical progress over the time period of the sample and an increasing technical inefficiency over time. All production units are technically inefficient, although wine cooperatives are less inefficient than private firms.

We would like to express our gratitude to the anonymous referees and to the Editor for their valuable comments and suggestions for improvements in the paper.
Citation: Pedro Macedo and Elvira Silva, (2010) "A stochastic production frontier model with a translog specification using the generalized maximum entropy estimator", *Economics Bulletin*, Vol. 30 no.1 pp. 587-596.
Submitted: Apr 24 2009. **Published:** February 17, 2010.

1. Introduction

In the last decade, the work of Golan et al. (1996) inspired the development of the Maximum Entropy (ME) econometrics. The fundamental reason that explains the increasing interest on this methodology was pointed out by Edwin Jaynes, cited by Golan et al. (1996) in the authors' preface: "in economics a truly well posed problem is virtually unknown". Since ill-posed problems seem to be the rule rather than the exception in econometrics, the Generalized Maximum Entropy (GME) estimator has acquired importance in the wide group of econometric techniques. A review of the ME formalism, with several contributions that have been emerging in the literature, can be found in Golan (2002, 2007).

The stochastic frontier analysis (SFA) is central in efficiency measurement. The stochastic frontier model has been proposed by Aigner et al. (1977), Meeusen and van den Broeck (1977) and Battese and Corra (1977). In SFA, a functional form for the frontier production function must be specified and a composed error structure takes into account the random error and a component accounting for inefficiency. Details of SFA can be found in Kumbhakar and Lovell (2000) and Greene (1993, 2001).

The purpose of this paper is twofold. The estimation of a stochastic frontier model with a translog specification by the GME estimator, that is unusual in the efficiency analysis literature. The analysis of technical efficiency is performed in the wine sector of the "Bairrada" region in Portugal for wine cooperatives and private firms, in the time period 1993-2002.

In the estimation procedure, some issues need to be improved but the preliminary results indicate a good performance of the GME estimator. It seems that this is a promising approach to stochastic frontier models by avoiding distributional assumptions (only some restrictions in the domains of random variables are needed in the optimization structure) and because the GME estimator is not sensitive to collinearity. Note that collinearity is usual in SFA and is responsible for inflating the variance associated with the regression coefficients, as well as affecting the signs of coefficients.

The empirical results indicate that all production units are technically inefficient, yet wine cooperatives are less inefficient than private firms. The results also indicate technical progress over the time period of the sample and an increasing technical inefficiency over time, implying production units are moving away from the production frontier.

The paper proceeds as follows. In section 2, the GME estimator is briefly presented. Description of the data and discussion of the empirical model are presented in section 3. The empirical results are discussed in section 4 and the final section concludes.

2. The GME estimator

Consider the following general linear model

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u} \tag{1}$$

where $\mathbf{y}$ is the $(N \times 1)$ vector of observations on the dependent variable, $\boldsymbol{\beta}$ is the $(K \times 1)$ vector of unknown parameters, $\mathbf{X}$ is a known $(N \times K)$ matrix and $\mathbf{u}$ is the $(N \times 1)$ vector of random errors. Given that the economic processes are typically stochastic and the available economic data are often composed of limited and non-experimental observations, the usual economic models may be ill-posed (Golan et al. (1996)). Given heed to this problem, Golan et al. (1996) specify a set of support values for each unknown parameter and the error term and use the ME to estimate the unknown probabilities associated with the support values, i.e, they transform the linear regression model such that the unknown parameters and the random unknown errors are in the form of probabilities. The random error vector is considered as another set of unknown parameters to be estimated simultaneously with the coefficients $\boldsymbol{\beta}$. Golan et al. (1996) treat each β_k as a discrete random variable with a compact support and $2 \leq M \leq \infty$ possible outcomes and each u_n as a finite and discrete random variable with $2 \leq J \leq \infty$ possible outcomes. Assuming that both the unknown parameters and the unknown error terms may be bounded *a priori*, the linear model in (1) can be presented as

$$\mathbf{y} = \mathbf{XZp} + \mathbf{Vw}, \tag{2}$$

where

$$\boldsymbol{\beta} = \mathbf{Zp} = \begin{bmatrix} \mathbf{z}'_1 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{z}'_2 & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{z}'_K \end{bmatrix} \begin{bmatrix} \mathbf{p}_1 \\ \mathbf{p}_2 \\ \vdots \\ \mathbf{p}_K \end{bmatrix},$$

with $\mathbf{Z}$ a $(K \times KM)$ matrix of support values and $\mathbf{p} \gg \mathbf{0}$ a $(KM \times 1)$ vector of unknown weights, and

$$\mathbf{u} = \mathbf{Vw} = \begin{bmatrix} \mathbf{v}'_1 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{v}'_2 & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{v}'_N \end{bmatrix} \begin{bmatrix} \mathbf{w}_1 \\ \mathbf{w}_2 \\ \vdots \\ \mathbf{w}_N \end{bmatrix},$$

with $\mathbf{V}$ a $(N \times NJ)$ matrix of support values and $\mathbf{w} \gg \mathbf{0}$ a $(NJ \times 1)$ vector of unknown weights. The ME methodology is used to estimate the unknown $\mathbf{p}$ and $\mathbf{w}$ vectors that maximize

$$H(\mathbf{p}, \mathbf{w}) = -\mathbf{p}' \ln \mathbf{p} - \mathbf{w}' \ln \mathbf{w}, \tag{3}$$

subject to the model constraint and the additivity constraints on $\mathbf{p}$ and $\mathbf{w}$,

$$\mathbf{y} = \mathbf{XZp} + \mathbf{Vw},$$

$$\mathbf{1}_K = (\mathbf{I}_K \otimes \mathbf{1}'_M)\mathbf{p},$$
$$\mathbf{1}_N = (\mathbf{I}_N \otimes \mathbf{1}'_J)\mathbf{w},$$

where $\otimes$ represents the Kronecker product. The GME estimator generates the optimal probability vectors $\hat{\mathbf{p}}$ and $\hat{\mathbf{w}}$ that can be used to form point estimates of the unknown parameter vector (and the unknown random errors).

The GME estimator generalizes the ME formalism of Jaynes (1957*a*,*b*), which uses the Shannon (1948) entropy measure in the objective function. The Shannon entropy measure reflects the uncertainty about the occurrence of a collection of events. Jaynes (1957*a*,*b*) proposes maximizing the entropy, subject to the moment-consistency conditions and additivity requirements, to recover the unknown probabilities that characterize a given data set. For further details of the GME estimator, see Golan et al. (1996).

3. Data and the Empirical Model

The information collected on wine production units consists in five wine cooperatives and two of the most important private firms in the "Bairrada" region in Portugal, in the time period of 1993-2002. Table I presents summary statistics of the data.

Output (Y) is measured by liters of wine produced. Four inputs are distinguished: labour, grapes, energy and capital. The number of workers (L) is converted into annual working hours

Table I: Summary statistics of the variables.

variable		mean	std. dev.	maximum	minimum
Foundation year	I	1951	11.8	1962	1926
Workers	L	38	33.7	138	13
Grapes [Kg]	G	3 022 893	1 978 436.4	8 663 580	531 260
Wine [liters]	Y	2 221 314	1 461 062.9	6534 195	432 000
Energy [Eur]	E	46 851	63 996.7	252 455	6 562
Capital [Eur]	K	94 116	103 395.1	445 151	20 727

for each production unit and the variable G are kilograms of grapes used in the wine production. Energy (E) is measured in constant prices of 1993 using the *Törnqvist-Theil* index (Caves et al. (1982)). Energy consists of electricity and oil. Electricity represents 70% of the energy costs and oil represents 30%. The prices of electricity and oils are obtained from *Energias de Portugal* and *Direcção Geral de Geologia e Energia*. The information available on capital (K) is very poor. The capital variable is defined as a user cost of capital or an opportunity cost of capital[1]. Depreciation is not considered in the user cost since there is no information on the different components of K. Due to lack of information on the acquisition cost of equipment and the maintenance and repair costs, the user cost of capital is calculated as the product of the stock of capital times the Treasury Bond rate. Given the heterogeneity of the stock of capital and lack of information on the prices of each component, the industrial producer price index is used to deflate the variable K. The industrial producer price index is obtained from *Instituto Nacional de Estatística* and *Eurostat*.

The panel data model used in this study is inspired on the model proposed by Battese and Coelli (1995) and Puig-Junoy (2001). The stochastic production frontier model is defined by

$$\mathbf{Y}_{it} = f(\mathbf{X}_{it}; t; \boldsymbol{\alpha}_t) \exp(\mathbf{V}_{it} - \mathbf{U}_{it}); \quad \mathbf{U}_{it} = \mathbf{z}_{it}\boldsymbol{\delta} + \mathbf{W}_{it} \tag{4}$$

where $\mathbf{Y}_{it}$ is the wine produced in time period t by the production unit i, $f(\cdot)$ represents the production technology, $\mathbf{X}_{it} = (\mathbf{G}_{it}, \mathbf{L}_{it}, \mathbf{E}_{it}, \mathbf{K}_{it})$ is the input vector for each production unit i in the time period t, t is a time trend representing exogenous technical change ($t = 1, \ldots, 10$), $\boldsymbol{\alpha}_t$ is a vector of unknown parameters, $\mathbf{V}_{it}$ is a random error, $\mathbf{U}_{it}$ is the nonnegative error component representing technical inefficiency in production, $\mathbf{z}_{it}$ is a vector of explanatory variables associated with technical inefficiency, $\boldsymbol{\delta}$ is a vector of unknown parameters and $\mathbf{W}_{it}$ is a random variable. In the optimization program, it must be ensured that $\mathbf{U}_{it}$ is nonnegative with $\mathbf{W}_{it} \geq -\mathbf{z}_{it}\boldsymbol{\delta}$. Note, for instance, that $\mathbf{V}_{it}$ is usually assumed to be i.i.d. random error that have a Normal distribution, $N(0, \sigma_V^2)$, and is independently distributed of the component $\mathbf{U}_{it}$. Also, $\mathbf{W}_{it}$ is usually defined by the truncation of a Normal distribution, $N(0, \sigma^2)$, such that the truncation point is $-\mathbf{z}_{it}\boldsymbol{\delta}$ (Battese and Coelli (1995)). The usual distributional assumptions for $\mathbf{V}_{it}$ and $\mathbf{W}_{it}$ are not used. We will discuss below the assumptions for $\mathbf{V}_{it}$ and $\mathbf{W}_{it}$ under the GME estimator.

Assuming a translog specification for the stochastic production frontier, the model in (4) can be presented as

$$\mathbf{y}_{it} = \alpha_0 + \alpha_t t + \alpha_{tt}\frac{t^2}{2} + \sum_{j=1}^{4} \alpha_j \mathbf{x}_{jit} + \frac{1}{2}\sum_{j=1}^{4}\sum_{h=1}^{4} \alpha_{jh}\mathbf{x}_{jit}\mathbf{x}_{hit} + \sum_{j=1}^{4} \alpha_{jt}\mathbf{x}_{jit} t +$$
$$+\mathbf{V}_{it} - (\delta_0 + \delta_1 \mathbf{I}_{it} + \delta_2 \mathbf{C}_{it} + \delta_3 \mathbf{A}_t + \mathbf{W}_{it}), \tag{5}$$

[1]An alternative measure of capital (a stock measure) was also used in the empirical application. However, the empirical results are qualitatively similar to the ones reported in this study.

where $\mathbf{y}$ and $\mathbf{x}$ represents the natural logarithm of wine produced and the natural logarithms of the four inputs ($j, h = G$, L, E, K).

The inefficiency component specification includes three variables: two firm-specific variables and time. The age of the production unit (I) is used as a proxy of experience and know-how. It is expected that the higher the experience, the lower the inefficiency level. The variable C (cooperative *versus* private sector) attempts to capture eventual differences in the degree of inefficiency between the cooperative sector and the private sector. C is equal to 0 if i is a wine cooperative or is equal to 1 if i is a private firm. The variable A (year) allows for a time-varying specification of technical inefficiency.

The stochastic production frontier model in (5) accounts for both technical change and time-varying inefficiency effects. The time variable t in the stochastic production function represents shifts in the production technology. Technical progress (regress) occurs if the rate of technical change is positive (negative). The year variable (A) in the technical inefficient component postulates that inefficiency may vary linearly with respect to time.

The production frontier model in (5) can be defined as follows

$$\mathbf{y}_{it} = \mathbf{x}_{it}\boldsymbol{\alpha}_k - \mathbf{z}_{it}\boldsymbol{\delta}_{k'} + \mathbf{V}_{it} - \mathbf{W}_{it} \tag{6}$$

where the vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\delta}$ are expressed by

$$\boldsymbol{\alpha}_k = \sum_{m=1}^{M} p_{km} z_{km} \quad \text{and} \quad \boldsymbol{\delta}_{k'} = \sum_{m'=1}^{M'} q_{k'm'} r_{k'm'} \tag{7}$$

with p_{km} and $q_{k'm'}$ being the probability vectors to be estimated and z_{km} and $r_{k'm'}$ representing the support values. We also define

$$\mathbf{V}_{it} = \sum_{n=1}^{N} w_{itn} v_{itn} \quad \text{and} \quad \mathbf{W}_{it} = \sum_{n'=1}^{N'} w'_{itn'} v'_{itn'} \tag{8}$$

where w_{itn} and $w'_{itn'}$ are the probability vectors to be estimated and v_{itn} and $v'_{itn'}$ are the support values. The Chebychev's inequality is used as a conservative means of specifying sets for the error bounds. We adopt the usual 3σ rule, where $\sigma = 0.68$ represents the standard deviation of the dependent variable. The 3σ rule is not a very strong and restrictive assumption when used in the context of maximum entropy (e.g., Golan et al. (1996) and Campbell and Hill (2006)). In general, the number of the support values (M, M', N, N') is arbitrary depending usually on computational aspects. However, an increase in the number of the support values, keeping the distance among them constant, decreases the variance of the uniform distribution. Since the estimation is not usually improved by choosing more than five points in supports, we define M and M' equal to five, and N and N' equal to three. As will be mentioned below, other numbers of the support points were tested.

Maximum entropy is used to estimate the unknown parameters through the probability vectors $\mathbf{p}$, $\mathbf{q}$, $\mathbf{w}$ and $\mathbf{w}'$. The problem can be defined by the maximization of the following entropy function

$$\begin{aligned} H(\mathbf{p}, \mathbf{q}, \mathbf{w}, \mathbf{w}') \;=\; & -\sum_{k=1}^{21}\sum_{m=1}^{5} p_{km} \ln p_{km} - \sum_{k'=1}^{4}\sum_{m'=1}^{5} q_{k'm'} \ln q_{k'm'} - \\ & -\sum_{i=1}^{7}\sum_{t=1}^{10}\sum_{n=1}^{3} w_{itn} \ln w_{itn} - \sum_{i=1}^{7}\sum_{t=1}^{10}\sum_{n'=1}^{3} w'_{itn'} \ln w'_{itn'} \end{aligned} \tag{9}$$

subject to

$$\begin{aligned}\mathbf{y}_{it} &= \mathbf{x}_{it}\boldsymbol{\alpha}_k - \mathbf{z}_{it}\boldsymbol{\delta}_{k'} + \mathbf{V}_{it} - \mathbf{W}_{it} \\ &= \sum_{k=1}^{21}\sum_{m=1}^{5} p_{km} z_{km} x_{kit} - \sum_{k'=1}^{4}\sum_{m'=1}^{5} q_{k'm'} r_{k'm'} z_{k'it} + \\ &\quad + \sum_{n=1}^{3} w_{itn} v_{itn} - \sum_{n'=1}^{3} w'_{itn'} v'_{itn'}\end{aligned} \tag{10}$$

and

$$\sum_{m=1}^{5} p_{km} = 1\ ;\ \sum_{m'=1}^{5} q_{k'm'} = 1\ ;\ \sum_{n=1}^{3} w_{itn} = 1\ ;\ \sum_{n'=1}^{3} w'_{itn'} = 1, \forall\, k, k', i, t. \tag{11}$$

In this empirical application, there is not a "pure" ill-posed problem, as defined by Golan et al. (1996), since the number of parameters does not exceed the number of observations. However, in this study the condition numbers of matrixes **X** and **z** are high, which means that this is an ill-conditioned problem. The GME estimator is suitable to this kind of problem since is not sensitive to collinearity. It is well-known that collinearity is responsible for inflating the variance associated with the regression coefficients and may also affect the signs of coefficients, as well as statistical inference. Other estimators, such as Ridge Regression, Restricted Least Squares and Maximum Entropy Leuven estimator, can also be applied.

4. Empirical results

The estimates (Est) obtained for the coefficients (Coef) of the stochastic production frontier model in (5), using the GME estimator formalized in (9), are presented in Table II. Three different intervals for the support values with different widths (Support) are used. The last two rows of Table II present the number of iterations (NIter) and the value of the infinity norm of the residual vector (Norm).

The stability of the estimates is verified and is in accordance with other experimental results (e.g., Golan et al. (1996) and Campbell and Hill (2006)). Other support intervals and other number of the support points not reported in Table II were used with similar results to the ones presented in this table. Two reasons justify the choice of the support intervals presented in Table II: (1) these intervals are consistent with the suggestions of Golan et al. (1996) for the case of minimum prior information, and (2) the first support interval (Support1) is consistent with the analysis of the mean values of the variables employed by Lansink et al. (2000, 2001). In these preliminary results, the infinity norm of the residual vector is used as an indicator of the quality of the estimation. Following Campbell and Hill (2006) we can use the percentile method to obtain confidence intervals and evaluate the precision of the GME estimator. Alternatively, confidence intervals based in the Normal distribution, after the estimation of the coefficients' standard errors with 1000 bootstrap data samples, were used to evaluate the estimates. All the parameters are significantly different from zero at 1% significance level. We use an optimization program to solve the maximization problem and all the statistics are computed by us. However, note that the GME estimator with tests and statistics is already available in SAS and Limdep software. The confidence intervals and standard errors are not given in this paper due to space limitations, but they can be obtained from the authors upon request.

The estimated coefficients in the inefficiency model are also reported in Table II. The coefficient δ_1 associated with the variable I (age of the production unit) is negative suggesting that the

Table II: Estimated parameters of the model.

Coef	Est1	Support1	Est2	Support2	Est3	Support3
α_0	0.0414	$z = [-16, -8, 0, 8, 16]$	0.0420	$z = [-8, -4, 0, 4, 8]$	0.0422	$z = [-8, -4, 0, 4, 8]$
α_t	−0.1451	$z = [-16, -8, 0, 8, 16]$	−0.1169	$z = [-8, -4, 0, 4, 8]$	−0.1460	$z = [-8, -4, 0, 4, 8]$
α_{tt}	−0.0053	$z = [-16, -8, 0, 8, 16]$	−0.0053	$z = [-8, -4, 0, 4, 8]$	−0.0053	$z = [-8, -4, 0, 4, 8]$
α_G	0.0345	$z = [-16, -8, 0, 8, 16]$	0.2471	$z = [-8, -4, 0, 4, 8]$	0.2474	$z = [-8, -4, 0, 4, 8]$
α_L	0.5020	$z = [-16, -8, 0, 8, 16]$	0.3892	$z = [-8, -4, 0, 4, 8]$	0.3898	$z = [-8, -4, 0, 4, 8]$
α_E	0.5077	$z = [-16, -8, 0, 8, 16]$	0.3916	$z = [-8, -4, 0, 4, 8]$	0.3960	$z = [-8, -4, 0, 4, 8]$
α_K	0.4080	$z = [-16, -8, 0, 8, 16]$	0.3240	$z = [-8, -4, 0, 4, 8]$	0.3193	$z = [-8, -4, 0, 4, 8]$
α_{GG}	0.0435	$z = [-16, -8, 0, 8, 16]$	0.0297	$z = [-8, -4, 0, 4, 8]$	0.0291	$z = [-8, -4, 0, 4, 8]$
α_{LL}	−0.1933	$z = [-16, -8, 0, 8, 16]$	−0.1375	$z = [-8, -4, 0, 4, 8]$	−0.1364	$z = [-8, -4, 0, 4, 8]$
α_{EE}	0.0244	$z = [-16, -8, 0, 8, 16]$	0.0341	$z = [-8, -4, 0, 4, 8]$	0.0328	$z = [-8, -4, 0, 4, 8]$
α_{KK}	0.1266	$z = [-16, -8, 0, 8, 16]$	0.1152	$z = [-8, -4, 0, 4, 8]$	0.1160	$z = [-8, -4, 0, 4, 8]$
α_{GL}	0.1073	$z = [-16, -8, 0, 8, 16]$	0.0898	$z = [-8, -4, 0, 4, 8]$	0.0894	$z = [-8, -4, 0, 4, 8]$
α_{GE}	−0.0492	$z = [-16, -8, 0, 8, 16]$	−0.0398	$z = [-8, -4, 0, 4, 8]$	−0.0386	$z = [-8, -4, 0, 4, 8]$
α_{GK}	−0.0388	$z = [-16, -8, 0, 8, 16]$	−0.0305	$z = [-8, -4, 0, 4, 8]$	−0.0301	$z = [-8, -4, 0, 4, 8]$
α_{LE}	0.0502	$z = [-16, -8, 0, 8, 16]$	0.0302	$z = [-8, -4, 0, 4, 8]$	0.0301	$z = [-8, -4, 0, 4, 8]$
α_{LK}	−0.0517	$z = [-16, -8, 0, 8, 16]$	−0.0544	$z = [-8, -4, 0, 4, 8]$	−0.0548	$z = [-8, -4, 0, 4, 8]$
α_{EK}	−0.0589	$z = [-16, -8, 0, 8, 16]$	−0.0499	$z = [-8, -4, 0, 4, 8]$	−0.0507	$z = [-8, -4, 0, 4, 8]$
α_{Gt}	0.0186	$z = [-16, -8, 0, 8, 16]$	0.0148	$z = [-8, -4, 0, 4, 8]$	0.0142	$z = [-8, -4, 0, 4, 8]$
α_{Lt}	0.0120	$z = [-16, -8, 0, 8, 16]$	0.0079	$z = [-8, -4, 0, 4, 8]$	0.0074	$z = [-8, -4, 0, 4, 8]$
α_{Et}	0.0111	$z = [-16, -8, 0, 8, 16]$	0.0097	$z = [-8, -4, 0, 4, 8]$	0.0097	$z = [-8, -4, 0, 4, 8]$
α_{Kt}	−0.0173	$z = [-16, -8, 0, 8, 16]$	−0.0125	$z = [-8, -4, 0, 4, 8]$	−0.0120	$z = [-8, -4, 0, 4, 8]$
δ_0	−0.0414	$r = [-16, -8, 0, 8, 16]$	−0.0420	$r = [-8, -4, 0, 4, 8]$	−0.0236	$r = [-6, -3, 0, 3, 6]$
δ_1	−0.0088	$r = [-16, -8, 0, 8, 16]$	−0.0074	$r = [-8, -4, 0, 4, 8]$	−0.0073	$r = [-6, -3, 0, 3, 6]$
δ_2	0.0266	$r = [-16, -8, 0, 8, 16]$	0.0267	$r = [-8, -4, 0, 4, 8]$	0.0185	$r = [-6, -3, 0, 3, 6]$
δ_3	0.1451	$r = [-16, -8, 0, 8, 16]$	0.1169	$r = [-8, -4, 0, 4, 8]$	0.0790	$r = [-6, -3, 0, 3, 6]$
NIter		12 iterations		7 iterations		7 iterations
Norm		0.4183		0.1369		0.1373

younger production units are more inefficient than the older ones. The coefficient δ_2 associated with the variable C (cooperative *versus* private sector) is positive meaning that wine cooperatives tend to be less inefficient than private firms. However, the difference in the technical efficiency values must be analyzed with some precaution since the panel includes only two private firms. The empirical results for the inefficiency model also indicate an increasing technical inefficiency over the ten-year period of the panel. This means that production units are moving away from the production frontier. During the time period of 1993-2002, cooperatives faced difficulties in selling their wine both in the national market and in the foreign markets, yet no information is available for private firms.

Table III reports the output elasticities with respect to the inputs (E_G, E_L, E_E and E_K), as well as the returns to scale (RS) and the rate of technical change (TC), by year and by firms. Output elasticities, RS and TC are computed at the mean values of inputs and output. The parameter estimates used correspond to the support interval that provides the lowest value of the infinity norm of the residual vector. The output elasticities are all positive implying that the marginal product of each input is positive. The differences in the output elasticities of labour, energy and capital suggest a different production structure for cooperatives and private firms and probably the need of a different production frontier specification for each type of firms. The value of returns to scale is, on average, greater than one for each year and increases over time from 1.66 (1993) to 1.87 (2002). The values of TC in Table III indicate technical progress, since all values are positive in every year of the panel. The inclusion of the time variable t in the model (5) with linear, quadratic and cross-product terms allows the rate of technical change to be decomposed into pure technical change $(\alpha_t + \alpha_{tt}t)$ and non-neutral technical change $(\sum_j \alpha_{jt}\mathbf{x}_{jit})$. The values of TC reported in Table III include these two components.

Table III: Output elasticities, returns to scale and technical change by year and firms.

Year	$\mathbf{E_G}$	$\mathbf{E_L}$	$\mathbf{E_E}$	$\mathbf{E_K}$	RS	TC
1993	0.94	0.54	0.05	0.13	1.66	0.17
1994	0.98	0.59	0.05	0.09	1.71	0.18
1995	0.98	0.60	0.05	0.09	1.72	0.17
1996	1.01	0.63	0.06	0.05	1.76	0.17
1997	1.02	0.61	0.08	0.06	1.76	0.16
1998	0.96	0.47	0.13	0.14	1.70	0.12
1999	1.05	0.60	0.09	0.05	1.80	0.15
2000	1.05	0.59	0.11	0.05	1.81	0.14
2001	1.08	0.60	0.12	0.03	1.83	0.13
2002	1.11	0.63	0.13	0.01	1.87	0.13
Cooperatives	1.02	0.62	0.06	0.08	1.79	0.15
Private firms	1.00	0.47	0.27	0.01	1.68	0.16

The technical efficiency levels for the wine cooperatives and private firms are presented in Table IV. The efficiency levels are computed as follows (Battese and Coelli (1995)):

$$TE_{it} = \exp(-\mathbf{U}_{it}) = \exp(-\mathbf{z}_{it}\boldsymbol{\delta} - \mathbf{W}_{it}). \tag{12}$$

Results in Table IV indicate technical inefficiency for both cooperatives and private firms in all years of the study, implying it is possible to increase the production of wine with the same input factors given the production technology. The average value of the technical efficiency for all production units is approximately 0.80. This means that, on average, the quantity of wine produced is only 80% of the possible maximum amount, given the inputs used and the production technology.

Table IV: Technical efficiency levels by year and firms.

Year	Cooperatives	Private firms	Mean
1993	0.82	—	0.82
1994	0.84	—	0.84
1995	0.82	0.73	0.80
1996	0.80	0.73	0.79
1997	0.81	0.82	0.81
1998	0.80	0.66	0.77
1999	0.81	0.77	0.80
2000	0.80	0.74	0.78
2001	0.81	0.72	0.78
2002	0.79	0.72	0.77

5. Conclusions

In this paper, we illustrate the application of the GME estimator to investigate technical efficiency in the wine sector of the "Bairrada" region in Portugal, in the time period of 1993-2002.

The GME estimator reveals a good performance in the estimation of the stochastic production frontier model. The economic analysis indicates that all production units are technically inefficient but wine cooperatives are less inefficient than private firms. The empirical results also indicate technical progress over the time period of the sample and an increasing technical inefficiency over time, implying production units are moving away from the production frontier.

Topics of future research include, in the estimation procedure, a comparison of the GME estimator with other traditional competitors under collinearity, improvements in the optimization program with more flexible restrictions in the error terms, and programming other basic statistics in addition to confidence intervals and coefficients' standard errors. In the efficiency analysis, since the differences in the output elasticities of labour, energy and capital suggest a different production structure for cooperatives and private firms, and probably the need of a different production frontier specification for each type of firms, this will be the subject of detailed study.

References

Aigner, D. J., Lovell, C. A. K. and Schmidt, P. (1977), 'Formulation and estimation of stochastic frontier production function models', *Journal of Econometrics* **6**(1), 21–37.

Battese, G. E. and Coelli, T. J. (1995), 'A model for technical inefficiency effects in a stochastic frontier production function for panel data', *Empirical Economics* **20**, 325–332.

Battese, G. E. and Corra, G. S. (1977), 'Estimation of a production frontier model: with application to the pastoral zone of Eastern Australia', *Australian Journal of Agricultural Economics* **21**(3), 169–179.

Campbell, R. C. and Hill, R. C. (2006), 'Imposing parameter inequality restrictions using the principle of maximum entropy', *Journal of Statistical Computation and Simulation* **76**(11), 985–1000.

Caves, D. W., Christensen, L. R. and Diewert, W. E. (1982), 'The economic theory of index numbers and the measurement of input, output and productivity', *Econometrica* **50**(6), 1393–1414.

Golan, A. (2002), 'Information and entropy econometrics - editor's view', *Journal of Econometrics* **107**, 1–15.

Golan, A. (2007), 'Information and entropy econometrics - volume overview and synthesis', *Journal of Econometrics* **138**, 379–387.

Golan, A., Judge, G. and Miller, D. (1996), *Maximum Entropy Econometrics: Robust Estimation with Limited Data*, John Wiley and Sons, Chichester.

Greene, W. (1993), The econometric approach to efficiency analysis, *in* H. Fried, C. Lovell and S. Schmidt, eds, 'The measurement of Productive Efficiency: Techniques and Applications', Oxford University Press, Oxford.

Greene, W. (2001), New developments in the estimation of stochastic frontier models with panel data. Department of Economics, New York University.

Jaynes, E. T. (1957*a*), 'Information theory and statistical mechanics', *Physical Review* **106**, 620–630.

Jaynes, E. T. (1957*b*), 'Information theory and statistical mechanics. II', *Physical Review* **108**, 171–190.

Kumbhakar, S. C. and Lovell, C. A. K. (2000), *Stochastic Frontier Analysis*, Cambridge University Press, Cambridge.

Lansink, A. O., Silva, E. and Stefanou, S. (2000), 'Decomposing productivity growth allowing efficiency gains and price-induced technical progress', *European Review of Agricultural Economics* **27**(4), 497–518.

Lansink, A. O., Silva, E. and Stefanou, S. (2001), 'Inter-firm and intra-firm efficiency measures', *Journal of Productivity Analysis* **15**, 185–199.

Meeusen, W. and van den Broeck, J. (1977), 'Efficiency estimation from Cobb-Douglas production functions with composed error', *International Economic Review* **18**(2), 435–444.

Puig-Junoy, J. (2001), 'Technical inefficiency and public capital in U.S. states: a stochastic frontier approach', *Journal of Regional Science* **41**(1), 75–96.

Shannon, C. E. (1948), 'A mathematical theory of communication', *The Bell System Technical Journal* **27**, 379–423.