Positive Mathematical Programming with Multiple Data Points: A Cross-Sectional Estimation Procedure

Thomas HECKELEI
Wolfgang BRITZ

Cahiers d'économie et sociologie rurales, n° 57, 2000

Thomas HECKELEI*, Wolfgang BRITZ*

Programmation mathématique positive avec observations multiples : estimation en coupes transversales

Résumé – Cet article présente une approche visant à spécifier des fonctions de coût non linéaires dans le cadre de modèles de programmation régionaux. La méthodologie en question peut être considérée comme une application de la programmation mathématique positive (PMP) à des observations multiples. L'application de la PMP dans les modèles d'offre des produits agricoles s'est sensiblement développée au cours de ces dix dernières années. Cependant, beaucoup de modélisateurs n'ont pas fait état du comportement arbitraire et potentiellement invraisemblable des modèles résultant de l'application standard de l'approche PMP. Paris et Howitt (1998) interprètent la PMP comme étant l'estimation d'une fonction de coût non linéaire et généralisent sa spécification en utilisant un procédé de «maximisation de l'entropie». Néanmoins, leur approche manque d'une base empirique suffisante. Elle comporte toujours une paramétrisation nécessaire pour imposer les bonnes conditions de coubure de la fonction de coût, ce qui pose d'importants problèmes dans les applications. La méthodologie que nous proposons est conçue pour exploiter l'information contenue dans un échantillon de données en coupes pour spécifier des fonctions de coût quadratiques régionales avec des effets croisés entre les activités. L'approche apporte également une solution au problème de la courbure de la fonction de coût. Elle est appliquée ici à des modèles de programmation régionaux sur 22 régions françaises. Une simulation *a posteriori* de la réforme de la Politique agricole commune de 1992 produit des résultats plausibles. Des prolongements de cette méthode ainsi que des améliorations possibles sont également identifiés.

Mots-clés :
programmation mathématique positive, méthode du maximum d'entropie, validation *a posteriori*, courbure de la fonction de coût

Positive Mathematical Programming with Multiple Data Points: A Cross-Sectional Estimation Procedure

Summary *– This paper introduces an approach to the specification of non-linear cost functions in regional programming models. It can be characterised as an application of positive mathematical programming (PMP) to multiple observations. The application of PMP in policy relevant agricultural supply models as a mean for calibration has significantly increased during the last ten years. However, many modellers have not reflected the arbitrary and potentially implausible response behaviour of the resulting models implied by standard applications of the approach. Paris and Howitt (1998) interpret PMP as the estimation of a non-linear cost function and generalize the specification by employing a « Maximum Entropy (ME) » procedure. However, their approach still lacks a sufficient empirical base and involves a parameterisation to enforce correct curvature of the cost function, which induces significant problems in applications. The suggested methodology is designed to exploit information contained in a cross sectional sample to specify — regionally specific — quadratic cost functions with cross effects for crop activities. It also provides a solution to the curvature problem. The approach is applied to regional programming models for 22 regions in France. An ex-post simulation across the 1992 CAP-reform shows plausible results with respect to the simulation behaviour of the resulting models. Paths for extensions and improvements of this methodology are identified.*

Key-words: *positive mathematical programming, maximum entropy estimation, curvature restrictions, ex-post validation*

* *Institute of Agricultural Policy, Market Research and Economic Sociology, University of Bonn, Nussallee 21, 53115 Bonn, Germany.*
e-mail : heckelei@agp.uni-bonn.de ; britz@agp.uni-bonn.de

This research was pursued within the project « Common Agricultural Policy Regional Impact » (CAPRI) which was partly funded under the FAIR program of the European Commission. The paper considerably benefited from the comments of two anonymous referees and a co-editor. Even their extensive and constructive input probably left numerous errors for which the authors assume full responsability.

THE project «Common Agricultural Policy Regional Impact» (CAPRI) aims at regionally differentiated EU-wide analysis of the CAP[1]. The concept of the underlying comparative static modelling system combines a supply component comprising about 200 regional programming models with a multi-commodity market model in an iterative fashion to endogenously determine regional supply and national demand quantities, net trade at Member State and EU-level, and equilibrium market prices. The basic question that initiated the research presented in this article is how to specify the regional programming models such that they offer an empirically valid supply response for a large number of crop activities (up to 20).

Aggregate programming models are still widely used for policy relevant analysis of agricultural supply behaviour. Their ability to easily incorporate important policy measures such as quotas and per hectare premia at a highly differentiated product level, the implied consistency with primary factor constraints during simulations, and the possibility to use explicit assumptions on technology renders this methodological choice preferable to the use of duality based econometric models for many analysts. However, these advantages come at the price of enormous data requirements – which often exclude the compilation of time series – and a typical lack of empirical validation. The CAPRI database offers average yields, average use of variable inputs by production activities, and activity levels at least for the years 1990 to 1995 based on the REGIO database of Eurostat[2] and complementary national statistics [3]. It currently lacks, however, regional stocks on labour and capital and their activity differentiated use as well as a representation of the heterogeneous soil qualities in the EU regions. Consequently, the specification of the regional production technology is not sufficient to avoid overspecialisation of model solutions and to guarantee plausible simulation behaviour based on a typical linear programming formulation. The use of – at the aggregate level – weakly justified rotational constraints or direct bounds on activity levels to better match observed land allocation cannot be seriously considered for policy simulation exercises.

Positive mathematical programming (PMP, see Howitt, 1995a and 1995b) promises a remedy: it allows to calibrate insufficiently specified programming models to observed behaviour in an elegant fashion without restricting the model's simulation behaviour by unjustified bounds. Consequently, the application of PMP in policy relevant agricultural supply models – which started already in the eighties (for example: Howitt and Gardner, 1986; House, 1987; Kasnakoglu and Bauer, 1988) – has si-

[1] For detailed information on the CAPRI project, consult Heckelei and Britz (2001) or the internet page *http://www.agp.uni-bonn.de/agpo/rsrch/capri/capri_e. htm*.

[2] Statistical Office of the European Communities.

[3] The CAPRI database is currently updated until the year 2000. Completeness at this point, however, is only guaranteed until 1995.

gnificantly increased during the last ten years (for example: Horner *et al.*, 1992; Schmitz, 1994 ; Arfini and Paris, 1995; Barkaoui and Butault, 1999; Cypris, 2000 ; Graindorge *et al.*, 2001; Helming *et al.*, 2001).

However, many modellers have not reflected the arbitrary and potentially implausible response behaviour of the resulting models implied by standard applications of the approach (Heckelei, 1997). Paris and Howitt (1998) interpret PMP as the estimation of a non-linear cost function and generalize the specification by employing a «Maximum Entropy» (ME) procedure. This paper presents an approach which overcomes some of the drawbacks involved in their analysis, providing a useful tool for calibration but – more importantly – for the specification of a plausible crop allocation response of aggregate programming models based on observed behaviour. This paper is organised as follows: the first section reminds the reader of the general PMP-approach, introduces the use of the ME technique in this context and identifies problems associated with the approach by Paris and Howitt. The second section describes an ME-PMP approach for crop production which is designed to exploit information contained in a cross sectional sample to specify – regionally specific – quadratic cost functions with cross effects for crop activities. The approach is applied to CAPRI's regional programming models in France and validated in an ex-post simulation exercise. The last section draws conclusions and identifies possible directions for further research.

THE MAXIMUM ENTROPY APPROACH TO POSITIVE MATHEMATICAL PROGRAMMING

Reminder on PMP

First we remind the reader of the two steps involved in PMP to calibrate typical linear programming models to observed activity levels (see Howitt, 1995a or Heckelei, 1997 for a more detailed description and Gohin and Chantreuil, 1999 for a very accessible introduction and discussion). The general idea of PMP is to use information contained in dual variables of a linear programming (LP) problem [(4)] bounded to observed activity levels by calibration constraints (Step 1), in order to specify a non-linear objective function such that observed activity levels are reproduced by the optimal solution of the new programming problem without bounds (Step 2).

[(4)] The method can be applied to non-linear programming problems as well. In order to ease the understanding, a simple but general layout of a LP is discussed here.

Using a simplified LP formulation designed to determine the profit maximizing crop mix, **Step 1** of this procedure is formally described in the following way[5]:

$$\underset{x}{Max}\ Z = \boldsymbol{p}'\boldsymbol{y} - \boldsymbol{c}'\boldsymbol{x}$$

subject to

$$\boldsymbol{A}\begin{bmatrix}\boldsymbol{x}\\ \boldsymbol{y}\end{bmatrix} \leq \boldsymbol{b}\ [\pi] \qquad (1)$$

$$\boldsymbol{x} \leq (\boldsymbol{x}^{o} + \varepsilon)\ [\lambda]$$

$$\boldsymbol{x} \geq [0]$$

where Z denotes the objective function value, $\boldsymbol{c}$ and $\boldsymbol{x}$ are $(n \times 1)$ vectors of variable cost per unit of activity and production activity levels, respectively, $\boldsymbol{p}$ and $\boldsymbol{y}$ are (1×1) vectors of (expected) output prices and sales activity levels, respectively, $\boldsymbol{A}$ represents a $(m \times (n + 1))$ matrix of coefficients in resource/policy constraints, $\boldsymbol{b}$ and π are $(m \times 1)$ vectors of available resource quantities and their dual variables, respectively, λ are dual variables associated with the calibration constraints[6], $\boldsymbol{x}^{o}$ is a $(n \times 1)$ vector of *observed* production activity levels and ε denotes a vector of small positive numbers.

The addition of the calibration constraints forces the optimal solution of the LP model (1) to almost perfectly reproduce the observed base year activity levels $\boldsymbol{x}^{o}$, given that the specified resource constraints allow for this solution (which they should if the data are consistent). « Almost perfectly » is defined by the range of the positive perturbations of the calibration constraints, ε, which are introduced to prevent linear dependencies between resource and calibration constraints. The latter would provoke degenerate dual solutions with marginal values arbitrarily distributed across resource and calibration constraints.

In **Step 2** of the procedure, the vector λ is employed to specify a non-linear objective function such that the marginal cost of the preferable activities are equal to their respective revenues at the base year activity levels $\boldsymbol{x}^{o}$. Given that the implied variable cost function has the right curvature properties (convex in activity levels) the solution to the resulting programming problem will be a « boundary point, which is the combination of binding constraints and first order conditions » (Ho-

[5] Matrices and vectors are printed in bold letters.

[6] The calibration constraints are expressed as *upper* bounds on activity levels. This is sufficient as long as the realisation of the activity provides a positive contribution to the objective function. This should be the case for *expected* profits if positive activity levels are observed. When using realised yields and prices of a calibration year, however, negative profits per activity may occur so that calibration constraints must be formulated as lower bounds as well.

witt, 1995a, p. 330) and equal to the results of (1) with respect to activity levels and dual values on the resource constraints, π.

For reasons of computational simplicity and lacking strong arguments for other types of functions, we will illustrate the specification of the parameters in the objective function with the following general version of a quadratic variable cost function [7]:

$$C^v = \boldsymbol{d}'\boldsymbol{x} + \frac{1}{2}\boldsymbol{x}'\boldsymbol{Q}\boldsymbol{x} \qquad (2)$$

where C^v denotes variable costs, $\boldsymbol{d}$ is a ($n \times 1$) vector of parameters associated with the linear term, and $\boldsymbol{Q}$ is a ($n \times n$) symmetric positive definite matrix of parameters associated with the quadratic term of C^v.

The parameters of (2) need to be specified such that

$$\frac{\partial C^v(\boldsymbol{x}^o)}{\partial \boldsymbol{x}} = \boldsymbol{MC}^v = \boldsymbol{d} + \boldsymbol{Q}\boldsymbol{x}^o = \boldsymbol{c} + \boldsymbol{\lambda} \qquad (3)$$

This specification problem is «ill-posed», because the number of parameters to be specified ($= n + n(n + 1)/2$) is greater than the number of observations ($= n$ observations on marginal cost). Traditional econometric approaches could handle this type of problem if an appropriate number of *a priori* restrictions on the parameters leave enough degrees of freedom. Most applications of PMP go without any type of estimation by setting all off-diagonal elements of $\boldsymbol{Q}$ to zero and calculating the remaining parameters by some standard approach (see Heckelei, 1997 for a discussion). Although these approaches work perfectly well with respect to the calibration property of PMP by setting appropriate first order derivatives of the objective function according to (3), the resulting simulation behaviour is completely arbitrary and potentially unsatisfactory, (see Cypris, 2000 and the last section). This is because the response behaviour of the calibrated model depends to a large extent on the *second order derivatives* of the objective function, *i.e.* on the *change* in marginal cost when activity levels are changing. However, just one observation on dual values of the calibrations constraints does not provide any information on this.

Maximum Entropy specification of the cost function

Paris and Howitt suggest to use Maximum Entropy (ME) estimation [8] which allows for a more objective specification of the parameters

[7] Paris and Howitt (1998) show the general applicability of their approach also with respect to other functional forms. Compared to equation (2) they choose, however, a somewhat restricted quadratic functional form by excluding linear parameters.

[8] See Golan, Judge and Miller (1996) for a comprehensive introduction to Maximum Entropy Econometrics, or Mittelhammer, Judge, and Miller (2000), chapter E3, in the context of a general Econometrics textbook.

of the non-linear cost function based on an « econometric type » criterion. Moreover, it has the potential of incorporating more than one observation on activity levels into the specification of the parameters and decreases the need to decide on exact *a priori* restrictions on the parameters. The application of ME to the calibration of programming models comes at a time of significantly increased general interest in entropy techniques by agricultural economists after the comprehensive introduction by Golan *et al.* (1996). Their framework based on probability supports of parameters and error terms allowed to apply the entropy criterion to ill-posed problems in econometrics. Studies in the realm of production economics often focus on the estimation of input allocation to products and estimation of production technologies (for example: Lence and Miller, 1998a and 1998b; Léon *et al.*, 1999; Zhang and Fan, 2001). Applications to dual behavioural models are, so far, less frequently observed (Oude Lansink, 1999). Note that this article should rather be seen in the context of the PMP literature and consequently does not focus on contributions to the application of entropy techniques in general. However, below we draw upon various of the already mentioned publications when specifying the calibration approach.

To make ME-estimation of the variable cost function (2) operational [9], we first need to define support points for the parameter vector $\boldsymbol{d}$ and the matrix $\boldsymbol{Q}$. One could centre the linear parameters $\boldsymbol{d}$ around the observed accounting cost per unit of the activity, $\boldsymbol{c}$. For example, we could choose 4 support points for each parameter by setting [10]

$$\boldsymbol{zd}_i = \begin{bmatrix} -2 \cdot c_i \\ 0 \cdot c_i \\ +2 \cdot c_i \\ +4 \cdot c_i \end{bmatrix} \quad \forall i \tag{4}$$

In the case of the $\boldsymbol{Q}$ matrix we have to distinguish the diagonal (= change in marginal cost of activity i with respect to the level of activity i) from the off-diagonal elements (= change in marginal cost of activity i with respect to the level of activity j). Given that the *a priori* expectation for the linear parameter vector $\boldsymbol{d}$ are the accounting costs (supports centred around c_i in equation (4)), it is consistent with condition (3) to centre the support points for q_{ii} around λ_i/x_i^0 and the off – diagonal elements q_{ij} around zero. The centre of the support points λ_i/x_i^0 for the diagonal elements are positive, a necessary condition for convexity of C^v.

[9] See Paris and Howitt (1998) for further details and a more extensive motivation of the approach.

[10] The variance of the maximum entropy estimates is negatively correlated with the number of support points defined and has a limit value for an infinite number of support points (see Golan *et al.*, 1996, p. 139). There is no general rule for the « right » number of support points, but tests with our models have shown that choosing more than 4 support points does not change the numerical results of the calculated parameter expectations by an extent of any practical relevance.

A suitable specification for the support points of $\boldsymbol{Q}$ would then be

$$\boldsymbol{zq_{i,i}} = \begin{bmatrix} 0 \cdot \lambda_i / x_i^0 \\ 2/3 \cdot \lambda_i / x_i^0 \\ 4/3 \cdot \lambda_i / x_i^0 \\ 2 \cdot \lambda_i / x_i^0 \end{bmatrix} \quad \forall i \text{ and } \boldsymbol{zq}_{i,j} = \begin{bmatrix} -3 \cdot \lambda_i / x_j^0 \\ -1 \cdot \lambda_i / x_j^0 \\ +1 \cdot \lambda_i / x_j^0 \\ +3 \cdot \lambda_i / x_j^0 \end{bmatrix} \quad \forall i \neq j \tag{5}$$

Denoting the probabilities for the K support points $\boldsymbol{zd_i}$, $i = 1,\ldots,n$, and $\boldsymbol{zq_{ij}}$, $i, j = 1,\ldots,n$, as $pd_{k,i}$ and $pq_{k,i,j}$, respectively, the estimated values of the corresponding parameters are calculated as

$$\begin{aligned} d_i &= \sum_{k=1}^{K} pd_{k,i}\, zd_{k,i}, \quad \forall i \\ q_{i,j} &= \sum_{k=1}^{K} pq_{k,i,j}\, zq_{k,i,j}, \quad \forall i, j \end{aligned} \tag{6}$$

The ME formulation of estimating the parameters then looks like the following:

$$\max_{p} H(p) = -\sum_{k=1}^{K}\sum_{i=1}^{n} pd_{k,i}\, ln\, pd_{k,i} - \sum_{k=1}^{K}\sum_{i=1}^{n}\sum_{j=1}^{n} pq_{k,i,j}\, ln\, pq_{k,i,j}$$

Subject to

$$\begin{aligned} & d_i + \sum_{j=1}^{n} q_{i,j}\, x_j^0 = c_i + \lambda_i, \quad \forall i \\ & d_i = \sum_{k=1}^{K} pd_{k,i}\, zd_{k,i}, \quad \forall i \\ & q_{i,j} = \sum_{k=1}^{K} pq_{k,i,j}\, zq_{k,i,j}, \quad \forall i, j \\ & \sum_{k=1}^{K} pd_{k,i} = 1, \quad \forall i \\ & \sum_{k=1}^{K} pq_{k,i,j} = 1, \quad \forall i, j \\ & q_{i,j} = q_{j,i} \quad \forall i, j \end{aligned} \tag{7}$$

The entropy criterion in the objective function of (7) looks for the set of probabilities which adds the least amount of information – *i.e.* deviates the least from a uniform distribution over the support points –

but satisfies the explicitly shown « data constraint » of the estimation problem being the marginal cost condition (3).

At this point we need to hold for a moment and need to address the question under what conditions a ME formulation for estimating the parameters of the quadratic cost function deems useful. If we have only a $(1 \times n)$ vector of marginal cost available (from calibrating one linear programming problem to one base year solution), the outcome of the estimation and hence the simulation behaviour of the resulting model will be heavily dominated by the supports. Such an application of ME should hence be interpreted as calibrating a cost function based on prior expectations on the parameter values to observed values according to condition (3). The entropy criterion works here as a penalty function for the deviation from the prior expectations (centre of supports) and the term « estimation for the calibration process » may be misleading. The approach with just one observation on marginal cost could be sensibly applied to derive a cost function based on specific prior information, for example a full matrix of elasticities[(11)] or to exogenously given yield functions (Howitt, 1995a).

The example defined according to (4) and (5) is however not a meaningful application for such a calibration process as supports were defined without any valuable prior information on the cost function. Specifically, the ME problem will reach its optimum when the probabilities follow a uniform distribution, since the centres of the support values already satisfy the data constraints. The resulting parameter estimates will be exactly the ones implied by the « standard approach » as defined in Heckelei (1997), *i.e.* linear parameters of the cost function are equal to the respective activity's accounting costs c_i, the off-diagonal elements of the $\boldsymbol{Q}$ matrix are zero, and the diagonal elements are equal to λ_i / x_i^0. The simulation behaviour of the resulting model is arbitrary as it completely depends on the arbitrary specification of the support values.

For similar reasons, the approach of Paris and Howitt (1998) who reparameterize the $\boldsymbol{Q}$ matrix based on a $\mathbf{LDL}'$ (Cholesky) decomposition to ensure appropriate curvature properties of the estimated cost function should – in our view – only be seen as a demonstration on how to combine ME and PMP. The choice of their support values is not based on prior information. They centre the elements of $\boldsymbol{D}$ around the value for the diagonal elements of $\boldsymbol{Q}$ which would satisfy the marginal cost condition and the elements of $\boldsymbol{L}$ around zero. Due to the complex (and even order-dependent) relationship between the matrices $\boldsymbol{L}$, $\boldsymbol{D}$ and $\boldsymbol{Q}$, this implies rather nontransparent *a priori* expectations for the parameters of $\boldsymbol{Q}$. The nonzero cross costs effects of activities obtained from their ME solution is merely based on this technically motivated choice of support points.

(11) Parameters can be easily defined for the diagonal elements of $\boldsymbol{Q}$ without the need to apply a ME approach if just own-price elasticities are given, see for example Helming *et al.* (2001).

In contrast to the examples given above, we now suggest an approach based on a cross-sectional sample of marginal cost from a set of regional programming models. We apply the term «estimation» for this procedure, since several regional vectors of marginal costs are used to specify the cost functions. The choice of ME instead of other estimators is motivated by the fact that we have still negative degrees of freedom. The discussion will mostly concentrate on necessary parametric restrictions across regions to accommodate for regions with different sizes and crop rotations. Additionally, we provide a solution to the curvature problem which allows the definition of support points for the actual parameters to be estimated by incorporating a $\mathbf{LL}'$ decomposition as direct constraints of the estimation problem.

A PMP-ME APPROACH BASED ON A CROSS SECTIONAL SAMPLE

The first part of this section presents a rationale for the approach and introduces the most important parts of the mathematical formulation. The second part delivers some details on an application for CAPRI's regional programming models for France and presents results of an ex-post validation for the simulation behaviour of the specified model across the CAP-reform of 1992.

Rationale

Our objective here is to estimate a quadratic cost function with cross cost effects (full $\boldsymbol{Q}$ matrix) between *crop production activities*. Suppose one can generate R ($1 \times n$) vectors of marginal costs from a set of R regional programming models by applying the first step of PMP. In our example, n represents 18 crops and R the 22 French NUTS-2 regions. In order to exploit this information for the specification of quadratic cost functions for all regions, we need to define appropriate restrictions on the parameters across regions, since otherwise no informational gain is achieved.

Consider the following suggestion for a regional vector of marginal cost:

$$\begin{aligned} \boldsymbol{MC}_r^v &= \boldsymbol{d}_r + \boldsymbol{Q}_r \boldsymbol{x}_r \\ \boldsymbol{Q}_r &= (cpi_r)^g \, \boldsymbol{S}_r \, \boldsymbol{B} \boldsymbol{S}_r' \quad \text{with} \quad s_{r,i,i} = \sqrt{1/x_{r,i}^o} \end{aligned} \tag{8}$$

where $\boldsymbol{d}_r$ is a ($n \times 1$) vector of linear cost function parameters in region r, $\boldsymbol{Q}_r$ represents a ($n \times n$) matrix of quadratic cost term parameters in region r, cpi_r stands for regional «crop profitability index» defined as the relation between the regional and average revenue per hectare $(\boldsymbol{p}'\boldsymbol{y}_r/L_r)/(\sum_r \boldsymbol{p}'\boldsymbol{y}_r/\sum_r L_r)$ where L_r is land available, g is a parameter deter-

mining the influence of the crop profitability index, S_r constitutes ($n \times n$) diagonal scaling matrices for each region r, and finally $\boldsymbol{B}$ is a ($n \times n$) parameter matrix related to $\boldsymbol{Q}_r$.

The rationale for (8) can best be inferred from a didactic example shown in table 1, based on a LP model with 2 crops, 2 regions and a land constraint. Rows 1-3 present the observed base year data – total revenues, accounting costs and activity levels – from which the duals of the calibration constraints, marginal costs (rows 4-6) as well as average revenues and the crop profitability index (row 7-8) can be deducted.

Contrary to the ultimate application, the matrix $\boldsymbol{B}$ as shown in the last columns of row 11 is given and not estimated. It is defined such that the relative increase of marginal costs for a 1 % increase in levels is equal to 5 % at national level. In order to motivate the scaling with S_r, we have a look at the implied elasticities of marginal costs to changes in activity levels as shown in row 16 if the scaling vectors $\boldsymbol{S}$ are left out. In that case, elasticities are a direct function of observed activity levels: the smaller the level, the smaller the elasticity. Including the scaling vectors, as shown in row 17, provides a more plausible parameter restriction.

The term $(cpi_r)^g$ which reflects *differences in regional profitability* is supposed to capture the economic effect of differences in soil, climatic conditions etc. The magnitude of the effect on the marginal cost function estimated by the exponent g. A negative g, for example, would imply that specialising in a certain crop is penalised less in a region with cropping conditions above average since, *ceteris paribus*, $\boldsymbol{Q}_r$ is smaller than average in this case.

The overall specification implies that – apart from the effect of the crop profitability index – the $\boldsymbol{Q}_r$'s are identical for regions with the same crop rotation. We motivated the use of more than one observation by the fact that second order derivatives of the cost function strongly influence the simulation behaviour of the model. Where does this information hide in equation (8)? Observed rotations and marginal costs recovered by the calibration step differ between regions. The matrix $\boldsymbol{B}$ – common across regions – is estimated as to describe the differences in marginal costs depending on the differences in levels. The parameters are now estimated such that changing region i's rotation to the rotation in region j causes changes in marginal cost matching the observed differences between the two regions (again apart from the effect of the crop profitability index). This is the important contribution of the cross-sectional analysis: the simulation behaviour resulting from the ME problems is not longer depending in an arbitrary way on the support points, but is based on a clear hypothesis about the relation between crop rotation and marginal costs.

Table 1. **2-region/2-crop example for cost function specification**

	Items	Symbols in Text	Unit	Region 1		Region 2		National average	
				Cereals	Other Crops	Cereals	Other Crops	Cereals	Other Crops
1	Revenue	$p'y$	EURO/ha	1 000.00	900.00	800.00	700.00	995.12	842.86
2	Accounting cost	c	EURO/ha	550.00	500.00	450.00	400.00	547.56	471.43
3	Activity level	l	ha	40.00	10.00	1.00	4.00	41.00	14.00
4	Dual on calibration constraint		EURO/ha	50.00	0.00	50.00	0.00		
5	Dual on land constraint		EURO/ha		400.00		300.00		
6	Marginal Costs	mc	EURO/ha	600.00	500.00	500.00	400.00	597.56	471.43
7	Average Revenue		EURO/ha	980.00		720.00		956.36	
8	Crop profitability index	cpi_r		1.02		0.75		1.00	
9	Average Marginal cost	amc						597.56	534.49
								534.49	471.43
10	Scaled supports for uniform B	zbs						5.00	0.00
								0.00	5.00
11	*A priori* expectation for uniform B	$E[zb] = E[zbs]^* amc$						2987.80	0.00
								0.00	2 357.14
12	Elements of scaling matrix S	$S_{ii} = (1/x_i)^{0.5}$		0.16	0.05	1.00	0.50	0.16	0.04
				0.05	0.32	0.50	0.50	0.04	0.27
13	Exponent of crop profitability index	g						1.00	
14	Influence of crop profitability index	cpi^g		1.02		0.75		1.00	
15	*A priori* expectation Q_r	$E[Q_r] = cpi^{g*}S^*E[zb]^*S'$		76.54	0.00	2249.37	0.00	72.87	0.00
				0.00	241.54	0.00	443.64	0.00	168.37
	Percentage change of marginal costs for a 1 % increase in levels								
16	using B only			199.19	47.14	5.98	23.57	205.00	70.00
17	using SBS'			4.98	4.71	5.98	5.89	5.00	5.00
18	using Q_r			5.10	4.83	4.50	4.44	5.00	5.00

The general formulation of the corresponding ME problem is now straightforward:

$$\max H(p) = -\sum_{p}^{K}\sum_{k=1}^{n}\sum_{i=1}^{R} pd_{k,i,r}\, ln\, pd_{k,i,r} - \sum_{r=1}^{K}\sum_{k=1}^{n}\sum_{i=1}^{n} pb_{k,i,j}\, ln\, pb_{k,i,j} - \sum_{j=1}^{K} pg_k\, ln\, pg_k$$

subject to

$$d_{i,r} + cpi_r^g \cdot \sum_{j=1}^{n} s_{i,i}\, s_{j,j}\, b_{i,j}\, x_{j,r}^o = c_{i,r} + \lambda_{i,r}, \quad \forall i, r$$

$$d_{i,r} = \sum_{k=1}^{K} pd_{k,i,r}\, zd_{k,i,r}, \quad \forall i, r$$

$$b_{i,j} = \sum_{k=1}^{K} pb_{k,i,j}\, zb_{k,i,j}, \quad \forall i, j$$

$$g = \sum_{k=1}^{K} pg_k\, zg_k \tag{9}$$

$$\sum_{k=1}^{K} pd_{k,i,r} = 1, \quad \forall i, r$$

$$\sum_{k=1}^{K} pb_{k,i,j} = 1, \quad \forall i, j$$

$$\sum_{k=1}^{K} pg_k = 1$$

$$b_{i,j} = b_{j,i}, \quad \forall i, j$$

The current formulation in (9) does not guarantee that a *positive definite* matrix $\boldsymbol{B}$ – and consequently – positive definite matrices $\boldsymbol{Q}_r$ will be recovered. A violated curvature property might result in a specification of the objective function that does not calibrate to the base year, since only first order but not second order conditions for a maximum are satisfied at the observed activity levels. In order to circumvent the problems with the $\boldsymbol{LDL'}$ reparameterisation of Paris and Howitt described above, a «classic» Cholesky decomposition of the form $\boldsymbol{B} = \boldsymbol{LL'}$ is used *indirectly* as additional constraints of the ME problem (9) in the form of [12]

[12] The two different forms of the Cholesky decompositions are related in the following way: Replacing the «ones» on the diagonal triangular matrix $\boldsymbol{L}$ of $\boldsymbol{Q} = \boldsymbol{LDL'}$ with the square roots of the corresponding diagonal elements of D allows to write $\boldsymbol{Q} = \boldsymbol{LL'}$.

$$l_{i,j} = \sqrt{E[b_{i,j}] - \sum_{k=1}^{i-1} l_{i,k}^2} \quad \forall i, j$$

$$l_{i,j} = \left(E[b_{i,j}] - \sum_{k=1}^{i-1} l_{i,k}\, l_{j,k}\right) \Big/ l_{i,i} \quad \forall i, j \text{ and } j > i \tag{10}$$

Because $\boldsymbol{B}$ is supposed to be a symmetric and positive definite matrix, the $l_{i,i}$ must always be positive and real (Golub and van Loan, 1996). Appropriate lower bounds on $l_{i,i}$ deviating from zero avoid zero divisions during estimation. Due to the properties of positive definite matrices, the regional matrices $\boldsymbol{Q}_r$ calculated according to equation (8) are positive definite if $\boldsymbol{B}$ exhibits this property. A separate enforcement of curvature for each $\boldsymbol{Q}_r$ would be computationally infeasible which potentially restricts the type of alternative parameter restrictions across regions if this curvature solution is employed [13].

An application to crop production in France

In this section we describe an application and ex-post validation of the suggested approach for the regional programming models of the CAPRI system for France. Before turning to the results, the specification of the support points for the parameters is presented:

The support points for the exponent g of the crop profitability index cpi_r in (9) are defined as

$$\boldsymbol{zg} = \{-2, -{}^2/_3, +{}^2/_3, +2\} \tag{11}$$

so that the influence of the crop profitability index covers the range from $1/cpi_r^2$ to cpi_r^2 and the support of g is centred around 0. The estimation came out with a slightly negative value which implies that cropping conditions above average allow crop specialisation with marginal cost increases below average.

The crop and region specific linear terms $\boldsymbol{d}$ reflect marginal costs when all production activity levels $\boldsymbol{x}$ are zero. Since an interpretation in economic terms is hardly possible and irrelevant – especially as « fallow land » is one of the production activities – the spread of the support points $\boldsymbol{zd}$ is

[13] In earlier tests, a pragmatic solution was chosen for the curvature problem by forcing the first and second order minors of $\boldsymbol{B}$ to have the appropriate sign and restricting all off-diagonal elements to be smaller than diagonal elements during the ME step. The resulting matrix was then – if necessary – treated by a so-called « modified » Cholesky-decomposition which ensures definiteness by employing optimal correction factors to the diagonal elements (Gill *et al.*, 1989, p. 108). This procedure has proven to be operational for very large matrices.

consequently set to a very wide interval around the observed costs. The spread is 180 times the national average in revenue per ha[14].

$$\boldsymbol{zd} = c_r + \{-90, -30, 30, 90\} \sum_r p'y_r \,/\, \sum_r L_r \qquad (12)$$

Let $\overline{MC_i}$ be the land-weighted average of marginal cost for crop i across regions. The support points for $\boldsymbol{B}$ are then defined as follows (see rows 9-11 in Table 1 as well):

$$\boldsymbol{zb}_{i,j} = \boldsymbol{zbs}_{i,j}\, \boldsymbol{amc}_{i,j}$$

where

$$\boldsymbol{zbs}_{i,j} = \begin{bmatrix} \{0.001, 3.3, 6.66, 10\}\ \forall i = j \\ \{-2, -{}^{2}/_{3}, {}^{2}/_{3}, 2\}\ \forall i \neq j \end{bmatrix} \text{ and} \qquad (13)$$

$$\boldsymbol{amc}_{i,j} = 1/2\,(\overline{MC_i} + \overline{MC_j})$$

According to the spread defined by ***zbs***, the supports ***zb*** for $\boldsymbol{B}$ are defined such that changing the activity level of crop i by 1 % increases own costs between zero and ten percent. The cross effects are symmetrically centred around zero and allow for a change between – 2 % and +2 % of the average marginal costs of crop i and j, ***amc***. This support point definition clearly introduces prior information. The elements of $\boldsymbol{B}$ will be drawn towards the centre of the support intervals by the entropy criterion as much as the data constraints allow. In addition we excluded (the theoretically impossible) negative values for the diagonal elements and restricted the cross effects to be small relative to the own activity level effects on marginal cost. Nevertheless, the spread of the support points specification leaves considerable freedom for obtaining a wide range of implied elasticities.

The determination of support points in the context of ME and GME (Generalised ME which includes error terms in data constraints) is a delicate problem and therefore deserves some further discussion: there seems to be a great desire to determine support points objectively and to avoid prior information as much as possible. Léon *et al.* (1999), for example, employ the normalised entropy measure to judge the « superiority » of different (predefined) symmetric and asymmetric support point specifications[15]. This measure reaches its maximum when the estimated parameters do not deviate at all from the *a priori* expectations defined by the support values[16]. Consequently, it allows to compare different sup-

[14] With this support point formulation, the linear terms $\boldsymbol{d}_r$ can actually be viewed as the sum of a predetermined parameter vector c_r and a crop and region specific error term which is centred around zero. Consequently, the specification is numerically equivalent to a generalised ME formulation with error terms. We opted for the representation above, because the « error term » is ultimately kept in the specification of the objective function so that the resulting programming models calibrate exactly to observed activity levels.

[15] See also Golan *et al*. for a discussion of « Normalised Entropy » and its use in various applications.

[16] Generally, prior expectations are defined as a weighted average of support values. In the ME case, the weights are probabilities following a uniform distribution. In the CE case, the weights are the probabilities as defined by the reference distribution.

port point specifications with respect to their compatibility with the data constraints. The measure does not allow, however, to identify an optimal set of support values for an underdetermined estimation problem. Just as there is an infinite number of parameter vectors satisfying the data constraints, there is as well an infinite number of support definitions with prior expectation equal to these parameter vectors. All these support point specifications obtain the same value of the normalised entropy measure, but not the same parameter estimates. Therefore, we did not consider this measure for the choice of support values here.

Other researches focussing on the idea of using purely « data-based » supports (van Akkeren *et al*., 2001) show advantages over classical estimation techniques in some ill-conditioned (*e.g*. multicollinear) data situations but well posed with respect to the number of observations. Those techniques obviously cannot make up for limited data information. From our point of view it should simply be accepted that a small number of observations relative to the number of parameters imply little information and that ME and GME succeed in these situations, only because they allow to flexibly incorporate prior information by restricting the parameter space. There is certainly the danger of introducing a strong bias if the prior is formulated very tight and far off the true value. In the GME context, however, it can be taken as some comfort that the estimator is consistent under general regularity conditions as long as the true value of the parameters is within the support range (see Mittelhammer and Cardell, 2000).

Above, we tried to make our *a priori* information as transparent as possible and chose to use a uniform distribution where the centres are the prior expectations. Note that this is numerically equivalent to a cross entropy (CE) approach with this uniform distribution serving as the reference distribution. Other possibilities to represent prior information include differentiated prior weights in the CE reference distribution or asymmetric support point spacing in ME and CE contexts. These methods provide flexibility in expressing just the prior information that is available, but – to our knowledge – there is no objective criterion that makes one approach generally superior to the others. At some point, there might be measures to compare the penalty involved for deviating from the prior information for the different approaches and this will improve transparency (see Preckel, 2001 for looking at the entropy criterion from a penalty view).

Returning from this general support point discussion to our specific case, we repeat that the specifications in (12) and (13) imply some prior information, but the support spread leaves considerable ranges for the parameters. Also the influence of the support points on the estimation outcomes becomes considerably smaller with an increasing number of observations and allowing for this to happen is a major objective of our approach in contrast to previous PMP applications.

The approach discussed in the previous section estimates a non-linear cost function depending on crop production activity levels based on observed regional differences in marginal costs at just one point in time. Naturally, doubt may be raised if that cross-sectional information can be just mapped in the time domain by assuming that changes in crop rotation over time in each single region have a similar effect on variable costs as the differences in observed crop rotations for a set of regions at one point of time. We consequently check the resulting simulation behaviour of the models in an ex-post simulation exercise.

We took three year averages both for the calibration and simulation year based on data in the CAPRI data base for the 22 NUTS-2 regions in France. Given data availability, we used years 1989 to 1991 (« 1990 ») for the calibration and 1993 to 1995 (« 1994 ») for the simulation. The move from 1991 to 1994 has the advantage that the 1992 CAP-reform lays just in between which offers a good opportunity to test the model under a significant policy change. However, some restrictions apply : We had no data on the participation in voluntary set-aside programs before the CAP-reform – therefore important information was left out in the calibration step. Naturally, no data on obligatory set-aside and non-food production, both introduced by the 1992 CAP-reform, entered the calibration for 1990. We therefore had to make some assumptions regarding these activities :

• The parameters in $\boldsymbol{d}$ and $\boldsymbol{B}$ relating to *voluntary and obligatory set aside* were set equal to the ones obtained for *fallow land* in 1990, assuming that they have the same rotational effects as represented by the cost function. Nevertheless, voluntary and obligatory set-aside are still treated in simulation according to the policy formulation in the CAP-reform, *i.e.* they are linked to the production of *« grandes cultures »* in the appropriate way (see below).

• The driving forces of *non-food production on set-aside* were unknown to us with respect to hard quantitative information. Therefore, we fixed non-food production to known levels in 1994. As non-food has a share around 10 % on oilseeds in total, the resulting improvement in the model's fit is not dramatic. We also applied this assumption to the other approaches which are compared to our ME-PMP calibrated model.

The set-aside regulation is modelled by constraints : the obligation must be fulfilled by an appropriate level of obligatory set-aside or non-food production on set-aside. Voluntary set-aside may be added as long as the sum of total set-aside including non-food production does not exceed 33 % of the endogenously determined *« grandes cultures »* area. Premia are cut if regional base areas are exceeded. As the presented ME-PMP approach is only suitable for annual crops, we fixed animal production and perennials to observed levels in 1994. Apart from the sugar beet quota and the land restriction, no other constraints enter the model specification.

The ME problem (9) was successfully solved with the General Algebraic Modelling System (GAMS) using the solver CONOPT2. It should be noted here that a powerful solver for this type of optimisation problem is necessary, especially due to the considerable non-linearity introduced by the Cholesky decomposition constraints (10).

We started the evaluation of the results with simulation experiments based on partial 10 % increases of product prices and calculated the aggregated national percentage change in area related to the price change. Table 2 shows selected elasticities which are somewhat comparable to the « classical » econometric estimates provided by Guyomard *et al*. (1996) with respect to product differentiation and scope. The estimates of *own price* elasticities are on average larger than their econometric counterparts (reported in brackets), but not uniformly so. The own price response of maize and soybeans is considerably below the values of Guyomard *et al.*. Generally, the estimated own price elasticities are smaller than the typical supply responses implied by LP's or standard PMP-procedures (see for example Cypris, 2000 and the subsequent simulation exercise). *Cross price* elasticities are also within the general magnitude of the econometric estimates, but they clearly show different structures of substitution between the crops. For example, with an increase of the soft wheat price, barley and rapeseed show the strongest (percentage) reductions in table 2. Those responses are rather small in the case of Guyomard *et al*., where maize is the main crop substituted by the increasing wheat production. One should not forget, however, that the theorical structure of the two underlying models (fixed versus variable input and output coefficients) as well as the employed data base (cross sectional versus time series) differ between the two sets of estimates which limits their comparability.

Table 2. **Price elasticities of supply for selected crops – National aggregate France**

	Soft Wheat	Maize	Barley	Rapeseed	Sunflower	Soya
Soft Wheat	**1.322**	-0.075	-0.443	-0.076	-0.039	-0.003
	[0.715]	[-0.303]	[-0.010]	[-0.007]	[-0.008]	[-0.001]
Maize	-0.165	**0.653**	-0.056	-0.004	-0.009	-0.003
	[-0.624]	[1.630]	[-0.041]	[-0.031]	[-0.038]	[-0.002]
Barley	-1.555	-0.105	**2.647**	-0.144	-0.075	-0.008
	[-0.042]	[-0.097]	[0.351]	[-0.002]	[-0.003]	[-0.000]
Rapeseed	-0.939	-0.041	-0.453	**1.457**	-0.065	-0.005
	[-0.079]	[-0.033]	[-0.025]	[0.428]	[-0.091]	[-0.017]
Sunflower	-0.540	-0.066	-0.216	-0.047	**1.126**	-0.006
	[-0.111]	[-0.046]	[-0.036]	[-0.048]	[0.223]	[-0.024]
Soya	-0.302	-0.224	-0.218	-0.036	-0.057	**1.861**
	[-0.351]	[-0.144]	[-0.112]	[-0.152]	[-0.403]	[3.701]

Supply in rows and changed prices in columns. Reported elasticities are calculated as average percentage supply change (change in land allocation due to fixed yields) per one percent price change. The simulations are based on a 10 % increase in the respective crop prices. Values in brackets are the (rounded) supply elasticity estimates reported in Table 2 of Guyomard *et al.* (1996).

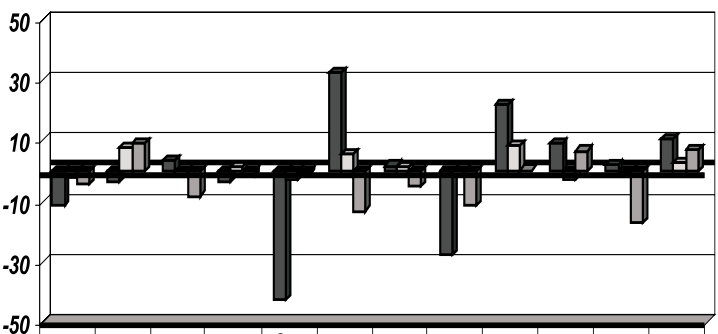
The comparison of the resulting partial supply responses with values of estimated behavioural functions is certainly interesting. However, the assessment of the simulation behaviour across larger economic and policy changes is closer to the ultimate purpose of our model. Therefore, we designed an ex-post simulation experiment as described above, results of which are rarely generated in the context of programming models, but are very informative from our point of view.

In order to judge if the new methodology has comparative advantages, we included a « standard PMP » approach in the ex-post validation as well. Here, only diagonal elements of $\boldsymbol{B}$ are specified such that the linear and quadratic terms for each production activity i implicitly define *average variable cost* matching the observed accounting cost c_i for the base year. In the case of the quadratic cost function this implies that

$$b_{i,i} = \frac{2\lambda_i}{x_i^0} \text{ and } d_i = c_i - \lambda_i \quad \forall\, i = 1, \ldots, n \tag{14}$$

Furthermore, we defined an « intelligent no change » forecast by taking 1990 levels of annual crops reducing them – where applicable – by set-aside obligations. The resulting areas were then made consistent to the available land in 1994.

Figure 1. **Percentage deviation of simulated from observed production activity levels for France**

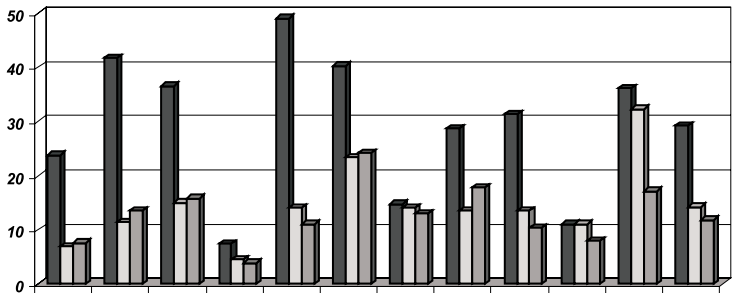


	Wheat	Barley	Maize	Cereals	Sun flower seeds	Rape seeds	Oil seeds	Pulses	Pota-toes	Arable Fodder	Fallow Set Aside	Total
Standard	-11	-3	3	-4	-42	33	2	-28	22	9	2	11
ME-PMP	-2	8	-2	1	-3	6	1	-1	8	-3	0	3
No change	-4	9	-8	-2	-1	-13	-5	-12	0	7	-17	7

Figure 1 shows the percentage national deviation of « simulated » production activity levels from the observed activity levels in 1994 for the three approaches. The « standard approach » shows rather high deviations for some major crops. Somewhat surprising, the « no-change » forecast is comparatively close to observed production activity levels. Apparently – the 1992 CAP-reform had – at least in France – a relatively small impact on the aggregate crop rotation apart from the set-aside

effect. With this in mind, the fit of the ME-PMP approach based on the cross sectional sample is rather promising: apart from sunflowers and potatoes it provides better simulated values than the « no-change » results. The sum of absolute deviation in levels weighted by the observed levels amounts just up to 3 % (see « Total »).

Figure 2.
Mean absolute percentage deviations of simulated from observed production activity levels across regions



	Wheat	Barley	Maize	Cereals	Sun flower seeds	Rape seeds	Oil seeds	Pulses	Potatoes	Arable Fodder	Fallow Set Aside	Total
Standard	24	42	37	7	49	40	15	29	31	11	36	29
ME-PMP	7	12	15	5	14	23	14	14	14	11	32	14
No change	8	14	16	4	11	24	13	18	10	8	17	12

However, the variation in regional forecasting errors is at least of the same importance. The results in figure 1 could be a rather « lucky » outweighing of regionally large under- and overestimating of activity levels. Therefore, we additionally checked the regional fit by calculating mean absolute percentage deviations over regions presented for the most important activities and aggregates in figure 2. The standard approach was again no real competitor. However, the performance of the ME-PMP approach is about the same as « no-change » apart from the aggregate of fallow land and set-aside. As explained above, problems could be expected here as no substantial information entered the calibration step.

CONCLUSIONS

So far, most PMP application in aggregate programming models suffered from a rather arbitrary specification of the non-linear objective function. Classical econometric approaches cannot be applied to this typically underdetermined estimation problem, a problem overcome by using the Maximum Entropy criterion as proposed in Paris and Howitt (1998). Their application, however, included just one observation on marginal cost and additionally suffered from a non-intuitive definition of supports based on their specific approach to ensure the correct curvature of the estimated cost function.

These problems were addressed by the ME-PMP approach presented in this paper which uses a cross-sectional sample in order to derive

changes in marginal cost based on observed differences between regions with different crop rotations and provides a solution for the curvature problem with limited computational burden and direct definition of support points for the parameters of interest.

An ex-post validation of the resulting model specification simulated the 1992 CAP-reform for crop production in France. The results show a promising fit of observed production activity levels – not only for the national aggregate, but as well in the regional dimension. The ex-post simulation exercise – rarely executed or published in the context of aggregate programming models – shows the validity of the calibration procedure for regional programming models. The allocation behaviour of the resulting models is clearly superior to a standard application of PMP.

Nevertheless, the general approach leaves ample opportunities for further research: Specific issues on our research agenda are the additional use of time series observations to extend the information base and to estimate time dependant parameters, the introduction of estimated parameters to describe the relationship between crop rotation and changes in marginal cost and, last but not least, an explicit elaboration on the links between PMP and duality based econometric models with explicit allocation of land to different production activities. The last issue could potentially improve the theoretical understanding of PMP which originated as a rather pragmatic solution to calibration problems with agricultural programming models.

REFERENCES

ARFINI (F.), PARIS (Q.), 1995 — A positive mathematical programming model for regional analysis of agricultural policies, *in*: SOTTE (E.) (ed.), *The Regional Dimension in Agricultural Economics and Policies*, EAAE, Proceedings of the 40th Seminar, June 26-28, Ancona, pp. 17-35.

BARKAOUI (A.), BUTAULT (J.-P.), 1999 — Positive mathematical programming and cereals and oilseeds supply within EU under Agenda 2000, Paper presented at the 9th European Congress of Agricultural Economists, Warsaw, August.

CYPRIS (Ch.), 2000 — Positiv Mathematische Programmierung (PMP) im Agrarsektormodell RAUMIS, Dissertation, University of Bonn.

GILL (P. E.), MURRAY (W.) and WRIGHT (M. H.), 1989 — *Practical Optimisation*, London, Academic Press.

GOHIN (A.), CHANTREUIL (F.), 1999 — La programmation mathématique positive dans les modèles d'exploitation agricole. Principes et importance du calibrage, *Cahiers d'économie et sociologie rurales,* 52, pp. 59-78.

GOLAN (A.), JUDGE (G.) and MILLER (D.), 1996 — *Maximum Entropy Econometrics*, Chichester, Wiley.

GOLUB (G. H.), VAN LOAN (C. F.), 1996 — *Matrix Computations*, 3rd edition, Baltimore, John Hopkins University Press.

GRAINDORGE (C.), DE FRAHAN (B.-H.) and HOWITT (R. E.), 2001 — Analysing the effects of Agenda 2000 using a CES calibrated model of Belgian agriculture, *in:* HECKELEI (T.), WITZKE (H. P.) and HENRICHSMEYER (W.) (eds), 2000, *Agricultural Sector Modelling and Policy Information Systems*, Proceedings of the 65th EAAE Seminar, March 29-31, 2000, Kiel, Wissenschaftsverlag Vauk, pp. 177-186.

GUYOMARD (H.), BAUDRY (M.) and CARPENTIER (A.), 1996 — Estimating crop supply response in the presence of farm programmes: application to the CAP, *European Review of Agricultural Economics*, 23, pp. 401-420.

HECKELEI (T.), 1997 — Positive mathematical programming: Review of the standard approach, CAPRI-working paper 97-03, University of Bonn.

HECKELEI (T.), BRITZ (W.), 2001 — Concept and explorative application of an EU-wide regional agricultural sector model (CAPRI-Projekt), *in*: HECKELEI (T.), WITZKE (H. P.), and HENRICHSMEYER (W.) (eds), *Agricultural Sector Modelling and Policy Information Systems*, Proceedings of the 65th EAAE Seminar, March 29-31, 2000, Kiel, Wissenschaftsverlag Vauk, pp. 281-290.

HELMING (J. F. M.), PEETERS (L.) and VEENDENDAAL (P. J. J.), 2001 — Assessing the consequences of environmental policy scenarios in Flemish agriculture, *in*: HECKELEI (T.), WITZKE (H. P.), and HENRICHSMEYER (W.) (eds), *Agricultural Sector Modelling and Policy Information Systems*, Proceedings of the 65th EAAE Seminar, March 29-31, 2000, Kiel, Wissenschaftsverlag Vauk, pp. 237-245.

HORNER (G. L.), CORMAN (J.), HOWITT (R. E.), CARTER (C. A.) and MACGREGOR (R. J.), 1992 — The Canadian regional agriculture model: Structure, operation and development, Agriculture Canada, Technical Report 1/92, Ottawa.

HOUSE (R. M.), 1987 — USMP regional agricultural model, Washington DC, USDA, National Economics Division Report, ERS, 30.

HOWITT (R. E.), 1995a — Positive mathematical programming, *American Journal of Agricultural Economics*, 77 (2), pp. 329-342.

HOWITT (R. E.), 1995b — A calibration method for agricultural economic production models, *Journal of Agricultural Economics,* 46, pp. 147-159.

HOWITT (R. E.), GARDNER (D. B.), 1986 — Cropping production and resource interrelationships among California crops in response to the 1985 food security act, *in: Impacts of Farm Policy and Technical Change on US and Californian Agriculture*, Davis, pp. 271-290.

KASNAKOGLU (H.), BAUER (S.), 1988 — Concept and application of an agricultural sector model for policy analysis in Turkey, *in:* BAUER (S.), HENRICHSMEYER (W.) (eds), *Agricultural Sector Modelling,* Kiel, Wissenschaftsverlag Vauk.

LENCE (H. L.), MILLER (D.), 1998a — Estimation of multioutput production functions with incomplete data: A generalized maximum entropy approach, *European Review of Agricultural Economics*, 25, pp. 188-209.

LENCE (H. L.), MILLER (D.), 1998b — Recovering output specific inputs from aggregate input data: A generalized cross-entropy approach, *American Journal of Agricultural Economics*, 80 (4), pp. 852-867.

LÉON (Y.), PEETERS (L.), QUINQU (M.) and SURRY (Y.), 1999 — The use of maximum entropy to estimate input-output coefficients from regional farm accounting data, *Journal of Agricultural Economics,* 50, pp. 425-439.

MITTELHAMMER (R. C.), CARDELL (S.), 2000 — The data-constrained GME-estimator of the GLM: Asymptotic theory and inference, Working paper of the Department of Statistics, Washington State University, Pullman.

MITTELHAMMER (R. C.), JUDGE (G. G.) and MILLER (D. J.), 2000 — *Econometric Foundations*, New York, Cambridge University Press.

OUDE LANSINK (A.), 1999 — Generalized maximum entropy and heterogeneous technologies, *European Review of Agricultural Economics*, 26, pp. 101-115.

PARIS (Q.), HOWITT (R. E.), 1998 — An analysis of ill-posed production problems using maximum entropy, *American Journal of Agricultural Economics*, 80-1, pp. 124-138.

PARIS (Q.), MONTRESOR (E.), ARFINI (F.) and MAZZOCCHI (M.), 2000 — An integrated multi-phase model for evaluating agricultural policies through positive information, *in*: HECKELEI (T.), WITZKE (H. P.), and HENRICHSMEYER (W.) (eds), *Agricultural Sector Modelling and Policy Information Systems*, Proceedings of the 65th EAAE Seminar, March 29-31, 2000, Kiel, Wissenschaftsverlag Vauk, pp. 101-110.

PRECKEL (P. V.), 2001 — Least squares and entropy: A penalty function perspective, *American Journal of Agricultural Economics*, 83-2, pp. 366-377.

SCHMITZ (H. J.), 1994 — Entwicklungsperspektiven der Landwirtschaft in den neuen Bundesländern Regionaldifferenzierte Simulationsanalysen, Alternativer Agrarpolitischer Szenarien, Studien zur Wirtschafts — und Agrarpolitik, Bonn, Witterschlick.

VAN AKKEREN (M.), JUDGE (G. G.) and MITTELHAMMER (R. C.), 2001 — Coordinate based empirical likelihood-like estimation in ill-conditioned inverse problems, Working paper of the University of California, Berkeley and Washington State University, Pullman.

ZHANG (X.), FAN (S.), 2001 — Estimating crop-specific production technologies in Chinese agriculture: A generalized maximum entropy approach, *American Journal of Agricultural Economics*, 83-2, pp. 378-388.