# Revenue and Cost Functions in PMP: a Methodological Integration for a Territorial Analysis of CAP

**F. Arfini*, M. Donati*, L. Grossi**, Q. Paris*****

*** University of Parma, ** University of Verona, *** University of California, Davis**
**filippo.arfini@unipr.it**

**Paper prepared for presentation at the 107th EAAE Seminar *"Modelling of Agricultural and Rural Development Policies"*. Sevilla, Spain, January 29th -February 1st, 2008**

## Abstract

*An integrated policy evaluation tool is proposed for assessing the effects of agricultural policy measures using all the information available at farm level. The tool combines the positive mathematical programming methodology with the cluster analysis technique by using the same panel of data. The PMP model proposed here allows to measure the effects of policy in term of agricultural supply responses including output market price variations. The novel procedure by which the PMP model is articulated permits to recover the set of farm level demand functions for agricultural products and the cost function characterizing the given sample of farms. Cluster analysis is useful for better appreciating the behaviour of farms before and after the policy scenario analysis by considering the transfers of farms among clusters. A decoupling scenario assessment presents the responses that the integrated tool can provide for evaluating agricultural policy instruments.*

**Key words:** Positive mathematical programming, Cluster analysis, Integrated tool, Agricultural policies, Policy evaluation

## 1. Introduction

Positive Mathematical Programming (PMP) is today widely used for evaluating the effects of the CAP instruments on the dynamics of the agricultural processes and farm economic variables, both for ex-post and ex-ante analysis. The main contribution of this methodology to the agricultural economics is due to its capacity to maximize the information contents in the agricultural datasets available at European level, as FADN, REGIO, IACS (Arfini et al., 2003; Paris and Howitt, 1998). Thanks to the farm decision variable recovering, by the way of the total variable cost estimation, PMP is capable to reproduce the exact observed farm allocation plan and the decision variables (total specific variable costs) that led farmers to decide for such a production plan.

Many papers have adopted the PMP methodology for developing models able to assess the impact of proposed or already implemented CAP reforms. Also in European research projects, this approach is used with micro-based information, like FADN[1]. In most cases, the PMP is proposed in the so-called "classical" form, where the procedure is articulated in three phases: the differential costs recovering, the estimation of the non-linear cost function and, finally, the calibration by using a non constrained production model with non-linear objective function (Howitt, 1995). Applications of this base version are the most diffused, e.g. for evaluating CAP's reform impacts (Arfini et al., 2005; Judez et al., 2002).

One attempt to introduce innovations in the basic approach is due to Heckelei and Wolff (2003) that proposed a methodology that overcomes the first phase for calibrating the base situation by directly imposing the first order conditions in the cost function estimation phase. This approach was also used with cross-section data in order to enhance the consistency of the cost estimation (Heckelei and Britz, 2000). More advanced extensions of the PMP are due to Paris (2001) that generalizes the method adopting an equilibrium model in a static framework and in a dynamic price expectation approach.

[1] Several European research projects have developed and applied models based on the Positive Mathematical Programming methodology, as CAPRI (Heckelei, 1997; Heckhelei and Britz, 2000) and EUROTOOLS (Paris and Arfini, 2000) in the VFP, GENEDEC (contract no. SSPE-CT-2004-502184 and CARERA (contract no. SSPE-CT-2005-022653) in the VIFP.

The demand for an assessment of agricultural policy measures rose with force during this last decade and contributed to the development of a set of economic tools that would respond to such needs using all the available information. In this field, the PMP plays a first order role. This methodology can provide useful results to policy makers even in the presence of a limited set of information as it generally happens when European agricultural databases are adopted. PMP can responds with flexibility and in a consistent way to a large spectrum of policy issues, typically concerning the land use change, production dynamics, variation in gross margin and in the other main economic variables (costs, subsidies, gross saleable production, etc.). However, all these applications are developed exploring the supply side of the agricultural sector while avoiding to implement an evaluation of the demand side, by measuring the effects on the output market prices. Indeed, the literature about the PMP models application seems to indicate that such class of models are just developed for investigating the supply side of the agricultural sector, delegating the demand issues side to well-posed problems solved by econometric techniques.

For improving the analysis, some studies integrate PMP models by other approaches, as cluster analysis (Buysse et al., 2007; Arfini et al., 2005) and convergence evaluation (Arfini et al., 2005). This allows researchers to reach a more readable, comparable and synthetic results by assessments based on very detailed information[2].

The objective of this paper is to present a new quantitative tool for assessing the CAP instruments effects on the agricultural supply dynamics and on the market price modifications, using a PMP approach based on individual farm information. This tool is projected for responding to specific demand of policy makers on the issues related to the impact of CAP measures with respect to land allocation, production levels, price variations and farm revenue modifications. The PMP model that represents the core of such a tool is integrated by a cluster analysis that is incorporated inside the mathematical structure of the same tool. The cluster analysis is useful in evaluating the degree of homogeneity of agriculture sector before and after the application of policy scenarios, by using the same set of farm data adopted for PMP evaluation.

This work is articulated as follows: the first section presents the organization of the quantitative tool, explaining how the PMP methodology is integrated with the cluster analysis; the second section focuses on the estimation of the novel PMP approach proposed in this paper, where the calibration of the model is obtained considering also the information about the farm level demand functions for agricultural products that characterizes the given group of farms; the third section is dedicated to describe the cluster analysis technique integrated to the PMP model; the fourth section concerns an application of the model on a group of farms collected from the IACS database and integrated by FADN information; and the last section concludes with some remarks.

## 2. The integrated approach

The core of the tool is represented by a PMP model able to capture the farm decision variables in order to simulate the impact of policy instruments as realistically as possible. The PMP model described in

[2] The cluster analysis is used when massive information deriving from farm model solutions have to be systematized in order to form group of farms similar in relation with the variables assumed as relevant in measuring the degree of homogeneity in and among the groups. Two examples of application of such a techniques are included in Paris et al. (2000) and Buysse et al. (2007).

the next section keep into account the farm level demand functions that characterize the agricultural outputs produced by the group of farms under evaluation. The PMP model is suitable for using individual farm information and to have a solution at the farm level. To achieve this goal it requires a database collecting farm variables at the individual level, as FADN or IACS databases that represent a sector or a region. Before using the farm information for the policy analysis, it is important to know the characteristics of the sample. A statistical technique that is useful for analysing a cross-section panel of data, with respect the degree of homogeneity of the elements (farms) composing the sample, is the multivariate analysis and, more specifically, the cluster analysis (CA). A tool performed by GAMS integrates such a technique with the PMP.

This kind of approach was used for analysing the impact of the CAP reform on the state of cohesion in EU (Arfini et al., 2005) and, more recently, for evaluating the sugar CMO reform (Buysse, 2007). The combined use of the cluster analysis and the PMP approach allow the analyst to portray the situation before the modification or the introduction of a policy measures and to infer the likely changes inside the groups of farms identified by CA after the application of such measures. When the evaluation considers a large number of farm observations, while the PMP can foresee the variation in the main agricultural variables, the CA allows to understand the behaviour of the farms with respect to their response to agricultural policy. Indeed, the groups identified by this statistical approach represent groups of farms characterized by a similarity with respect to the variables under evaluation (land allocation, gross margin, gross saleable production, variable costs, etc.). The classification obtained by the CA evaluates the degree of homogeneity among the farms in a dynamic perspective, before and after the policy scenarios.

As we have explained above, the dataset feeds the CA in the first step, before the agricultural policy scenario evaluation, and the PMP model. The results generated by the PMP model are the new data for a second CA run. This second run carries out a new configuration of the group of farms, highlighting how the groups have react to the policy instruments. The integrated tool is, thus, composed by different modules that are developed, in a unique modelling environment, using the specific algebraic software GAMS (Brooke et al., 2005).

## 3. Revenue and cost functions in PMP model

The PMP approach presented by Howitt and Paris (1995, 1998) is originally articulated in three sequences: 1) the recovering of the marginal costs associated with the agricultural processes present in a farm allocation plan by solving a linear programming model; 2) the estimation of a non-linear cost function able to capture the information about the substitution and complementarity among farm processes obtained by a consistent method of estimation (e.g. ME and OLS); 3) calibration of the base allocation farm plan using the cost function derived in the previous phase. The idea behind this method is to consider that farmers take their decisions not just considering the explicit variable accounting costs of the inputs used inside the production process, but also the part of variable costs that are connected with the farmers' knowledge about their own farm system. The PMP approach estimates this adding farm costs, the so-called differential costs (Paris and Howitt, 1998), for using them inside the calibration phase. Thus, the objective function that is maximized at the last PMP stage is the farm gross margin that takes into account those combined costs. In this perspective, the maximized gross margin can be considered the "economic" gross margin, instead of the accounting definition of this term.

The methodology approach proposed in this paper considers the problem of estimating the farm level demand functions associated with a group of farms selected for a policy scenarios evaluation inside a PMP framework. More specifically, the approach is articulated in four phases: 1) cross-section estimation of farm level demand functions using individual data; 2) recovering of the differential marginal costs that lead farmers to choose the observed production plan, considering inside the objective function a non-linear revenue function; 3) estimation of a quadratic cost function; 4) calibration of the base observed situation (the observed production plan) maximizing an objective function composed of the non-linear revenue function estimated in the first phase and the non-linear cost function derived in the third phase.

*Phase I – Estimation of farm level demand functions*

The farm level demand functions that we want to estimate have the following linear form:

(1) $$\mathbf{p} = \mathbf{d} - \boldsymbol{D}\mathbf{x}$$

or, in a sample formulation $$p_{n,j} = d_j + \sum_{j'=1}^{J'} D_{j,j'} x_{n,j} + u_{n,j}$$

where $\mathbf{p}$, $\mathbf{d}$ and $\mathbf{x}$ are vectors with dimensions (*Jx1*) and $\boldsymbol{D}$ a matrix with dimension (*JxJ*); $\mathbf{p}$, $\mathbf{d}$ and $\mathbf{x}$ are the vectors of agricultural product prices, the vector of intercepts of demand function and the vector of production quantities, respectively; $\boldsymbol{D}$ is a symmetric positive semidefinite matrix of quantity slopes. *J (j=1,…,J)* is the number of agricultural processes.

Economic theory assumes that market prices paid to producers vary in relation with the aggregated demand function. Under this assumption, a set of demand functions can be estimated on the basis of a sample of *N* farms. The term $u_{n,j}$ in (1) represents the deviation of the *n*-th farm from the regional *j*-th demand function. If the sample of farms concerns a given geographical region or a sector, it is possible to estimate a set of demand functions for the agricultural products of such a region or a sector. The objective is thus to obtain the set of demand function (1) using the information of a sample of individual farms.

The relevant information required for estimating (1), consists of prices paid for selling the farm products at the farm level and about the of output quantities introduced into market. Both types of information are generally available inside the most used agricultural database, as FADN. The methods of estimation that one can - implement- varies from generalized least squares, to maximum likelihood, to maximum entropy, etc. In this work, we choose the maximum entropy approach to estimate a well-posed problem. Furthermore, the choice of ME[3] is related to our empirical experience demonstrating

[3] After the publishing of the famous book of Golan, Judge and Miller (1996), the maximum entropy approach has known a new interest among agricultural economists. The idea is to use a physical concept applied to communication technology by Shannon (1948) and in economics by Jaynes (1957) in order to derive parameters when the information is poor and where the traditional econometric techniques prefer not to intervene. For a complete revision of maximum entropy theory, see Fang et al. (1998). For a detailed discussion about the maximum entropy estimator applied to economics see the book of Golan, Judge and Miller (1996), the paper of Paris and Howitt (1998), Heckelei and Britz (2000), Lansink (1998), Léon et alt. (1999), Lence and Miller (1998).

that a maximum entropy estimator seems to obtain parameters that provide very realistic results in a simulation phase[4].

The estimation carried out in the present section consists in recovering the demand functions (1) governing the output markets of a sample of 50 farms. The first group of parameters to estimate is that owning to the intercept $\mathbf{d}$, while the second group is related with the matrix $\boldsymbol{D}$. According to the generalized maximum entropy theory of Golan, Judge and Miller, each parameters to recover is equal to the product between a set of probabilities and a set support values. The objective of the problem is to identify the probability distribution that maximize the maximum entropy function. The support values are chosen by the researcher[5].

Thus, the intercept can be written as:

(2)
$$\mathbf{d}_j = \sum_{p=1}^{P} \mathbf{zd}_{j,p} \mathbf{pd}_{j,p}$$

where, $\mathbf{zd}_{j,p}$ is the vector of support values, while $\mathbf{pd}_{j,p}$ is the vector of the *p (p=1,...,P)* probabilities.

We assume that the matrix $\boldsymbol{D}$ is symmetric, positive semidefinite. The simplest and most efficient way to respect those properties is to decompose the matrix $\boldsymbol{D}$ in three components according to the Cholesky factorization method (Paris and Howitt, 1998). On the basis of this method the matrix $\boldsymbol{D}$ is divided in three matrices as follows:

(3)
$$\boldsymbol{D} = \boldsymbol{LHL'}$$

where, $\boldsymbol{D}$ is equal to the product among a unit lower triangular matrix $\boldsymbol{L}$, a non-negative diagonal matrix $\boldsymbol{H}$ and the transposed of $\boldsymbol{L}$. The decomposition guarantee in every cases to obtain a symmetric, positive and semidefinite matrix. This same decomposition can be rewritten in a more compact form, so that:

(4)
$$\boldsymbol{D} = \boldsymbol{LHL'} = \boldsymbol{RR'}$$

where the matrix $\boldsymbol{R} = \boldsymbol{LH}^{1/2}$.

In order to estimate the parameters of $\boldsymbol{L}$ and $\boldsymbol{H}$, it is required to specify a suitable set of support values to associate to an unknown probability distribution, as presented in the following equations:

(5)
$$L_{j,j'} = \sum_{p=1}^{P} Zl_{j,j',p} Pl_{j,j',p} \qquad \forall \quad j \neq j'$$

(6)
$$H_{j,j'} = \sum_{p=1}^{P} Zh_{j,j',p} Ph_{j,j',p} \qquad \forall \quad j = j'$$

Equation (5) states the relation about the unitary triangular matrix $L_{j,j'}$ and the product between the matrix of support values $Zl_{j,j',p}$ and the matrix of probability distribution $Pl_{j,j',p}$. The matrix $\boldsymbol{L}$ is a triangular matrix with unitary values on the diagonal and null values above the diagonal. In equation (6), the matrix $H_{j,j'}$ is equal to the product of the support values $Zh_{j,j',p}$ and the unknown matrix of probability distribution $Ph_{j,j',p}$. $\boldsymbol{H}$ is a non-negative diagonal matrix with null values outside the diagonal.

[4] The results achieved applying the ME estimator confirm the important role of this estimator in other fields of the applied sciences (Paris and Howitt, 1998; Shannon, 1948).

[5] One of the main criticism addressed to the maximum entropy methods concerns the choice of support values that are submitted to the subjective decision of the researcher (Lansink, 1997).

Keeping into account the statements above, the maximum entropy problem that recover the demand function (1) starting from a cross-section panel of individual farms is the following one:

(7)
$$\begin{aligned}\max_{p(g)} \quad Hd(p) = \; & -\sum_{j=1}^{J}\sum_{p=1}^{P} pd_{j,p}\log pd_{j,p} - \sum_{j=1}^{J}\sum_{j'=1}^{J'}\sum_{p=1}^{P} Pl_{j,j',p}\log Pl_{j,j',p} \\ & -\sum_{j=1}^{J}\sum_{j'=1}^{J'}\sum_{p=1}^{P} Ph_{j,j',p}\log Ph_{j,j',p} - \sum_{n=1}^{N}\sum_{j=1}^{J}\sum_{p=1}^{P} pe_{n,j,p}\log pe_{n,j,p}\end{aligned}$$

Subject to:

(8)
$$pr_{n,j} = \sum_{p=1}^{P} ze_{n,j,p}pe_{n,j,p} + \sum_{p=1}^{P} zd_{j,p}pd_{j,p} - \sum_{k=1}^{K}\sum_{j'=1}^{J'} R_{j,j'}R_{k,j'}\bar{x}_{n,k} \quad ,\forall n \forall j$$

(9)
$$R_{j,j'} = \left(\sum_{k=1}^{K}\sum_{p=1}^{P} Zl_{j,k,pp}Pl_{j,k,pp}\right)\left(\sum_{k=1}^{K}\sum_{p=1}^{P} Zd_{j,k,pp}Pd_{j,k,pp}\right)^{1/2} \quad ,\forall j \forall j'$$

(10)
$$0 = \sum_{n=1}^{N}\sum_{p=1}^{P} ze_{n,j,p}pe_{n,j,p} \quad ,\forall j$$

(11)
$$\begin{cases} 1 = \sum_{p=1}^{P} pd_{j,p} & ,\forall j \\ 1 = \sum_{p=1}^{P} Pl_{j,j',p} & ,\forall j \neq j' \\ 1 = \sum_{p=1}^{P} Pd_{j,j',p} & ,\forall j = j' \\ 1 = \sum_{p=1}^{P} pe_{n,j,p} & ,\forall n \forall j \end{cases}$$

The entropic objective function of the problem (7)-(11) is maximized with respect to the unknown probability distributions associated with the support values identified by the researcher. Equation (8) states that the observed prices $pr_{n,j}$ are equal to unique demand function plus a farm deviation, $ze_{n,j,p}pe_{n,j,p}$, that measures the distances between *n*-th observed farm price and the common/regional demand function. Equation (9) performs the Cholesky's decomposition rule established inside the relation (4). The constraint (10) concerns the summation to zero of the farm deviations and the set of constraints (11) state the adding-up relations for the probability distributions. This problem estimates the demand functions of the agricultural market generating the output prices of each farm.

*Phase II – Recovering of differential marginal costs*

The second phase of PMP is devoted to estimating the  the marginal costs borne by farmers in their input allocation process. When information about accounting variable costs is available, the estimation deals with the differential amount leading to a true economic marginal cost.

The novelty of the proposed PMP approach consists in defining an objective function that depends on the set of farm level demand functions estimated in phase I.

This revenue functions is derived integrating the demand function with respect the output levels, so:

(12) $$\int_0^{\bar{x}} (\mathbf{d} - \boldsymbol{D}\mathbf{x})\,dx = \mathbf{d}\bar{\mathbf{x}} - \frac{1}{2}\bar{\mathbf{x}}\boldsymbol{D}\bar{\mathbf{x}}'$$

The maximization problem of this phase II is usually improperly call as PMP calibration phase. In reality, this stage needs for calibrating the base situation through the differential marginal costs hidden inside the observed production quantities. The objective of this phase is to maximize a non-linear gross margin function subject to typical farm structural constraints (i.e. land) and to calibrating constraints that force the model to reproduce the observed production plan. In algebraic terms, the problem for the *n*-th farm is written as follows:

(13) $$\max_x \quad GM_0(x) = \sum_{j=1}^{J}\sum_{j'=1}^{J'} \hat{v}_{n,j}x_{n,j} + \hat{d}_j x_{n,j} - \frac{1}{2}x_{n,j}\hat{D}_{j,j'}x_{n,j'} - c_{n,j}x_{n,j}$$

subject to:

(14) $$\sum_{j=1}^{J} A_{n,j,i}x_{n,j} \le b_{n,i} \quad ,\forall i \quad [y_{n,i}]$$

(15) $$x_{n,j} \le \bar{x}_{n,j} + \varepsilon \quad ,\forall j \quad [\lambda_{n,j}]$$

(16) $$x_{n,j} \ge 0 \quad ,\forall j \quad [\mu_{n,j}]$$

where $\hat{v}_j$ is the deviation of each farm process from the demand function estimated on the sample of farms. The vectors of deviations is obtained by the previous phase as:

(17) $$\mathbf{v}_{n,j} = \mathbf{ze}_{n,j}\overline{\mathbf{pe}}_{n,j}$$

$c_{nj}$ is the explicit accounting variable cost associated with each output unit at *n*-th farm level; while $A_{n,j,i}$ and $b_{n,i}$ are respectively the matrix of technology, that is the matrix with the coefficients of input use for obtaining one unit of product, and the vector of input farm capacity $i$ (i.e. land acreage), for i=1,…,I. The coefficients $\hat{d}_j$ and $\hat{D}_{j,j'}$ are the estimates of the corresponding parameters obtained in phase I.

Problem (13)-(16) is optimized when the difference between total revenue and total variable cost is maximized with respect the level of output *x*. The solution of this problem is known before solving it, because the calibrating constraint (15) imposes that each variable *x* cannot exceed the observed level of those outputs $\bar{x}$ plus a terms very small $\varepsilon$[6]. The tautological problem (13)-(16) leads to obtain the dual information linked to the calibrating constraint (15), that is $\lambda_j$. $\lambda_j$ is the differential costs to add to the accounting marginal costs $c_j$ in order to obtain a total marginal cost needed for estimating the non-linear cost function of the third phase.

[6] The meaning of ε is to avoid the linear dependency between the structural constraint and calibrating constraint. For a deeper explanation about the role of e see Howitt (1995), Paris and Howitt (1998) and Gohin and Chantreuil (2000).

*Phase III – Non-linear cost function estimation*

The objective of the third phase is to estimate the farm cost function starting from the vector of marginal costs estimated in phase II using the shadow prices associated with the calibration constraints. The chosen functional form of the cost function is:

$$C(x) = (\boldsymbol{\lambda} + \mathbf{c})\overline{\mathbf{x}} = \boldsymbol{\alpha}\overline{\mathbf{x}} + \frac{1}{2}\overline{\mathbf{x}}'\boldsymbol{Q}\overline{\mathbf{x}} \tag{18}$$

where $\lambda$ and $c$ are, respectively, the vector of the dual values identified in the previous phase and the vector of the farm accounting costs, $\overline{x}$ is the vector of the known production levels and $\boldsymbol{Q}$ the matrix of the non-linear cost function. $\boldsymbol{\alpha}$ is the vector of intercepts for the marginal cost associated to farms processes. In (18) the elements for matrix $\boldsymbol{Q}$ are still unknown and must be obtained through suitable estimation methods. In the literature (see Paris et al., 2000) estimation of cost function through application of the principle maximum entropy is preferred. On the basis of these concepts and the arrangement given by Paris and Howitt (1998), the parameters of vector $\alpha$ and matrix $\boldsymbol{Q}$ can be recovered by maximizing the probability distribution associated with an interval of specified support values. The non linear program of maximum entropy is presented here in the form derived by Cholesky's decomposition according to which the matrix $\boldsymbol{Q} = \boldsymbol{\Gamma W \Gamma}' = \boldsymbol{TT}'$, where $\boldsymbol{\Gamma}$ is a triangular matrix, $\boldsymbol{W}$ a diagonal matrix and $\boldsymbol{T} = \boldsymbol{\Gamma W}^{1/2}$. The problem can then be solved by maximizing a probability distribution for which we know the expected value, which corresponds to the marginal cost $(\lambda + \mathbf{c})$ determined in the second phase. The objective function of the problem of maximum entropy is thus presented as follows:

$$\begin{aligned} \max_{p(g)} \quad Hc(p) = \; & -\sum_{j=1}^{J}\sum_{p=1}^{P} p\alpha_{j,p} \log p\alpha_{j,p} - \sum_{j=1}^{J}\sum_{j'=1}^{J'}\sum_{p=1}^{P} P\varphi_{j,j',p} \log P\varphi_{j,j',p} \\ & -\sum_{j=1}^{J}\sum_{j'=1}^{J'}\sum_{p=1}^{P} Pw_{j,j',p} \log Pw_{j,j',p} - \sum_{n=1}^{N}\sum_{j=1}^{J}\sum_{p=1}^{P} pu_{n,j,p} \log pu_{n,j,p} \end{aligned} \tag{19}$$

where $p\alpha_{j,p}$ are the unknown probability distributions of the intercepts of the cost function, $p\varphi_{j,j',p}$ and $pw_{j,j',p}$ are the probability of the distribution associated with elements of the triangular matrix $\boldsymbol{\Gamma}$ and of the diagonal matrix $\boldsymbol{W}$ respectively. $pu_{n,j,p}$ are elements of the probability of errors. The objective function (19) is maximized considering the information about the process marginal costs at farm level, as follows:

For $x > 0$ at farm level:

$$\lambda_{n,j} + c_{n,j} = \sum_{p=1}^{P} p\alpha_{j,p} z\alpha_{j,p} + \sum_{j'=1}^{J'} \left\{ \sum_{k=1}^{K} \left( T_{j,k} T_{k,j'} \right) \right\} \overline{x}_k + \sum_{p=1}^{P} pu_{n,j,p}\, zu_{n,j,w} \quad ,\forall n \forall j \tag{20}$$

For *x* <u>not</u> activated at farm level:

$$\lambda_{n,j} + c_{n,j} \le \sum_{p=1}^{P} p\alpha_{j,p} z\alpha_{j,p} + \sum_{j'=1}^{J'} \left\{ \sum_{k=1}^{K} \left( T_{j,k} T_{k,j'} \right) \right\} \overline{x}_k + \sum_{p=1}^{P} pu_{n,j,p}\, zu_{n,j,w} \quad ,\forall n \forall j \tag{21}$$

The equations (20-21) state that the total marginal cost $(\boldsymbol{\lambda}_{(\cdot)} + \mathbf{c}_{(\cdot)})$ is equal/less or equal to a new marginal cost function common for all the farms sample plus a farm error. $\boldsymbol{T}_{(\cdot)}$ is an element of the matrix $\boldsymbol{T}$ obtained through Cholesky's decomposition. In fact:

(22)
$$T_{j,j'} = \sum_{j'=1}^{J} \left\{ \sum_{p=1}^{P} \left( p\varphi_{j,j',w} z\varphi_{j,j',w} \right) \sum_{p=1}^{P} \left( pw_{j,j',p} zw_{j,j',p} \right)^{1/2} \right\}$$

The relations inserted in (22) clarify the role of the support values in the process of estimating the cost matrix. The components $z\varphi_{(g)}$ and $zw_{(g)}$ are the appropriately selected support values (Paris and Howitt, 1998). Associated with the distribution of probability, $p\varphi_{(g)}$ and $zw_{(g)}$, they define the elements of the triangular matrix $\boldsymbol{\Gamma}$ and of the diagonal matrix $\boldsymbol{W}$. It must be pointed out that the matrix $\boldsymbol{Q}$ is unique and is derived from the marginal costs.
All the probability distributions referred to above must meet the following condition:

(23)
$$\begin{cases} \sum_{p=1}^{P} p\alpha_{(\cdot)} = 1 \\ \sum_{p=1}^{P} p\varphi_{(\cdot)} = 1 \\ \sum_{p=1}^{P} pw_{(\cdot)} = 1 \\ \sum_{p=1}^{P} pu_{(\cdot)} = 1 \end{cases}$$

Problem (19)-(23) provides the probability distribution values for the elements of the triangular matrix $\boldsymbol{\Gamma}$, the diagonal matrix $\boldsymbol{W}$ and for the vector of the residual marginal variable costs for each farm in the sample. The cost function specified according to the above method preserves the technical information regarding the calibration constraints.

*Phase IV – Calibrating observed situation*

Finally, after having estimated the revenue and cost functions, we can develop a problem very similar to those in the second phase of the procedure, where a new cost function is inserted and the calibrating constraints are not considered. The problem can be build as follows:

(24)
$$\begin{aligned} \max_x \quad GM_1(x) = & \sum_{j=1}^{J} \sum_{j'=1}^{J'} \left\{ \hat{v}_{n,j} x_{n,j} + \hat{d}_j x_{n,j} - \frac{1}{2} x_{n,j} \hat{D}_{j,j'} x_{n,j'} \right\} \\ & - \sum_{j=1}^{J} \sum_{j'=1}^{J'} \left\{ \hat{u}_{n,j} x_{n,j} + \hat{\alpha}_j x_{n,j} + \frac{1}{2} x_{n,j} \hat{Q}_{j,j'} x_{n,j'} \right\} \end{aligned}$$

subject to:

(25)
$$\sum_{j=1}^{J} A_{n,j,i} x_{n,j} \le b_{n,i} \quad ,\forall i \quad \left[ y_{n,i} \right]$$

(26)
$$x_{n,j} \ge 0 \quad ,\forall j \quad \left[ \mu_{n,j} \right]$$

The error terms $\hat{v}_j$ and $\hat{u}_j$ are derived from the first and third phase of the procedure respectively, and they are specific to each farm. In other terms, they measure the distance between the prices and the costs observed at $n$-th farm level and the prices and costs estimated for the region considered by the analyst.

Inside the objective function (24) the new quadratic cost function takes the place of the calibrating constraints, establishing the economic bound for the activity allocation choice. In other terms, the latent decision variables revealed in the second phase enter inside the objective function (24) providing an economic calibrating constraint instead of a technical constraint such as the equation (15). The gross margin maximized in (24) is less than the gross margin specified in (13), $GM_0 < GM_1$, because the $GM_1$ also integrates the dual values associated to the farm activities. For this reason, we can say that the objective function (24) should be considered an economic profit in the sense of the economic theory.
The problem (24)-(26) permits to exactly reproduce the base situation without specific calibrating constraints. Furthermore, applying policy scenario simulations, the non-linear revenue function provide information on the likely variation in agricultural product prices in relation with changes in production levels.

## 4. Non-hierarchical cluster analysis: the k-means procedure

The aim of partitioning methods is to get a single partition of *n* points in *p* dimensions into *k* clusters ($k < n$), following a optimizing criteria and where *k* is chosen by the researcher. The *k*-means algorithm is, even if according to various versions, the best-known and applied partitioning method (for a review, see Atkinson *et al.*, 2004). This procedure leads to classify the *n* units into *k* distinct clusters, with *k* chosen *a priori* by the analyst, according to an iterative method whose steps can be resumed as follows:
1) *k* initial cluster centres are selected, that is *k p*-dimensional points which are the cluster centroids in the initial partition. Centres can be detected using different methods, but usually are such that they are as much as possible distant each other. The initial *k*-clusters partition is then built, adding each element to the closest cluster.
2) For each unit the distance to the *k* cluster's centroids is computed: if the minimum distance is different from that gained by the centroid of the group to which the unit belongs, the unit is moved and included to the closest group. When one units is reallocated the new and old cluster's centroids are re-estimated.
3) Step 2 is iterated until to the convergence of the algorithm, that is until clusters and centroids remain stable and unchanged with respect to the previous iteration.
Alternatively, when the computational burden of the procedure is an issue, the stopping rule introduced in step 3 can be replaced by less restrictive rules which ends the procedure when one of the following events happen:
a) the algorithm comes to convergence in the previously stated sense;
b) the distance of each centroid at current iterations to the corresponding centroid at previous iteration is not greater than a given threshold;
c) maximum number of iterations has reached.
To get a partition with a different number of clusters, for instance $k^*$, all the steps must be repeated, starting from the first stage where *k* is replaced by $k^*$.

In order to apply the above suggested procedure, the distance between each unit and the centroids must be iteratively computed and a suitable metric must be chosen. The most frequently applied metrics is the Euclidean distance, because it usually ensures the convergence of the iterative procedure is achieved (Rencher, 1997). Thus, at iteration *t*, the distance between unit *i* and the centroid of group *l* (*i*=1,2, …, *n*; *l*=1,2, …, *k*) is given by:

(27)
$$d(x_i, \bar{x}_l^{(t)}) = \sqrt{\sum_{s=1}^{p} \left(x_{is} - \bar{x}_{s,l}^{(t)}\right)^2}$$

where $\bar{x}_l^{(t)} = \left[\bar{x}_{1l}^{(t)}, \bar{x}_{2l}^{(t)}, \ldots, \bar{x}_{p,l}^{(t)}\right]'$ is the centroid of group *l* computed at iteration *t*.

From a practical point of view, it is not possible to enumerate the whole set of possible partitions of *n* elements in *k* groups, so that the optimal classification obtained by applying the *k*-means method can actually lead to a local minima of the objective function. As a consequence, the initial choices done by the researcher are crucial and must be examined carefully, particularly with reference to: a) choice of the number of clusters *k*; b) selection of the initial cluster centres.

The choice of the number of clusters *k* is a primary concern by the researcher. From (27) is clear that the main goal of the *k*-means partitioning method, with Euclidean distance, is to find a partition (with *k* clusters) which satisfies a criteria of internal homogeneity based on the minimization of the within deviance. A natural measure of the goodness of fit of the procedure is then given by the following index:

(28)
$$R^2 = 1 - \frac{W}{T} = \frac{B}{T}$$

where *W*, *B* and *T* are, respectively, within-groups, between-groups and total deviance and $T = W + B$.

A good partition usually presents a small within-groups deviance. Index introduced by equation (28) is contained in the interval [0,1] and can be used to compare partitions with a different number of groups. When $R^2$ is close to 1, the corresponding partition turns out to be homogeneous because units belonging to the same cluster are very similar ( $W_l \approx 0$ , for each $l = 1,2,\ldots,k$ ) and clusters are strongly separated ( $B \approx T$ ). Nevertheless, a trade-off can be observed between the number of clusters and the internal homogeneity, because $R^2$ is not-decreasing for increasing values of *k*. A good compromise between good separation of clusters and reduction of complexity of the partition can be achieved selecting the number of clusters *k* which produce a very high gain in internal cluster homogeneity with respect to a partition with *k*-1 clusters.

Once the number of clusters has been chosen, the initial cluster centres must be selected. A very simple criteria is to select the first *k* observations of the dataset, while a lightly more sophisticated method is to pull out a random sample of size *k* from the *n* units of the dataset.

## 5. Policy evaluation

The integrated tool presented in the previous sections is applied to a sample of farms belonging to the Emilia-Romagna region. The sample is composed by 50 farms placed in the provinces of Parma, Reggio-Emilia, Modena and Bologna and it is extracted from the IACS database, that is the dataset concerning the demand for subsidy payments that farmers must submit every year to the national agency charged of the communitarian agricultural subsidies payments[7]. The IACS information, concerning the crop area of each farm, is completed with the information deriving from Italian FADN. More specifically, the information concerning the yields, prices and specific variable costs are obtained by the national FADN[8]. 2003 is the reference year. The sample presents a production set of ten crops: cereal mix, alfa-alfa, sugarbeet, durum wheat, fodder crops, maize, barley, silage, soya and soft wheat.

**Table 1.** Characteristics of the sample

| Main information | |
|---|---|
| Number of farms | 50 |
| Incidence of cerelas (in %) | 64.5 |
| Incidence of oilseeds (in %) | 4.9 |
| Incidence of fodder crops (in %) | 19.4 |
| Incidence of sugarbeet (in %) | 11.2 |
| Revenue by ha (in euros) | 2,001 |
| Variable costs by ha (in euros) | 1,466 |

The aims of the analysis is to get a response on the effects of the single farm payment introduced by the EU regulation 1782/2003 with respect of farms behaviour. More in detail, the integrated tools is applied the policy scenario concerns the total decoupling of the COP crops. The reform of sugarbeet support system is not considered.

5.1. PMP outcomes

Thanks to the reconstruction of the revenue function and the cost function, the PMP model is able to provides dynamics about the supply side and demand side of the considered sample. The first aspect concerns the changes of land allocation operated by the farms in relation with the decoupling. The table 2 presents the variation of each crops after the decoupling implementation. The separation between payments and quantity of agricultural product seems to lead farms to abandon part of the cereal acreage for investing on fodder crops, oilseeds and sugarbeet.

The variation in land use has consequences on the production levels and, thus, on market prices. This PMP approach is capable to capture the price signals in relation to the output variations. This is the second relevant aspect of the model: the simulation can provide variation about market prices of each product. From table 2, it is possible to note the negative variation in the hectares of cereals that leads to an increase in market prices for such products. For example, maize reduces of around 15% its acreage, while its prices improve of 19%. Similarly, the fodder crops see a strong increasing in the number of hectares (+48%), while the prices is foreseen to dramatically decrease (-40%).

[7] The Italian agency charged of EU payments is AGEA (AGenzia per le Erogazioni in Agricoltura).
[8] For further details on the method of merging IACS with FADN database, see Arfini et al. (2005).

**Table 2.** PMP simulation results – Land allocation and Prices

| Activities† | Land use | | Prices | |
|---|---|---|---|---|
| | Baseline (ha) | Scenario (Var. %) | Baseline (euros/ton) | Scenario (var. %) |
| SoftWheat | 503.9 | -16.5 | 145.4 | +8.2 |
| Durum Wheat | 10.1 | -26.2 | 204.5 | +4.5 |
| Maize | 386.3 | -14.8 | 149.6 | +18.9 |
| Barley | 130.1 | -24.5 | 131.9 | +9.6 |
| Cereals Mix | 49.6 | -34.8 | 144.1 | +4.5 |
| Silage | 58.2 | -9.9 | 40.2 | +11.4 |
| Soya | 86.5 | +11.5 | 231.7 | -7.2 |
| Alfa Alfa | 338.4 | +0.5 | 100.9 | -9.9 |
| Other fodder | 3.7 | +48.4 | 12.4 | -39.9 |
| Sugarbeet | 197.7 | +6.5 | 43.1 | -10.0 |

† The model considers also the possibility to activate agricultural area submitted to good practices. The model results indicates that around 10% of the agricultural area would be dedicated to such non-productive activity.

The new production plan due to decoupling has effects on the main farm economics variables. The table 3 presents a situation where the decreasing in revenues and costs leads to improve the farm gross margin (+2%). This is due to a much more intensive reduction of the variable costs (-8.8%) that the farm revenues (-5,9%). The farm strategy within decoupling seems addressed to minimize as much as possible the production costs.

**Table 3.** PMP simulation results – Main economic variables

| Economic variables | Baseline (euros/ha) | Scenario (var. %) |
|---|---|---|
| Revenues (gsp+subs.) | 2,001 | -5.9 |
| Costs | 1,466 | -8.8 |
| Gross Margin | 536 | +2.1 |

The responses of the model in term of quantities, prices and economic variable dynamics depend in large part to the estimated matrices $\hat{\boldsymbol{Q}}$ and $\hat{\boldsymbol{D}}$ (see Appendix), that integrate the information about the degree of substitution and complementarity among activities.

5.2. Cluster analysis

The *k*-means approach for the cluster analysis carried out to identify three groups of farms, internally homogeneous with respect the following variables: the incidence of each type of crop on the total agricultural surface, the yields for each type of crop, the revenue per hectare and the total variable cost per hectare. The first clusterization developed in the base situation, before the policy scenario implementation, is presented by the table 4. The first cluster is composed by 7 farms with the least average surface, if compared with the others clusters, with a low revenue by hectare and low average cost by hectare. The difference between the unitary revenue and the unitary cost is the lowest among the groups. The presence of an important incidence of fodder crops is significant for explaining the low level of farming intensity inside such a group.

The second group composed by 18 farms highlights an average acreage of 35 higher than the previous group and the revenue by hectare and the variable cost per hectare are much higher than every other group. This group present the largest difference between the revenue per hectare and the cost per hectare, demonstrating that the farms belonging to such group are more intensive in producing cereals, that represent the major incidence, and with a relevant quota (15%) of land invested in harvesting sugarbeet.
The last group is very similar to the second one, but with a lower economic margin per hectare. Those farms are specialized in producing cereals (68% of the total acreage). The quota of land dedicated to sugarbeet is not so large as in the second group, but it represents a non marginal investment (10.4%).

**Table 4.** Cluster analysis – Base situation

| Clusters | No. Farms | Average acreage (ha) | Average revenue (euros/ha) | Average Costs (euros/ha) | Cereals (%) | Oilseeds (%) | Fodder crops (%) | Sugarbeet (%) |
|---|---|---|---|---|---|---|---|---|
| 1 | 7 | 21 | 1215 | 959 | 48.4 | 1.9 | 48.3 | 1.4 |
| 2 | 18 | 35 | 2126 | 1538 | 62.6 | 7.0 | 15.6 | 14.8 |
| 3 | 25 | 40 | 2039 | 1495 | 68.1 | 4.0 | 17.5 | 10.4 |

**Table 5.** Cluster analysis – Policy scenario

| Clusters | No. Farms | Average acreage (ha) | Average revenue (euros/ha) | Average Costs (euros/ha) | Cereals (%) | Oilseeds (%) | Fodder crops (%) | Sugarbeet (%) |
|---|---|---|---|---|---|---|---|---|
| 1 | 7 | 21 | 1117 | 848 | 40.7 | 2.2 | 53.7 | 3.4 |
| 2 | 27 | 42 | 1967 | 1356 | 63.8 | 5.2 | 18.3 | 12.7 |
| 3 | 16 | 30 | 1920 | 1444 | 53.7 | 9.2 | 19.4 | 17.7 |

The impact of decoupling has been relevant for every farms. As we can observe in table 5, the decoupling has produced an increase of the economic margin by hectares (revenues/ha – variables costs/ha) in every group, but he most relevant effect concerns the third group. Indeed, 9 farms that in the base situation belonged in the third group move toward the second group when the decoupling is applied. This means that for such farms, the decoupling amplifies the gross margin per hectare in relation to the process of minimization of the costs explained above. Only the farms of the first group don’t move their place: the decoupling improve significantly their economic results but the gap with respect the other clusters is too high for transfers.

## 6. Conclusions

The integrated tool proposed in this paper combines the PMP methodology with the cluster analysis technique. The aim of the first approach is to recover the hidden decision variables of farmers in order to estimate their behaviour in presence of agricultural policy changes. The implemented PMP model introduces a generalization of the traditional methodology. Indeed, the model is able to derive both the demand function that characterizes the agricultural market product of the sample of farms considered and the cost function kept in account by farmers during the production plan definition. The unknown

parameters of the revenue and cost functions are recovered by adopting the maximum entropy approach. The last calibration phase maximizes the difference between the farm revenue and cost functions derived by a procedure articulated on four phases.
The results achieved by using the PMP model in assessing policy scenarios can give responses on the supply side, providing the likely modification of the land use and the production level, and on the demand side, providing information about the dynamics of prices. This is why this method can be considered a generalization of the method firstly proposed by Howitt and Paris (1995, 1998).
The use of the cluster analysis for evaluating the behaviour of farms with respect a new policy scenario is useful in order to better understand the driving forces leading farmers to adopt a given strategy to respond to new policy measures. Indeed, the cluster analysis groups farms according to a homogeneity criteria with respect to variables assumed to be relevant for explaining the main characteristics of the considered farms. The picture given by the cluster analysis in the base situation can change when a simulation is carried out. The cluster analysis is important to portray the movement of farms among clusters.
So, when the analysis is developed using a sample of individual farms, the cluster analysis became a natural policy analysis component to integrate with the PMP model. In our work, the PMP model and the cluster analysis technique have been used in a same policy evaluation environment, using a common algebraic language package (GAMS).
The policy assessment presented in this paper shows the added value that an integrated tool can give to policy makers in order to evaluate the effects of the policy measures using farm database information at the maximum degree of extension. All the information used inside the PMP model and the CA technique concerns individual farms: the policy scenario simulation and the evaluation of the degree of homogeneity among farms are carried out with respect of each individual farm.
This work highlights that the PMP approach is not just a calibration techniques, as frequently is affirmed, but it is an efficient methodology able to reveal relevant information about the farm decision process in order to evaluate the farm behaviour when changes in base observed variables intervene. Moreover, the information generated by the PMP model can be enhanced and increased by using in an integrated way the CA technique. The tool proposed is consistent because uses all the available information, explicit and implicit, included in a dataset and also because it is capable to respond in a very complete and detailed level of analysis to the agricultural policy evaluation needs.

## References

Arfini F., Donati M., Montresor, E., Mazzocchi M., Bettocchi A., Fanfani F., Brasili C., Sassi M. and Bernini Carri C. (2005), Study on the CAP's contribution to EU economic and social cohesion, Final report, EU Commission Tender AGRI-G2-02.

Arfini F. Donati M. and Paris Q. (2003), A national PMP model for policy evaluation in agriculture using micro data and administrative information, contribute paper to international conference: Agricultural policy reform and the WTO: Where are we heading?, Capri, June 2003.

Arfini F., Donati M. and Zuppiroli M. (2005), Agrisp: un modello di simulazione regionale per valutare gli effetti per l'Italia di modifiche delle politiche agricoli, in Anania G. (Ed.), La riforma delle politiche agricoli dell'UE ed il negoziato WTO, Franco Angeli Editore, Milano.

Atkinson A. C., Riani M. and Cerioli A. (2004), Exploring Multivariate Data with the Forward Search, Springer, New York.

Brooke A., Kendrick D, Meeraus A. and Raman R. (2005),GAMS: a user's guide, GAMS Development Corporation.

Buysse J., Fernagut B, Harmignie O., De Frahan B.H., Lawers L., Polomé P. Van Huylenbroeck G. and Van Meensel J. (2007), Farm-based modelling of the EU sugra reform: impact on Belgian sugar beet suppliers, *European review of agricultural economics*, Vol. 34(1), pp. 21-52.

Fang S-C., Rajasekera J.R. and Tsao H.S.J. (1998), Entropy optimization and mathematical programming, Management Sciences-Operations Research, Kluwer's INTERNATIONAL SERIES.

Gohin A. and Chantreuil F. (1999), La programmation mathématique positive dans les modèles d'exploitation agricole, Cahiers d'économie et sociologie rurales, n. 52, pp. 59-78.

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

Hecklei T. (1997), Positive Mathematical Programming: Review of the standard approach, *CAPRI Working Paper*, University of Bonn.

Heckelei T. and Britz W. (2005), Model based on positive mathematical programming: state of the art and further estension, in Arfini F. (Ed.), Modelling agricultural policies: state of the art and new challenges, MUP, Parma, pp. 48-74.

Heckelei T. and Britz W. (2000), Positive mathematical programming with multiple data points: a cross-sectional estimation procedure, *Cahiers d'économie et sociologie rurales*, n. 57, 2000, pp. 27-50.

Heckelei T. and Wolff H. (2003), Estimation of constrained optimization models for agricultural supply analysis based on generalized maximum entropy, *European review of agricultural economics*, Vol. 30, pp. 27-50.

Howitt R.E. (1995), Positive Mathematical Programming, in *American Journal of Agriculture Economics* , Vol. 77, pp. 329-342.

Jaynes E.T. (1957), Information Theory and Statistical Mechanics II, *Phisics Review*, Vol.108, pp. 171-190.

Judez L., Gonzalez A., Ibanez M., De Andres R., Urzainqui E., Chaya C. and Fuentes-Pila J. (2000), Application of a model of PMP to analyse the effects of the measures of Agenda 2000 in Spain, EAAE Seminar on Agricultural Sector Modelling, Bonn.

Lence S.H and Miller D.J. (1998), Estimation of multi-output production functions with incomplete data : A generalised maximum entropy approach, *European Review of Agricultural Economics*, n.25, pp. 188-209.

Léon Y., Peeters L., Quinqu M. and Surry Y. (1999), The Use of Maximum Entropy to Estimate Input-Output Coefficients from Regional Farm Accounting Data, *Journal of Agricultural Economics*, Vol.50, n.3, pp. 425-439.

Oude Lansink A. (1999), Generalized maximum entropy estimation and heterogeneous technologies, *European review of agricultural economics*, Vol. 26(1), pp. 101-115.

Paris Q. (2001), Symmetric positive equilibrium problem: a framework for rationalizing the economic behaviour with limited information, *American journal of agricultural economics*, Vol. 83(4), pp.1049-1061.

Paris Q. and Arfini F. (2000), Frontier Cost Functions, Self-Selection, Price Risk, PMP and Agenda 2000, Programme CT97-3403 "Eurotools", *Working Papers Series*, n.20.

Paris Q., Montresor E., Arfini F. and Mazzocchi M. (2000), An Integrated Multi-phase Model for Evaluating Agricultural policies Trough Positive Information, EAAE Seminar on Agricultural Sector Modelling, Bonn, 29-31 March 2000.

Paris Q. and Howitt R.E. (1998), An Analysis of Ill-Posed Production Problems Using Maximum Entropy, *American Journal of Agricultural Economics*, n.80, pp. 124-138.

Rencher A. C. (1997), Multivariate Statistical Inference and Applications. Methods of Multivariate Analysis, Volume 2, Wiley-Interscience, New York.

Shannon C.E. (1948), A Mathematical Theory of Communication, *The Bell System Technical Journal*, Vol.27, pp. 379-423, 623-656.

## Appendix
## Revenue and cost function matrices

**Demand function matrix $\hat{\boldsymbol{D}}$**

| | | J' | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Cereals Mix | Alfa-alfa | Sugarbeet | Durum Wheat | Fodder crops | Maize | Barley | Silage | Soya | SoftWheat |
| | Cereals Mix | **0.7296** | 0.0378 | 0.0002 | 0.0791 | 0.1206 | -0.0742 | -0.0402 | -0.0028 | -0.2774 | -0.0195 |
| | FodderCrops | | **0.1171** | 0.0098 | -0.0157 | 0.0228 | -0.0592 | -0.0325 | -0.0165 | 0.0140 | -0.0151 |
| | Sugarbeet | | | **0.0199** | 0.0014 | -0.0004 | -0.0167 | 0.0110 | -0.0039 | -0.0126 | -0.0158 |
| | DurumWheat | | | | **2.0305** | -0.0244 | 0.0034 | -0.0251 | -0.0170 | -0.1762 | 0.0032 |
| J | AlfaAlfa | | | | | **1.7535** | 0.0114 | -0.0423 | -0.0516 | -0.1563 | -0.0155 |
| | Maize | | | | | | **0.2615** | 0.0302 | -0.0013 | -0.0442 | -0.0750 |
| | Barley | | | | | | | **0.7192** | 0.0079 | -0.0107 | -0.1155 |
| | Silage | | | | | | | | **0.0961** | 0.0289 | 0.0028 |
| | Soya | | | | | | | | | **1.4435** | -0.0061 |
| | SoftWheat | | | | | | | | | | **0.2184** |

**Cost function matrix $\hat{\boldsymbol{Q}}$**

| | | J' | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Cereals | Alfa-alfa | Sugarbeet | Durum Wheat | Fodder crops | Maize | Barley | Silage | Soya | SoftWheat |
| | Cereals | **0.0526** | 4.87E-5 | -0.0006 | 4.96E-6 | -5.85E-6 | -0.0001 | 2.74E-8 | -0.0002 | -4.21E-5 | -2.59E-6 |
| | FodderCrops | | **0.0043** | -5.57E-7 | -4.94E-9 | -6.00E-9 | -6.16E-8 | 4.21E-9 | -1.60E-7 | -4.26E-8 | -4.15E-9 |
| | Sugarbeet | | | **0.0006** | 5.44E-8 | 6.44E-8 | 7.87E-8 | -3.0E-10 | 1.93E-6 | 4.63E-7 | -2.85E-8 |
| | DurumWheat | | | | **0.5185** | 0.0543 | -0.0034 | 0.0282 | 0.0028 | 0.0105 | 0.0012 |
| J | AlfaAlfa | | | | | **0.0287** | -0.0004 | 0.0030 | 0.0003 | 0.0011 | 0.0001 |
| | Maize | | | | | | **0.0072** | -0.0002 | -1.84E-5 | -0.0001 | -7.91E-6 |
| | Barley | | | | | | | **0.0309** | 0.0002 | 0.0006 | 0.0001 |
| | Silage | | | | | | | | **0.0025** | 0.0001 | 6.37E-6 |
| | Soya | | | | | | | | | **0.1042** | -0.0029 |
| | SoftWheat | | | | | | | | | | **0.0074** |