Integrating risk and uncertainty in PMP models

Athanasios Petsakos
Agricultural University of Athens, Department of Agricultural Economics & Rural Development, Iera Odos 75 11855, Athens, Greece.
e-mail: petsakos@aua.gr

Stelios Rozakis
Agricultural University of Athens, Department of Agricultural Economics & Rural Development, Iera Odos 75 11855, Athens, Greece.
e-mail: rozakis@aua.gr

Integrating risk and uncertainty in PMP models

Abstract

Positive Mathematical Programming (PMP) is one of the most commonly used methods of calibrating activity linear programming (LP) models in agriculture. PMP applications published thus far focus on the estimation of a farm's nonlinear cost or profit function and rely on the recovery of unobserved or implicit information that can explain the initial model's inability to calibrate. In this paper we use the PMP procedure to calibrate an expected utility model under the assumption that this implicit information can reveal a farmer's profit expectations and risk attitude. The perfect calibration shows that PMP can be applied not only to LP models, but also to models that incorporate risk and this provides an interesting alternative to the traditional PMP methodology.

Keywords: E-V analysis; expected utility; farm model; Positive Mathematical Programming; risk.

1 Introduction

Positive Mathematical Programming (PMP) was originally introduced by Howitt (1995a) as an efficient method for calibrating LP models and is now considered one of the mainstream methodologies employed for building farm activity models. The PMP methodology is based on the assumption that the economic agent's observed behaviour is the optimal one, because production choices rely not only on the observed parameter set appearing in the objective function, but also on additional implicit information that cannot be observed when examining farm data. This information is revealed by initially solving an LP model that exactly reproduces base year activity levels via a set of calibration constraints and is used to specify a nonlinear, alternative objective function that allows for a model that calibrates.

PMP models published thus far follow the standard procedure of substituting for a nonlinear term either the cost (e.g. Petsakos and Rozakis, 2009; Heckelei and Wolff, 2003; Heckelei and Britz, 2000) or the production/revenue part (e.g. Júdez et al., 2002; Howitt, 1995a; Howitt, 1995b) of the objective function in order to recover the "true" objective function of an optimization problem. However, as far as we know, no paper has examined how calibration can be achieved by taking into account farmer's risk considerations because of uncertainty on prices and yields, since all publications use PMP inside a deterministic context. It is very important, especially these days when price volatility joined with decoupling policies increase systematic risk for farmers, to attempt integration of risk approaches and calibrating state-of-the-art methods, such as PMP.

In this paper we show how the rationale of PMP can be extended to cases where some of the model's parameters are stochastic and the farmer is risk averse. Our approach is drawn

1

from the mean-variance (E-V) analysis. In this context, we consider a misspecified variance-covariance matrix to be responsible for the model's inability to calibrate and thus we use the PMP approach to estimate the "true" matrix so that the final model exactly reproduces base year observations.

After a short presentation of the PMP methodology in the next section, we review the techniques of incorporating risk and uncertainty into programming models in section 3. A PMP method based on maximizing expected utility instead of profit is then presented in section 4, where we look into the first order conditions of the underlying optimization problem and provide the correspondence with traditional PMP approach. A simple farm-level application is then used in section 5 to illustrate the proposed approach.

2   Calibrating with PMP

The PMP algorithm usually involves three phases during which the linear objective function is gradually transformed into a nonlinear one. The reason for this is that due to the linearity of the objective function, the first order optimality conditions do not depend on the final values of the unknown variables, which means that the model could fail to exactly calibrate at the observed activity levels. In this case, Howitt (1995a) proves that a necessary and sufficient condition for calibration is that the objective function be nonlinear in at least some of the activities. The nonlinearity is usually sought in the cost term and is introduced in the model by replacing the linear cost function with a quadratic one, defined as:

$$VC(\mathbf{x}) = \mathbf{d}'\mathbf{x} + \frac{1}{2}\mathbf{x}'\mathbf{Q}\mathbf{x}$$

where $\mathbf{x}$ is the $I \times 1$ unknown vector of activity levels, $\mathbf{d}$ is an $I \times 1$ vector of linear terms and $\mathbf{Q}$ is an $I \times I$, positive, semi-definite matrix that is either diagonal or fully specified.

As mentioned above, the first phase of PMP involves using a simple LP model with additional calibration constraints that bind all activities at the observed level, $\hat{\mathbf{x}}$, thus forcing the model to exactly reproduce base year observations. This can be written in vector form as:

$$\begin{aligned}
\underset{\mathbf{x} \geq 0}{\text{Max}} \quad & Z = \mathbf{r}'\mathbf{x} - \mathbf{c}'\mathbf{x} \\
\text{s.t.} \quad & \mathbf{A}\mathbf{x} \leq \mathbf{b} && [\theta] \\
& \mathbf{x} \leq \hat{\mathbf{x}} + \varepsilon && [\lambda]
\end{aligned}$$

Symbol Z denotes farm's gross margin to be maximized, $\mathbf{r}$ is the $I \times 1$ vector of net profits and $\mathbf{c}$ the $I \times 1$ vector of average costs for the I activities respectively. $\mathbf{A}$ is the $J \times I$ matrix of technical coefficients, while resource availability is given by the vector $\mathbf{b}$ and the corresponding dual values by vector θ (both of $J \times 1$ dimension). The second set of constraints represents the additional calibration constraints that bound each activity to its observed level, λ is the corresponding $I \times 1$ dual vector and ε is a small perturbation term used to prevent model degeneration caused by linear dependency among calibration and resource constraints. It is argued that λ embodies any type of marginal implicit information, such as model misspecifications, data errors, price expectations and farmer's risk attitude (Heckelei, 2002). In this context, Paris and Howitt (1998) interpret λ as a "differential" marginal cost vector that, together with the observed "accounting" variable average cost ($\mathbf{c}$), reveals the actual variable marginal cost of

the activities at the observed $\hat{\mathbf{x}}$ production level. Hence, the derivative of the quadratic cost function at $\hat{\mathbf{x}}$ should be equal to $\boldsymbol{\lambda} + \mathbf{c}$:

$$\mathbf{d} + \mathbf{Q}\hat{\mathbf{x}} = \boldsymbol{\lambda} + \mathbf{c} \tag{1}$$

The estimation of the quadratic cost function is based on equation (1) that constitutes an underdetermined system with I equations and $2I$ or $I + I(I+1)/2$ unknown parameters, depending on the form of the $\mathbf{Q}$ matrix; If $\mathbf{Q}$ is diagonal, several ad hoc methods have been proposed, summarized by Petsakos and Rozakis (2009), while for a fully specified matrix, Paris and Howitt (1998) propose the use of the maximum entropy criterion. This involves finding a discrete probability distribution over a vector of support values in order to maximize the entropy of the system and consequently the uncertainty on the value of the unknown parameters. In either case, the final optimization in the third phase of PMP is specified as:

$$\underset{\mathbf{x} \geq 0}{\text{Max}} \quad Z = \mathbf{r}'\mathbf{x} - \mathbf{d}'\mathbf{x} - \frac{1}{2}\mathbf{x}'\mathbf{Q}\mathbf{x}$$
$$\text{s.t.} \quad \mathbf{A}\mathbf{x} \leq \mathbf{b}$$

The final model with the estimated nonlinear objective function is now able to exactly reproduce the observed activity levels.

3   Uncertainty and risk in MP models

The most common method for introducing risk and uncertainty into programming models is the E-V approach, which is based on Markowitz's (1959) pioneering work on portfolio theory. The latter has been used extensively in farm management (Hardaker et al., 2004) in order to define the optimal production plan, for example, the allocation of land to arable crops (Brink and McCarl, 1978) or to aid the design of feeding rations in animal nutrition (Torres-Rojo, 2001). The original formulation of the E-V problem that first appeared in Markowitz (1952), dictates that a farmer selects an activity mix that minimizes the variance of revenue, V, for a given expected revenue E:

$$\underset{\mathbf{x} \geq 0}{\text{Min}} \quad V = \mathbf{x}'\boldsymbol{\Sigma}\mathbf{x}$$
$$\text{s.t.} \quad \mathbf{g}'\mathbf{x} = \xi,$$
$$\mathbf{A}\mathbf{x} \leq \mathbf{b}$$

where $\boldsymbol{\Sigma}$ is the symmetric $I \times I$ variance-covariance matrix of average activity profits, $\mathbf{g}$ and $\mathbf{x}$ are $I \times 1$ vectors of average profits and activity levels respectively and ξ is the level of expected revenue. The model is then solved parametrically for different levels of ξ, in order to form the E-V efficient frontier which is the locus of points that represent Pareto optimal choices in the E-V space.

Mean-standard deviation (E-σ) is a widely used derivative of the E-V analysis that uses the standard deviation of random profits instead of their variance, but still yields the same efficient frontier. An E-σ model maximizes an objective function $L = \mathrm{E} - \varphi\sigma$ that can be interpreted as a farmer's expected utility function, where σ is the standard deviation and φ a measure of risk aversion. One interesting feature with the E-σ approach is that if profits are normally

distributed, φ helps to identify the $a\%$ income fractile, i.e. the value of income which, for a given distribution will be exceeded the $(1 - a)\%$ of the time.

An alternative E-V model that is based on maximization of expected utility ($E[U]$) on a mathematical programming framework was presented by Freund (1956). This approach relies on the assumption that the profits from each activity g_i (where $i \in I$) are normally distributed with mean $\bar{g}_i$ and variance σ_i^2 and that the farmer's revenue utility function has the following exponential form:

$$U(R) = 1 - e^{-\varphi R}$$

Symbol φ represents the farmer's constant absolute risk aversion and R his total revenue that is also normally distributed with mean $\mu = \bar{\mathbf{g}}'\mathbf{x}$ and variance $\sigma^2 = \mathbf{x}'\Sigma\mathbf{x}$. Expected utility, can then be calculated by the integral:

$$E[U(R)] = \int_{-\infty}^{\infty} -e^{-\varphi R} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(R-\mu)^2/2\sigma^2} dR$$

It can be shown that the maximization of expected utility is equivalent to maximizing the function $T = \mu - (\varphi/2)\sigma^2$. The final model is then specified as:

$$\max_{\mathbf{x} \geq 0} \quad T = \mathbf{g}'\mathbf{x} - \frac{1}{2}\varphi\mathbf{x}'\Sigma\mathbf{x}$$
$$\text{s.t.} \quad \mathbf{Ax} \leq \mathbf{b}$$

The use of a utility function that includes E and V terms has caused a long debate in the literature as to if and under which conditions the E-V model can give solutions that are consistent to expected utility maximization (e.g. Tsiang 1972; Borch, 1974; Levy, 1974). Initially it was argued that an E-V model maximizes expected utility when either the underlying distributions are normal or the utility function is quadratic. These assumptions are obviously problematic since a quadratic utility function has several undesired properties (Pratt, 1964), whereas a normal distribution may not always be the case for random events in agriculture. Yet, the lack of extended data on profits over long time periods, which is usually the case for agricultural activities, does not allow for the rejection of the normality hypothesis (Hazell and Norton, 1986).

It is now accepted that E-V analysis and expected utility give equivalent solutions when the utility function used is of exponential form and the corresponding distributions are normal (like the model of Freund described previously), or when the underlying distributions satisfy Meyer's (1987) location and scale condition. This condition states that an E-V model can in fact maximize expected utility when the model's different activity variables are distributed differently only with respect to location and scale. Said differently, whenever the alternative production activities (variables) can be written as a positive linear function of a random parameter, then the distributions of all possible activity variables differ from one another only by location and scale parameters and the resulting farmer's E-V efficient frontier represents choices that maximize expected utility.

Several authors (e.g. Levy and Markowitz, 1979; Pulley, 1983) have shown that an E-V model derived from expected utility approximations of an individual's stochastic income, R, using second order Taylor series expansion around its mean, μ, can yield preference orderings that are closely correlated to the ones given by expected utility maximization:

$$U(R) = U(\mu) + U'(\mu)(R - \mu) + \frac{U''(\mu)}{2!}(R - \mu)^2 + \ldots + \frac{U^{(n)}(\mu)}{n!}(R - \mu)^n \qquad (2)$$

4

By omitting central moments higher than 2, equation (2) can be written as

$$U(R) = U(\mu) + U'(\mu)(R - \mu) + \frac{1}{2}U''(\mu)(R - \mu)^2$$

The expected utility function is then approximately equal to

$$E[U(R)] \equiv U(\mu) + \frac{1}{2}U''(\mu)V \tag{3}$$

A desirable property of (3) pointed out by both Tsiang (1972) and Pulley (1983) is that the approximation of expected utility improves as the ratio of non-invested capital increases. In agricultural economics, this means that for given production choices, the approximation becomes more accurate when the chosen production plan includes activities with non stochastic profits, i.e. decoupled subsidies. One important remark drawn from all papers that treated the issue of approximating expected utility with an E-V formulation is that (3) holds for both empirical and theoretical distributions and thus it is not subject to the assumption presented in Freund's model that revenues are normally distributed.

4   E-V models in a PMP framework

Calibration of programming models under risk is usually performed indirectly during the elicitation of the risk aversion coefficient, φ, used in model specifications as the ones described in the previous section. More precisely, imputing values of φ by solving farm models for their efficient set of plans, finally allows for the selection of that value of φ that gives the closest fit between the actual and predicted farm plans (Hazell and Norton, 1986). This, however, does not always lead to exact calibration because the first order conditions of the maximization problem are not taken into account.

A more consistent approach for calibrating is described in the following paragraphs. It uses the standard PMP three-step procedure and takes advantage of the dual vector $\boldsymbol{\lambda}$ that is now interpreted as implicit information on price expectations and farmer's risk attitude.

We begin with the assumption that a farmer maximizes a logarithmic utility function. The logarithmic is one of the utility functions that gave approximations consistent to expected utility maximization in Levy and Markowitz (1979) and Pulley (1983). It has received much credit in the discipline of financial economics as an appropriate function to represent investors' preferences (Rubinstein, 1976) and is quoted by Hardacker et al. (2004, pp. 109) as "*everyone's utility function*". Additionally, from a theoretical viewpoint, it has all the necessary properties to qualify as an appropriate utility function for a risk-averse individual, since it is concave (shows a decreasing marginal income utility), which implies a decreasing absolute risk aversion. Compared to the exponential utility function proposed by Freund, the logarithmic specification has the advantage of not requiring the elicitation of an individual's absolute risk aversion coefficient, as this is a decreasing function of income and not just a constant parameter. The farmer's maximization problem of income utility under the usual linear constraints can thus be written as:

$$\text{Max} \quad U(\mathbf{x}) = \ln(w + \mathbf{g}'\mathbf{x})$$
$$\text{s.t.} \quad \mathbf{Ax} \le \mathbf{b}$$
$$\mathbf{x} \ge 0$$

where w is the farmer's initial wealth, or within the CAP setting, subsidies decoupled from production, and $\mathbf{g}$ is the vector of activity profits, defined as the element-wise product of the vectors for prices ($\mathbf{p}$) and yields ($\mathbf{y}$) minus the vector of variable costs ($\mathbf{c}$).

Since the logarithmic utility function is just a monotonic transformation of R, the above maximization problem is equivalent to just maximizing R under the same constraints, which leads us back to the classic farm LP problem. Hence, if the underlying LP model fails to calibrate, so will its utility maximizing counterpart. According to the theory of PMP, this is due to unknown implicit information (expressed by the dual vector λ) that may capture unobserved farmer's risk preferences and attitude towards uncertain prospects. In this context, if $\mathbf{p}$ and $\mathbf{y}$ (and consequently $\mathbf{g}$ as well) are treated as stochastic variables, for which an empirical distribution over T years is known, it can be said that the previous model does not calibrate because the farmer is not maximizing his income utility but his expected income utility.

Ideally, an expected utility formulation should include a variance term that would allow for the representation of the uncertainty on the evolution of prices and yields. As discussed previously, a simple way of doing so is to approximate the utility function using second order Taylor series around the mean (or the expected income) of the empirical distribution, which finally yields an approximation of expected utility, given in equation (3). By inserting the logarithmic utility function in (3) and setting $V = \mathbf{x}'\Sigma\mathbf{x}$, where Σ is the symmetric variance-covariance matrix of activity revenues, the expected utility maximization problem can be written as:

$$\text{Max} \quad E[U(\mathbf{x})] = \ln\left(w + \bar{\mathbf{g}}'\mathbf{x}\right) - \frac{1}{2}\left(\frac{\|\bar{\mathbf{g}}\|}{w + \bar{\mathbf{g}}'\mathbf{x}}\right)^2 \mathbf{x}'\Sigma\mathbf{x}$$

$$\text{s.t.} \quad \mathbf{Ax} \leq \mathbf{b} \quad [\theta]$$
$$\mathbf{x} \geq 0$$

The first order conditions define that

$$\frac{\bar{\mathbf{g}}}{w + \bar{\mathbf{g}}'\mathbf{x}} - \|\bar{\mathbf{g}}\|^2 \left[\frac{\Sigma\mathbf{x}}{(w + \bar{\mathbf{g}}'\mathbf{x})^2} - \frac{\bar{\mathbf{g}}\mathbf{x}'\Sigma\mathbf{x}}{(w + \bar{\mathbf{g}}'\mathbf{x})^3}\right] - \mathbf{A}'\theta = 0 \tag{4}$$

Among the various parameters in the previous equation, the one that is most prone to misspecification — and may therefore lead to calibration failure — is the variance-covariance matrix Σ. More precisely, the elements of Σ are calculated using empirical distributions for prices and yields over a period of T years. The problem with this approach is that such time series data is usually available only at the national level and are average values, which makes them unsuitable for farm models. This means, the variability of profit variances among different farms can be substantial, since the different microclimatic conditions and farming techniques can affect both yield and product quality, according to which prices are formulated. This means that the matrix Σ that the analyst has used to build his model may be different than the farm's true variance-covariance matrix, presumably known to the farmer (which we will denote by $\mathbf{S}$) that needs to be estimated so that equation (4) is satisfied at $\hat{\mathbf{x}}$.

Although we focus on a farm-level model, it should be noted that a misspecified Σ matrix may cause calibration problems even in sector models. As Hazell and Norton (1986) explain, sector models under stochastic yields and prices are more complex to specify than their farm-level counterparts, since market equilibrium requires a negative covariance relation between

prices and yields. Furthermore, the variance-covariance matrix should be an "appropriate aggregate" of the one perceived by farmers and for this reason the utilized time series should be de-trended in order to eliminate the underlying trends of prices and yields. This may result in a matrix that is different from the Σ that was calculated from national statistics.

For the estimation of $\mathbf{S}$ we follow a three step procedure, similar to that in classical PMP applications described in section 2. In our case, we use an initial expected utility model with additional constraints that bind activities at their observed level and a Σ matrix derived from national statistics. One point that needs to be taken into account is the direction of the inequality constraint. More specifically, the negative sign in front of the variance term renders expected utility a decreasing function of $\mathbf{x}$, which means that the calibration constraint must be written as $\mathbf{x} \geq \hat{\mathbf{x}} - \varepsilon$ (instead of $\mathbf{x} \leq \hat{\mathbf{x}} + \varepsilon$) in order to reproduce base year observations. This will yield negative λ values and possibly negative expected utility values as well. However, this doesn't pose any theoretical problems since expected utility deals with preference orderings and hence its absolute value under a specific production plan has a meaning only when compared to the corresponding value of another production plan. The expected utility maximization problem can thus be written as:

$$\text{Max} \quad E[U(\mathbf{x})] = \ln\left(w + \bar{\mathbf{g}}'\mathbf{x}\right) - \frac{1}{2}\left(\frac{\|\bar{\mathbf{g}}\|}{w + \bar{\mathbf{g}}'\mathbf{x}}\right)^2 \mathbf{x}'\Sigma\mathbf{x}$$

$$\text{s.t.} \quad \mathbf{Ax} \leq \mathbf{b} \qquad [\theta]$$
$$\mathbf{x} \geq \hat{\mathbf{x}} - \varepsilon \qquad [\lambda]$$
$$\mathbf{x} \geq 0$$

where λ is negative. The first order conditions for an optimum can be written as:

$$\frac{\bar{\mathbf{g}}}{w + \bar{\mathbf{g}}'\mathbf{x}} - \|\bar{\mathbf{g}}\|^2 \left[\frac{\Sigma\mathbf{x}}{(w + \bar{\mathbf{g}}'\mathbf{x})^2} - \frac{\bar{\mathbf{g}}\mathbf{x}'\Sigma\mathbf{x}}{(w + \bar{\mathbf{g}}'\mathbf{x})^3}\right] - \mathbf{A}'\theta - \lambda = 0$$

Assuming that the farmer's expectations on activity profits coincide with the actual profits achieved in the base year, the first order conditions of an expected utility model that uses an estimated $\mathbf{S}$ matrix and calibrates perfectly, will be satisfied at $\hat{\mathbf{x}}$, and the same will apply for the bounded model that uses an Σ matrix, formed by national averages of prices and yields. Using a PMP rationale, this means that

$$\|\bar{\mathbf{g}}\|^2 \left[\frac{\mathbf{S}\hat{\mathbf{x}}}{(w + \bar{\mathbf{g}}'\hat{\mathbf{x}})^2} - \frac{\bar{\mathbf{g}}\hat{\mathbf{x}}'\mathbf{S}\hat{\mathbf{x}}}{(w + \bar{\mathbf{g}}'\mathbf{x})^3}\right] = \|\bar{\mathbf{g}}\|^2 \left[\frac{\Sigma\hat{\mathbf{x}}}{(w + \bar{\mathbf{g}}'\hat{\mathbf{x}})^2} - \frac{\bar{\mathbf{g}}\hat{\mathbf{x}}'\Sigma\hat{\mathbf{x}}}{(w + \bar{\mathbf{g}}'\hat{\mathbf{x}})^3}\right] + \lambda \qquad (5)$$

Obviously, equation (5) requires that the dual values of the resource constraints, θ, are the same in both models. This is a point that has received severe critique in classical PMP models because it is argued that the bounded model imposes these duals on the final quadratic model (Heckelei and Wolff, 2003). In this case however, the dual values have a very different interpretation, since the objective function does not yield specific cardinal results (e.g. profits) but instead it provides an ordering of the decision maker's production preferences. More precisely, the resource duals express the increase in expected utility achieved by a marginal right-hand side increase in the resource constraints, which means that the elements of θ don't represent monetary values (i.e. the opportunity cost of resources) and thus it is questionable whether the previous critique holds in this case.

The use of the E-V model produces an underdetermined system of I equations with $I(I + 1)/2$ unknowns and thus the estimation of the unknown parameters should be based on a maximum entropy specification, as proposed by Paris and Howitt (1998). Maximum entropy was first introduced as an estimation method for econometric analysis by Golan et al. (1996b), following the innovative work of Shannon (1948) in information theory and later Jaynes (1957) in physics and has been mostly used in production economics in order to estimate production technologies or input allocation to different activities (e.g. Zhang and Fan, 2001).

When national data on average prices and yields over a sample period of T years are used for each crop i, the farm's profit in year $t \in T$ should be equal to that year's statistic plus a small error term that represents the deviation between the two values. Therefore, if we denote the national per hectare profit for activity i at year t as γ_i^t, the farm profit will be equal to $g_i^t = \gamma_i^t(1 + \epsilon_i^t)$, where ϵ_i^t is the corresponding error term. A maximum entropy model is then formulated in order to estimate the error terms in the variances and covariances in the $\mathbf{S}$ matrix. For this, every error term is expressed as the expected value of an unknown discrete probability distribution $\boldsymbol{\pi}_i^t = [\pi_{i1}^t, \pi_{i2}^t, \dots, \pi_{iK}^t]$ over a vector of discrete support values $\mathbf{z} = [z_1, z_2, \dots, z_K]$. Unlike the arbitrary selection of support values in traditional PMP applications (which is also the case in the original article by Paris and Howitt), the chosen support values for the elements of a single farm's $\mathbf{S}$ matrix are now defined around proxy values that come from national averages on activity profits and constitute relevant prior information. Assuming that $\boldsymbol{\Sigma}$ is known, the entropy maximization problem can thus be written as:

$$\text{Max} \quad H(\pi_{ik}^t) = -\sum_{t=1}^{T} \sum_{i=1}^{I} \sum_{k=1}^{K} \pi_{ik}^t \ln \pi_{ik}^t$$

$$\text{s.t.} \quad \epsilon_i^t = \sum_{k=1}^{K} z_k \pi_{ik}^t \quad \text{and} \quad \sum_{k=1}^{K} \pi_{ik}^t = 1 \quad \text{with} \quad \pi_{ik}^t \geq 0$$

$$\mu_i = \frac{1}{T} \sum_{t=1}^{T} \gamma_i^t (1 + \epsilon_i^t)$$

$$s_i^2 = \frac{1}{T-1} \sum_{t=1}^{T} [\gamma_i^t (1 + \epsilon_i^t) - \mu_i]^2$$

$$s_{ij} = \frac{1}{T} \sum_{t=1}^{T} \left\{ [\gamma_i^t (1 + \epsilon_i^t) - \mu_i] \left[\gamma_j^t \left(1 + \epsilon_j^t\right) - \mu_j \right] \right\}$$

$$\|\bar{\mathbf{g}}\|^2 \left[\frac{\mathbf{S}\mathbf{x}}{(w + \bar{\mathbf{g}}'\mathbf{x})^2} - \frac{\bar{\mathbf{g}}\mathbf{x}'\mathbf{S}\mathbf{x}}{(w + \bar{\mathbf{g}}'\mathbf{x})^3} \right] = \|\bar{\mathbf{g}}\|^2 \left[\frac{\boldsymbol{\Sigma}\mathbf{x}}{(w + \bar{\mathbf{g}}'\mathbf{x})^2} - \frac{\bar{\mathbf{g}}\mathbf{x}'\boldsymbol{\Sigma}\mathbf{x}}{(w + \bar{\mathbf{g}}'\mathbf{x})^3} \right] + \lambda$$

where H denotes the system's entropy, μ_i is average profit of activity i over T periods, s_i^2 is the corresponding profit variance (the diagonal elements of matrix $\mathbf{S}$) and s_{ij} is the covariance of profits for activities i and j (the off-diagonal elements of $\mathbf{S}$, for $i > j$).

5   An illustrative example

To provide an example of calibrating farm models under uncertainty, we select a farm in Thessaly, Greece, whose production choices and other information for year 2002 are presented in Table 1.

Table 1. Example farm characteristics

Data input options	Cotton	D. Wheat	S. Wheat	Maize
Land (ha)	82	30	55	10
Yield (t/ha)	3.4	4.0	3.9	13.0
Price (€/t)	293	132	147	145
Production subsidy (€/t)	590	-	-	-
Variable expenses (€/ha)	1023.7	480.0	478.0	1207.3
Land subsidy (€/ha)	-	300	300	500

Source: Petsakos et al. (2009)

To calculate the initial Σ matrix, we used national averages on prices and yields for the decade 1992-2001 from the site of the Greek Ministry of Rural Development and Food [1]. Prices were deflated in order to represent constant 2002 prices. Since we had no information on annual variable costs and subsidies for the decade under consideration, we used the variable expenses and subsidies from the farm observations of year 2002. This allowed annual differences of crops'gross margins to be attributed solely to the two random parameters, prices and yields. Additionally, we set the initial wealth, w, equal to 10,000€. For this example, the selection of w is in fact arbitrary but in a post 2003 CAP reform scenario, it may represent the single payment received by the farmer. However, the inclusion of w turned out to be necessary for solving the model in GAMS (Brook et al., 1998) with the CONOPT3 solver, since it guaranteed a positive value inside the logarithm in the initial bounded expected utility model.

The dual vector λ obtained in the first phase of PMP was then used to estimate the "real" variance-covariance matrix $\mathbf{S}$ of the farm's activities gross margins. We selected five support values [−0.4, −0.2, 0, 0.2, 0.4], so that the farm's profit from each activity would be located in a range ±40% of the corresponding national average in the same year. We should note that there is no consensus as to the appropriate number of discrete support points. Golan et al. (1996a) state that although the increase in the number of support points improves the estimation results, this improvement diminishes as K increases. The choice is therefore left to the analyst but examples in the literature include among others $K = 5$ as in Paris and Howitt (1998), $K = 4$ as in Heckelei and Britz (2000) and $K = 2$ as in Zhang and Fan (2001). The estimation produced a farm's variance-covariance matrix that is shown in table 2, together with the initial Σ matrix.

Table 2. Average and farm variance-covariance matrices

	Σ matrix				$\mathbf{S}$ matrix			
	Cot	*Swt*	*Dwt*	*Mze*	*Cot*	*Swt*	*Dwt*	*Mze*
Cotton	1567.66				116.37			
S. Wheat	22.44	15.87			4.97	16.33		
D. Wheat	3.77	8.94	14.97		3.61	9.38	15.44	
Maize	-122.71	23.48	31.79	256.52	-122.15	9.15	22.60	330.79

Cot = Cotton, Swt = Soft Wheat, Dwt = Durum Wheat and Mze = Maize

We observe that with the exception of the cotton variance and soft wheat covariance with

cotton and maize, all other terms differ only a little between the two matrices. For cotton and soft wheat, this should be expected since these two crops constitute the "preferred" crops in our PMP formulation, i.e. they are bounded by the calibration constraint and thus have a nonzero λ_i dual.

To verify that our alternative PMP procedure works we set up the final expected utility model with an **S** variance-covariance matrix:

$$\text{Max} \quad E[U(\mathbf{x})] = \ln(w + \bar{\mathbf{g}}'\mathbf{x}) - \frac{1}{2}\left(\frac{\|\bar{\mathbf{g}}\|}{w + \bar{\mathbf{g}}'\mathbf{x}}\right)^2 \mathbf{x}'\mathbf{Sx}$$

$$\text{s.t.} \quad \mathbf{Ax} \leq \mathbf{b}$$
$$\mathbf{x} \geq 0$$

Not surprisingly, the final model exactly reproduced base year observations.

6 Conclusions

In this paper we present an alternative PMP approach that is based on the assumption that production choices differ from the ones suggested by a normative LP model, not because the farmer's true cost function is quadratic, but because he is risk averse responding to stochastic yields and prices. Therefore the farmer maximizes expected utility instead of profits and his objective function includes a variance term derived from second order Taylor approximation of logarithmic utility.

Following the usual PMP procedure, we set up an initial bounded model that reproduces base year observations and also provides a vector of duals that are treated as implicit information. We then search for mispecification problems in the variance-covariance matrix that is commonly built with national (or regional) data. A maximum entropy problem is then formulated to estimate deviations between the farm's activities gross margins and the ones calculated at the national level. This finally yields the "true" variance-covariance matrix of the farm with which perfect calibration can be achieved.

Our approach provides another contribution to the PMP literature, which is dominated thus far by nonlinear cost and production functions within a deterministic context. We show that the rationale of PMP, which is to find misspecified parameters in the objective function, by equating first order conditions of a calibrating with a non-calibrating model, is not limited to just LP but can be extended to all kinds of programming models.

References

[1] Brink, L., and McCarl, B. (1978). The Tradeoff Between Expected Return and Risk About Corn-Belt Farmers. *American Journal of Agricultural Economics* 60: 259-263.

[2] Brooke, A., Kendrick, D., Meeraus, A. and Raman, R. (1998). *GAMS: A User's Guide.* GAMS Development Co.

[3] Borch, K. (1974). The Rationale of the Mean-Standard Deviation Analysis: Comment. *The American Economic Review* 64(3): 428-430.

[4] Freund, R. J. (1956). The Introduction of Risk into a Programming Model. *Econometrica* 24(3): 253-263.

[5] Golan, A., Judge, G. and Perloff, J. M. (1996a). A Maximum Entropy Approach to Recovering Information From Multinomial Response Data. *Journal of the American Statistical Association* 91(434): 841-853.

[6] Golan, E., Judge, G. and Miller, D. (1996b). *Maximum Entropy Econometrics*. Wiley, Chichester, UK.

[7] Hardacker, J. B., Huirne, J. R., Anderson, J. R. and Lien, G. (2004). Copying with Risk in Agriculture. CAB International, Wallingford, UK.

[8] Hazell, P. B. R. and Norton, R. D. (1986). *Mathematical Programming for Economic Analysis in Agriculture*. Macmillan, New York.

[9] Heckelei, T. and Wolff, H. (2003). Estimation of constrained optimisation models for agricultural supply analysis based on generalized maximum entropy. *European Review of Agricultural Economics* 30(1): 27-50.

[10] Heckelei, T. (2002). *Calibration and estimation of programming models for agricultural supply analysis*. Habilitation Thesis, University of Bonn, Germany.

[11] 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.

[12] Howitt, R. E. (1995a). Positive Mathematical Programming. *American Journal of Agricultural Economics* 77: 329-342.

[13] Howitt, R. E. (1995b). A Calibration Method for Agricultural Economic Production Models. *Journal of Agricultural Economics* 46(2): 147-159.

[14] Jaynes, E. T. (1957). Information Theory and Statistical Mechanics. *Physics Review* 106: 620-630.

[15] Júdez, L., de Miguel, J. M., Mas, J. and Bru, R. (2002). Modeling Crop Regional Production Using Positive Mathematical Programming. *Mathematical and Computer Modelling* 35: 77-86.

[16] Levy, H. (1974). The Rationale of the Mean-Standard Deviation Analysis: Comment. *The American Economic Review* 64(3): 434-441.

[17] Levy, H. and Markowitz, H. M. (1979). Approximating Expected Utility by a Function of Mean and Variance. *The American Economic Review* 69(3): 308-317.

[18] Markowitz, H. M. (1959). *Portfolio Selection: Efficient Diversification of Investments*. Wiley, New York.

[19] Markowitz, H. M. (1952). Portfolio Selection. *The Journal of Finance* 7(1): 77-91.

[20] Meyer, J. (1987). Two-moment Decision Models and Expected Utility Maximization. *American Economic Review* 77(3): 421-430.

[21] Paris, Q. and Howitt, R. E. (1998). An analysis of ill-posed production problems using maximum entropy. *American Journal of Agricultural Economics* 80(1): 124-138.

[22] Petsakos, A. and Rozakis, S. (2009). Critical review and state-of-the-art of PMP models: an application to Greek arable agriculture. In: Rezitis, A. (Ed.), *Research Topics in Agricultural and Applied Economics* (Vol. 1, pp. 36-61), E-Book. Bentham Science Publishers.

[23] Petsakos, A., Rozakis, S. and Tsiboukas, K. (2009). Risk optimal farm plans in the context of decoupled subsidy payments: the case of cotton production in Thessaly. *Journal of Farm Management* 13(7): 467-483.

[24] Pratt, J. W. (1964). Risk Aversion in the Small and in the Large. *Econometrica* 32(1/2): 122-136.

[25] Pulley, L. B. (1983). Mean-Variance Approximations to Expected Logarithmic Utility. *Operations Research* 31(4): 685-696.

[26] Rubinstein, M. (1976). The Strong Case for the Generalized Logarithmic Utility Function as the Premier Model of Financial Markets. *Journal of Finance* 31: 555-571.

[27] Shannon, C. E. (1948). A Mathematical Theory of Communication. *Bell system Technical Journal* 27(July and October): 379-423, 623-656.

[28] Torres-Rojo, J. M. (2001). Risk management in the design of a feeding ration: a portfolio theory approach. *Agricultural Systems* 68(1): 1-20.

[29] Tsiang, S. C. (1972). The Rationale of the Mean-Standard Deviation Analysis, Skewness Preference, and the Demand for Money. *The American Economic Review* 62(3): 354-371.

[30] Zhang, X. and Fan, S. (2001). Estimating Crop-Specific Production Technologies in Chinese Agriculture: A Generalized Maximum Entropy Approach. *American Journal of Agricultural Economics* 83(2): 378-388.