Farm production costs estimation trough PMP Models: an application in three Italian Regions

Arfini F.[1], Donati M.[1] , Cesaro L.[2] and Marongiu S.[2]

[1] University of Parma, Department of Economics, Parma, Italy
[2] Istituto Nazionale di Economia Agraria, Roma, Italy

michele.donati@unipr.it

Paper prepared for presentation at the 1st AIEAA Conference
'Towards a Sustainable Bio-economy: Economic Issues and Policy Challenges'

4-5 June, 2012
Trento, Italy

Summary

The objective of this paper is to present a Generalised Positive Mathematical Programming model suitable for the estimation of variable cost of production associated with different farm activities. This work present, discuss and demonstrates that the Generalised PMP model is a useful theoretical framework for the representation of farm choice, including for the description of costs related to the production function chosen by each entrepreneur. For this characteristics the model can be used for the farms belonging the FADN sample providing a powerful tools for researcher that would like to know variable costs of production for agricultural activities or estimate the impact of agricultural policy and market reform at regional and sectorial level. The main feature of the generalised PMP model is its independence from any "external" information, included the support value of the GME parameters and the abandon of the "tautology" problem always present in the standard PMP models. The paper also present the results of the cost estimation process and validate it comparing the observed variable cost with the estimated variable cost related to a sample of 738 farms belonging the FADN data base of three Italian regions.

Keywords: variable cost of production, positive mathematical programming, farm accountancy data network

JEL Classification codes: Q12, C61, Q18, C38

Farm production costs estimation trough PMP Models: an application in three Italian Regions

Arfini F.[1], Donati M.[1] , Cesaro L.[2] and Marongiu S.[2]

[1] University of Parma, Department of Economics, Parma, Italy
[2] Istituto Nazionale di Economia Agraria, Roma, Italy

1. INTRODUCTION

It is well known that Farm Accountancy Data Network (FADN) does not collect information at the European level about the variable costs associated with different farm activities, but rather compiles the total variable costs at the farm level (FADN, 2010). The problem of determining the cost of production for all activities within the FADN database is an important issue when the production cost derivation is finalised to evaluate policies using micro-based economic models.

In this respect, it is evident that all analyses that aim to evaluate production allocation decisions under policy or market pressures cannot be performed without knowing the specific variable costs. These costs can be derived from external sources (engineering information, literature, experts etc.), with the risk that these costs may not be able to be differentiated according to the technology level, farm specialisation and farm size. Thus, a model misspecification becomes a problem when applied to single farms or a small group of farms within panel data, such as that collected by FADN. An alternative approach is to estimate variable cost using econometrics (Butault et al. 2011, Butault et al. 2001) or mathematical programming models. The most recent development concerning this latter methodology is the proposal and discussion of Positive Mathematical Programming by Howitt (1985), Paris and Howitt (1998), Paris and Arfini (2000), Heckelei and Britz (2005), and Paris (2011). All these methodologies estimate the parameters of a total variable cost function represented by a symmetric positive semi-definite matrix (named the Q matrix). In any case, the Q matrix represents a set of parameters but does not allow for the exact estimation of the variable cost of activities at the farm level for each observed farm.

For this reason, the debate in the literature (Heckelei, 2002; Arriaza and Gomez-Limon, 2003; Heckelei and Britz, 2005; Henry de Frahan et al., 2006; Buysse et al., 2007a; Britz and Adenauer, 2009; Kallenopoulos et al., 2010) does not fully recognise the economic meaning of the Q matrix, which is only considered based on its diagonal elements. Thus, PMP is usually only considered an efficient tool for the calibration of mathematical programming models, especially when misspecification problems occur.

It should also be considered that little information from the FADN database is used and that the data typically refer to a sample of farms that group together farms belonging to the FADN region (Britz and Adenauer, 2009; Kallenopoulos et al., 2010), creating one "average" farm. Only a few cases have considered single farms within the FADN region (Paris and Howitt, 1998; Paris and Arfini, 2000). The implication of using average farms to model farm typology is that the model does not capture differences between farms in the same region in terms of technologies and related costs.

The main criticisms of considering the average farm include the following: i) not all the information is considered; ii) it is difficult to take economies of scale into account; iii) changing the year of observation also changes the composition of the farms belonging to the FADN database, and the "farm type" that the model considers does not represent the observed reality. In this sense, the “positive” feature of PMP no longer exists.

This paper proposes a Generalised PMP model that recovers the dual information linked to each farm production activity and estimates the specific variable accounting cost per activity without using external information.

Within this framework, the Generalised PMP methodology is able to recover and validate the specific variable accounting costs related to activities with data collected by the FADN.

The paper is divided into three parts. The first part presents the state of the art related to the use of PMP models. The second part presents the Generalised PMP model, the theoretical justification of the cost estimation and the economic meaning of the estimated cost. The third part of the paper presents the validation of specific estimated variable costs considering a sample of individual farms belonging to three Italian FADN regions: Veneto, Piedmont and Lombardy. The paper will end with a discussion of the strengths and weaknesses of the proposed methodology.

2. THE PMP AND RELATED PROBLEMS

The introduction of PMP by Howitt (1995) and Paris and Howitt (1998) was perceived as one of the most important innovations in the field of Mathematical Programming. PMP has provided researchers in the field of agricultural economics with powerful new tools reviving mathematical programming and creating a bridge to econometrics (Heckelei and Britz, 2005). In brief, PMP has opened a new research frontier and has created new opportunities for investigating land allocation under the pressure of new market and policy scenarios.

The seminal works of 1995 and 1998 was criticised and discussed in many aspects, with the main areas of discussion as follows:

- The PMP approach (Cafiero, 2004);
- The introduction of the positive constraints (Heckelei and Wolff, 2003; Heckelei and Britz, 2005);
- The estimation of the dual value linked to each production activity (Heckelei, 2002; Judez et al, 2001, 2008);
- The use of Maximum Entropy and support values for the estimation of the Q matrix (Henry de Frahn et al., 2006);
- The characteristics of the Q matrix (Diagonal or full) (Kanellopulos et al., 2010; Severini and Cortignani, 2011);
- The use of single observations compared with multiple observations (Heckelei and Britz, 2000);
- The possibility of introducing new activities (Röhm, O. and Dabbert, 2003; Blanco et al., 2008).

For the purposes of this paper, the literature has two main implications related to the effective implementation of PMP methodology. The first is associated with the objective of PMP methodology. PMP was perceived by many researchers as a good tool to calibrate LP models, especially when dealing with problems of over-specialisation (Helming et al., 2001; Heckelei, 2002; Hekelei, 2003; Helming, 2005; Buysse et al., 2007a; Buysse et al., 2007b; Kanellopoulos, 2010). This is because many researchers interpret the dual variables associated with the calibration constraints as parameters able to capture any type of model misspecification, data error, aggregation bias, risk behaviour and price expectations (Heckelei et al, 2005). As a consequence, the so-called standard approach has led to the lack of a clear economic meaning for the dual variables associated with calibration constraints and, thus, for the Q matrix. A related issue is the “tautology problem” that was perceived as a negative element of the PMP methodology (Paris, 2011).

The second most important implication of the standard PMP approach is the difficulty to estimate the Q matrix that considers all the observed activities when no information is available related to the activity costs, $\boldsymbol{c}$. The problem of implementing the PMP model without knowing $\boldsymbol{c}$ is related to the fact that the imposition of calibration constraints generates at least one associated shadow value equal to zero; otherwise, the shadow price for the structural constraint (land) will be equal to zero (Paris and Howitt, 1998) and will be missed an observed activity in the Q matrix.
To clarify these two important points (the economic meaning of the calibration constraints and the difficulty to estimate the total variable cost per activity), let us recall the basic form of the PMP methodology (Paris and Arfini, 2000) with N farms articulated in two phases:

$$\max_{\boldsymbol{x}_n \geq 0} \left(\boldsymbol{p}_n' \boldsymbol{x}_n - \boldsymbol{c}_n' \boldsymbol{x}_n \right) \tag{1}$$

subject to

$$\boldsymbol{A}_n \boldsymbol{x}_n \leq \boldsymbol{b}_n \tag{2}$$

$$x_{nj} \leq x_{Rnj}, \qquad \text{for } x_{Rnj} > 0, \quad j = 1, \ldots, J_n \tag{3}$$

The first phase is composed of a model where $\boldsymbol{n}$ is the number of farms, $\boldsymbol{p}_n$ is the vector of output prices faced by the n-th farm, $\boldsymbol{c}_n$ is the vector of observed accounting costs per unit of output, $\boldsymbol{A}_n$ is the matrix of fixed technical coefficients involving limiting allocable inputs, $\boldsymbol{b}_n$ is the vector of availability of limiting allocable inputs and $\boldsymbol{x}_{Rn}$ is the vector of realised output levels. The vector $\boldsymbol{x}_n$ is nonnegative. Each farm exhibits I allocable inputs and J_n products. The vector of realised land allocation decisions is indicated by $\boldsymbol{h}_{Rn}$. Land is assumed to be the only limiting input. The n-th matrix $\boldsymbol{A}_n$ of technical coefficients is defined as $\boldsymbol{A}_n = \left[a_{nij} \right]$, where $a_{nij} = h_{Rni} / x_{Rnj_n}$.
Constraints (2) are called structural constraints while constraints (3) are called calibration constraints. The vector of shadow prices, $\boldsymbol{y}_n$, is associated with the allocable input constraints (2). The vector of differential marginal costs, $\boldsymbol{\lambda}_n$, corresponds to the calibration constraints (3) and its economic interpretation is the usual interpretation given to the dual variables of linear constraint.
The dual of model (1)-(3) can be stated as

$$\min_{y \geq 0, \boldsymbol{\lambda} \geq 0} = (\boldsymbol{b}_n' \boldsymbol{y}_n + \boldsymbol{\lambda}_n' \boldsymbol{x}_{Rn}) \tag{4}$$

subject to

$$\boldsymbol{A}_n' \boldsymbol{y}_n + \boldsymbol{\lambda}_n + \boldsymbol{c}_n \geq \boldsymbol{p}_n \tag{5}$$

where vectors $\boldsymbol{y}_n$ and $\boldsymbol{\lambda}_n$ are nonnegative. A detailed interpretation of both the primal and dual models is given in Paris and Howitt (1998: 126-127).
It is important to note that the sole purpose of the first stage of the PMP methodology is to obtain an accurate and consistent measure of the marginal cost associated with the vector of the realised level of activities, $\boldsymbol{x}_{Rn}$.
The second stage of the PMP approach deals with the reconstruction of the marginal cost function using a specification that is linear in its parameters. The linearity aspect of the model becomes important when the number of farms is large. The integration of the marginal cost function with respect to the output variables within the admissible domain will produce the desired total variable cost function. The cost function is hypothesised to be a quadratic function in output quantities (input prices are not available from the farm

survey and are assumed to be fixed): $C(\boldsymbol{x}) = \boldsymbol{x}'\boldsymbol{Q}\boldsymbol{x}/2$, where the $\boldsymbol{Q}$ matrix is positive symmetric semi-definite.

Given the LP specification discussed above, the associated marginal cost function for all the observed farms can be represented as $mc(\boldsymbol{x}) \equiv \bar{\boldsymbol{\lambda}}_{LP} + \bar{\boldsymbol{c}} = \boldsymbol{Q}\bar{\boldsymbol{x}}_R$. From the above formulations, it is clear that the economic interpretation of $(\boldsymbol{\lambda}_n + \boldsymbol{c}_n)$ for all $\boldsymbol{n}$ farms is the marginal cost for each activity. This expression represents the marginal cost perceived by the entrepreneur and is composed of two elements: the specific explicit accounting cost $\boldsymbol{c}$ and the differential marginal cost λ**.** Due to the estimation process of the standard PMP approach, it is necessary to know the accounting cost $\boldsymbol{c}$ for each activity.

Problem (1)-(5) demonstrates that the calibration constraint avoids the degeneration of the problem and leads to a positive shadow price for the structural constraints. According to the formulation of Paris and Howitt (1998), in the second phase of PMP, the presence of a shadow price equal to zero can lead to a misspecification of the non-linear cost function and impede the correct estimation of the cost function. Thus, the lack of accounting cost per activity at the farm level means that we cannot derive the cost function parameters for the marginal product because its marginal cost value is null. By contrast, if λ is equal to zero for a certain process, the component *c* permits a positive value to be maintained for the cost associated with an activated process. In other words, to guarantee that the standard approach works properly, the accounting cost per activity should always be known and present in the model.

$\boldsymbol{c}$ is only known by researchers in a few cases because while it is collected in national farm datasets[1], information related to specific accounting costs is often lacking as in the case of FADN.

To solve the impasse due to the lack of information that can be related to the countable cost, the literature provides a number of contributions that modify the standard PMP formulation. In particular, Heckelei (2002) and Heckelei and Wolff (2003) offer a wide range of instruments to assess the cost function starting from the observed production level. In particular, they propose to overcome the “first phase” of calibrating the observed situation by directly imposing the first-order conditions on the cost function estimation phase. This procedure obliges the use of external information from the FADN dataset provided by experts or by regional investigations concerning the dual value of the fixed factor. The main advantage of this procedure is that it makes PMP a very useful tool for all EU FADN regions. At the same time, the main disadvantage is that the constructed models represent only one farm (usually at the regional level) and that external data do not always fit with the characteristics of the observed farms collected in the FADN sample. In reality, the value of the rent of the land may change within the region according to several factors, and the dual price of the land may also be quite different for different farms typology according to their size, level of specialisation and the specific characteristics of each farm holder. In sum, the value of the rented land is not easy to identify and can lead to an incorrect estimation of the PMP models.

Many papers (Kanellopoulos, 2010; Gotch, 2005; Gotch and Britz, 2011) and, in particular, many European researchers (Kanellopoulos, 2010; Gotch and Britz, 2011) have adopted the suggestions of Heckelei for developing PMP models capable of assessing the impact of CAP policies. These studies use aggregated FADN information for different years. Only aggregate farm information allows for the collection of consistent dual information from the home territory of the farms. At the same time, farms are considered as panel data containing several years of observation. This strategy makes it possible to better consider the yield variation and, thus, the technology characteristics.

[1] For example, in the Italian FADN information on the specific accounting costs is always present because INEA (National Institute of Agricultural Economics) collects this information or estimates it by attributing the accounting costs associated with each process for each surveyed farm.

Given the methodological setting described in the previous paragraphs, an alternative PMP approach is proposed in this paper with the objective to use only the endogenous information available for all farms belonging to the FADN database and thus maintain flexibility in terms of creating models able to describe and represent different situations according to the research questions and then consider different farm types at different territorial levels, starting from individual farms. For this reason a "generalised" model is proposed to allow for the new formulation of PMP to be used in a different context.

The model, also discussed and justified in the KKT conditions by Paris (2011), uses the information available in the European FADN archive as a guide for the correct estimation of the explicit variable activity costs and proposes to merge the first phase with the second phase through the dual properties of the PMP approach.

The model can be presented as follows: assume a sample of farms composed of N farms and consider that information about the production plan, prices and technical coefficients (the quantity of factors used to obtain one unit of each farm product) are known at the farm level. We also consider only one limited factor, the land available at the farm level, b_n. The use of this factor per unit output is represented by the technology matrix A_n. The known levels of production for each farm are indicated by the vector $\overline{\mathbf{x}}_n$, while output market prices are represented by the vector $\mathbf{p}_n$ and exogenous marginal costs related to each activity are represented by the vector $\mathbf{c}_n$. This latter variable can be viewed as the cost originating from the farm accountancy and is observed.

The objective of a PMP model is to recover the part of the information that cannot be directly collected at the farm but contributes to the decision-making process of farmers in a more or less conscious way. The information takes on different meanings as price expectations, specific production preferences and technological skills of the individual farmers. This part of the information, which is obviously lacking inside European FADN databases, can be derived through the PMP properties. The implicit information that we want to reveal is the vector $\boldsymbol{\lambda}_n$, which contains the adding marginal cost for each farm considered by farmers in defining a certain production plan with the explicit cost $\mathbf{c}_n$. Adopting the generalised PMP approach the following problem can be introduced:

$$\min_{\mathbf{u}_n, y_n, \boldsymbol{\lambda}_n, \mathbf{Q}} \left\{ \sum_{n=1}^{N} \frac{1}{2} \mathbf{u}_n' \mathbf{u}_n + \sum_{n=1}^{N} \left(b_n y_n + \boldsymbol{\lambda}_n' \overline{\mathbf{x}}_n + \mathbf{c}_n' \overline{\mathbf{x}}_n - \mathbf{p}_n' \overline{\mathbf{x}}_n \right) \right\} \tag{6}$$

subject to

$$A_n' y_n + \boldsymbol{\lambda}_n + \mathbf{c}_n \geq \mathbf{p}_n \qquad (\mathbf{w}_n) \tag{7}$$

$$\mathbf{c}_n + \boldsymbol{\lambda}_n = \mathbf{Q}\overline{\mathbf{x}}_n + \mathbf{u}_n \qquad (\mathbf{z}_n) \tag{8}$$

where $y_n \geq 0$, $\boldsymbol{\lambda}_n \geq 0$ and $\mathbf{Q}$ is a symmetric positive semi-definite matrix, as stated by Paris and Howitt (1998) and Paris (2011). $\mathbf{w}_n$ and $\mathbf{z}_n$ are the shadow prices associated with equations (7) and (8), respectively. $\mathbf{u}_n$ is the vector of marginal cost deviations per farm, that is, the distance between the marginal cost $\mathbf{c}_n + \boldsymbol{\lambda}_n$ and the marginal cost $\mathbf{Q}\overline{\mathbf{x}}_n$ of a non-linear cost function such that $\mathbf{c}_n + \boldsymbol{\lambda}_n - \mathbf{Q}\overline{\mathbf{x}}_n = \mathbf{u}_n$. The estimated parameters of Q are part of a quadratic cost function aiming to provide flexibility to model responses towards farm simulations. The model is optimised by a combined objective function, (6), that considers a least-squares technique and minimises the difference between the total revenue, $\mathbf{p}_n' \overline{\mathbf{x}}_n$, and the total cost, $b_n y_n + \boldsymbol{\lambda}'_n \overline{\mathbf{x}}_n + \mathbf{c}'_n \overline{\mathbf{x}}_n$. This latter expression identifies the optimal condition for the standard PMP approach, or in general terms, states that under optimal conditions, the primal objective function should be equal to the dual function.

The above model integrates the first and second phases of the standard PMP approach using the PMP dual properties. In this model, there is no explicit trace of both the calibrating constraints and the epsilon terms that help to break the linear dependence between structural and calibration constraints. The constraints of the model (7)-(8) concern the equilibrium conditions with marginal costs greater than or equal to marginal revenue and the relationship by which a linear cost function is shifted to a quadratic cost function. The model does not repeat the tautological procedure of the standard approach of deriving information about the output levels, which were already known before the model was developed, but rather reveals hidden information about the differential marginal costs within the production levels and makes this information available for the simulation phase.

To better understand the significance of this problem and the properties of the solution, we can transform the model in its alternative Lagrangean representation, as follows:

$$\begin{aligned} L = & \sum_{n=1}^{N} \frac{1}{2} \mathbf{u}'_n \mathbf{u}_n + \sum_{n=1}^{N} \left(b_n y_n + \boldsymbol{\lambda}'_n \overline{\mathbf{x}}_n + \mathbf{c}'_n \overline{\mathbf{x}}_n - \mathbf{p}'_n \overline{\mathbf{x}}_n \right) \\ & + \sum_{n=1}^{N} \mathbf{w}'_n (\mathbf{p}_n - A'_n y_n - \boldsymbol{\lambda}_n - \mathbf{c}_n) + \sum_{n=1}^{N} \mathbf{z}'_n (\boldsymbol{\lambda}_n + \mathbf{c}_n - \mathbf{Q} \overline{\mathbf{x}}_n - \mathbf{u}_n) \end{aligned} \tag{9}$$

From the Lagrangean function we can obtain the following relevant KKT conditions:

$$\frac{\partial L}{\partial \mathbf{u}_n} = \mathbf{u}_n - \mathbf{z}_n = \mathbf{0} \tag{10}$$

$$\frac{\partial L}{\partial \boldsymbol{\lambda}_n} = \overline{\mathbf{x}}_n - \mathbf{w}_n + \mathbf{z}_n \geq \mathbf{0} \tag{11}$$

$$\frac{\partial L}{\partial y_n} = b_n - A_n \mathbf{w}_n \geq 0 \tag{12}$$

The partial derivatives (10) indicate that the deviation terms, $\mathbf{u}_n$, are equal to the dual values, $\mathbf{z}_n$, linked to the equation (8). Because the problem attempts to minimise the squares of the farm cost, the deviations $\mathbf{u}_n$ and $\mathbf{z}_n$ should assume very small values close to zero. The KKT condition (11) can be rewritten as $\mathbf{w}_n - \mathbf{z}_n \leq \overline{\mathbf{x}}_n$, showing that the difference between the two shadow prices associated with equations (7) and (8) should be less than or equal to the realized outputs. In this respect, if we consider that the shadow price of the equation representing the equilibrium condition can be interpreted as the shadow output quantity, we can state that $\mathbf{w}_n \approx \overline{\mathbf{x}}_n$. Furthermore, as we have affirmed for the KKT condition (11), $\mathbf{z}_n$ can be viewed as a small term close to zero, and thus, we can state that $\mathbf{z}_n \approx \boldsymbol{\varepsilon}$. Rearranging this information, the KKT condition (11) becomes $\mathbf{w}_n \leq \overline{\mathbf{x}}_n + \mathbf{z}_n$, corresponding to the calibration constraint of the standard approach, which implies that models (6)-(8) correctly replicate the standard PMP specification without the explicit calibration constraints. Taking the previous considerations into account, the KKT condition (12) can be interpreted as the structural constraint related to land use. Moving b_n to the right-hand side of equation (12) and changing the sign, we obtain $A_n \mathbf{w}_n \leq b_n$ corresponding to equation (7).

In turn, the Generalised PMP approach overcomes the tautological procedure of the standard PMP approach and provides all the necessary information on the total marginal cost that is useful for the simulation phase.

3. PMP COST ESTIMATION APPROACH USING TOTAL VARIABLE FARM COSTS

The proposed Generalised PMP approach assumes knowledge of information related to the accounting cost $\boldsymbol{c}$, but it is well known that this information is lacking at the European level. In addition, to properly represent the observed land use for each farm in the sample, the "self selection" problem should be considered (Paris and Arfini, 2000). According to this objective, models (6)–(8) take the information related to the total variable costs available at the farm level in the European FADN as a guide for the accounting cost estimation and is modified in the following manner:

$$\min_u LS = \frac{1}{2}\mathbf{u}'\mathbf{u} \tag{13}$$

subject to

$$\boldsymbol{\alpha} + \boldsymbol{\lambda} = \mathbf{R}'\mathbf{R}\overline{\mathbf{x}} + \mathbf{u} \qquad \text{se } \overline{x} > 0 \tag{14}$$

$$\boldsymbol{\alpha} + \boldsymbol{\lambda} \leq \mathbf{R}'\mathbf{R}\overline{\mathbf{x}} + \mathbf{u} \qquad \text{se } \overline{x} = 0 \tag{15}$$

$$\boldsymbol{\alpha}'\overline{\mathbf{x}} \leq TVC \tag{16}$$

$$\mathbf{u}'\overline{\mathbf{x}} + \frac{1}{2}\overline{\mathbf{x}}'\left(\mathbf{R}'\mathbf{R}\right)\overline{\mathbf{x}} \geq TVC \tag{17}$$

$$\boldsymbol{\alpha} + \boldsymbol{\lambda} + \mathbf{A}'\mathbf{y} \geq \mathbf{p} \tag{18}$$

$$\mathbf{b}'\mathbf{y} + \boldsymbol{\lambda}'\overline{\mathbf{x}} = (\mathbf{p} - \boldsymbol{\alpha})'\overline{\mathbf{x}} \tag{19}$$

$$\mathbf{R} = \mathbf{L}\mathbf{D}^{1/2} \tag{20}$$

$$\sum_{n=1}^{N} u_{n,j} = 0 \tag{21}$$

The objective of models (13)-(21) is to estimate a non-linear cost function including the unknown accounting variable cost $\boldsymbol{\alpha}$. The restrictions (14) and (15) define the relationship between marginal costs derived from a linear function and marginal costs derived from a quadratic cost function. $\boldsymbol{\alpha}+\boldsymbol{\lambda}$ defines the sum of the unknown (or estimated) accounting variable costs and the differential variable marginal costs. The latter are implicit in the decision-making process of the entrepreneur and are not accounted for in the holding's bookkeeping. Both components are endogenous variables within the minimisation problem. The restrictions (14) and (15) also guarantee that the self-selection rule is followed, enabling farmers to select possible production activities from all activities present in the region (represented by the sample dimension) but restricting activities to those observed in the first phase of the PMP methodology (Paris and Arfini, 2000). Moreover, to guarantee consistency between the estimated accounting variable costs and those effectively recorded by the farm accounting system, constraint (16) requires that the total estimated variable cost is not greater than the total variable cost observed in the FADN databank at the farm level. Equation (17) states that the costs estimated by the model by means of a non-linear cost function must be at least equal to the value of the observed total variable cost (TVC). To guarantee consistency between the estimation process and the optimal conditions, restriction (18) introduces the traditional condition of economic equilibrium, where total marginal costs must be greater than or equal to marginal revenues. The total marginal costs also consider the use cost of the factors of production defined by the product of the technical coefficients matrix $\mathbf{A}'$ and the

shadow price of the restricting factors ***y***; while the marginal revenues are defined by the sum of the products' selling prices, ***p***, and any associated public coupled subsidies. The additional constraint (19) defines the optimal condition where the value of the primal function corresponds exactly to the value of the objective function of the dual problem. To ensure that the matrix of the quadratic cost function is symmetric positive semi-definite, the model adopts Cholesky's decomposition method (20). Finally, restriction (21) establishes that the sum of the errors **u** must be equivalent to zero.
The cost function estimated with the model (13)-(21) may be used in a model of maximisation of the farm gross margin, ignoring the calibration restrictions imposed during the first phase of the standard PMP approach. In this case, the dual relations entered in the preceding cost estimation model guarantee the reproduction of the observed situation. The model, therefore, appears as follows:

$$\max_{x\geq 0} ML = \mathbf{p'x} - \left\{ \frac{1}{2}\mathbf{x'\hat{Q}x} + \mathbf{\hat{u}'x} \right\} \tag{22}$$

subject to

$$\mathbf{Ax} \leq \mathbf{b} \tag{23}$$

$$A_j x_j - h_j = 0 \qquad \forall j = 1,...,J \tag{24}$$

Models (22)-(24) calibrate the observed farming system, thanks to the non-linear objective function, which preserves the (economic) information on the levels of production effectively attained. The estimated matrix $\boldsymbol{Q}$ is reconstructed using Choleschy's decomposition: $\mathbf{\hat{Q}} = \mathbf{\hat{R}}'\mathbf{\hat{R}} = \mathbf{\hat{L}\hat{D}\hat{L}}'$. Constraint (23) represents the restriction on the structural capacity of the farm, while relation (24) enables us to obtain information on the hectares of land (or number of animals) associated with each process *j*. Once the initial situation has been calibrated through the maximisation of the farm gross margin, it is possible to introduce variations in the public aid mechanisms and/or in the market price levels to evaluate the farm's reactions to various policy conditions. The reaction of the farm production plan will take into account the information used during the estimation phase of the cost function, where it is possible to identify a real, true matrix of firm choices, i.e., **Q**.
The use of the Least Squares approach to estimate the cost function is an alternative to the Maximum Entropy method and has the advantages of avoiding the unsolved problems of the arbitrary use of support values (Golan and other, 1996; Lence and Miller, 1998; Henry de Frahan, 2006).

4. THE COST ESTIMATION AND MODEL STRUCTURE

The Generalised PMP model described here can be used in three different contexts: i) the estimation of the accounting variable costs (α) related to each activity with data (output price, yields, farm production level, land use and total variable cost at the farm level) collected by the FADN; ii) the estimation of the total variable cost per crop perceived by the farmers ($\alpha + \lambda$) (this allows for a set of information that is useful for evaluating farm behaviour by means of the definition of a new profit function); iii) the representation of farm behaviour under market and policy pressures as a simulation of the impact of CAP reform (Arfini and Donati, 2008; Arfini and Donati, 2010).
Within this framework, the PMP methodology described in this section will be implemented to recover the variable costs related to the process where data are collected by the FADN.
To validate the proposed methodology, the results are compared with observed information recovered from the same FADN database. In this respect, the Italian FADN liaison office, INEA, collects the specific variable costs for each crop, including the costs of seeds, fertilisers, pesticides and services provided by third parties. This information arises from the process of accounting attribution starting from the farm invoice

information collected by the local FADN interviewer. This information is computed each year but is not transferred to the European database. It is clear that the result of the process of cost distribution among activities can lead to an imperfect evaluation of specific farm costs, but it is the closest possible approximation of the real information. For our purposes, the accounting variable costs detected by the Italian FADN represent a benchmark for the validation of the estimated accounting variable costs for the same activity observed in the same farm.

For the purpose of the present analysis, the cost estimation is then developed for each farm belonging to a specific farm type stratification to maintain a sufficient degree of homogeneity with respect to farm technology. Not all farm types have been investigated but only the most numerous, such as farms with arable crops.

4.1. Model architecture

The Generalised PMP model is developed as part of an articulated elaboration system and is fully developed using GAMS with the support of GDX features. This system is divided into different modules, each devoted to a specific task and interfaced with the others to provide input information. Four modules can be distinguished: i) data entry; ii) stratification; iii) estimation and calibration; and iv) the output module.

The scope of this articulation is to facilitate researchers' management of large databases that require several treatments to check the presence of outliers, select relevant information for analysis and organise the data for the subsequent processes. To improve the degree of homogeneity inside the data, the stratification module is comprised of groups of farms with similar characteristics. The "estimation and calibration" module aims to estimate the specific costs calibrating the observed situation. Finally, the output module stores and manages all the model outputs with the main objective of validating the specific variable cost per activity using a statistical test based on the Student's *t*-test.

4.2. Data description: Veneto, Lombardy and Piedmont samples

The Italian regions selected for model validation are in Northern Italy (north of the Po River) and are characterised by highly specialised and intensive agricultural practices. The most important activities in this region are arable crops and livestock production, mainly dairy and beef cattle (Eurostat, 2009).

The farm sample considered in this analysis is composed of 738 farms belonging to FT1 (Farm Type – arable crops). The average size of each farm in the sample is 50 ha. The FADN farms in Piedmont are the largest in terms of hectares. On average, cereals occupy 43% of the total UAA in the sample. The average GSP per hectare is 1,774 Euros, while the total variable cost per hectare is 600 Euros (Table 1).

Table 1. Description of the Italian FADN sample – Farm type 1

Area	N. of farms	Av. UAA (ha)	Cereals/tot (%)	GSP/ha (€)	Total Variable Costs/ha (€)
Veneto	220	44	62	1956	656
Lombardy	165	46	40	1763	370
Piedmont	353	56	36	1689	661
Total	738	50	43	1774	600

Source: own elaboration

Considering the entire sample, rice covers 39% of the total land area, followed by maize with 25% and winter wheat with 15%. In Veneto maize is the main crop, while in Lombardy and Piedmont rice occupies the greatest area. Another important crop is soya, accounting for 17% of the entire acreage in the Veneto sample. Indeed, Veneto specialises in producing maize and soya due to the presence of dairy and beef farms and important foodstuff industries.

All the crops described above are considered in the analysis conducted by the Generalised PMP model, and the related accounting variable production cost is estimated for each crop. As described in the previous section, the estimation is performed using the information on acreage, yields, and prices for each crop and the total variable costs at the farm level.

4.3. The homogeneity of the data and the treatment of outliers

To achieve a good fit between the estimation and reality, it is important to avoid the presence of outliers, and it is also useful to utilise a homogeneous sample of farms with respect to the main variable that influences the production function and dynamics of production cost, such as yields and output prices. Figure 1 presents some descriptive information on the prices and yields of four main crops included in the FADN sample.

Figure 1: Distribution of the ratio between price and yield in the FT1 sample.

Figure 1.a: Winter Wheat

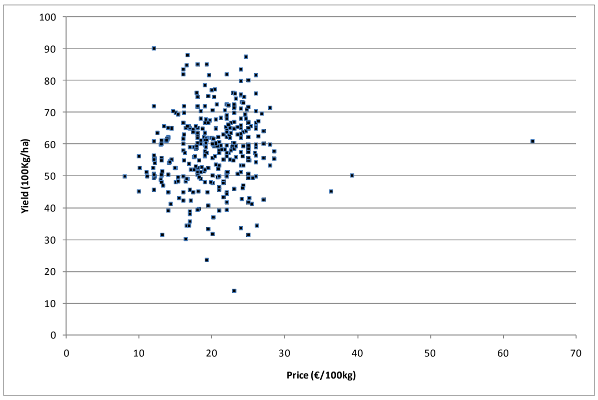


Fig. 1.b: Maize

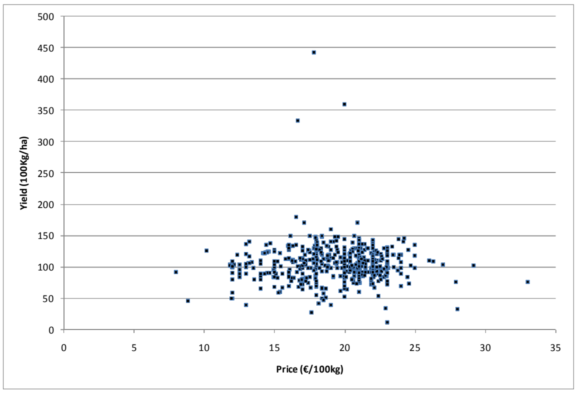


Figure 1.c: Soya

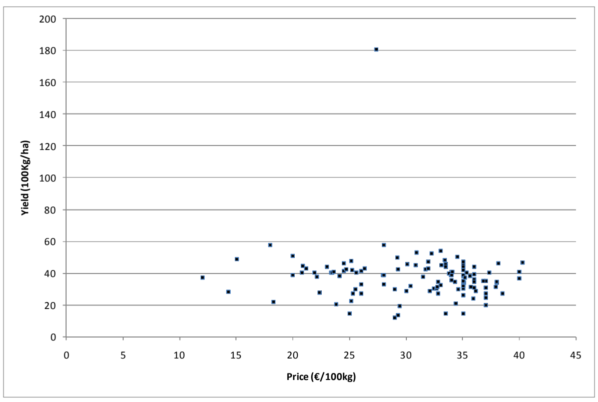
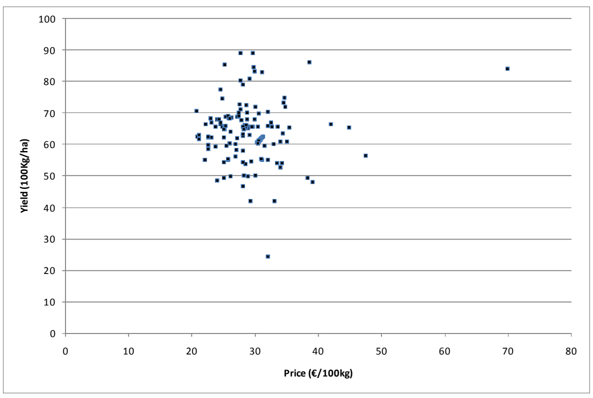


Figure 1.d: Rice



Source: own elaboration

For some crops, these observations are less dispersed for some crops, such as for common wheat and rice, while for other crops, such as maize and soya, the dispersion is very high. The main factor that influences the dispersion of the observations is the variation in yields. In fact, for maize, the standard deviation is very high, equal to 31, representing a variation with respect to the mean of 3.1 tons per hectare. In contrast, the dispersion in yields of rice is more restrained, with 0.9 tons per hectare.

The high level of dispersion also masks the presence of outliers, which can strongly influence the estimation results for some crops. For example, maize contains several observations that are out of range. In fact, Figure 1.b shows a cluster of points surrounded by several out-of-range observations. These points represent outliers that influence the capacity of the model to correctly estimate production costs and should thus be eliminated from the estimation procedure.

The disturbing information represented by the outliers can be appreciated at both the process and farm levels. Considering the farm information, the bad information is also present for the normalised variables concerning the gross saleable production (GSP) and farm total variable cost (TVC). Figure 2 and Figure 3 shows the farms on a scatter plot with GSP per hectare plotted against TVC per hectare. Some points are extremely far from the average observations and can be considered as outliers. Reducing the scale to observe the same sample in detail clearly shows that it is possible to adopt statistical techniques to detect a homogeneous set of observations.

Figure 2: Farm distribution of ratio between GSP/ha and TVC/ha (Veneto-Lombardy-Piedmont)- standard scale

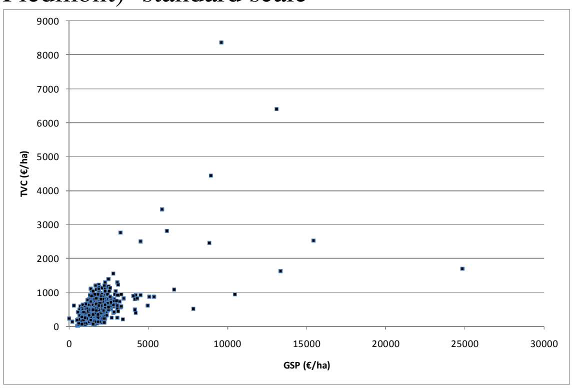
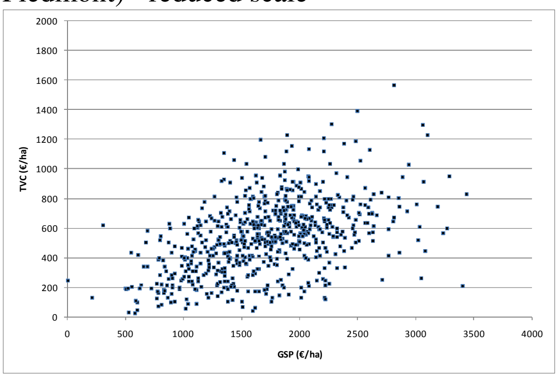


Figure 3: Farm distribution of ratio between GSP/ha and TVC/ha (Veneto-Lombardy-Piedmont) - reduced scale



Source: own elaboration

The estimation strategy is to consider all the activities present on the observed farm and to exclude those farms that are out of range because they cultivate a single crop or carry out many activities that are not homogeneous with respect to the characteristics of the sample. The homogeneity is evaluated using Principal Component Analysis (PCA) and Cluster Analysis (CA). This latter technique is implemented using the K-mean methodology. Only the clusters with the highest number of homogeneous farms are used for the cost estimation by the PMP model to guarantee sufficient numbers of observations of crops for estimation. Clusters with a low number of observations are considered as groups of dispersed farms, i.e., outliers.

4.4. The output of the model

Before discussing the analysis of the results, it is useful to recall that the Generalised PMP model allows for the estimation of two types of specific variable costs for each activity: the accounting cost (α) and the differential marginal cost (λ). These costs are estimated under the economic constraints of the model (13)-

(21), yielding the *estimated accounting cost* and the *estimated differential marginal cost* for each product. The estimated accounting cost can be interpreted as the part of the total marginal cost that can be explained by the accounting values, while the estimated differential marginal cost can be considered as the opportunity cost associated with each activity. This last statement is very close to the interpretation provided by Paris and Howitt (1998) for the shadow price associated with the calibration constraints. In turn, the sum of the estimated accounting cost and the estimated differential marginal cost provides the exact measure of the total variable (marginal) cost associated with each activity.

The estimated differential marginal costs are defined in this work as "*hidden costs*" to indicate the part of the estimated marginal costs that are considered by farmers in defining their production plans but that are absent from farm accounting sheets. As mentioned above, these are the part of marginal costs related to the specific and individual opportunity costs that each farmer has considered to decide to introduce a given crop to the production plan. We can consider this category of cost as “pure economic costs” because it depends on profit maximisation logic (expressed by the observed price) and on the characteristics of the production function (expressed by the observed yields).

Researchers must be aware that the *observed variable accounting cost* registered by FADN can theoretically contain errors for several reasons, including data collection and imputation. In particular, farmers may wrongly specify some costs related to a production technique. One example is irrigation costs, which are difficult for farmers to properly record because they are often not explicit. For these reasons comparing the estimated accounting cost with the observed accounting cost can fail when some costs are not explicit even for farmers.

The estimation of specific variable costs is conducted on a sub-sample of FADN farms belonging to Veneto, Lombardy and Piedmont selected using a cluster analysis. The cluster analysis has been developed using the *K*-mean method, the best-known and most widely applied partitioning method (for a review, see Atkinson et al., 2004). This procedure classifies the *n* units into *k* distinct clusters, with *k* chosen *a priori* by the analyst, according to a step-by-step iterative method that reaches the optimal distribution of observations into the defined groups.

The analysis of Table 2 highlights that some crops are missing, such as durum wheat and rice. These two crops are grown on farms that are not present in the largest cluster. The degree of homogeneity is thus reliant on the level of farm specialisation such that farms specialised in rice production with technologies that are quite different from those of the other farms are not captured by the largest group. The same issue arises for tomato production, which is also missing from that cluster.

The comparison of the observed and estimated accounting costs for common wheat and barley exhibits an excellent fit with a deviation of 8.6% from the observed accounting costs, which is also a sign of the high uniformity in the technology used to cultivate these two crops. All the estimates obtained for common wheat and barley have given results close to the observed reality. For maize, the deviation is quite restrained, +16% on the observed information; for soya, the variation is approximately 11%, while for sugarbeet and alfalfa, the results are more satisfactory, with deviations of 8% and 0.6%, respectively. Only sunflower exhibits a high difference from the observed value, equal to -41% (see Figures 4 and 5).

Table 2. Specific cost estimates obtained from the PMP model, Veneto-Lombardy-Piedmont.

Crop	n. of obs.	Observed variable accounting cost	Std. Error	Estimated Accounting Cost	Std. Error	Hidden cost	Std. Error	Total Marginal Cost	Std. Error
		€/t.		€/t.		€/t.		€/t.	
Soft wheat	197	0.07113	0.00231	0.06501	0.00383	0.04752	0.00718	0.11254	0.00511
Maize	311	0.06106	0.00133	0.07100	0.00175	0.04258	0.00267	0.11357	0.00230
Barley	62	0.06208	0.00504	0.05673	0.00699	0.01929	0.00367	0.07602	0.00612
Protein crops	11	0.09320	0.01502	0.08331	0.00980	0.01383	0.00897	0.09714	0.00948
Soya	74	0.11812	0.00659	0.10452	0.00804	0.02489	0.00608	0.12941	0.00784
Sugarbeet	17	0.01452	0.00079	0.01569	0.00207	0.00179	0.00043	0.01747	0.00065
Rape	6	0.11845	0.02814	0.11202	0.03581	0.09389	0.03235	0.20591	0.03229
Sunflower	8	0.12248	0.02720	0.07227	0.03161	0.00601	0.00245	0.07828	0.02434
Temporary grass	5	0.03504	0.00623	0.03290	0.01416	0.00040	0.00025	0.03331	0.00323
Alfalfa	6	0.01432	0.00296	0.01423	0.00168	0.01149	0.00760	0.02572	0.00750
Pasture	1	0.00571	0.00000	0.00000	0.00000	0.00104	0.00022	0.00104	0.00022
Meadows	77	0.01404	0.00113	0.02553	0.00407	0.01834	0.00230	0.04387	0.00331

Source: own elaboration

Figure 4: Total marginal cost distribution, Veneto-Lombardy-Piedmont (€/t)

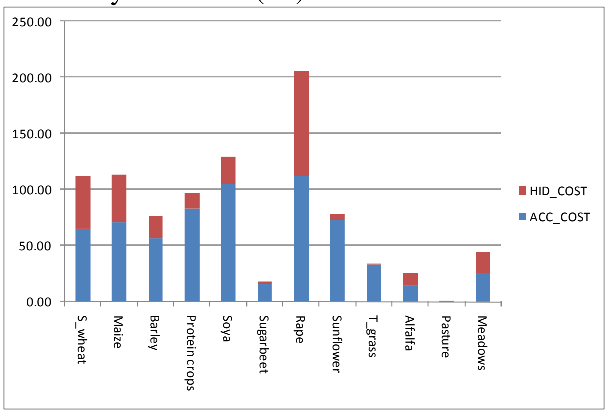
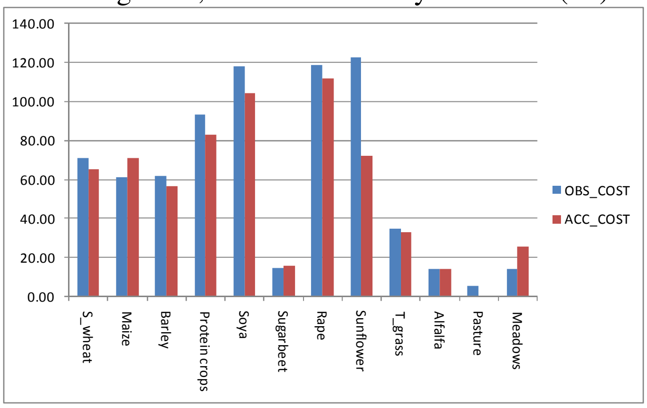


Figure 5: Comparison between observed and estimated accounting costs, Veneto-Lombardy-Piedmont (€/t)



Source: own elaboration

The *t*-test provides significance values for these estimates. Estimated accounting costs for some crops (such as sugarbeet) with poor *t*-test significance present a good fitness with respect to the observed accounting cost. For other crops, such as maize, the estimates remain in an unacceptable range of significance. Common wheat shows a significance level of 30%. Other tests conducted in preparation of this work and not reported here for reasons of brevity include the finding that the significance level for common wheat increases to 90% when the basic information is extended to the entire sample of farms across the three regions. This finding shows that the specific internal characteristics of the sample can affect the estimation results and that the target crops used for the development of the estimation control the characterisation of the sample.

The Generalised PMP model provides a good estimation of the variable accounting costs for some crops, such as sugarbeet, barley, oilseeds and common wheat, while the statistical tests for some other crops, such as maize, remain unsatisfactory due to a lack of information in some specific cost component, e.g., irrigation costs.

Table 3. T-test for estimated and observed accounting costs, Veneto-Lombardy-Piedmont.

Crops	Paired Differences					t	Sig. (2-tailed)
				95% Confidence Interval of the Difference			
	Mean	Std. Deviation	Std. Error Mean	Lower	Upper		
S_wheat	-0.0044612	0.0517577	0.0042689	-0.0128980	0.0039756	-1.045	0.298
Maize	0.0099957	0.0233423	0.0013499	0.0073391	0.0126522	7.405	0.000
Barley	-0.0058595	0.0642430	0.0099129	-0.0258790	0.0141600	-0.591	0.558
Protein crops	-0.0185571	0.0733837	0.0277364	-0.0864257	0.0493115	-0.669	0.528
Soya	-0.0104111	0.0820993	0.0103435	-0.0310876	0.0102653	-1.007	0.318
Sugarbeet	0.0012933	0.0082950	0.0021417	-0.0033003	0.0058869	0.604	0.556
Rape	-0.0375500	0.0942819	0.0471409	-0.1875735	0.1124735	-0.797	0.484
Sunflower	-0.0508286	0.1304534	0.0493068	-0.1714779	0.0698207	-1.031	0.342
T_grass	-0.0061500	0.0577706	0.0408500	-0.5251985	0.5128985	-0.151	0.905
Alfalfa	-0.0048000	0.0008485	0.0006000	-0.0124237	0.0028237	-8.000	0.079
Meadows	0.0111474	0.0277703	0.0045049	0.0020195	0.0202752	2.474	0.018

Source: own elaboration

The discussion of the results above has been limited to the estimated accounting cost at the sample level, without considering the deeper level of information. Nevertheless, the General PMP model is a micro-based model that uses farm information at the individual level to enable the achieved results to be inferred at the individual level. Thus, the results can be aggregated in different ways according to the research objectives. The model can provide the estimated variable accounting cost for each crop from the farm level to a more aggregated level, such as a specific territorial area or farm characteristics including altitude, economic size, or physical size. Table 4 shows the results of variable accounting costs for common wheat achieved for the three Italian regions aggregated according to physical size.

Table 4. Estimated and observed variable accounting costs for soft wheat according to size (€/t).

Region	Class of size (ha)	Variable acounting Cost			N. of obs.
		Estimated	Observed	Var. %	
Veneto	<10	71.44	72.63	*-1.6*	23
	10-20	62.53	77.13	*-18.9*	14
	20-50	67.81	71.55	*-5.2*	36
	50-100	82.01	67.48	*21.5*	20
	100-200	68.04	63.85	*6.6*	13
	>200	48.67	53.95	*-9.8*	4
	Total	68.53	70.20	*-2.4*	110
Lombardy	<10	28.94	31.17	*-7.2*	6
	10-20	45.55	61.71	*-26.2*	10
	20-50	79.03	48.34	*63.5*	15
	50-100	35.58	45.89	*-22.5*	8
	100-200	48.35	42.30	*14.3*	4
	>200	157.30	54.30	*189.7*	3
	Total	55.46	48.45	*14.5*	46

Source: own elaboration

Table 4 (continue). Estimated and observed variable accounting costs for soft wheat according to size (€/t).

Region	Class of size (ha)	Variable acounting Cost			N. of obs.
		Estimated	Observed	Var. %	
Piedmont	<10	68.04	63.07	*7.9*	50
	10-20	67.85	80.31	*-15.5*	39
	20-50	69.62	80.08	*-13.1*	50
	50-100	74.81	78.25	*-4.4*	22
	100-200	81.11	81.38	*-0.3*	11
	>200	207.00	101.72	*103.5*	6
	Total	72.30	75.94	*-4.8*	178
Veneto-Lombardy-Piedmont	<10	65.55	62.81	*4.4*	80
	10-20	58.30	76.84	*-24.1*	65
	20-50	63.18	71.72	*-11.9*	99
	50-100	64.26	69.50	*-7.5*	49
	100-200	74.45	68.07	*9.4*	29
	>200	146.76	77.20	*90.1*	13
	Total	66.02	70.16	*-5.9*	335

Source: own elaboration

The estimation variability increases when moving from an aggregated to a less aggregated result. In particular, observing the results for the three regions considered as a whole, it is evident that the stratification leads to an amplification of the estimation errors for some size classes. For instance, the class containing large farms (>200 ha) exhibits a high divergence between the estimated and observed accounting costs, while most of the other classes show differences that are greater than the average value calculated for the entire sample. This estimation behaviour is repeated separately for the three regions; the worst results generally correspond to the size classes with few observations, indicating that the estimation process tends to estimate cost based on the average information.

5. CONCLUSIONS

PMP is primarily used to predict farm behaviour at the micro and regional levels. For these reasons, many researchers consider PMP as "only" a method to calibrate mathematical models. This work demonstrates that the Generalised PMP model developed based on the theoretical foundation of PMP and on mathematical programming and production cost theory is a useful theoretical framework for the representation of farm choice, including for the description of costs related to the production function chosen by each entrepreneur. In this context the Generalised PMP model provides good estimates of the variable accounting costs, especially when the observations are numerous and homogeneous. The possibility to estimate variable accounting costs becomes very important if differences in production costs between farm type and region are important criteria for the definition of farm strategies and specific public policies aimed at supporting the agricultural sector. At the same time, this possibility is also relevant when researchers develop regional or sector models according to the specific characteristics of the region or according to the level of farm specialisation, where the possibility of collecting homogeneous external information is rather difficult.

The main feature of the generalised PMP model is its independence from "external" information and the fact that researchers do not choose the support value when the GME is adopted to maximise the gross margin. It is true that the model is sensitive to the quality of the data and that outliers can negatively influence the quality of the estimations. However, this feature is rather common in all econometric models.

In this respect, this work has proposed and tested a methodology for improving the estimation and validating the estimated variable cost. This objective was achieved by adopting a cluster analysis technique applied to the basic data. This procedure helped to control the outliers and improves the estimates. Moreover, the results were validated by adopting the Student's *t*-test; these results were compared with the observed accounting costs collected by FADN in three Italian regions.

The Generalised PMP model provides robust estimations in different contexts and is sufficiently flexible to allow for adaption to represent and simulate different production patterns and different policy and market scenarios.

AKNOWLEDGMENTS

The research activity on which this paper is based benefited from the financial support received from the European Commission (Research project FACEPA — Farm Accountancy Cost Estimation and Policy Analysis of European Agriculture, FP7/2007-2013 contract No. 212292).

REFERENCES

Arfini F. and Donati M. (2008), Health Check ed efficienza delle aziende agricole: una valutazione comparativa su quattro regioni agricole europee. *Rivista di Economia Agraria,* vol. N.1; p. 65-92

Arfini F. and Donati M. (2010), Impact of the Health Check on Structural Change and Farm Efficiency: a Comparative Assessment of three EU Agricultural Regions, Proceedings of the OECD Workshop on the Disaggregated Impacts of CAP Reform, Paris, 10-11 March 2010.

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

Arriaza, M. and Gómez-Limón, J.A. (2003). Comparative performance of selected mathematical programming models. Agricultural Systems, 77(2): 155-171.

Blanco M., Cortignani R., Severini, S. (2008). Evaluating Changes in Cropping Patterns due to the 2003 CAP Reform. An Ex-post Analysis of Different PMP Approaches Considering New Activities. Paper prepared for presentation at the 107th EAAE Seminar „Modeling of Agricultural and Rural Development Policies". Sevilla, 2008

Britz, W. and Adenäuer, M. (2009). Calibration of Regional Programming Models to Exogenous Supply Elasticities by a Variable Cost Function: The CAPRI-EXPAMOD Example. In Integrated Assessment of Agriculture and Sustainable Development (AgSAP), March 10-12, 2009, Egmont an Zee (The Netherlands), S. 52-53.

Butault, J.P., Chantry, E., Pollet, P. (2001), The agricultural production cost model, ESTAT-2001-03195-00-00-EN-TRA-00 (FR)

Butault J.P., Zardet G., Loïc M., Delame N., Desbois D., Rousselle J.M., Kleinhanss W., Offermann F., (2011), The FACEPA Model Software, Estimating Costs of Production using the EU FADN database, User guide, FACEPA Deliverable No. D4.2 - September 2011

Buysse J., Van Huylenbroeck G., Lauwers L. (2007a). *Normative, positive and econometric programming as tools for incorporation of multifunctionality in agricultural policy modelling*. Agricultural, Ecosystems and Environment 120 (2007) 70-81.

Buysse, J., Fernagut, B., Harmignie, O., Henry de Frahan, B., Lauwers, L., Polomé, P., Van Huylenbroeck, G., and Van Meensel J. (2007b). Farm-based modelling of the EU sugar reform: impact on Belgian sugar beet suppliers. European Review of Agricultural Economics, 34: 21 - 52.

Cafiero, C. (2004) "Ci si può fidare dei modelli matematici nelle analisi di politica agraria?" Politica Agricola Internazionale. 2:1-18.

Eurostat (2009), http://epp.eurostat.ec.europa.eu/portal/page/portal/region_cities/regional_ statistics/data/database .

Gocht A. (2005). Assessment of simulation behaviour of different Mathematical Programming approaches. In Arfini, F. (Ed.): *Modelling agricultural policies: state of the art and new challenges*. Monte Università Parma Editore, Parma, Italy, pp. 166-187.

Gocht, A. and Britz, W. (2011). EU-wide farm type supply models in CAPRI - How to consistently disaggregate sector models into farm type models. *Journal of Policy Modeling* 33(1): 146-167.

Golan A., Judge, G.G. Miller, D., (1996). Maximum Entropy econometrics: Robust estimation with limited data. John Wiley & Sons, pp 324.

Heckelei, T. and Britz, W. (2000). Positive mathematical programming with multiple data points: a cross-sectional estimation procedure. Cahiers d'Economie et Sociologie Rurales, 57: 28-50.

Heckelei, T. (2002). Calibration and Estimation of Programming Models for Agricultural Supply Analysis. Habilitation Thesis, University of Bonn, Germany (http://www.ilr1.uni- bonn.de/agpo/ staff/heckelei/heckelei_hab.pdf).

Heckelei, T. (2003). Positive Mathematical Programming: Review of the standard approach. CAPRI working paper (http://www.ilr1.unibonn.de/agpo/rsrch/capri/wrkpape.htm).

Heckelei T., and Wolff H. (2003): "Estimation of Constrained Optimisation Models for Agricultural Supply Analysis Based on Generalised Maximum Entropy", European Review of Agricultural Economics, 30 (1): 27-50.

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

Helming, J. F. M. (2005). A Model of Dutch Agriculture Based on Positive Mathematical Programming with Regional and Environmental Applications. Ph.D. thesis, Wageningen University, The Netherlands.

Helming, J.F.M., Peeters, L., Veendendaal P.J.J. (2001). Assessing the consequences of environmental policy scenarios in Flemish agriculture. In: Heckelei, T., Witzke, H.P., and Henrichsmeyer, W. (Ebs): Agricultural sector Modelling and policy information systems. of the 65th 598 EAAE seminar, March 29-31, 2000 at Bonn University, Vauk Verlag Kiel,pp. 237-245

Henry de Frahan, B., Buysse, J., Polomé, P., Fernagut, B., Harmignie, O., Lauwers, L. Van Huylenbroeck, G. and Van Meensel, J. (2006). Positive mathematical programming for agriculture and environmental policy analysis: review and practice. In: Weintraub, A., Bjorndal T., Epstein, R. and Romero, C. (Eds.). Management of Natural Resources: A Handbook of Operations Research Models, Algorithms, and Implementations. Kluwer Academic Publishers, Kluwer´s International Series in Operations Research and Management Science (Series Editor: Hillier, F.S.).

Howitt R.E. (1995). Positive Mathematical Programming. 590 American Journal of Agricultural Economics, 77: 329-342.

Judez, L., Chaya, C., Martinez, S. and Gonsalez, A.A. (2001). Effects of the measures envisaged in 'Agenda 2000' on arable crop producers and beef and veal producers: an application of Positive Mathematical Programming to representative farms of a Spanish region. Agricultural Systems, 67: 121-138.

Júdez, L., Andres, R., Ibanez M., de Miguel, J. M. and Miguel, J. L. (2008). Impact of the CAP Reform on the Spanish Agricultural Sector. Paper prepared for the 109th EAAE Seminar " The CAP after the Fischler Reform: National Implementations, Impact Assessment and the Agenda for Future Reforms". Viterbo, Italy, November 20-21st, 2008.

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

Kanellopoulos, A., Berentsen, P., Heckelei, T., Van Ittersum, M. and Oude Lansink, A. (2010). Assessing the forecasting performance of a generic bio-economic farm model calibrated with two different PMP variants. *Journal of Agricultural Economics* 61(2): 274-294.

Paris Q (2011), Economic Foundation of Symmetric programming, Cambridge University Press, New York, pp. 1- 550

Paris Q. and Arfini F., (2000). Funzioni di Costo di Frontiera, Auto-selezione, Rischio di Prezzo, Pmp e Agenda 2000. Rivista di Economia Agraria. vol. n. 3, LV, n. 2, Giugno, pp. 211-242.

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.

Röhm, O. and Dabbert, S. (2003). Integrating Agri-Environmental Programs into Regional Production Models: An extension of positive mathematical programming. American Journal of Agricultural Economics, 85(1): 254-65.

Severini, S. and Cortignani R. (2011). Modeling farmer participation to a revenue insurance scheme by means of Positive Mathematical Programming. Paper prepared for presentation at the EAAE 2011 Congress. http://purl.umn.edu/116001.