Journal of Quantitative Economics, Vol. 10 No.2, July 2012

ESTIMATION OF JOINT COSTS ALLOCATION COEFFICIENTS USING THE MAXIMUM ENTROPY: A CASE OF MEDITERRANEAN FARMS

RUI MANUEL DE SOUSA FRAGOSO[1]

MARIA LEONOR DA SILVA CARVALHO[2]

Abstract

This paper aims to estimate the farm joint costs allocation coefficients from whole farm input costs. An entropy approach was developed under a Tobit formulation and it was applied to a sample of farms from the 2004 Farm Accounting Data Network base for the Alentejo region, Southern Portugal. Five alternative model specifications respecting error bounds, the central value of the uniform prior support and the generalized cross entropy were tested. Model results were assessed in terms of their precision and estimation power and were compared with real data. The entropy estimation showed a high degree of precision and its practical validity was guaranteed to allocate joint costs, even in the specific context of Mediterranean farms.

Keywords: Generalized maximum entropy; allocation costs; join costs; Alentejo region, FADN data base.

JEL Classification: Q12; C51.

1. Introduction

In spite of the trend towards increased specialization, most farms still have more than one enterprise. Yet, standard farm-accounting information is typically restricted to aggregate or whole farm input expenditures, without revealing production costs per unit of each enterprise's output or by type of cost. Obtaining information on the per-unit production costs for the individual enterprises is particularly important, though, both from a business management and agricultural-policy perspective.

Farm input costs by crop or livestock activities are critical to calculate the output profits and returns to resources, getting an insight of the cost structure of each output activity and

---

[1] **Corresponding author:** University of Évora, Department of Management, CEFAGE/ICAAM, Apartado 94, 7000 Évora, Portugal, Fax: + 351 266 740 807, Phone: +351 266 740 892; Email: rfragoso@uevora.pt

[2] University of Évora, Department of Economics, ICAAM/CEFAGE, Apartado 94, 7000 Évora, Portugal, Email: leonor@uevora.pt

allowing comparisons between technologies, farms, regions and even at an international level (Lips, 2009). Standard farm-accounting information is typically restricted to whole farm input expenditures, without revealing production costs per unit of each enterprise's output or by type of cost. In this situation, there is a non-separation of costs and processes associated to the production of more than one output activity (Trenchard and Dixon, 2003). Hence, calculating the activity costs is not allowed and one can set a management accounting problem of joint costs that invites to use product linked allocations.

The problem of joint costs allocation is frequently solved assuming some behavioural rules for input allocations, such as equal fixed costs per activity (Just et al., 1983), crop acreage or gross outputs. Hemmer (1996) developed methods to allocate joint costs according to the profit maximization, using the net revenue and the statistical variance associated. However, some restrictive assumptions have to be imposed.

In the agricultural sector there are several approaches for estimating crop-specific inputs (Just *et al*., 1983; Shumway *et al*., 1984; and Lence & Miller, 1998a and 1998b) and most of them are based on the relationship existing between input allocation and production coefficients under the assumption of profit maximization. However, this assumption is often rejected by empirical production analysis when micro level data are used (Love, 1999; and Zhang & Fan, 2001).

Traditionally, the tools used in these approaches are linear regression techniques (LRT), Bayesian estimation techniques (BES) and linear programming (LP), but their use creates some practical problems.

According to Peeters and Surry (2002), for the estimation of farm joint costs allocation coefficients, it is necessary to ensure the accounting balance between total revenue and total costs. Once, error terms of the various input-demand equations are not independent from each other, the system of input-demand equations is singular and invalidates the use of LRT techniques (Bewley, 1986). Another well known result of the LRT technique is the negative input-demand coefficients that could be estimated.

The non-negativity of the estimated input-demand coefficients can be ensured using BES techniques or inequality least squares methods (Moxey & Tiffin, 1994). However, the application of these alternative methods, being very heavy, does not allow incorporating the accounting constraints needed to be treated separately in the system.

The entropy approach, namely the generalized maximum entropy (GME) method proposed by Golan et al. (1996), is a good alternative to estimate economic data that may be ill-behaved or ill-posed. This is an efficient approach that is superior to other traditional estimation methods because it does not need strong parametric or behavioural assumptions, works well with small samples and the covariates are highly correlated. In addition, the entropy approach allows the incorporation of out sample information on both the multinomial probabilities and the response parameters (Soofi, 1992).

Lence & Miller (1998a) proposed a generalized maximum entropy (GME) approach for simultaneously estimating multi-output production function and recovering input allocations. This approach was also used by Zhang & Fan (2001) to estimate crop-specific production technologies in Chinese agriculture. Leon *et al*. (1999) and more recently Peeters & Surry (2005) used maximum entropy and FADN base to derive farm crop coefficients in Brittany, France.

Garvey & Britz (2002) estimated agricultural input allocation from EU farm accounting data. Hansen & Surry (2006) estimated input quantities for different production branches based on regional economic accounts. Fragoso *et al.* (2008) have demonstrated the adherence of a minimum cross entropy model as a process leading to dynamic spatial information generation and spatial information allocation in the Alentejo region.

The objective of this paper is to estimate the farm cost allocation coefficients from whole farm input costs using a model based on the GME approach. This model follows the research of Peeters & Surry (2002) and is applied to a sample of 30 farms from the Farm Accounting Data Network (FADN) base in the Alentejo region, South of Portugal. The study also assesses the quality of the GME approach by looking at the precision of the coefficients extracted from the whole farm joint costs and compares them with the observed input-output coefficients. In order to improve the GME model results, a cross entropy formulation is also tested.

In Mediterranean regions, farms have a great diversity of crops and production technologies, which makes it hard for any attempt to estimate farm cost allocation coefficients from incomplete data, such as that observed in whole farm input cost from the FADN base. Hence, there are few studies with the purpose of estimating joint cost allocation coefficients in the multi-crop context of Mediterranean areas, being this study the first one done in Portugal using an entropy approach.

There are several reasons to choose Portugal for this study and namely the Alentejo region. This region, with a Mediterranean climatic pattern, has similar agriculture, farm structure and production technologies to other Mediterranean areas, such as the ones in the south of Spain and France.

This paper is organized in four more sections. Section 2 describes the model used to estimate the input cost coefficients. Section 3 reports the data and the empirical model implementation to the sample of farms. Section 4 relates to the presentation and discussion of results and section 5 provides the conclusion and final remarks.

2. The Generalized Maximum Entropy Tobit Model

The estimation of joint costs allocation coefficients from whole farm accounting data is often based on the derived demand of each farm input as a function of several farm outputs (Peeters & Surry, 2002). In this approach, inputs and outputs are both expressed as costs and revenues in a system of linear equations, where they are treated as dependent variable and independent variable, respectively.

Considering *I* input types, a sample of *T* farms producing *K* outputs, the estimation problem of cost allocation coefficients can be expressed as the following linear system of equations:

$$x_i^t = \sum_{k=1}^{K} \alpha_{k,i} y_k^t + u_i^t \qquad \text{for i= 1,2,…I and t=1,2,…T} \qquad ...(1)$$

where x_i^t is the total joint cost by input type *i* and farm *t;* $\alpha_{i,k}$ is the unknown cost allocation coefficient by input *i* and output farm *k*; y_k^t is the gross output value of activity *k* in farm *t;* and u_i^t is the noisy component specified by input *i* and by farm *t*.

The estimation of the cost allocation coefficients $\alpha_{k,i}$ can be formulated as a ME problem. The basis for the entropy formulation can be seen in Shannon (1948), Jaynes (1957a; 1957b), Kullback (1959), Gokhale and Kullback (1978), Levine (1980), Jaynes (1984), Csiszár (1991) and Golan *et al.* (1996a).

The aim of the ME problem is to maximize the Shannon's entropy measure, $H(p) = \sum_k p_k \ln(p_k)$, where p_k is the probability of observing outcome *k*. To assigning or recovering the unknown probabilities p_k, the entropy measure should be subject to appropriated information-moment relations and adding up constraints for the probabilities (Jaynes, 1957a; 1957b):

$$\text{Max } H(p) = -\sum_k p_k \ln(p_k) \quad ...(2)$$

s.t.

$$\sum_k p_k y_k = x \quad ...(3)$$

$$\sum_k p_k = 1 \quad ...(4)$$

$$p_k \geq 0 \quad ...(5)$$

where p_k is the probability of observing outcome *k*, y_k is the independent variable and *x* is the dependent variable.

The maximum value of the entropy function in (2) is subject to the constraints of information-moment relations (3), adding-up constraints (4) and non-negativity conditions (5). According to the ME principle, there is a unique probability distribution that maximizes Shannon's entropy measure and is consistent with the available information contained in the data. The selected probability distribution is the one that, among others, satisfies the condition of information consistency, with the minimum information content. The Lagrangian function corresponding to the former formulation is the following:

$$L = -\sum_k p_k \ln(p_k) + \lambda\left(x - \sum_k p_k y_k\right) + \delta\left(1 - \sum_k p_k\right) \quad ...(6)$$

The first order conditions are:

$$\frac{\partial L}{\partial p} = -\sum_k \ln(p_k) - 1 + \sum_k y_k \hat{\lambda} - \delta = 0 \quad ...(7)$$

$$\frac{\partial L}{\partial \lambda} = x - \sum_k p_k y_k = 0 \quad ...(8)$$

$$\frac{\partial L}{\partial \delta}=1-\sum_k p_k=0 \qquad ...(9)$$

If *H* is strictly concave, there is a unique interior solution to the problem. Solving the first order conditions, $\hat{p}_k$ is given by a function of the Langrange multiplier $\hat{\lambda}$ on the moment constraint (3):

$$\hat{p}_k=\frac{\exp\left(-y_k\hat{\lambda}\right)}{\sum_k \exp\left(-y_k\hat{\lambda}\right)}=\frac{\exp\left(-y_k\hat{\lambda}\right)}{\Omega\left(\hat{\lambda}\right)} \qquad ...(10)$$

This result shows that there is a unique non-linear relation between $\hat{p}_k$ and y_k through $\hat{\lambda}$, and as $\hat{p}_k(\lambda)$ is given by an exponential function, the non-negativity requirement, $p_k \geq 0$, is satisfied.

The associated information matrix is:

$$I(\lambda)=\sum_k p_k(\lambda)y_k^2-\left(\sum_k p_k(\lambda)y_k\right)^2=Var(y) \qquad ...(11)$$

The GME formulation allows to treat the noisy component u_i^t. The unknown and unobservable coefficients $\alpha_{k,i}$ and u_i^t from equation (1) are rewritten as the expected value of a probability distribution, defined over a set of known and discrete "support values" (Golan & Judge, 1996 and Golan et *al*., 1996a):

$$\alpha_{k,i}=\sum_{m=1}^{M} p_{k,i}^m . z_{k,i}^m , \quad \forall k\in\{1,2,\ldots K\} \text{ e } \forall i\in\{1,2,\ldots I\} \qquad ...(11)$$

$$u_i^t=\sum_{n=1}^{N} v_{i,t}^n . w_{i,t}^n \quad \forall i\in\{1,2,\ldots I\} \text{ e } \forall t\in\{1,2,\ldots T\} \qquad ...(12)$$

where $p_{k,i}^m$ and $w_{i,t}^n$ are the corresponding unknown probability vectors for $\alpha_{k,i}$ and u_i^t, respectively; $z_{k,i}^m$ are the support vectors of dimension *M* for *k* production coefficients associated with input *i*; and $v_{i,t}^n$ are the support vectors of dimension *N* associated with *i* inputs for the *T* errors terms.

Thus, under the GME formulation the problem is written as:

$$\text{Max } H(p,w)=-p' \ln(p) - w'\ln(w) \qquad ...(13)$$

s.t.

$$x=\alpha y+u=\left(p'z\right)y+wv' \qquad ...(14)$$

$$z=1; \quad w=1; \quad \text{and} \quad \alpha=p'z \qquad ...(15)$$

Solving the first order conditions of the Lagrangian function, $\hat{p}$ and $\hat{w}$ are given by:

$$\hat{p}_{k,m}=\frac{q_{k,m}\exp\left(z_{k,m}y_k\hat{\lambda}\right)}{\sum_t q_{k,t}\exp\left(z_{k,t}y_k\hat{\lambda}\right)}=\frac{q_{k,m}\exp\left(z_{k,m}y_k\hat{\lambda}\right)}{\Omega_k\left(\hat{\lambda}\right)} \quad ...(16)$$

$$\hat{w}_{i,n}=\frac{u_{i,n}\exp\left(v_{i,n}\hat{\lambda}\right)}{\sum_t u_{i,t}\exp\left(v_{i,t}\hat{\lambda}\right)}=\frac{u_{i,n}\exp\left(v_{i,n}\hat{\lambda}\right)}{\Omega_i\left(\hat{\lambda}\right)} \quad ...(17)$$

The unknown Lagrange parameters λ, which link $p_{k,m}$ to y_k , are associated to the optimal solution of $\hat{p}_{k,m}$ and $\hat{w}_{i,n}$ that satisfy the constraints (14) and (15). As referred before, the exponential forms guarantee that $\hat{p}$ and $\hat{w}$ are always positive. The partition functions Ω are the normalization factors, which ensure that $\hat{p}$ and $\hat{w}$ have the propreties of probabilities and provide information on the distribution probabilities.

The information matrix is given by the following Hessian matrix of the objective function:

$$\nabla_{(p,w)(p',w')}I(p,w)=\begin{bmatrix}P^{-1} & 0\\ 0 & W^{-1}\end{bmatrix} \quad ...(18)$$

where ∇ is a positive definite matrix for *p>0* and *w>0*, which satisfies the sufficient condition for strict convexity and hence the problem solution is unique.

To avoid problems related with the existence of cost items with zero in some farms of the sample, a Tobit formulation is recommended (Golan *et al.*, 1996a; and Peteer & Surry, 2002).

Under this formulation the observations are ordered as follows:

$$\begin{bmatrix}x_1>0\\ x_2=0\end{bmatrix}=\begin{bmatrix}y_1\\ y_2\end{bmatrix}p^m_{k,i}.z^m_{k,i}+\begin{bmatrix}v_1'w_1\\ v_2'w_2\end{bmatrix} \quad ...(19)$$

Thus, given the support vectors $z^m_{k,i}$, $v^n_{1,t,i}$, and $v^n_{2,t,i}$ with the sample data on inputs x^t_i and outputs y^t_k, to find the positive probability vectors $\hat{p}^m_{k,i}$, $\hat{w}^n_{1,t,i}$ and $\hat{w}^n_{2,t,i}$, the GME formulation was re-parameterized in the following GME-tobit problem:

$$\text{Max}(p,w_1,w_2)=-\sum_{i=1}^{I}\sum_{k=1}^{K}\sum_{m=1}^{M}p^m_{k,i}\ln\left(p^m_{k,i}\right)-$$

$$-\sum_{t=1}^{T}\sum_{i=1}^{I}\sum_{n=1}^{N}w^n_{1,t,i}\ln\left(w^n_{1,t,i}\right)-\sum_{t=1}^{T}\sum_{i=1}^{I}\sum_{n=1}^{N}w^n_{2,t,i}\ln\left(w^n_{2,t,i}\right) \quad ...(20)$$

s.t.

$$x^t_{1,i}=\sum_{k=1}^{K}\sum_{m=1}^{M} p^m_{k,i}.z^m_{k,i}.y^t_{1,k}+\sum_{n=1}^{N} v^n_{1,i,t}.w^n_{1,i,t}, \quad \forall t_1 \text{ and } \forall i\in\{1,..,$$ ...(21)

$$0=\sum_{k=1}^{K}\sum_{m=1}^{M} p^m_{k,i}.z^m_{k,i}.y^t_{2,k}+\sum_{n=1}^{N} v^n_{2,i,t}.w^n_{2,i,t}, \ \forall t_2 \text{ and } \forall i\in\{1,..,I\}$$ ...(22)

$$\sum_{m=1}^{M} p^m_{k,i}=1, \quad \forall k\in\{1,2,\ldots K\} \text{ and } \forall i\in\{1,2\ldots I\}$$ ...(23)

$$\sum_{n=1}^{N} w^n_{1i,t}=1, \quad \forall i\in\{1,2,\ldots I\} \text{ and } \forall t_1\in\{1,2\ldots T_1\}$$ ...(24)

$$\sum_{n=1}^{N} w^n_{2,i,t}=1, \quad \forall i\in\{1,2,\ldots I\} \text{ and } \forall t_2\in\{1,2\ldots T_2\}$$ ...(25)

$$\sum_{i=1}^{I}\sum_{m=1}^{M} p^m_{k,i}.z^m_{k,i}=1, \quad \forall k\in\{1,2,\ldots K\}$$ ...(26)

where T_1 are the farms with positive observations and T_2 are the farms with zero observations for input *i*.

The data consistency constraints (21)-(22) treat the relations in equation (1) as an interdependent system of equations in which all inputs *i* are taken in account simultaneously and positive and null observations are separated. These constraints guaranty, in the optimization model, the balance between total costs and revenues. However, it is also necessary to consider that $\sum_{i=1}^{I}\alpha_{k,i}=1$ for all *k* and $\sum_{i=1}^{I}u_i^t=0$ for all *t*. This is brought into the model in equation (26), which ensures that the sum by each input *i* of the probability vectors $p^m_{k,i}$, weighted by the support vector of dimension *M,* is equal to 1 for all *k*.

Equations (23)-(24) are the adding-up constraints that are related with the probability properties and make in the model the normalization of the probability values of $\hat{p}^m_{k,i}$ and $\hat{w}^n_{i,t}$ concerning the dimensions *M* and *N*, respectively.

When the GME results are unrealistic, the entropy approach allows solving the problem by introducing out sample prior information in the model. This is done by minimizing the cross entropy (CE) between the unknown coefficients and the prior information. In the new formulation, called general cross entropy (GCE), the structure of the GME model remains equal, and the objective of maximum entropy becomes the minimization of the cross entropy. Thus, the distance between the estimates of unknown coefficients *p* and the previous out sample information is minimized. The formulation of the new objective function is given by:

$$\text{Min}(p,w_1,w_2)=\sum_{i=1}^{I}\sum_{k=1}^{K}\sum_{m=1}^{M} p_{k,i}^{m}\ln\left(p_{k,i}^{m}/q_{k,i}^{m}\right)+$$

$$+\sum_{t=1}^{T}\sum_{i=1}^{I}\sum_{n=1}^{N} w_{1,t,i}^{n}\ln(w_{1,t,i}^{n})+\sum_{t=1}^{T}\sum_{i=1}^{I}\sum_{n=1}^{N} w_{2,t,i}^{n}\ln(w_{2,t,i}^{n}) \quad ...(27)$$

where $q_{k,i}^{m}$ is the vector of prior information known that is used to bring model results closer to the observed data.

Operationally, the optimal solution can be achieved using non-linear software. In this case, the GAMS (General Algebraic Modeling System) software was used.

3. Data and Empirical Model Implementation

The data used were from the 2004 FADN base for the Alentejo region. This is a management accounting database that contains general information for a regional sample of farms, such as gross output value, crops acreage and joint costs of crops and livestock.

The gross output value is defined in euros by activity (crops and livestock). Crops acreage is presented in hectares and livestock in heads. The joint costs are given per item for the whole farm and it is not possible to know directly from data the costs of each output activity.

Table 1 presents the distribution in the Alentejo 2004 FADN base of gross margin and number of farms by technical and economic specialization. Farms specialized in big crops and permanent crops represent almost 40% of the regional gross margin and of the number of farms. Farms specialized in herbivores are also representative in Alentejo region, once they have about 30% of the gross margin and of farms. Hence, from the Alentejo 2004 FADN base was extracted a convenience sample of 30 farms, 24 of them being specialized in big crops (80%) and 6 in permanent crops (20%).

Table 2 presents the main characteristics of the sample such as average, maximum and minimal values of gross output, crops acreage and joint cost items.

The sample of farms includes 8 output crops and 5 joint cost items. The crops are wheat, maize, rice, other cereals, horto-industrials and melons, oilseeds, olive trees and vineyards, and represent an important share on the agricultural value in the Alentejo region. According to the methodology used by the Ministry of Agriculture to calculate the standard gross margins, the joint cost items considered are plants and seeds, fertilizers, pesticides, other costs with crops and gross margin. The values of these items were taken directly from the FADN base, but the gross margin had to be calculated as the difference between the whole farm production value and input costs.

Table 1. Percentage of gross margin and number of farms in the Alentejo region in 2004

Technical and economic specialization	*Gross margin*	*Number of Farms*
Big crops	30.3	23.0
Vegetables	0.6	1.5
Permanent crops	6.2	16.7
Herbivores	29.4	30.9
Livestock without land	1.6	0.7
Mixed farming of crops	6.9	4.8
Mixed farming of livestock	10.3	5.2
Mixed farming of crops and livestock	14.7	17.1
Total	100.0	100.0

Source: FADN, 2004

Table 2. Description of the farm sample data

	Average	*Maximum*	*Minimum*	*Percentage*
Gross output (Euros)				
Wheat	3331	12308	0	5.3
Maize	6089	107384	0	9.7
Rice	1930	22899	0	3.1
Other cereals	2752	47276	0	4.4
Horto-industrials and melons	21269	352796	0	33.8
Oilseeds	954	13398	0	1.5
Olive trees	1480	12833	0	2.4
Vineyards	25107	522302	0	39.9
Acreage (ha)				
Wheat	19.3	67.0	0	38.0
Maize	3.6	47.3	0	7.0
Rice	2.1	21.5	0	4.1
Other cereals	8.1	78.4	0	16.0
Horto-industrials and melon	4.5	5.0	0	8.8
Oilseeds	6.9	55.0	0	13.7
Olive trees	3.6	24.9	0	7.2
Vineyards	2.6	28.7	0	5.2
Joint cost items (Euros)				
Seeds and plants	4715	40301	0	7.5
Fertilizers	7437	71573	0	11.8
Pesticides	5388	55105	0	8.6
Other costs	5791	81821	0	9.2
Gross margin	39581	440767	-6019	62.9

Source: FADN, 2004

Most of the gross output value is allocated to vineyards (40%) and horto-industrial and melons (34%). The average and maximum values are €25,107 and €21,269, and €522,302 and €352,769, respectively. However, their acreage shares are low relatively to the ones of gross output value. Vineyards represent only 5% of the total acreage and horto-industrials and melons are 9%.

Maize is the third crop in terms of gross output value, representing 10% of total, with an average of €6,089 and a maximum of €107,384. Its acreage (7% of total) is less than the acreage of wheat (38%), oilseeds (14%) and other cereals (16%), which have lower gross output values.

Contrary to what is frequently argued, the use of simple methods for allocating farm joint costs, based on gross output and crops acreage, are not required in this case. In our sample, these two items are uncorrelated the costs of each activity so its use to allocate farm joint costs will lead to erroneous results.

Regarding cost items, gross margin represents 63% of gross output value. Their average and maximum values are the highest among cost items, €39,581 and €440,767, respectively. The cost items of fertilizers, other costs, pesticides and seeds and plants, represent 11.8%, 9.2%, 8.6% and 7.5% of gross output value, respectively. Among them, the average varies between €4,715 in seeds and plants and €7,437 in fertilizers, and the maximum value varies from €40,301 in seeds and plants to €81,821 in other costs.

The gross output maximum value is far from its average in vineyards, maize, other cereals and horto-industrials and melons. In these cases, the maximum values are greater than the average almost twenty times. Relatively to cost items, these differences are smaller, and vary from 8.5 times in seeds and plants to 14 times in other costs with crops.

In the empirical implementation of the GME-tobit model, special attention was given to the definition of the error term and of appropriate indicators to assess the statistical properties of the GME estimators. The practical validity of the model was also discussed.

In the GME formulations, the bounds of the error term are defined as some multiple of the standard deviation of the dependent variable (Golan, et al. 1996b). In this case, the two error support vectors (v_1 and v_2) were defined with the center on zero and the endpoints of intervals $[-3\sigma, 3\sigma]$ based on the “3σ rule” (Pukelsheim, 1994). The σ is obtained assuming a uniform and censored data distribution (Golan et al., 1997). Then, the uniform variance (s^2) is used as an estimator of σ such as:

$$\ldots\ \sigma_{t,i} \equiv \sqrt{s^2} = \sqrt{[\mathrm{Max}(x_i^t) - \mathrm{Min}(x_i^t)]^2/12}, \quad \forall t \in \{1,\ldots T\} \text{ and } \forall i \in \{1,\ldots I\} \qquad \ldots(28)$$

Golan et *al.* (1996a and 2001) studied the properties of the ME estimators of constrained and unconstrained system of equations. If the ME estimators $\alpha_{k,i}$ are consistent and asymptotically normal, it is possible to show that entropy ratio statistic (*ER*) for different parameters of unknown distribution has a limiting distribution (Peeters & Surry, 2002).

The null-hypothesis considers that the sum of the output coefficients $\alpha_{k,i}$, for a given output *k* is equal to one $\left(H_0 = \sum_k \alpha_{k,i} = 1\right)$. Given a system of linear input-demand equations, *ER* is defined as follows:

$$ER=\left\{\sum_k \alpha_{k,i}=1\right\}=2S_{UN}\left(\sum_k \alpha_{k,i}\neq 1\right)-2S_R\left(\sum_k \alpha_{k,i}=1\right) \qquad ...(29)$$

where S_{UN} and S_R are the values of unrestricted and restricted ME functions, respectively.

The *ER* follows a χ^2 distribution with *k* degrees of freedom and the null-hypothesis H_0 cannot be rejected if the ER statistic is smaller than the critical value of χ^2 at a given significance level.

The information content in the estimates can be assessed using the normalized entropy (Golan *et al.*, 1996a, and 1996b). The entropy reaches its maximum value if the uncertainty is also maximal, which is obtained when the information moment constraints are unrestricted and the distribution of probabilities is uniform over all states. Any information added will reduce the uncertainty and the proportion of the remaining uncertainty is measured by the normalized entropy as:

$$S(\hat{p})=\frac{-\sum_{i=1}^{I}\sum_{k=1}^{K}\sum_{m=1}^{M}\hat{p}_{k,i}^{m}\ln\left(\hat{p}_{k,i}^{m}\right)}{IK\ln(M)} \qquad ...(30)$$

where $S(\hat{p})$ is the normalized entropy and $I\times K$ is the total number of coefficients that have to be estimated, considering the *M* support values.

The $S(\hat{p})$ value can vary between zero and one. Before adding any information or theoretical constraint the uncertainty is maximal, the value of p^m is *1/M* and the entropy of its probability distribution is *ln(M)*. For the $K\times I$ joint entropy the maximum value is equal to *IKln(M)* and $S(\hat{p})=1$. When the $S(\hat{p})=0$, there is no uncertainty and the information content of the estimates from the data is maximal.

The normalized entropy indicator for the probability distributions associated to the error terms w_1 and w_2 is calculated as:

$$S(\hat{w})=\frac{-\left[\sum_{i=1}^{I}\sum_{t=1}^{T}\sum_{n=1}^{N}\hat{w}_{1,t,i}^{n}\ln\left(\hat{w}_{1,t,i}^{n}\right)+\hat{w}_{2,t,i}^{n}\ln\left(\hat{w}_{2,t,i}^{n}\right)\right]}{IT\ln(N)} \qquad ...(31)$$

Here, $S(\hat{w})$ can also vary between zero and one and *ITln(N)* is the entropy level of the uniform distribution that represents the maximum uncertainty.

Another indicator to assess both the model precision and validity, frequently used, is the empirical mean squared error (MSE), given by:

$$MES=\frac{\sum_k\sum_I\left(\hat{\alpha}_{k,i}-\alpha_{k,i}\right)^2}{KI} \qquad ...(32)$$

where $\hat{\alpha}_{k,i}$ and $\alpha_{k,i}$ are the estimated and observed cost allocation coefficients; $K\times I$ is the number of estimated coefficients.

Many different indicators can be used to assess the predictive power of the GME-tobit model. Like in Peeters & Surry (2002), the "*pseudo* R^2", defined as the square of the correlation between predicted and observed values for each cost item, was used:

$$R_i^2=\frac{\left[\sum_{t=1}^{T} x_t^i \hat{x}_t^i\right]^2}{\sum_{t=1}^{T} \hat{x}_t^{i^2} x_t^{i^2}}, \text{ with } \hat{x}_t^i=\sum_{k=1}^{K} \hat{\alpha}_{k,i}\, y_k^t,\ \forall i\in\{1,\ldots I\} \quad ...(33)$$

where $\hat{x}_t^i$ are the estimated values of inputs costs.

The R^2 was computed by cost item and its value can vary between zero and one. The closer this value is to one, the better is the predictive power of the model.

After studying the statistic and econometric properties of estimates, the estimated costs and the observed costs were compared. The observed costs were obtained crossing the total values of the sample data with the input costs structure that are used by the Portuguese Ministry of Agriculture to calculate the standard gross margins per output crop activity. In order to measure the in-sample prediction performance, the Percentage of Absolute Deviation (PAD) was used. This PAD was calculated by cost item and output crop for all farms, as the difference between estimated and observed cost items divided by the total number of farms in the sample:

$$PAD_k^i=\frac{1}{T}\sum_{t=1}^{T}\left(C_{k,t}^i-\hat{C}_{k,t}^i\right)\times 100, \quad \text{with} \quad \hat{C}_{k,t}^i=\alpha_k^i y_k^t \quad ...(34)$$

where $\hat{C}_{k,t}^i$ and $C_{k,t}^i$ are the estimated and observed cost items per crop and farm.

4. Results

Model results were obtained for a sample of 30 farms from the Alentejo 2004 FADN database, under five simulations of the model specification. Results are presented in terms of the precision estimates and prediction power, and then, a discussion about its validation is done.

The implementation of the GME-Tobit model requires the choice of the support vector *z* of dimension *M* and the support vector *v* of dimension *N,* which are informative uniform distributions to be considered as priors when any prior information is not available (Howitt and Reynauld, 2003). Several studies have shown that GME estimates are fairly sensitive to the choice of the bounds of the support values, particularly regarding the error terms (Fraser, 2000; Leon et al., 1999; Paris & Caputo, 2001; Preckel, 2001; and Huang et al., 2007).

The natural bounds of z^m are zero and one. Fragoso *et al.* (2008), Martins *et al*. (2011) and Howitt and Reynauld (2003) fixed *M* equal to 3 and considered for the vector of support *z* the set of {0.0, 0.5, 1.0}.

In this research, model results were tested under five simulation scenarios, including changes on the error bounds and on the central value of the uniform prior, and the use of prior information for estimates.

In Table 3, the characteristics of support vectors z and v and the simulation results concerning the indicators of model precision, such as entropy ratio statistics, normalized entropy and empirical mean square error, are presented.

Table 3. Results of simulations for indicators of model precision

Simulations	$z_{k,i}^{m}$	$v_{t,i}^{n}$	*ER*	*S($\hat{p}$)*	*S($\hat{w}$)*	*MSE*
GME-1	{0.0,0.5,1.0}	[-1,1]	22.8*	0.688	0.994	0.596
GME-2	{0.0,0.5,1.0}	[-0.5,0.5]	22.8*	0.663	0.990	0.545
GME-3	{0.0,0.25,1.0}	[-1,1]	17.2**	0.814	0.993	0.596
GME-4	{0.0,0.25,1.0}	[-0.5,0.5]	17.2**	0.784	0.989	0.541
GCE	{0.0,0.25,1.0}	[-1,1]		0.551	0.997	0.029

Notes: * - > critical value of χ2 at a significance level of 1%

** - > critical value of χ2 at a significance level of 5%

Source: Models results

Two simulations consider a prior distribution equal to {0.0, 0.5, 1.0}, and error bounds of [-1, 1] (model GME-1) and [-0.5, 0.5] (model GME-2). Other two simulations consider the same error bounds, but the central value of the uniform prior *z* was changed to 0.25 in order to approximate the values of estimates to zero, that is, with a z={0.0, 0.25,1.0} and error bounds of [-1,1] (model GME-3) and [-0.5,0.5] (model GME-4). Considering the use of prior information known about the cost allocation coefficients, in order to have more realistic estimates an additional simulation was done. In this case the initial GME model was transformed in its GCE formulation (model GCE), considering *z*={0.0, 0.25,1.0} and *v* in the interval [-1,1].

The entropy ratio allows to assess the consistent estimation of the cost allocation coefficients $\alpha_{k,i}$, by testing the null hypothesis, which is rejected when the ER statistic is greater than the critical value of χ^2 at a given significance level. The null hypothesis stays that the sum of the output coefficients $\alpha_{k,i}$ for a given output *k* is equal to one.

For model simulations GME-1 and GME-2, having the uniform prior *z* centred in 0.5, the ER statistic is 22.8. This value is greater than 15.51 and 20.09, the critical values of χ^2 for a significance level of 5% and 1%, respectively. Then the null hypothesis is rejected in both cases. Model simulations GME-3 and GME-4, having the prior close to zero (0.25), present a ER statistic of 17.2, which also allows rejecting the null hypothesis, but only at a significance level of 5%.

For the GCE model simulation the null hypothesis was not tested, because in the CE formulation the entropy instead of being maximized is minimized and the theoretical principles and assumptions adopted in the GME approach are also applied in the GCE model.

The normalised entropy assesses the information content in a model, measuring the remaining uncertainty. This is the amount of new information that is generated by the GME estimators. A greater value means that the model solution for the recovered cost allocation is closer to the prior distribution indicated in the support values.

All model simulations have high values of $S(\hat{p})$, showing that the proportion of new information generated is relevant. The highest values were obtained for model simulations GME-3 and GME-4. In these models the information generated reaches 81.4% in the first one and 78.4% in the last one. For the other model simulations the obtained values are smaller.

The GCE simulation is the one presenting the lowest $S(\hat{p})$, since cross entropy copes with additional out-sample information in the estimation process. Therefore, the uncertainty of the estimates is lower and the generated information is smaller.

These results suggest that more information is generated when the uniform prior is centred, closer to zero, on 0.25, specially when is considered the range of [-1,1] for the error bound.

Regarding the noise ratio $S(\hat{w})$, all model simulations have values above 90%, but they do not vary with changes in support vectors, even when the error bounds are reduced. The invariant results to the choice of the support set are similar to the ones obtained by other authors (Peeters & Surry, 2002).

The MSE results are consistent with the values obtained for the $S(\hat{p})$ indicator. In the four GME simulations, the MSE varies between 0.596 and 0.541. The best values are obtained with model simulations GME-2 and GME-4, which have the support error bounded at [-0.5, 0.5]. However the choice of the central value of the uniform prior *z* has a slight influence in the model results.

In general, the results are not quite different from other authors. Golan et al. (1996b) had, as results of a sampling experiment, values of MSE for GME estimators between 0.41 and 0.183, and for the MSE of the ME-ML logit estimator 0.555.

However the possibility of incorporating prior information known about the estimates in the model through a CGE formulation allows decreasing the value of MSE to 2.9%.

In Table 4, the results relatives to the values of "pseudo-R^2" computed for each joint cost item are presented.

Table 4. Results of simulations for the "*pseudo-R^2*" statistic

Simulations	*Seeds and plants*	*Fertilizers*	*Pesticides*	*Other costs*	*Gross margin*
GME-1	0.943	0.983	0.968	0.863	0.993
GME-2	0.953	0.983	0.972	0.895	0.995
GME-3	0.941	0.983	0.967	0.850	0.994
GME-4	0.952	0.983	0.972	0.888	0.995
GCE	0.957	0.975	0.965	0.895	0.995

Source: Models results

The "pseudo R^2" statistic was used to assess the predictive power of the GME-Tobit models. For all cost items, the "pseudo R^2" is close to one. The cost item having the lower results is "other costs", where the "pseudo R^2" varies between 0.85 (model GME-3) and 0.895 (models GME-2 and GCE). For the other costs items, the "pseudo R^2" is always above 94%, being greater than 99% in the case of gross margin.

From these results the conclusion is that all model specifications have a high prediction power and the choice of the support sets does not have any influence.

The estimated and observed cost allocation coefficients ($\alpha_{k,i}$) are presented in Table 5. Table 6 shows the values of the PAD indicator.

Table 5. Estimated and observed cost allocation coefficients

	Model simulations					
	GME-1	*GME-2*	*GME-3*	*GME-4*	*GCE*	*Observed*
Seeds and plants						
Wheat	0.196	0.206	0.196	0.204	0.165	0.152
Maize	0.120	0.090	0.125	0.093	0.082	0.094
Rice	0.192	0.173	0.192	0.176	0.067	0.066
Other cereals	0.160	0.128	0.162	0.131	0.108	0.114
Horto-industrials	0.115	0.110	0.115	0.110	0.108	0.060
Oilseeds	0.189	0.180	0.189	0.179	0.061	0.061
Olive trees	0.197	0.191	0.197	0.192	0.062	0.061
Vineyards	0.020	0.007	0.026	0.010	0.006	0.060
Fertilizer						
Wheat	0.204	0.208	0.204	0.208	0.201	0.192
Maize	0.223	0.212	0.223	0.211	0.159	0.108
Rice	0.200	0.196	0.200	0.197	0.157	0.161
Other cereals	0.214	0.224	0.214	0.224	0.099	0.079
Horto-industrials	0.213	0.198	0.212	0.198	0.191	0.110
Oilseeds	0.204	0.207	0.204	0.208	0.076	0.071
Olive trees	0.201	0.204	0.201	0.204	0.085	0.085
Vineyards	0.048	0.018	0.053	0.022	0.017	0.110
Pesticides						
Wheat	0.196	0.184	0.197	0.186	0.116	0.114
Maize	0.190	0.160	0.190	0.161	0.155	0.136
Rice	0.201	0.203	0.201	0.204	0.074	0.070
Other cereals	0.212	0.239	0.211	0.238	0.133	0.104
Horto-industrials	0.161	0.149	0.160	0.149	0.146	0.086
Oilseeds	0.203	0.209	0.203	0.209	0.139	0.130
Olive trees	0.197	0.190	0.198	0.191	0.051	0.051
Vineyards	0.033	0.012	0.039	0.016	0.010	0.051
Other costs						
Wheat	0.197	0.191	0.197	0.191	0.058	0.060
Maize	0.183	0.153	0.182	0.150	0.131	0.164
Rice	0.200	0.207	0.199	0.204	0.172	0.174
Other cereals	0.190	0.169	0.190	0.169	0.142	0.156
Horto-industrial	0.073	0.032	0.080	0.040	0.029	0.080
Oilseeds	0.197	0.191	0.197	0.192	0.119	0.122
Olive trees	0.200	0.200	0.200	0.200	0.060	0.060
Vineyards	0.185	0.159	0.181	0.158	0.139	0.057
Gross margin						
Wheat	0.207	0.211	0.207	0.211	0.460	0.482
Maize	0.283	0.385	0.280	0.384	0.474	0.497
Rice	0.208	0.220	0.208	0.220	0.530	0.530
Other cereals	0.224	0.240	0.223	0.238	0.518	0.548
Horto-industrial	0.439	0.510	0.433	0.504	0.525	0.664
Oilseeds	0.207	0.212	0.207	0.212	0.605	0.616
Olive trees	0.204	0.214	0.204	0.213	0.742	0.742
Vineyards	0.714	0.803	0.701	0.794	0.827	0.721

Source: Models results.

Table 6. Percentage of absolute deviation on cost allocation coefficients

	Model simulations				
	GME-1	GME-2	GME-3	GME-4	GCE
Seeds and plants					
Wheat	29.0	35.4	28.7	34.1	8.4
Maize	27.7	4.8	32.6	0.9	12.5
Rice	190.3	162.2	191.3	166.4	1.9
Other cereals	40.8	12.3	42.2	14.9	5.1
Horto-industrials	91.1	83.5	91.4	82.8	80.5
Oilseeds	209.3	194.5	209.7	194.9	0.1
Olive trees	223.5	213.7	223.6	214.5	2.0
Vineyards	66.6	88.5	57.3	84.0	89.3
Fertilizer					
Wheat	6.3	8.2	6.3	8.4	4.7
Maize	106.8	96.1	106.2	95.4	46.7
Rice	23.9	21.9	24.1	22.5	2.6
Other cereals	170.7	183.5	170.6	183.7	25.0
Horto-industrial	93.9	80.2	92.8	79.6	73.7
Oilseeds	187.5	192.2	187.5	192.7	6.8
Olive trees	136.8	140.2	136.6	139.8	0.3
Vineyards	56.4	83.7	51.3	79.5	84.3
Pesticides					
Wheat	72.2	61.8	72.7	63.5	1.5
Maize	39.6	17.8	39.9	18.5	14.0
Rice	186.8	190.2	187.0	190.8	5.4
Other cereals	103.3	130.2	102.9	128.6	28.0
Horto-industrial	86.9	73.3	86.4	73.0	69.9
Oilseeds	56.1	61.1	56.1	61.0	7.3
Olive trees	287.0	273.3	287.3	274.6	0.7
Vineyards	35.0	75.7	23.1	68.6	80.5
Other costs					
Wheat	227.6	218.9	227.6	218.3	3.0
Maize	11.7	6.4	10.9	8.2	20.3
Rice	14.9	19.1	14.4	17.1	1.0
Other cereals	21.6	8.2	21.7	8.2	9.1
Horto-industrial	9.1	59.5	0.0	50.0	63.6
Oilseeds	61.7	56.9	61.7	57.0	2.5
Olive trees	233.3	233.3	233.4	233.4	0.6
Vineyards	224.3	179.6	216.7	176.6	143.9
Gross margin					
Wheat	57.1	56.2	57.1	56.2	4.5
Maize	42.9	22.5	43.5	22.7	4.7
Rice	60.7	58.4	60.8	58.5	0.0
Other cereals	59.1	56.2	59.2	56.5	5.4
Horto-industrial	33.	23.1	34.8	24.1	20.
Oilseeds	66.4	65.5	66.4	65.6	1.8
Olive trees	72.5	71.1	72.5	71.2	0.0
Vineyards	1.0	11.4	2.7	10.1	14.7

Source: Models results.

For the cost item of seeds and plants, results of the four GME models are reasonable for wheat, maize and other cereals. In these cases the PAD varies between 0.9% and 40%.

Regarding the cost item of fertilizers, results for wheat are close to the observed coefficients, and the estimates for rice are also within satisfactory ranges of PAD values (2% to 24%). The estimates of cost allocation for pesticides and other costs are not so closer to the observed coefficients, particularly in the cases of olive trees and vineyards.

Despite the limitations of the GME estimates regarding gross margin, the results are acceptable. In general, cereals are the output crops having the best results.

Comparing the PAD of the GME models with the PAD of the GCE simulation, the differences are clear. While GME models have in average 20 to 29 estimated parameters with a PAD above 30%, the GCE model has only 9 parameters with a PAD above 30%, and 26 of the 40 estimated parameters have a PAD value below 15%.

In the GCE model, the results for gross margin are very close to the observed data, being 20% and 14% the highest values of PAD for oilseeds and vineyards, respectively. The worst PAD results were obtained from oilseeds and vineyards for cost items of seeds and plants, fertilizers, pesticides and other costs.

In general, the GME model specifications present estimators having good statistical and econometric properties. However, the GCE model is the best alternative tested, allowing estimates that are very close to the real parameters.

Results suggest that entropy approach is a good alternative to the traditional methods and is a very useful tool to deal with joint costs problems and incomplete information.

5. Conclusion

Standard farm-accounting information is typically restricted to joint costs for whole farm input expenditures, without revealing production costs per enterprise's output. Alternative tools based on econometric techniques may offer an attractive alternative for obtaining reliable estimates of joint cost allocation coefficients at a significantly low cost. In this context maximum entropy approaches have been widely used.

In this paper the farm costs allocation to output crops from the information of whole farm joint costs using an entropy approach were estimated. A generalized maximum entropy Tobit model was used under five specifications, including two simulations of error bounds, two simulations of the central value of the uniform prior support and one simulation with generalized cross entropy using previous non sample information known.

The model was applied to a sample of 30 farms from the Alentejo 2004 FADN base. Its implementation had in account the assessment of the statistical and econometric properties of the generalized maximum entropy estimators, namely precision and prediction power and its practical validity.

The model precision was assessed by entropy ratio statistics and its critical qi-square values. The information content was measured calculating the normalized entropy and the empirical mean square error. For evaluating the model prediction power, the "*pseudo* R^2" for each input was calculated. The discussion about the model validity was driven by the analysis of the percentage of absolute deviation between estimated and observed coefficients.

Several interesting conclusions were achieved. Results showed that the generalized maximum entropy estimators are consistent and asymptotically normal.

The proportion of information generated by the entropy models is relevant, being the higher levels achieved when the uniform prior has its central value close to zero, and when error bounds are less restricted. However, the models generating new information are also those where the uncertainty of estimators is larger and the results are less coherent with the observed coefficients.

The good precision of generalized maximum entropy estimators was also verified in relation to its predictive power, which is independent from the choice of the support set.

Despite the general good statistic and econometric performance of generalized maximum entropy estimators, the results can always be improved by introducing some previous information known in a generalized cross entropy formulation. Under this specification, the entropy estimation showed a high degree of precision and its practical validity was guaranteed.

The entropy approach showed to be an important tool to deal with the problem of farm joint costs allocation in a context of incomplete information. Furthermore, the paper shows that this kind of flexible approach, namely under its cross entropy formulation disposing of out sample prior information, is well suited to be applied when the variation on data is great. This is the case of Mediterranean regions where crop acreage, gross output and production technologies have a high variability among farms.

In order to overstep some limitations on results, further research developments should foresee the application of the model to other samples and a more accurate test of new support values.

References

Bewley, R. (1986), *Allocation Model: Specification, Estimation and Applications*. Cambridge: Ballinger Publishing Company.

Csiszár, L. (1991), "Why Least Squares and Maximum Entropy? An Axiomatic Approach to Inference for Linear Inverse Problems". *The Annals of Statistics*, 19, 2032-2066

Fragoso, R., Martins, M.B., and Lucas, M.R. (2008), "Disaggregated soil allocation data using a Minimum Cross Entropy Model", *WSEAS Transactions on Environment and Development*, Issue 9, Vol. 4: 756-766.

Fraser, I. (2000), "An application of maximum entropy estimation the demand for meat in the United Kingdom", *Applied Economics*, 32, 45-59.

Garvey, E., and Britz, W. (2002), "Estimation of Input Allocation from EU Farm Accounting Data using Generalized Maximum Entropy", Working Paper, 02-01, University of Irland and University of Bonn.

Gokhale, D.V., and Kullback, S. (1978), *The Information in Contigency Tables*. New York: Mercel Dekker.

Golan, A., Perloff, M. and Shen, Z. (2001), “Estimating a demand system with the non-negativity constraints: Mexican meat demand”, *Review of Economics and Statistics*, LXXXIII:541-551

Golan, A., Karp, S., Perloff, M. (1997), Estimation and inference with censored and ordered multinomial response data. *Journal of Econometrics*, 73: 23-52.

Golan, A. and Judge, M. (1996), “A maximum entropy approach to empirical likelihood estimation and inference”, Working paper. University of California, Berkeley.

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

Golan, A., Judge, M., and Perloff (1996b), “A Maximum Entropy Approach to Recovering Information From Multinomial Response Data”, *Journal of the American Statistical Association*, Vol. 91, Nº434, Theory and Methods: 841-853.

Hansen, H., and Surry, Y. (2006), “Estimating the cost allocation for German agriculture: an application of the maximum entropy methodology”, Conference paper, 46th Annual Conference of German Association of Agricultural Economists, October 4-6.

Hemmer, T. (1996), “Allocations of sunk capacity costs in linear principal-agent model”, *Accounting Review*, 71: 419-432.

Howitt, R.E. and Reynaud, A. (2003), “Spatial Disaggregation of Agricultural Production Data by Maximum Entropy”, *European Review of Agricultural Economics*, 30(3):359-387.

Huang, Q., Rozelle, S. and Howitt, R. (2007), “Determining the Optimal Disaggregated Level or Policy Analysis”, Paper presented at the Annual Conference of Australian Agricultural and Resource Economics Society, February 13-17, Queenstown.

Jaynes, E.T. (1984), “Prior Information and Ambiguity in Inverse Problems”, In: McLaughlin (ed.), *Inverse Problems*, Providence RI: American Mathematical Society: 151-166.

Jaynes, E.T. (1957a), “Information theory and statistic mechanics”, *Physics Review*, 106: 620-630.

Jaynes, E.T. (1957b),” Information theory and statistic mechanics”, *Physics Review*, 108: 171-190.

Just R., Zilberman, D., and Hochman, E. (1983), “Estimation of Multicrop Production Functions”, *American Journal of Agricultural Economics*, 65 (November): 770-780.

Kullback, J. (1959), *Information Theory and Statistics*. New York: John Wiley.

Lence, H.L, and Miller, D. (1998a), “Estimation of Multi-Output Production Functions with Incomplete Data: A Generalized Cross Entropy Approach”, *European Review of Agricultural Economics*, 25(December): 188-209.

Lence, H.L, and Miller, D. (1998b), “Recovering Output-Specific Inputs from Aggregated Input Data: A Generalized Cross Entropy Approach”, *American Journal of Agricultural Economics*, 80(November): 852-867.

Leon, Y., Peeters, L., Quinqu, M. and Suury, Y. (1999), “The use of maximum entropy to estimate input-output coefficients from regional accounting data”, *Journal of Agricultural Economics*, 50: 425-439.

Levine, R. D. (1980), “An Information Theoretical Approach to Inversion Problems”, *Journal of Physics*, Ser A, 13:91-108.

Lips, M. (2009), “Full Product Costs on Base of Farm Accountancy Data by Means of Maximum Entropy”, Contributed Paper prepared for presentation at the International Association of Agricultural Economists Conference, Beijing, China, August 16-22.

Love, H. A. (1999), “Conflicts between Theory and Practice in Production Economics”, *American Journal of Agricultural Economics,* 81(August): 696-702.

Martins, M.B., Fragoso, R., and Xavier, A. (2011), “Spacial Disaggregation of Agricultural Data: A Maximum Entropy Approach”, *JP Journal of Biostatistics*, Vol 5, issue 1: 1-16.

Moxey, A, and Tiffin, R. (1994), “Estimating linear production coefficients from farm business survey data: A note”, *Journal of Agricultural Economics*, 45: 381-385.

Paris, Q., and Caputo, M. (2001), “Sensitivity of the GME Estimates to Support Bounds”, Department of Agricultural and Resource Economics, University if California, Davis, Working paper.

Peeters, L., and Surry, Y. (2005), “Estimation d’un modèle à parameters variables par la méthode d’entropie croisée généralisée et application à la répartition des couts de production en agriculture”, In : *Actes des Journées de Méthodologie Statistique 2005*.

Peeters, L. and Surry, Y. (2002), “Farm cost allocation based on the Maximum Entropy Methodology”. In Lorimer, B*. Agriculture and Agri-Food Canada Strategic Policy Branch – Research and Analysis Directorate*, Publication 2121/E.

Preckel, P.V. (2001), “Least Squares and Entropy: Penalty Function Perspective”, *American Journal of Agricultural Economics*, 83: 366-377.

Pukelsheim, F. (1994), “The Three Sigms Rule”, *American Statistician*, 48: 88-91.

Shannon, C.E. (1948), “A Mathematical Theory of Communication”, *Bell System Technical Journal*, 27: 379-423.

Shumway, C.R., Rope, R.D., and Nash, E.K. (1984), "Allocable Fixed Inputs and Jointness in Agricultural Production: Implication for Economic Modeling", *American Journal of Agricultural Economics*, 66: 72-78.

Trenchard, P.M., and Dixon, R. (2003), "The clinical allocation of joint blood product costs", *Management Accounting Research*, 14: 165-176.

Zhang, X., and Fan, S. (2001), "Crop-specific Production Technologies in Chinese Agriculture". *American Journal of Agricultural Economics*, 83(May): 378-388.

♦♦♦