European Review of Agricultural Economics Vol **30 (1)** (2003) pp. 27–50

Estimation of constrained optimisation models for agricultural supply analysis based on generalised maximum entropy

Thomas Heckelei and Hendrik Wolff

Washington State University, Pullman, WA, USA

Received January 2002; final version received January 2003

Abstract

The paper introduces a general methodological approach for the estimation of constrained optimisation models in agricultural supply analysis. It is based on optimality conditions of the desired programming model and shows a conceptual advantage compared with Positive Mathematical Programming in the context of well-posed estimation problems. Moreover, it closes the empirical and methodological gap between programming models and duality-based models with explicit allocation of fixed factors. Monte Carlo simulations are performed with a maximum entropy estimator to evaluate the functionality of the approach as well as the impact of empirically relevant prior information with small samples.

Keywords: agricultural supply analysis, mathematical programming models, maximum entropy estimation, prior information, constrained optimisation models

JEL classification: C51, C61, Q12

1. Introduction

Quantitative models of multi-output multi-input cropping decisions in agriculture typically belong to one of two main methodological types: either programming models or dual systems of supply and input demand equations. The former determine input allocation to various production activities using an explicit optimisation, the latter constitute closed-form solutions to economic optimisation models. Maintained economic hypotheses and objectives do not necessarily have to differ between these types.[1] However, in empirical work the structure and specification procedures are clearly distinguished: a programming model is chosen when the analyst sees the need to explicitly model complex technological or policy constraints under which behavioural functions cannot be derived easily or at all. This generally comes at a cost:

1 Programming models are often characterised as *normative* because they use an explicit optimisation. This neither reflects the original meaning nor is it a very useful distinction. The objective of normative analysis is to say 'what should be' and thus farm or regional planning models qualify for this category. However, programming models designed to explain or project behaviour are *positive* in nature. Furthermore, an integrable dual supply system could just as well be used as an explicit optimisation model for simulation and yield exactly the same results.

the inability to perform statistical estimation and validation for the whole model. Dual equation systems, on the other hand, allow well-established econometric techniques to be applied so that parametric specification can be based on observed supply and demand decisions of agricultural producers. This choice limits the model's complexity and potentially oversimplifies for the purpose of a differentiated analysis.

During the last decade these two methodological approaches seem to have moved a little closer to each other. Chambers and Just (1989) developed a dual supply model specification with explicit allocation of fixed factors. This allowed a previous deficiency in modelling crop supply to be overcome, by incorporating land constraints and the observable decision variable 'land allocated to production activities'. It also provided a useful framework for modelling the European policy instrument 'hectare premium' separately from product price effects (Guyomard *et al.*, 1996; Moro and Sckokai, 1999). Nevertheless, additional constraints cannot be easily incorporated and the choice of functional form is restricted because of analytical limitations in deriving the behavioural functions to be estimated. From the programming side, Howitt (1995a) presented 'Positive Mathematical Programming' (PMP), which allows models to be calibrated on observed behaviour of a base year. PMP established itself as a widely used approach for the specification of programming models designed for policy analysis.[2] The incorporation of several observations employing an econometric criterion was generally made possible by Paris and Howitt (1998) and applied to a cross-sectional data set by Heckelei and Britz (2000). However, the theoretical base of this approach is weak or at least not apparent.

This paper aims at moving the two methodological approaches closer together. We present a general approach to estimating parameters of constrained programming models for agricultural supply analysis based on optimality conditions of the desired model. This method provides a consistent alternative to PMP for the specification of programming models. Simultaneously, it allows the estimation of models with multiple constraints that cannot be solved to obtain behavioural functions. The methodology thus supports specifications of more complex models and a more flexible choice of functional form compared with previous econometric approaches with explicit allocation of fixed factors. The paper is organised as follows: Section 2 explains why PMP is not well suited to the estimation of programming models based on multiple observations. Section 3 describes a general alternative. Section 4 illustrates the approach for three optimisation models that stem from the programming and econometric literature. It provides Monte Carlo simulation results to demonstrate functionality with solely data-based estimates. In addition, approaches using prior information exploit the potential of maximum entropy techniques in this context and address the problem of limited sample sizes often confronted by differentiated

2 For examples of PMP application see Howitt and Gardner (1986), House (1987), Kasnakoglu and Bauer (1988), Arfini and Paris (1995), Cypris (2000) and Helming *et al.* (2001).

modelling exercises. Section 5 concludes, discusses limitations and identifies promising directions for further research.

2. Positive Mathematical Programming: short review and critique

The general idea of PMP is to employ dual values of calibration constraints that force the optimisation model to observed outcomes of endogenous variables (step 1). These dual values are used to specify additional non-linear terms in the objective function that allow the observed outcomes to be reproduced exactly without calibration constraints (step 2). Starting from a typical linear program (LP) in agricultural supply analysis, step 1 can be illustrated as

$$\max_{\mathbf{l}} Z = \mathbf{p}'\mathbf{l} - \mathbf{c}'\mathbf{l} \quad \text{subject to} \quad \mathbf{A}\mathbf{l} \leq \mathbf{b}\,[\boldsymbol{\lambda}], \quad \mathbf{l} \leq (\mathbf{l}^{\mathrm{o}} + \boldsymbol{\varepsilon})\,[\boldsymbol{\rho}], \quad \mathbf{l} \geq \mathbf{0} \tag{1}$$

where Z is the objective function value, $\mathbf{p}$, $\mathbf{l}$ and $\mathbf{c}$ are $(N \times 1)$ vectors of product prices, non-negative activity levels and variable costs per activity unit, respectively. $\mathbf{A}$ represents an $(M \times N)$ matrix of coefficients, $\mathbf{b}$ and $\boldsymbol{\lambda}$ are $(M \times 1)$ vectors of resource availability and their corresponding shadow prices. The $(N \times 1)$ vector $\mathbf{l}^{\mathrm{o}}$ contains observed activity levels in a base period, $\boldsymbol{\varepsilon}$ is an $(N \times 1)$ vector of small numbers and $\boldsymbol{\rho}$ $(N \times 1)$ contains the dual variables of the calibration constraints. In the second step of PMP, the dual values $\boldsymbol{\rho}$ are used to specify a non-linear variable cost function $C^{\mathrm{V}}(\mathbf{l}^{\mathrm{o}})$, such that the 'variable' marginal cost $\mathbf{MC}^{\mathrm{V}}(\mathbf{l}^{\mathrm{o}})$ of the activities is equal to the sum of the known cost $\mathbf{c}$ and the 'non-specified marginal cost' $\boldsymbol{\rho}$. In case of the frequently used quadratic functional form, the following condition for calibration is implied:

$$\mathbf{MC}^{\mathrm{V}} = \frac{\partial C^{\mathrm{V}}(\mathbf{l}^{\mathrm{o}})}{\partial \mathbf{l}} = \mathbf{d} + \mathbf{Q}\mathbf{l}^{\mathrm{o}} = \mathbf{c} + \boldsymbol{\rho} \tag{2}$$

where the $(N \times 1)$ vector $\mathbf{d}$ and the $(N \times N)$ symmetric positive definite matrix $\mathbf{Q}$ correspond to the linear and quadratic terms of $C^{\mathrm{V}}(\mathbf{l}^{\mathrm{o}})$, respectively. This condition does not include the opportunity cost of using fixed resources, because they are still accounted for by the resource constraints in the resulting model:

$$\max_{\mathbf{l}} Z = \mathbf{p}'\mathbf{l} - \mathbf{d}'\mathbf{l} - 0.5\mathbf{l}'\mathbf{Q}\mathbf{l} \quad \text{subject to} \quad \mathbf{A}\mathbf{l} \leq \mathbf{b}\,[\boldsymbol{\lambda}], \quad \mathbf{l} \geq \mathbf{0}. \tag{3}$$

To solve the underdetermined system (2) with $N + N(N+1)/2$ parameters and N equations, the literature suggests many approaches, which include simple *ad hoc* procedures with some parameters set *a priori* (Howitt, 1995a), the use of supply elasticities (Helming *et al.*, 2001) and the employment of a maximum entropy criterion (Paris and Howitt, 1998). As long as conditions (2) are satisfied, the calibration of the resulting optimisation model is guaranteed, but the different specifications of $\mathbf{d}$ and $\mathbf{Q}$ imply significant differences with respect to the simulation behaviour (see Cypris, 2000, or Heckelei and Britz, 2000).

However, this paper is concerned not with calibration but with *estimation* of programming models. Paris and Howitt (1998) and Paris (2001) already suggest the possibility that more than one observation on production outcomes could be incorporated, implying a set of N marginal cost conditions (2) for *each* observation. Heckelei and Britz (2000) use this idea for the estimation of regional cost functions based on a cross-sectional sample.

Here we want to argue that the PMP procedure is not well suited to exploiting the additional data information, because the derived marginal cost conditions do not allow consistent estimation of the parameters. For this purpose, it is useful to look at PMP from the perspective of an econometrician. This implies having some idea of a 'true' model, or at least the assumption that a specific model is capable of representing the true data generating process adequately. Many PMP modellers have apparently believed this with regard to the resulting non-linear model ultimately used to perform economic analysis.

To show the inconsistency of PMP we take, for example, the quadratic model (3) and assume exclusively positive activity levels and binding resource constraints at the optimal solution. The first-order conditions imply the shadow price values:

$$\boldsymbol{\lambda} = (\mathbf{AQ}^{-1}\mathbf{A}')^{-1}(\mathbf{AQ}^{-1}(\mathbf{p} - \mathbf{d}) - \mathbf{b}). \tag{4}$$

In the first step of PMP, equation (1), a different result is obtained: we partition $\mathbf{l}$ into two subvectors, an $((N - M) \times 1)$ vector of 'preferable' activities, $\mathbf{l}^{\mathrm{p}}$, bounded by the calibration constraints and an $(M \times 1)$ vector of marginal activities, $\mathbf{l}^{\mathrm{m}}$, bounded by the resource limits. Then the dual values

$$\text{(a)}\ \boldsymbol{\lambda} = (\mathbf{A}^{\mathrm{m}\prime})^{-1}(\mathbf{p}^{\mathrm{m}} - \mathbf{c}^{\mathrm{m}}), \quad \text{(b)}\ \boldsymbol{\rho}^{\mathrm{p}} = \mathbf{p}^{\mathrm{p}} - \mathbf{c}^{\mathrm{p}} - \mathbf{A}^{\mathrm{p}\prime}\boldsymbol{\lambda}, \quad \text{(c)}\ \boldsymbol{\rho}^{\mathrm{m}} = \mathbf{0} \tag{5}$$

can be derived. Notice that in (2) $\boldsymbol{\lambda}$ is exclusively determined by objective function entries and technological coefficients of the marginal activities $\mathbf{p}^{\mathrm{m}}$, $\mathbf{c}^{\mathrm{m}}$ and $\mathbf{A}^{\mathrm{m}}$. Therefore, they are generally different from the values of $\boldsymbol{\lambda}$ implied by the quadratic model, because in (1) $\boldsymbol{\lambda}$ is determined by $\mathbf{p}$, $\mathbf{d}$, $\mathbf{A}$, $\mathbf{b}$ and $\mathbf{Q}$. Thus, the value of $\boldsymbol{\lambda}$ calibrated by step 1 of PMP (5), is expected to be different from the one implied by the model assumed to represent farmer behaviour and used as the final simulation model, i.e. model (2). Now, because step 1 of PMP sets $\boldsymbol{\rho}$ simultaneously with $\boldsymbol{\lambda}$ (5b) and step 2 uses $\boldsymbol{\rho}$ to specify $\mathbf{MC}^{\mathrm{V}}$, the latter vector is also generally inconsistent with model (2). Consequently, the set of equations (2) cannot be seen as unbiased estimating equations and will generally yield inconsistent parameter estimates if the true data generating process is correctly described by the quadratic model.

3. A general alternative

Our suggested 'general alternative' to PMP relies on a simple principle. It directly employs the optimality conditions of the desired programming model. No 'step 1' for the determination of dual values of calibration constraints is necessary. Instead, the simultaneous estimation of shadow prices and parameters avoids methodological inconsistencies.

The basic principle can be illustrated by writing the programming model as a general Lagrangian form with an objective function $h(\mathbf{y}|\boldsymbol{\alpha})$ to be optimised subject to a constraint vector $\mathbf{g}(\mathbf{y}|\boldsymbol{\beta}) = \mathbf{0}$:

$$L(\mathbf{y}, \boldsymbol{\lambda}|\boldsymbol{\alpha}, \boldsymbol{\beta}) = h(\mathbf{y}|\boldsymbol{\alpha}) + \boldsymbol{\lambda}'[\mathbf{g}(\mathbf{y}|\boldsymbol{\beta})]$$

where $\mathbf{y}$, $\boldsymbol{\lambda}$, $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ represent column vectors of endogenous variables, unknown dual values, parameters of the objective function and parameters of the constraints, respectively. The appropriate first-order optimality conditions are

$$\frac{\partial L}{\partial \mathbf{y}} = \frac{\partial h(\mathbf{y}|\boldsymbol{\alpha})}{\partial \mathbf{y}} + \boldsymbol{\lambda}' \frac{\partial \mathbf{g}(\mathbf{y}|\boldsymbol{\beta})}{\partial \mathbf{y}} = \mathbf{0}$$

$$\frac{\partial L}{\partial \boldsymbol{\lambda}} = \mathbf{g}(\mathbf{y}|\boldsymbol{\beta}) = \mathbf{0}.$$

For the case of inequality constraints $\mathbf{g}(\mathbf{y}|\boldsymbol{\beta}) \leq \mathbf{0}$, we need to substitute the gradient with respect to $\boldsymbol{\lambda}$ by the complementary slackness representation[3]

$$\frac{\partial L}{\partial \boldsymbol{\lambda}} = \mathbf{g}(\mathbf{y}|\boldsymbol{\beta}) \leq \mathbf{0}; \quad \boldsymbol{\lambda} \odot \mathbf{g}(\mathbf{y}|\boldsymbol{\beta}) = \mathbf{0}.$$

The unknowns $\boldsymbol{\lambda}$, $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ of these Kuhn-Tucker conditions can be estimated with some econometric criteria applied directly to these equations. Depending on the parametric specification, appropriate curvature restrictions (second-order conditions) might have to be enforced as well.

The direct use of optimality conditions for estimation is certainly not new in itself. In the context of investment models, for example, the Euler equations (dynamic equivalents of Kuhn–Tucker conditions) have been frequently used as estimating equations to overcome analytical and empirical problems for more complex models (Chirinko, 1993, p. 1893f). However, their use as an alternative to PMP or to the estimation of behavioural functions in the context of multi-output agricultural supply models has not been considered. One of the examples in the next section will show that not only is this approach useful for estimating typical primal agricultural programming models but that it also provides a flexible alternative for estimating parameters of duality-based behavioural functions with explicit allocation of fixed factors. In this context, the only difference between programming and econometric models is the model form used for simulation purposes.

It is perhaps not surprising that one of the most innovative PMP experts has already used this principle in part for the calibration of an agricultural supply model (Paris, 2001). Paris's 'Symmetric Positive Equilibrium Problem' calibrates a multi-input multi-output profit model on the basis of the marginal cost conditions. However, his approach still uses a 'first phase' to determine dual values of calibration constraints with the aforementioned inconsistency to the final model specification. Furthermore, unlike the models considered here, Paris's ultimate simulation model is not based on an optimisation

3 The symbol '$\odot$' represents the Hadamard or element-wise product of two matrices. If a_{ij} and b_{ij} are the elements of two matrices with equal dimension, **A** and **B**, then $A \odot B = C$, where **C** is of the same dimension as **A**, **B** and each element of **C** is defined as $c_{ij} = a_{ij} \times b_{ij} \, \forall \, i, j$.

hypothesis. Finally, our examples in the next section differ from Paris's in that our models all imply the existence of at least one fixed factor.[4]

4. Examples and Monte Carlo evidence

The models presented in this section are intended to illustrate our proposal for the estimation of constrained programming models. The models featured are not necessarily the most useful models for agricultural supply analysis, but are rather chosen to span the literature on programming models and econometric models with explicit allocation of fixed factors. Monte Carlo simulations based on artificial data and known parameters show the performance of the estimation approach through statistical evaluation of errors made in estimating the true parameters. The use of the Generalised Maximum Entropy (GME) estimator allows the influence of prior information on estimation results to be assessed in situations with limited data information. It also has the advantage of computational stability, especially in the non-linear model context encountered in the applications below.

4.1. Land allocation with quadratic cost function

This subsection deals with estimating the parameters of the optimisation model employing a quadratic cost function often used in the PMP context and already described above. For simplicity, we consider only the resource land as fixed, obtaining a quadratic programming model (QP model) with a scalar shadow price. In addition, we replace the vector of prices $\mathbf{p}$ by a vector of gross margins $\mathbf{gm}$[5] to obtain

$$\max_{\mathbf{l}} Z = \mathbf{gm}'\mathbf{l} - \mathbf{d}'\mathbf{l} - 0.5\mathbf{l}'\mathbf{Q}\mathbf{l} \quad \text{subject to} \quad \mathbf{u}'\mathbf{l} \leq b\,[\lambda], \quad \mathbf{l} \geq \mathbf{0} \tag{6}$$

with the $(N \times 1)$ summation vector $\mathbf{u}$, i.e. a vector of ones.

If we assume that the optimal land allocations satisfy the land constraint as an equality for every observation $t = 1, \ldots, T$, and that observed land allocations, $\mathbf{l}_t^{\mathrm{o}}$, are obtained from optimal values by adding an $(N \times 1)$ vector of stochastic errors $\mathbf{e}_t$ with mean zero and standard deviation σ_i, we can write the first-order conditions as[6]

$$\mathbf{gm}_t^{\mathrm{o}} - \lambda_t \mathbf{u} - \mathbf{d} - \mathbf{Q}(\mathbf{l}_t^{\mathrm{o}} - \mathbf{e}_t) = \mathbf{0}, \qquad \mathbf{u}'(\mathbf{l}_t^{\mathrm{o}} - \mathbf{e}_t) = b_t^{\mathrm{o}}. \tag{7}$$

This error specification can be motivated in several different ways. It might represent a measurement error of the variable or an optimisation error by the farmer, or stem from specific circumstances relevant to the optimal allocation of the respective economic unit unknown to the econometrician,

4 For a critical discussion of Paris (2001) see Britz *et al.* (2003).

5 The quadratic cost function represents 'some' unknown non-linear cost, which is independent of the variable inputs per activity unit. This lack of rationalisation in the model is analogous to many PMP applications. The illustration based on this model by no means indicates the preferability of the model.

6 Here and subsequently equations are valid for all elements in the respective indices, i.e. in this specific case for all $t = 1, \ldots, T$.

or some combination of these factors. It implies that the optimum for each observation is obtained at some unobserved level of the decision variable and that the deviation of the true optimum from the observed allocation is randomly distributed across observations.[7]

For estimation, we use the Generalised Maximum Entropy (GME) approach (Golan, Judge and Miller, 1996), which incorporates out-of-sample information in a flexible way.[8] We reparameterise the error vectors as expected values of a discrete probability distribution. The $(N \times (N \cdot 2))$ matrix $\mathbf{V}$ with $S = 2$ support points for each error term bounds the support to ± 5 standard deviations.[9] For the simulation experiments below we have $N = 3$ crops so that the error terms can be represented as the multiplication of $\mathbf{V}$ with an $((N \cdot S) \times 1)$ vector of probabilities $\mathbf{w}_t$ to obtain

$$\mathbf{e}_t = \mathbf{V}\mathbf{w}_t = \begin{bmatrix} -5\sigma_1 & 5\sigma_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -5\sigma_2 & 5\sigma_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & -5\sigma_3 & 5\sigma_3 \end{bmatrix} \begin{bmatrix} w_{11t} \\ w_{12t} \\ w_{21t} \\ w_{22t} \\ w_{31t} \\ w_{32t} \end{bmatrix}. \tag{8}$$

The complete GME formulation is then[10]

$$\max_{\mathbf{w}_t, \mathbf{Q}, \mathbf{L}, \lambda_t} \mathrm{H}(\mathbf{w}_t) = -\sum_{t=1}^{T} \mathbf{w}_t' \ln \mathbf{w}_t \tag{9}$$

subject to

$$\mathbf{gm}_t^{o} - \lambda_t \mathbf{u} - \mathbf{Q}(\mathbf{l}_t^{o} - \mathbf{V}\mathbf{w}_t) = 0 \tag{10}$$

$$\mathbf{u}'(\mathbf{l}_t^{o} - \mathbf{V}\mathbf{w}_t) = b_t^{o} \tag{11}$$

$$\mathbf{Q} = \mathbf{L}\mathbf{L}' \quad \text{with } L_{ij} = 0 \; \forall j > i \tag{12}$$

$$\sum_{s=1}^{S} \mathbf{w}_{its} = 1 \tag{13}$$

7 The stochastic specification of the optimisation model is maintained in the estimation, because the first-order conditions are used as estimating equations directly. It avoids potential inconsistency problems of many stochastic specifications of supply and input demand equations derived from dual indirect objective functions (on this issue see, for example, McElroy (1987), or more recently Pope and Just (2002)).

8 In this context of 'well-posed' estimation problems with more observations than parameters to be estimated, classical techniques such as least squares could have been applied as well.

9 The 'right' number of support points as well as the range of the support is an often discussed but not ultimately solved question. We chose two support points here mainly to restrict the computational demands in the already complex Monte Carlo simulation exercises, despite the fact that three or four support points promise a limited reduction of the mean estimation error (Golan, Judge and Miller, 1996, pp. 139–140). With respect to the support range, Golan *et al*. suggest the '3-sigma' rule. Preckel (2001) advocates a rather large range to approximate the behaviour of the least-squares estimator.

10 For the current case of just one resource constraint, the vector **d** is not identified. Therefore its elements are set to zero. See Appendix for further details.

where $\mathrm{H}(\mathbf{w}_t)$ denotes entropy, equation (12) guarantees the positive definiteness of $\mathbf{Q}$ based on a Cholesky factorisation,[11] and (13) ensures that the probabilities add up to one. Note that we do not need any reparameterisation of model parameters, because we consider only 'well-posed' problems with positive degrees of freedom in our simulations (Preckel (2001) denotes this estimator as $\mathrm{GME}_{-\beta}$). At this point, the parameters and shadow values are freely chosen so as to maximise the entropy related to the error terms. Later, we consider situations where support points for functions of parameters or for shadow prices of fixed factors are employed to incorporate prior information.

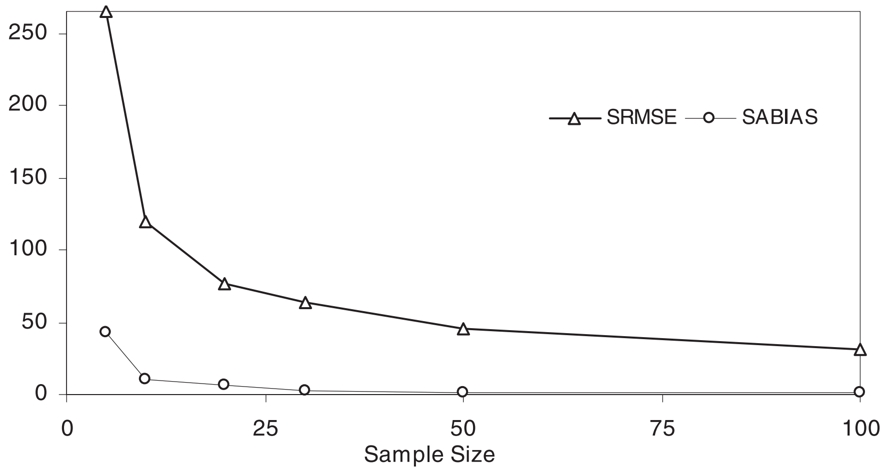
A Monte Carlo simulation experiment is used to test the estimator's precision. Based on the output and input differentiation in Howitt (1995b),[12] a data set with T observations is generated for T different random vectors $\mathbf{gm}_t$ and $\mathbf{b}_t$ for given parameters $\mathbf{Q}$. Normally distributed errors are added to the optimal land allocations ℓ^*_{1t} and ℓ^*_{2t} of the first two crops with a standard deviation of 2 per cent of the average land allocation, so that the 'observed' allocations are calculated as $\ell^o_{1t} = \ell^*_{1t} + e_{1t}$ and $\ell^o_{2t} = \ell^*_{2t} + e_{2t}$. To ensure that the land restriction is binding at the observed production activity levels we set $\ell^o_{3t} = b_t - \ell^*_{1t} - \ell^*_{2t}$.

An anonymous referee rightly suggested that our standard deviation for the land allocations is rather small and that correlation between the first two errors is also a more appropriate setup. We agree that a combination of the error types mentioned above would suggest a larger variance, and we present results for larger standard deviations below in Figure 2. However, we chose 2 per cent as the standard case across most simulations presented in the paper, because the stochastic specification of the following models is more consistent with the interpretation of the land allocation error as measurement error only (see footnote 15). We performed sensitivity analyses with correlation between the errors of the first two crops and concluded that it produced no changes of relevance for the conclusions drawn.

For each generated data set, the model parameters are estimated with the GME approach and the whole procedure is repeated 1000 times for each sample size. The quality of the estimation is evaluated using the measures *absolute bias* (ABIAS = absolute value of the difference between average estimate across samples and true value of the parameter) and root mean square error (RMSE = square root of the mean squared distance between estimates and true parameter). To summarise the results, the measures are summed over all estimated parameters (here all elements of $\mathbf{Q}$) to obtain 'SABIAS' and 'SRMSE'.

11 Lau (1978) introduced the Cholesky factorisation to ensure curvature in supply analysis by using a reparameterised model. Our use of it as an explicit constraint during estimation was originally suggested by Gallant and Golub (1984).

12 See the Appendix for the basic data set from Howitt, which reflects the values of non-random model variables and the means of random variables in the simulation exercises.



Source: Own calculations.

Figure 1. QP model—SRMSE and SABIAS without prior information.

Figure 1 presents the results for different sample sizes. SRMSE decreases with increasing sample size, indicating consistency of the estimator.[13] The bias (SABIAS) reaches negligible values already at a sample size of 20. Recalling that the MSE is the sum of the squared bias and the variance of the estimator, Figure 1 shows that the bias reflects only a small fraction of the RMSE and the much more important part of the MSE is given by the standard errors of the estimates. For small sample sizes, this could result in very poor estimates. In this case, the use of out-of-sample information is a potential remedy. Ideally, the use of prior information would reduce the estimator's variance at small sample sizes without introducing a strong additional bias.

To achieve a better feel for the required precision of the prior information and the general interplay between prior and data in our modelling context, we further extended the simulations. An often feasible procedure for incorporating out-of-sample information is the use of priors on elasticities, as the literature often provides some idea of their range. Elasticities can be reparameterised in the same way as the error terms. For the current model, we can employ the following analytical expression for the $(N \times 1)$ vector of land allocation elasticities with respect to own gross margins $\boldsymbol{\varepsilon}$:[14]

13 Mittelhammer and Cardell (2000) prove consistency of the GME approach for the general linear model under mild regularity conditions. No such general theoretical result is known to us for non-linear models except for the special case of the multinomial model (see Golan, Judge and Perloff, 1996).

14 The expression for the gradient $\partial \mathbf{l}/\partial \mathbf{gm}$ is obtained by first solving the first-order conditions (7)—ignoring the error terms—for the vector of land allocations $\mathbf{l}$ in terms of $\mathbf{gm}$ and $\mathbf{Q}$. Then, the (matrix) derivative of this expression with respect to $\mathbf{gm}$ is taken.

$$\boldsymbol{\varepsilon} = \mathrm{diag}\left(\frac{\partial \mathbf{l}}{\partial \mathbf{gm}} \odot \left[\frac{\overline{\mathbf{gm}}}{\overline{\mathbf{l}}^{\mathrm{o}}}\right]'\right)$$

$$= \mathrm{diag}\left((\mathbf{Q}^{-1} - \mathbf{Q}^{-1}\mathbf{u}(\mathbf{u}'\mathbf{Q}^{-1}\mathbf{u})^{-1}\mathbf{u}'\mathbf{Q}^{-1}) \odot \left[\frac{\overline{\mathbf{gm}}}{\overline{\mathbf{l}}^{\mathrm{o}}}\right]'\right) \qquad (14)$$

where $[\partial \mathbf{l}/\partial \mathbf{gm}]$ represents the $(N \times N)$ Jacobian matrix of the land demand functions and the i, jth element of the $(N \times N)$ matrix $[\overline{\mathbf{gm}}/\overline{\mathbf{l}}^{\mathrm{o}}]$ is defined as the sample mean of gross margin i, $\overline{gm}_i$, divided by the sample mean of observed land allocation to crop j, $\overline{\ell}_j^0$. Combined with the elasticity reparameterisation we have to add the constraint

$$\mathbf{V}^{\varepsilon}\mathbf{w}^{\varepsilon} = \mathrm{diag}\left((\mathbf{Q}^{-1} - \mathbf{Q}^{-1}\mathbf{u}(\mathbf{u}'\mathbf{Q}^{-1}\mathbf{u})^{-1}\mathbf{u}'\mathbf{Q}^{-1}t) \odot \left[\frac{\overline{\mathbf{gm}}}{\overline{\mathbf{l}}^{\mathrm{o}}}\right]'\right) \qquad (15)$$

with

$$\mathbf{V}^{\varepsilon} = \begin{bmatrix} v_{11}^{\varepsilon} & v_{12}^{\varepsilon} & 0 & 0 & 0 & 0 \\ 0 & 0 & v_{21}^{\varepsilon} & v_{22}^{\varepsilon} & 0 & 0 \\ 0 & 0 & 0 & 0 & v_{31}^{\varepsilon} & v_{32}^{\varepsilon} \end{bmatrix} \quad \text{and} \quad \mathbf{w}^{\varepsilon} = \begin{bmatrix} w_{11}^{\varepsilon} \\ w_{12}^{\varepsilon} \\ w_{21}^{\varepsilon} \\ w_{22}^{\varepsilon} \\ w_{31}^{\varepsilon} \\ w_{32}^{\varepsilon} \end{bmatrix}$$

to the previous (10)–(13), where v_{i1}^{ε} and v_{i2}^{ε} are the lower and upper support points of the ith elasticity, respectively, and w_{i1}^{ε} and w_{i2}^{ε} the corresponding probabilities. The objective function has to be modified to

$$\max_{\mathbf{w}_t, \mathbf{w}^{\varepsilon}, \mathbf{Q}, \mathbf{L}, \lambda_t} H(\mathbf{w}_t) = -\sum_{t=1}^{T} \mathbf{w}_t' \ln \mathbf{w}_t - \mathbf{w}^{\varepsilon\prime} \ln \mathbf{w}^{\varepsilon}. \qquad (16)$$

The intuition behind the objective function is as follows: the entropy criterion pulls towards the centre of the elasticity support range, in opposition to the error terms of the data constraints. The smaller the elasticity support range, the higher the penalty for deviating from the support centre. Consequently, the width of the support range reflects the precision of the *a priori* information. A necessary condition for consistency, however, is that the true elasticity remains within the support range. Only then is it possible that the increasing weight of the error probabilities in the objective function draws the parameter estimates to their true values as more observations become available.

The approach is analogous to the typical GME procedure, but the standard theoretical exposition (Golan, Judge and Miller, 1996) and agricultural economics applications (e.g. Lence and Miller, 1998a, 1998b; Oude Lansink, 1999; Zhang and Fan, 2001) have so far employed direct restrictions on the parameter space only to make the approach suitable for ill-posed and/or ill-conditioned problems. The restrictions on *functions* of parameters used here, however, are often more appropriate for incorporating available out-of-sample information.

Table 1. QP model—RMSE of one estimated gross margin elasticity with different priors

Prior information	Sample size (T)				
	5	10	20	30	50
'without'	0.187	0.110	0.071	0.055	0.045
'true'	0.158	0.105	0.063	0.055	0.045
'false'	0.163	0.105	0.063	0.055	0.045

Source: own calculations. The value of the true gross margin elasticity is 1.03.

Now we turn to the specific formulation of priors in our simulation experiments. The support point range for the elasticities is set to four, so that a rather large variation of the estimated elasticities is possible without strong penalties. Two different support centres are considered. In one case they are equal to the true elasticities, in the other case they are shifted upwards by 0.3.

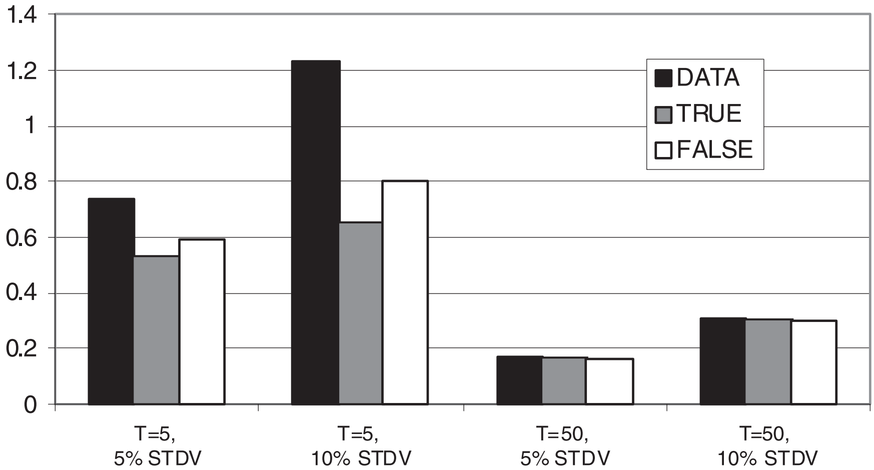
Table 1 presents the RMSE of the gross margin elasticity of one output at different sample sizes. We see that the large variation of parameter estimates displayed in Figure 1 is accompanied by rather stable elasticity estimates even with little data information. Nevertheless, one can infer the general advantage of incorporating the prior information: the estimation error decreases for small sample sizes for both formulations of the priors, although the 'true' prior shows some advantage at the sample size of five. Beyond a sample size of 20, no differences between the three variants exist and they all approach the true parameters as sample size increases.

The impact and usefulness of prior information is certainly also related to the noise in the data generation process. Figure 2 shows sums of root mean square errors across all three gross margin elasticities for two different standard deviations of the error terms (measured as a percentage of the mean land allocation). Clearly, the relative advantage of using priors at small sample sizes increases with the noise in the data generation process. However, for both versions, a sample size of 50 is enough to render the priors almost irrelevant for the quality of the estimates.

We note that the inclusion of prior information at small sample sizes can be seen as an intermediate approach between the calibration of the model to exogenous elasticities for some base year value and the estimation of model parameters with sufficient data information. Consequently, one can use at least the small amount of data information available without jeopardising the 'plausibility' of the estimation results.

4.2. Input allocation with crop-specific production functions

In this section we consider a programming model that allocates variable and fixed inputs to different production activities with a functional representation



Source: Own calculations.

Figure 2. QP model—SRMSE of estimated gross margin elasticities with different priors and noise components.

of crop-specific production technology. The general form of the desired profit maximisation model is given by

$$\begin{aligned} &\max_{x_{ik},b_{ij}} Z = \sum_{i=1}^{N} p_i f_i(x_{ik}, b_{ij} \mid \boldsymbol{\theta}_i) - \sum_{i=1}^{N} \sum_{k=1}^{K} q_k x_{ik} \\ &\text{subject to} \quad \sum_{i=1}^{N} b_{ij} = b_j \; [\lambda_j] \end{aligned} \tag{17}$$

where i, j, k are indices of outputs as well as fixed and variable inputs, respectively, and $\boldsymbol{\theta}_i$ is a vector of technological parameters. Prices and allocated variable inputs are denoted by q_k and x_{ik}, whereas b_{ij} and b_j represent allocated and total available quantities of the fixed inputs. The transformation of input to output quantities is given by

$$y_i = f_i(x_{ik}, b_{ij} \mid \boldsymbol{\theta}_i). \tag{18}$$

The first-order conditions comprise the resource constraints, the marginal value product conditions of variable inputs, and the shadow price equations of fixed factors:

$$\begin{aligned} &\sum_{i=1}^{N} b_{ij} = b_j, \qquad \frac{\partial Z}{\partial x_{ik}} = p_i \frac{\partial f_i(x_{ik}, b_{ij} \mid \boldsymbol{\theta}_i)}{\partial x_{ik}} - q_k = 0, \\ &\frac{\partial Z}{\partial b_{ij}} = p_i \frac{\partial f_i(x_{ik}, b_{ij} \mid \boldsymbol{\theta}_i)}{\partial b_{ij}} - \lambda_j = 0. \end{aligned} \tag{19}$$

Solving this system of first-order conditions for the input demand and output supply functions is very cumbersome, if not impossible. Instead, we can use equations (18) and (19) directly as data constraints for estimating the unknowns $\boldsymbol{\theta}_i$ and λ_j. This implies a considerable advantage with respect to the choice of functional form as well as to the degree of model complexity.

Again, we assume that the data generation process is disturbed by random errors around the endogenous model variables, here not only land allocations, but all input allocations x_{ik} and b_{ij} as well as supply quantities y_i. The corresponding errors e^x_{ikt}, e^b_{ijt} and e^y_{it} for each observation are reparameterised as

$$e^x_{ikt} = \mathbf{v}^x_{ik}\mathbf{w}^x_{ikt}, \qquad e^b_{ijt} = \mathbf{v}^b_{ij}\mathbf{w}^b_{ijt}, \qquad \text{and} \qquad e^y_{it} = \mathbf{v}^y_i\mathbf{w}^y_{it} \tag{20}$$

with the (1×2) vectors $\mathbf{v}^x_{ik}$, $\mathbf{v}^b_{ij}$ and $\mathbf{v}^y_i$ representing lower and upper support points and the (2×1) $\mathbf{w}^x_{ikt}$, $\mathbf{w}^b_{ijt}$ and $\mathbf{w}^y_{it}$ their corresponding probabilities for each observation. Adding indices for observations $t = 1, \ldots, T$ we obtain the complete GME program as

$$\max_{\mathbf{w}^x_{ikt}, \mathbf{w}^b_{ijt}, \mathbf{w}^y_{it}, \boldsymbol{\theta}_i, \lambda_{jt}} H(\mathbf{w}^x_{ikt}, \mathbf{w}^b_{ijt}, \mathbf{w}^y_{it})$$

$$= -\sum_{i=1}^{N}\left[\sum_{t=1}^{T}\sum_{k=1}^{K}\mathbf{w}^{x\prime}_{ikt}\ln\mathbf{w}^x_{ikt} + \sum_{t=1}^{T}\sum_{j=1}^{M}\mathbf{w}^{b\prime}_{ijt}\ln\mathbf{w}^b_{ijt} + \sum_{t=1}^{T}\mathbf{w}^{y\prime}_{it}\ln\mathbf{w}^y_{it}\right] \tag{21}$$

subject to

$$p_{it}\frac{\partial f_i((x^o_{ikt} - \mathbf{v}^x_{ik}\mathbf{w}^x_{ikt}), (b^o_{ijt} - \mathbf{v}^b_{ij}\mathbf{w}^b_{ijt}) \,|\, \boldsymbol{\theta}_i)}{\partial x_{ik}} - q_{kt} = 0 \tag{22}$$

$$\frac{\partial Z}{\partial b_{ij}} = p_{it}\frac{\partial f_i((x^o_{ikt} - \mathbf{v}^x_{ik}\mathbf{w}^x_{ikt}), (b^o_{ijt} - \mathbf{v}^b_{ij}\mathbf{w}^b_{ijt}) \,|\, \boldsymbol{\theta}_i)}{\partial b_{ij}} - \lambda_{jt} = 0 \tag{23}$$

$$(y^o_{it} - \mathbf{v}^y_i\mathbf{w}^y_{it}) = f_{it}((x^o_{ikt} - \mathbf{v}^x_{ik}\mathbf{w}^x_{ikt}), (b^o_{jt} - \mathbf{v}^b_{ij}\mathbf{w}^b_{ijt}) \,|\, \boldsymbol{\theta}_i) \tag{24}$$

$$\sum_{i=1}^{N}(b^o_{ijt} - \mathbf{v}^b_{ij}\mathbf{w}^b_{ijt}) = b_{jt} \tag{25}$$

$$\sum_{s=1}^{S} w^x_{ikts} = 1; \qquad \sum_{s=1}^{S} w^b_{ijts} = 1; \qquad \sum_{s=1}^{S} w^y_{its} = 1. \tag{26}$$

Again, the data constraints have to be satisfied at estimated values of the endogenous variables calculated as the observed values minus the estimated errors.[15]

Before presenting the Monte Carlo simulations for this model, we discuss the prior information used with this model to test its impact on the estimator's

15 The introduction of error terms around the endogenous variables x_{ikt} and b_{ijt} allows estimation of input allocations consistent with the economic model. The presumed quality of 'observed' input allocations can be taken into account by varying the size of the support range. The chosen specification implies that output does not depend on the errors in input demand, because they are subtracted from the observed values in the production function (24). This works well for the case of measurement errors, but optimisation errors are likely to influence production (see Pope and Just, 2002). For appropriate data in a real world empirical application, a more sophisticated error specification might allow for separate identification of error types.

accuracy. As a variation on the previous model, here we assume information is available on the mean value of shadow prices of the fixed factors.[16] The GME approach needs to be modified by adding a constraint with the reparameterised mean shadow prices for the two fixed factors

$$\mathbf{V}^{\lambda}\mathbf{w}^{\lambda} = \begin{bmatrix} v_{11}^{\lambda} & v_{12}^{\lambda} & 0 & 0 \\ 0 & 0 & v_{21}^{\lambda} & v_{22}^{\lambda} \end{bmatrix} \begin{bmatrix} w_{11}^{\lambda} \\ w_{12}^{\lambda} \\ w_{21}^{\lambda} \\ w_{22}^{\lambda} \end{bmatrix} = \begin{bmatrix} \frac{1}{T}\sum_{t=1}^{T} \lambda_{1t} \\ \frac{1}{T}\sum_{t=1}^{T} \lambda_{2t} \end{bmatrix}. \tag{27}$$

The objective function is extended by the additional probabilities to become

$$\max_{\mathbf{w}_{ikt}^{x},\mathbf{w}_{ijt}^{b},\mathbf{w}_{it}^{y},\mathbf{w}^{\lambda},\boldsymbol{\theta}_i,\lambda_{jt}} H(\mathbf{w}_{ikt}^{x}, \mathbf{w}_{ijt}^{b}, \mathbf{w}_{it}^{y}, \mathbf{w}^{\lambda})$$
$$= -\sum_{i=1}^{N}\left[\sum_{t=1}^{T}\sum_{k=1}^{K} \mathbf{w}_{ikt}^{x\prime} \ln \mathbf{w}_{ikt}^{x} + \sum_{t\,\bar{1}}^{T}\sum_{j=1}^{M} \mathbf{w}_{ijt}^{b\prime} \ln \mathbf{w}_{ijt}^{b} + \sum_{t=1}^{T} \mathbf{w}_{it}^{y\prime} \ln \mathbf{w}_{it}^{y}\right]$$
$$- \mathbf{w}^{\lambda\prime} \ln \mathbf{w}^{\lambda}. \tag{28}$$

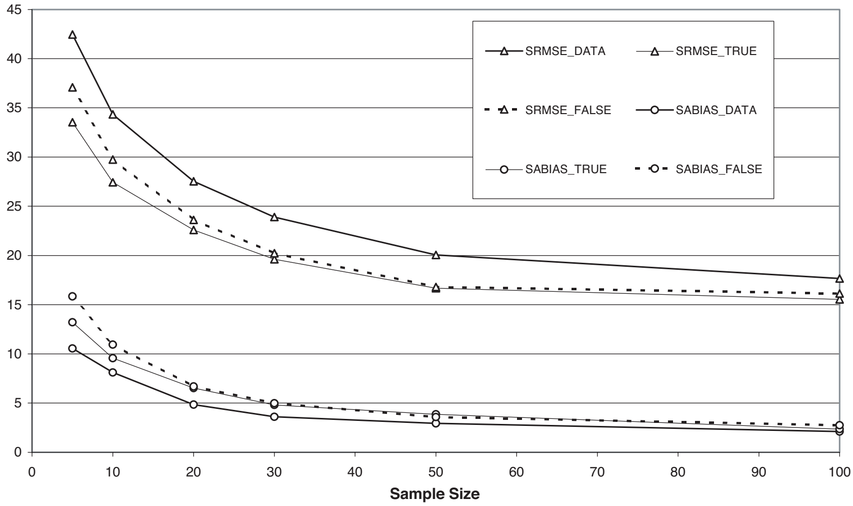
The functional form for the production technology chosen for the Monte Carlo simulations is the Constant Elasticity of Substitution (CES) function, which distinguishes between two variable inputs (chemicals and capital) and two fixed inputs (land and water).[17] This model structure is analogous to that of Howitt (1995b), which was treated using PMP. However, our model does not contain any additional non-linear cost terms in the objective function, and the estimation approach does not require the determination of dual values of calibration constraints from the 'step 1' of PMP. In addition, our model requires decreasing returns to scale, to allow for positive production levels of all crops.[18]

The data generation process adds normally distributed error terms to the optimal output and input quantities (with standard deviations of 10 and 2 per cent respectively from the mean quantities) to obtain 'observed' allocations, where again the 'incorrectly measured' allocated quantities of the fixed inputs add up exactly to the available and known total quantities. For

16 The use of prior information on elasticities for this model is also possible despite the fact that an analytical expression for the elasticities might not be available. One can use discrete approximations based on additional 'artificial' constraints that are simply copies of the data constraints, but with systematically varied exogenous prices and variable 'simulated' output and input quantities. Although conceptually simple, the mathematical representation of this approach is considered too cumbersome for this paper.

17 For the true parameter values, see the Appendix. Additional parametric restrictions included during estimation so as to obtain a well-defined production function are also reported in the Appendix.

18 Constant returns to scale (as in Howitt, 1995b) would result in specialisation, because the maximum profit per unit of land in each activity would be independent of the land allocation. Consequently, the number of positive activity levels at the optimum could not be larger than the number of fixed factors, as in a linear programming model.



Source: Own calculations.

Figure 3. SRMSE and SABIAS of parameter estimated with different prior information on shadow prices of fixed resources in the CES model.

the simulation we distinguish again between a 'true' and a 'false' prior. The former defines supports for the shadow prices of land and water around the true values at the mean of the observations. The latter is defined by a support centre that is 10 per cent below the true values. The size of supports is chosen to be 40 per cent of the true mean shadow price. This is well above 5 standard deviations of the mean shadow prices across samples so that the support contains the true mean shadow price for both types of priors with almost certainty.

Figure 3 shows the absolute bias and the root mean square error as sums over the parameters of all three production functions (SABIAS and SRMSE). SRMSE decreases with increasing sample size, suggesting consistency of the estimator. The use of both types of prior information again reduces the SRMSE compared with the data-only case. The reduction is relatively modest compared with the priors on elasticities for the QP model, but it is still relevant even for $T = 100$. However, the difference between the true and the false prior is negligible from $T = 30$ upwards. It is interesting to note that the bias of the true prior lies above the one for the false prior. This can certainly happen in the case of a biased estimator but should be seen as a lucky 'accident'. In fact, it can be shown that the result is reversed if we formulate the false prior such that the centre of the supports lies above the true values.

Generally, prior information on shadow prices could also be formulated for every observation t instead of for the shadow price mean, which might better reflect the type of data available (e.g. leasing rates for each observation). In this case, however, the number of associated probabilities in the objective function would increase with increasing observations. This may harm the convergence of the estimates to the true parameter values when (as inevitably occurs) the centres of the shadow price supports are not the true values. Additional simulations (not reported here) have supported this hypothesis. This effect, however, could be counteracted by including a factor in the objective function that continuously decreases the weight of the prior-related probabilities with increasing sample size.[19]

4.3. Allocation of fixed inputs with crop-specific profit functions

Our third example keeps the general model structure of the previous subsection with respect to assumptions about producer behaviour and crop-specific technologies with allocable inputs, but employs duality concepts for the determination of variable output and input quantities. We return to the case with only one fixed factor, and adopt the specification used by Guyomard *et al.* (1996) and Moro and Sckokai (1999), who base their analysis on econometrically estimated systems of supply and explicit land allocation functions. On the one hand, we want to point out the full equivalence of our approach with respect to parameter estimation. On the other hand, we want to illustrate the advantages with respect to flexibility in the choice of functional form as well as greater complexity of the model structure. The desired programming model is given by

$$\max_{\mathbf{l}} Z = \sum_{i=1}^{N} \pi_i(p_i, \mathbf{q}, \ell_i \,|\, \boldsymbol{\theta}_i) \quad \text{subject to} \quad \sum_{i=1}^{N} \ell_i = b\,[\lambda] \tag{29}$$

where

$$\pi_i(p_i, \mathbf{q}, \ell_i \,|\, \boldsymbol{\theta}_i) = \max_{y_i, \mathbf{x}_i} \left[p_i y_i - \sum_{k=1}^{K} q_k x_{ik} \text{ subject to } y_i = f_i(x_{ik}, \ell_i) \right] \tag{30}$$

is a restricted profit function of crop i for a given land allocation ℓ_i and $\boldsymbol{\theta}_i$ is now a vector of profit function parameters for product i. Model (29) distributes the available land b to the different production activities so as to maximise overall profit Z, where the profit of the single crops is determined by $\pi_i(p_i, \mathbf{q}, \ell_i \,|\, \boldsymbol{\theta}_i)$. Consequently, the optimal land allocation is obtained if the marginal profits of land in each use are equal, i.e. if the first-order conditions

$$\frac{\partial \pi_i(p_i, \mathbf{q}, \ell_i \,|\, \boldsymbol{\theta}_i)}{\partial \ell_i} - \lambda = 0 \tag{31}$$

are satisfied. For some functional forms of $\pi_i(.)$ a solution of system (31)—observing the land constraint in (29)—is possible and results in explicit land allocation equations depending on exogenous model parameters. Guyomard

19 In the context of the method of regularisation (see Golan, Judge and Miller, 1996, p. 129f), this would imply that the penalty function reflecting information about the plausible values of the parameters (here the objective function terms related to the shadow price probabilities) would receive less weight with increasing sample size.

et al. (1996) describe the derivation based on normalised quadratic profit functions and estimate a system of land allocation equations and supply functions

$$\frac{\partial \pi_i(p_i, \mathbf{q}, \ell_i \,|\, \boldsymbol{\theta}_i)}{\partial p_i} = y_i(p_i, \mathbf{q}, \ell_i \,|\, \boldsymbol{\theta}_i). \tag{32}$$

The resulting system is linear, but the regression coefficients have to satisfy non-linear constraints across equations. With our approach, no derivation of land allocation equations is necessary. Instead, the optimality conditions (31) are used directly in combination with (32) as data constraints in a GME approach, as in the previous two examples. As long as the same statistical model and econometric criterion are employed, the parameter estimates of this approach must be equal to those obtained by estimating the behavioural functions, because of the mathematical equivalence of the data constraints. This was confirmed on the basis of a GME and a non-linear least-squares approach.

There are at least three advantages in estimating the model using the optimality conditions even when a programming model is subsequently used for simulation purposes. First, the choice of the functional form for $\pi_i(p_i, \mathbf{q}, \ell_i)$ is wider, because a closed form solution for land allocation functions is not necessary. Second, and for the same reason, a more complex model structure with more than one fixed factor or general constraints on land allocation (e.g. quotas, base areas) is no longer an impediment to the econometric estimation of the parameters. Third, formulating the resulting simulation model as an explicit optimisation model allows the flexible incorporation of additional relevant constraints on allocation for the simulation horizon without necessarily obstructing the structural validity of the estimation results.

For this model also we performed simulation experiments based on an appropriate GME estimator. We mirrored the Guyomard *et al.* (1996) approach in the sense that we used only data on supply quantities and land allocations, disregarding any observations on allocated input quantities and the related input demand functions as data constraints. Reparameterising the errors of these endogenous variables of the programming model in the same way as for the CES production function model, we formulated the GME program for the estimation of the profit function parameters as

$$\max_{\mathbf{w}_{it}^{\ell}, \mathbf{w}_{it}^{y}, \boldsymbol{\theta}_i, \lambda_t} H(\mathbf{w}_{it}^{\ell}, \mathbf{w}_{it}^{y}) = -\sum_{i=1}^{N} \left[\sum_{t=1}^{T} \mathbf{w}_{it}^{\ell\prime} \ln \mathbf{w}_{it}^{\ell} + \sum_{t=1}^{T} \mathbf{w}_{it}^{y\prime} \ln \mathbf{w}_{it}^{y} \right] \tag{33}$$

subject to

$$\frac{\partial \pi_i(p_{it}, \mathbf{q}_t, (\ell_{it}^{o} - \mathbf{v}_i^{\ell} \mathbf{w}_{it}^{\ell}) \,|\, \boldsymbol{\theta}_i)}{\partial \ell_i} - \lambda = 0 \tag{34}$$

$$\frac{\partial \pi_i(p_{it}, \mathbf{q}, (\ell_{it}^{o} - \mathbf{v}_i^{\ell} \mathbf{w}_{it}^{\ell}) \,|\, \boldsymbol{\theta}_i)}{\partial p_i} = (y_{it}^{o} - \mathbf{v}_i^{y} \mathbf{w}_{it}^{y}) \tag{35}$$

$$\sum_{i=1}^{N} (\ell_{it}^{o} - \mathbf{v}_i^{\ell} \mathbf{w}_{it}^{\ell}) = b_t \tag{36}$$

Table 2. Profit function model—Monte Carlo results without prior information

Measures	Sample size (T)						
	4	5	10	20	30	50	100
SRMSE	2965	2888	1212	570	462	346	253
SABIAS	914	900	417	222	159	102	57
SSTD	2715	2672	1102	516	426	325	242

Source: own calculations.

$$\sum_{s=1}^{S} w_{its}^{\ell} = 1; \qquad \sum_{s=1}^{S} w_{its}^{y} = 1. \tag{37}$$

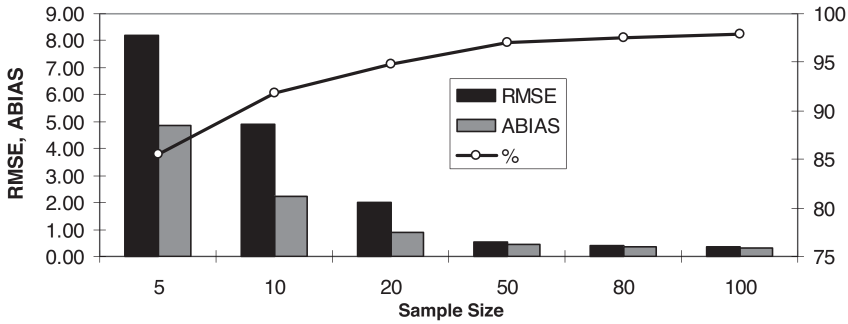
Again, for different sample sizes, artificially generated optimal supply quantities and land allocations were disturbed by normally distributed errors (with standard deviations of 10 and 2 per cent respectively of the mean variable values) and estimation was carried out without the use of prior information on parameters or functions of parameters. The Monte Carlo results are given in Table 2. The change in the estimation errors (summed over all estimated parameters of the profit functions) with sample size indicates consistent estimator behaviour.

The large share of the variance in the mean squared error again indicates the considerable potential of plausible prior information to improve estimator precision for small sample sizes. However, we wish to focus here on another issue of empirical relevance. Constraints on allocation, such as the land constraint, are frequently of the inequality type and across observations they may sometimes be binding as equalities and sometimes not. As long as the data tell us directly whether or not a constraint is binding for each observation, this is handled simply by setting the shadow prices to zero *a priori* for observations with non-binding constraints. But because of the noise in the data generation process, it is conceivable that the measured variable quantities are misleading. Apparently binding constraints might in fact not be binding for the true quantities and vice versa. In this case, we must allow the estimated or 'fitted' variable values to satisfy the constraints either in equality or inequality form. In principle, this can be easily accommodated by changing the land constraint (36) to the inequality type

$$\sum_{i=1}^{N} (\ell_{it}^{o} - \mathbf{v}_i^{\ell}\mathbf{w}_{it}^{\ell}) \leq b_t \tag{38}$$

and adding the appropriate complementary slackness condition

$$\sum_{i=1}^{N} (b_t - (\ell_{it}^{o} - \mathbf{v}_i^{\ell}\mathbf{w}_{it}^{\ell}))\lambda_t = 0 \tag{39}$$



Source: Own calculations.

Figure 4. ABIAS and RMSE of the dual values and percentage of correctly identified status of the constraints in the NQ model with inequality constraints.

to the GME program. Given the discontinuous nature of λ_t, the numerical stability might be hampered with solvers based on gradient methods. To test this for the relatively simple example above, we changed the data generation process for the simulation approach as follows.

First, the available mean level of land was increased in such a way that, on average, about 25 per cent of the optimal solutions of the data-generating model did not use all the land. Second, we did not force the errors disturbing the optimal land allocations to sum to zero, which generally implies a non-zero difference between the sum of 'observed' land allocations and the available total quantity of land. Third, we used equations (38) and (39) instead of (36) in the GME approach. All other details of the simulation remained the same. The results are rather promising: as in the above experiments, the SRMSE and SABIAS of $\boldsymbol{\theta}_i$ indicate consistent behaviour of the estimator. For more insight into the technique's performance with non-binding and binding resources, we focus in Figure 4 on the finite sample properties of the estimated dual values and on the ability of the approach to correctly identify the status of the constraint.

The RMSE of the dual values is calculated as the square root of the average absolute difference between the estimated and the true shadow prices across all observations and repetitions. Both the ABIAS and the RMSE diminish with increasing sample sizes. To provide further information we include, on the right-hand axis, the percentage of estimated observations that are correct concerning the binding or non-binding status of the land constraint. It is shown that, for small sample sizes, the estimation procedure is able to correctly identify binding and non-binding constraints for more than 85 per cent of observations (which is significantly higher than the 75 per cent obtained by assuming constraints are always binding). With increasing sample size, the success rate approaches 100 per cent, indicating that the estimates converge to the true data generation process as the amount of data information increases.

5. Conclusions

The paper introduces a general approach based on optimality conditions for the estimation of programming models, and shows its theoretical advantage compared with approaches based on PMP. The method simultaneously allows the specification of more complex models and a more flexible choice of functional form compared with previous estimation approaches of duality-based behavioural functions with explicit allocation of fixed factors. The procedure and application were demonstrated for three examples of programming models. Monte Carlo simulations with a maximum entropy criterion indicated consistent behaviour of the estimator. In this context, the potential benefit from prior information on elasticities and shadow prices in situations with small sample sizes, as well as the technical implementation, were also shown. Last but not least, the approach proved itself capable of estimating model parameters across binding and non-binding constraints in the data generation process.

Several limitations of the general approach, and of the specific applications, presented in this paper need to be mentioned. First, there is no need to base estimation on first-order conditions if the assumed optimisation model allows for a consistent derivation from an indirect objective function. In this case, well-established procedures for estimating systems of input demand and output supply equations are available. Second, the stochastic specifications of the models presented are somewhat limited, as the error could not be interpreted as optimisation error in all cases. Finally, although the GME procedure used is appropriate for utilising available data and prior information, the lack of easily applicable statistical test procedures might harm its prospects for widespread use in the near future.

Apart from different applications to large, 'real world' profit-maximising programming models, many other directions for future research can be identified: extensions of the approach to multi-output production technologies with non-allocable variable factors, or to expected utility models with risk, might increase the empirical potential of these types of models. From an econometric methodology point of view, there is still much scope for improving our current knowledge of the GME approach; a more systematic investigation with respect to the formulation of prior information and its impact on estimation quality in small sample situations is desirable. Of course, the use of GME is not a necessary requirement for the estimation of programming models in well-posed situations.

Acknowledgements

The authors thank two anonymous referees and the editor for their helpful contributions. Their constructive comments enhanced the clarity of the paper. All remaining deficiencies are the sole responsibility of the authors. A significant part of this research was undertaken while the authors were at the Institute for Agricultural Policy, Market Research and Economic Sociology, University of Bonn.

References

Arfini, F. and Paris, Q. (1995). A Positive Mathematical Programming model for regional analysis of agricultural policies. In F. Sotte (ed.), *The Regional Dimension in Agricultural Economics and Policies*. EAAE, Proceedings of the 40th Seminar, 26–28 June 1995, Ancona, 17–35.

Britz, W., Heckelei, T. and Wolff, H. (2003). Symmetric Positive Equilibrium Problem: a framework for rationalizing economic behavior with limited information: Comment. *American Journal of Agricultural Economics* (forthcoming).

Chambers, R. G. and Just, R. E. (1989). Estimating multioutput technologies. *American Journal of Agricultural Economics* 71(4): 980–995.

Chirinko, R. S. (1993). Business fixed investment spending: modeling strategies, empirical results, and policy implications. *Journal of Economic Literature* 31: 1875–1911.

Cypris, C. (2000). Positiv Mathematische Programmierung (PMP) im Agrarsektormodell RAUMIS. Dissertation, University of Bonn.

Gallant, A. R. and Golub, G. H. (1984). Imposing curvature restrictions on flexible functional forms. *Journal of Econometrics* 26: 295–321.

Golan, A., Judge, G. and Miller, D. (1996). *Maximum Entropy Econometrics*. Chichester: John Wiley.

Golan, A., Judge, G. and Perloff, M. (1996). A maximum entropy approach to recovering information from multinomial response data. *Journal of the American Statistical Association* 91(434): 841–853.

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: 401–420.

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.

Helming, J. F. M., Peeters, L. and Veendendaal, P. J. J. (2001). Assessing the consequences of environmental policy scenarios in Flemish agriculture. In T. Heckelei, H. P. Witzke and W. Henrichsmeyer (eds), *Agricultural Sector Modelling and Policy Information Systems*. Proceedings of the 65th EAAE Seminar, 29–31 March 2000, Bonn. Kiel: Vauk, 237–245.

House, R. M. (1987). USMP Regional Agricultural Model. National Economics Division Report, ERS, 30. Washington DC: USDA.

Howitt, R. E. (1995a). Positive Mathematical Programming. *American Journal of Agricultural Economics* 77(2): 329–342.

Howitt, R. E. (1995b). A calibration method for agricultural economic production models. *Journal of Agricultural Economics* 46: 147–159.

Howitt, R. E. and Gardner, B. D. (1986). Modeling 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 (USA), 271–290.

Kasnakoglu, H. and Bauer, S. (1989). Concept and application of an agricultural sector model for policy analysis in Turkey. In S. Bauer and W. Henrichsmeyer (eds), *Agricultural Sector Modelling*. Kiel: Vauk: 71–84.

Lau, L. J. (1978). Testing and imposing monotonicity, convexity, and quasi-convexity. In M. Fuss and D. McFadden (eds), *Production Economics: a Dual Approach to Theory and Applications*. Amsterdam: North-Holland, 409–453.

Lence, H. L. and Miller, D. (1998a). Estimation of multioutput production functions with incomplete data: a Generalised Maximum Entropy approach. *European Review of Agricultural Economics* 25: 188–209.

Lence, H. L. and Miller, D. (1998b). Recovering output specific inputs from aggregate input data: a generalized cross-entropy approach. *American Journal of Agricultural Economics* 80(4): 852–867.

McElroy, M. B. (1987). Additive general error models for production, cost, and derived demand or share systems. *Journal of Political Economy* 95(4): 737–757.

Mittelhammer, R. C. and Cardell, S. (2000). The data-constrained GME-Estimator of the GLM: asymptotic theory and inference. Working Paper of the Department of Statistics. Pullman, WA: Washington State University.

Moro, D. and Sckokai, P. (1999). Modelling the CAP arable crop regime in Italy: degree of decoupling and impact of Agenda 2000. *Cahiers d'Economie et Sociologie Rurales* 53: 50–73.

Oude Lansink, A. (1999). Generalised Maximum Entropy and heterogeneous technologies. *European Review of Agricultural Economics* 26: 101–115.

Paris, Q. (2001). Symmetric Positive Equilibrium Problem: a framework for rationalizing economic behavior with limited information. *American Journal of Agricultural Economics* 83(4): 1049–1061.

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.

Pope, R. D. and Just, R. E. (2002). Random profits and duality. *American Journal of Agricultural Economics* 84(1): 1–7.

Preckel, P. V. (2001). Least squares and entropy: a penalty function perspective. *American Journal of Agricultural Economics* 83(2): 366–377.

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.

Appendix

Basic data set

The data set shown in Table A1, from Howitt (1995b), provides the differentiation of all models presented with respect to outputs and inputs. Variable quantities are the means of random variables and values of fixed variables for all Monte Carlo simulations.

Information on Monte Carlo simulation with QP model

True parameter values of **Q**:

$$\mathbf{Q} = \begin{bmatrix} 500 & -20 & -10 \\ -10 & 60 & -2 \\ -10 & -2 & 200 \end{bmatrix}.$$

Table A1. Base year data on California agriculture

Crop	Price ($/bu)	Yield (bu/acre)	Input allocation			
			Land (10^6 acres)	Water (10^6 acre ft)	Capital (Index)	Chemicals (Index)
Cotton	2.924	220	1.49	4.47	3.96	2.64
Wheat	2.98	85	0.62	1.14	1.98	1.32
Rice	7.09	70.1	0.54	3.08	2.94	1.96
Variable input prices ($)				10	10	10
Resource constraints			2.65	8.69		

Source: Based on Howitt (1995b).

Information on Monte Carlo simulation with CES model

Functional form of production functions and true parameter values:

$$y_i = f(x_{ik}, b_{ij} \,|\, \boldsymbol{\theta}_i) = \alpha_i \left(\sum_{k=1}^{2} \beta_{ik} x_{ik}^{\gamma_i} + \sum_{j=3}^{4} \beta_{ij} b_{ij}^{\gamma_i} \right)^{\nu_i \gamma_i}.$$

where

$$\boldsymbol{\theta}_1 = \begin{bmatrix} \alpha_1 \\ \beta_{11} \\ \beta_{12} \\ \beta_{13} \\ \beta_{14} \\ \gamma_1 \\ \upsilon_1 \end{bmatrix} = \begin{bmatrix} 200 \\ 0.2 \\ 0.1 \\ 0.6 \\ 0.1 \\ -0.25 \\ 0.6 \end{bmatrix}; \quad \boldsymbol{\theta}_2 = \begin{bmatrix} \alpha_1 \\ \beta_{11} \\ \beta_{12} \\ \beta_{13} \\ \beta_{14} \\ \gamma_1 \\ \upsilon_1 \end{bmatrix} = \begin{bmatrix} 100 \\ 0.1 \\ 0.1 \\ 0.7 \\ 0.1 \\ -0.25 \\ 0.8 \end{bmatrix}; \quad \boldsymbol{\theta}_3 = \begin{bmatrix} \alpha_1 \\ \beta_{11} \\ \beta_{12} \\ \beta_{13} \\ \beta_{14} \\ \gamma_1 \\ \upsilon_1 \end{bmatrix} = \begin{bmatrix} 50 \\ 0.1 \\ 0.1 \\ 0.5 \\ 0.3 \\ -0.25 \\ 0.8 \end{bmatrix}.$$

Parametric restrictions enforced during estimation:

$$\alpha_i \geq 0; \qquad 0 \leq \beta_{ij} \leq 1; \qquad \sum_{j=1}^{4} \beta_{ij} = 1$$

$$\sigma_i = \frac{1}{1 - \gamma_i} \geq 0; \qquad 0 \leq \nu_i \leq 1.$$

Information on Monte Carlo simulation with profit function model

Functional form of the profit functions and true parameter values:

$$\pi_i(p_i, \mathbf{q}, l_i \,|\, \boldsymbol{\theta}_i) = \alpha_{0i} + \alpha_{1i} \frac{p_i}{q_2} + \alpha_{2i} \frac{q_1}{q_2} + \alpha_{3i} \ell_i + 0.5\beta_{1i} \left(\frac{p_i}{q_2} \right)^2 + 0.5\beta_{2i} \left(\frac{q_1}{q_2} \right)^2 + 0.5\beta_{3i} (\ell_i)^2 + \gamma_{1i} \frac{p_i q_1}{q_2^2} + \gamma_{2i} \frac{p_i}{q_2} \ell_i + \gamma_{3i} \frac{q_1}{q_2} \ell_i$$

where

$$\boldsymbol{\theta}_1 = \begin{bmatrix} \alpha_{01} \\ \alpha_{11} \\ \alpha_{21} \\ \alpha_{31} \\ \beta_{11} \\ \beta_{21} \\ \beta_{31} \\ \gamma_{11} \\ \gamma_{21} \\ \gamma_{31} \end{bmatrix} = \begin{bmatrix} -67.6571 \\ 115.023 \\ 1.914 \\ 87.836 \\ 24.854 \\ 1.167 \\ -60.321 \\ -9.391 \\ 144.229 \\ -2.883 \end{bmatrix}; \quad \boldsymbol{\theta}_2 = \begin{bmatrix} \alpha_{02} \\ \alpha_{12} \\ \alpha_{22} \\ \alpha_{32} \\ \beta_{12} \\ \beta_{22} \\ \beta_{32} \\ \gamma_{12} \\ \gamma_{22} \\ \gamma_{32} \end{bmatrix} = \begin{bmatrix} -16.2746 \\ -34.130 \\ -0.176 \\ 49.817 \\ 23.552 \\ 0.607 \\ -82.432 \\ -4.232 \\ 135.553 \\ -1.855 \end{bmatrix}; \quad \boldsymbol{\theta}_3 = \begin{bmatrix} \alpha_{03} \\ \alpha_{13} \\ \alpha_{23} \\ \alpha_{33} \\ \beta_{13} \\ \beta_{23} \\ \beta_{33} \\ \gamma_{13} \\ \gamma_{23} \\ \gamma_{33} \end{bmatrix} = \begin{bmatrix} -8.7735 \\ 4.656 \\ -0.650 \\ 27.335 \\ 6.115 \\ 0.882 \\ -53.635 \\ -2.611 \\ 58.284 \\ -2.446 \end{bmatrix}.$$

Parametric restrictions enforced during estimation:

$$\beta_{1i} > 0; \quad \beta_{2i} > 0; \quad \gamma_{i2} > 0; \quad \beta_{3i} < 0; \quad \gamma_{i1} < 0.$$

Parameters not estimated (only appear in profit function or input demand functions):

$$\alpha_{0i}, \quad \alpha_{2i}, \quad \beta_{2i}.$$

Parameters not identified relative to shadow price of land (same as in Guyomard *et al.* (1996), where those parameters are part of the composite estimated regression coefficients in the land allocation equations) and therefore fixed at true values:

$$\alpha_{31}, \quad \gamma_{31}.$$

Corresponding author: Thomas Heckelei, Impact Center, Hulbert 123C, Washington State University, Pullman, WA 99164-6214, USA. E-mail: heckelei@wsu.edu