IMPOSING REGULARITY CONDITIONS ON A SYSTEM OF COST AND COST-SHARE EQUATIONS: A BAYESIAN APPROACH

by

William E. Griffiths, Christopher J. O'Donnell

and Agustina Tan Cruz

Department of Econometrics

University of New England NSW 2351

AUSTRALIA

January 1997

Abstract

A problem with the use of flexible functional forms is that the estimated functions frequently violate the regularity conditions (eg. monotonicity, concavity) implied by economic theory. Sampling theory methods exist for imposing curvature conditions at all non-negative prices, but such methods may significantly reduce the flexibility properties of the functional forms. Bayesian methods can be used to maintain these flexibility properties, by imposing regularity conditions at a point, at several points, or over the region within which inferences will be drawn. We use this Bayesian approach to estimate a system of cost and input cost-share equations for the Australian merino-woolgrowing sector. The imposition of both monotonicity and concavity constraints at several sets of prices leads to sign reversals and significant changes in the magnitudes of a small number of estimated input-price elasticities.

KEYWORDS: cost functions, Markov Chain Monte Carlo, inequality constraints

1. INTRODUCTION

Estimated flexible functional forms frequently violate the monotonicity, concavity, convexity and quasi-convexity conditions implied by economic theory. Examples from the agricultural economics literature include the profit and cost function estimates of Berndt and Khaled (1979), Jorgensen and Fraumeni (1981) and O'Donnell and Woodland (1995). One solution to this problem is to impose these regularity conditions at the time of estimation; a number of methods currently exist for imposing at least some conditions at all non-negative prices. It is possible to impose global curvature restrictions, for example, using eigenvalue decomposition methods and methods involving Cholesky factorisation (see Wiley, Schmidt and Bramble, 1973; Talpaz, Alexander and Shumway, 1989; Coelli, 1996).

Unfortunately, the global imposition of curvature conditions forces many flexible functional forms to take on properties not implied by economic theory. For example, imposing global concavity on a translog cost function may lead to an upward bias in the degree of input substitutability, and imposing global concavity on a Generalised Leontief cost function will rule out complementarity between inputs (Diewert and Wales, 1987). An alternative approach which can be used to maintain the flexibility properties of flexible functional forms involves the imposition of regularity conditions only at a point, at several points, or over a region of interest, usually the region over which inferences will be drawn. Methods which can be used to impose curvature restrictions locally include the numerical methods of Lau (1978) and Gallant and Golub (1984). More recently, Chalfant and Wallance (1992) and Terell (1996) have used a Bayesian approach to impose both monotonicity and curvature restrictions locally. An advantage of the Bayesian approach is its ability to provide finite sample inference procedures for nonlinear functions of parameters.

The Bayesian approach of Terell exploits recent developments in Markov Chain Monte Carlo (MCMC) simulation methods. The use of MCMC methods has grown rapidly with the availability of inexpensive high-speed computers and with the further development of powerful computer-intensive statistical algorithms. These algorithms, which include the Gibbs sampler and the Metropolis-Hastings algorithm, can be used to generate random variables from a marginal distribution indirectly, without having to derive the density itself. Not surprisingly, MCMC methods have revolutionised Bayesian econometrics, where posterior marginal densities can be analytically difficult or impossible to derive.

In this paper we use MCMC methods to estimate a system of cost and cost-share equations for a sector of the Australian woolgrowing industry. This empirical application of the MCMC methodology is motivated by the large number of curvature violations reported in the study by O'Donnell and Woodland. Although we retain most of the features of the O'Donnell and Woodland model, and we use their data set, we estimate a system of cost and cost-share equations which has a less complex stochastic structure, and we focus only on one Australian woolgrowing sector (merino-woolgrowing) instead of three. These simplifications allow us to better illustrate the applicability and usefulness of the MCMC technique, and still allow us to validate the elasticity estimates obtained in O'Donnell and Woodland's earlier work.

The outline of the paper is as follows. In Section 2 we translate a standard economic model of producer behaviour into the system of cost and cost-share equations to be used in our empirical work. This empirical model takes the form of the Seemingly Unrelated Regression (SUR) model discussed in most elementary econometrics textbooks. In Section 3 we describe two alternative but equivalent iterative procedures for obtaining maximum likelihood estimates of the SUR model parameters. We also describe the Gibbs sampler and Metropolis-Hastings algorithms, and the manner in which monotonicity and curvature restrictions can be imposed. The data are described in Section 4 and the estimation results are presented in Section 5. The results are evaluated in terms of parameter estimates, predicted cost shares, estimated eigenvalues and estimated input-price elasticities. The paper is concluded in Section 6.

2. MODEL

Our model is predicated on the assumption that the technological possibilities faced by the firm can be summarised by the cost function

(1) $C(w, q) \equiv \min_{x} \{w^T x: f(x) \geq q, x \geq 0\}$

where x is an (Ix1) vector of inputs, w is an (Ix1) vector of input prices and q is scalar output. If the production function f(x) satisfies a standard set of relatively weak assumptions then the cost function will be nonnegative for all positive prices and output, linearly homogenous in prices, nondecreasing in prices (ie. monotonic) and concave and continuous in prices (Chambers, 1988). Moreover, the Hessian matrix of second order derivatives of the cost function will be symmetric. Our interest lies in the properties of monotonicity and concavity, and the manner in which these properties can be imposed on an estimated flexible functional form.

A functional form is flexible if it can provide a local 'second-order approximation' to an arbitrary functional form. An excellent discussion of the concept of a 'second-order approximation' can be found in Barnett (1983). The two most commonly used flexible functional forms are the generalised Leontief introduced by Diewert (1971) and the translog introduced by Christensen, Jorgensen and Lau (1971). We follow O'Donnell and Woodland and assume a constant returns to scale translog functional form, which implies a cost function of the form

(2) $\ln(C/q) = \alpha_0 + \alpha_T T + \sum_{i=1}^{I} \alpha_i \ln(w_i) + 0.5 \sum_{i=1}^{I} \sum_{j=1}^{I} \alpha_{ij} \ln(w_i) \ln(w_j)$

where C represents total costs, w_i represents the price of input i and T is a time trend which is used to capture the effects of exogenous technical change. The factor cost-share equations are obtained using Shephard's lemma:

(3) $s_i = \alpha_i + \sum_{j=1}^{I} \alpha_{ij} \ln(w_j) \qquad i=1, \ldots, I$

where s_i represents the cost share of input i. It is clear from equations (2) and (3) that our assumed form of technical change is Hicks neutral: factor shares are unaffected by technical change while unit costs decrease at a constant percentage rate.

Some of the theoretical properties of the cost function (1) can be expressed in terms of the parameters appearing in equation (2). Specifically, linear homogeneity and symmetry will be satisfied if the parameters satisfy the restrictions

$$(4) \qquad \sum_{i=1}^{I} \alpha_i = 1, \qquad \sum_{j=1}^{I} \alpha_{ij} = 0 \quad (i = 1, \ldots, I), \qquad \alpha_{ij} = \alpha_{ji} \quad (i, j = 1, \ldots, I)$$

Monotonicity will be satisfied if the estimated factor cost-shares are positive, while concavity will be satisfied if the the Hessian matrix of second order derivatives of the cost function is negative semi-definite. In turn, the Hessian matrix will be negative semi-definite if and only if its eigenvalues are non-positive.

Our empirical model is obtained by following the usual practice of embedding equations (2) and (3) in a stochastic framework. After incorporating stochastic terms, and after recognising that our data vary over time and cross-sectional units by introducing the firm and time subscripts n and t (n=1, ..., N and t=1, ..., T), our empirical model is given by:

$$(5) \qquad \begin{aligned} s_{int} &= \alpha_i + \sum_{j=1}^{I} \alpha_{ij} \ln(w_{jnt}) + \varepsilon_{int} \qquad i = 1, \ldots, I-1 \\ \ln(C_{nt}/q_{nt}) &= \alpha_0 + \alpha_T T_{nt} + \sum_{i=1}^{I} \alpha_i \ln(w_{int}) + 0.5 \sum_{i=1}^{I} \sum_{j=1}^{I} \alpha_{ij} \ln(w_{int}) \ln(w_{jnt}) + \varepsilon_{Int} \end{aligned}$$

where ε_{int} (i=1, ..., I) represents statistical noise. Note that we have adopted the usual practice of dropping one share equation to avoid singularity of the error covariance matrix. The share and cost equation errors are assumed to be independently and identically distributed over firms and time with properties:

$$(6) \qquad E\{\varepsilon_{int}\} = 0$$

$$(7) \qquad E\{\varepsilon_{int}\varepsilon_{mks}\} = \begin{cases} \sigma_{im} & \text{if } n=k \text{ and } t=s \\ 0 & \text{otherwise} \end{cases}$$

The model given by equations (4) to (7) has an identical deterministic structure and a similar stochastic structure to the model of O'Donnell and Woodland. Like O'Donnell and Woodland, our stochastic assumptions allow for within-firm contemporaneous correlation between the disturbances ε_{int}, i=1, ..., I. However, unlike O'Donnell and Woodland, our underlying economic model ignores yield uncertainty and, as a consequence, our cost function does not have an error components structure.

3. METHODS OF ESTIMATION

In this section we describe four methods for estimating the parameters of the model given by equations (4) to (7): two equivalent methods for obtaining maximum likelihood estimates, the Gibbs sampler and the Metropolis-Hastings algorithm. The maximum likelihood methods we describe do not allow for the imposition of monotonicity or concavity constraints. Nor does our Gibbs sampler: the Gibbs sampler is only used in this study to illustrate an alternative MCMC technique and to provide a benchmark by which to judge the results of the maximum likelihood and Metropolis-Hastings approaches. Our description of a standard Metropolis-Hastings algorithm provides details on the modifications necessary to ensure that monotonicity and concavity conditions are satisfied.

3.1 Maximum Likelihood Estimation

For a model consisting of four inputs the system of equations given by (5) can be more conveniently written:

(8) $$y_{int} = x_{int}'\beta_i + \varepsilon_{int} \qquad i = 1, ..., 4$$

where $$y_{int} = s_{int} \qquad i = 1, ..., 3$$

$$y_{4nt} = \ln(C_{nt}/q_{nt})$$

(9) $\beta_i = (\alpha_i, \alpha_{i1}, \ldots, \alpha_{i4})'$ $\qquad i = 1, \ldots, 3$

(10) $\beta_4 = (\alpha_0, \alpha_T, \alpha_1, \ldots, \alpha_4, \alpha_{11}, \alpha_{12}, \ldots, \alpha_{14}, \alpha_{22}, \alpha_{23}, \ldots, \alpha_{44})'$

and the definitions of $\mathbf{x}_{int}$ ($i = 1, \ldots, 4$) conform to the definitions of β_i ($i = 1, \ldots, 4$) and are obvious. Notice from equations (9) and (10) that the β_i ($i = 1, \ldots, 4$) have many elements in common. Indeed, the restrictions given by equation (4) and the restrictions implicit in equations (9) and (10) together mean that only 11 of the 31 parameters in the β_i ($i = 1, \ldots, 4$) are 'free'. Those that are redundant or 'not free' can be obtained from the other parameters and restrictions.

Stacking equation (8) by firm, time period and then by equation we obtain

$$(11)\quad \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \mathbf{y}_3 \\ \mathbf{y}_4 \end{bmatrix} = \begin{bmatrix} \mathbf{X}_1 & & & \\ & \mathbf{X}_2 & & \\ & & \mathbf{X}_3 & \\ & & & \mathbf{X}_4 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \end{bmatrix}$$

where $\mathbf{y}_i = (y_{i11}, y_{i21}, \ldots, y_{iN1}, y_{i12}, y_{i22}, \ldots, y_{iN2}, \ldots, y_{iNT})'$ is (NT×1) for all i and $\mathbf{X}_i$ and ε_i are similarly defined, although it is worth noting that $\mathbf{X}_i$ is (NT×5) for i=1, ..., 3 and $\mathbf{X}_4$ is (NT×16). Thus, we can write the empirical model more compactly as:

$$(12)\quad \mathbf{y} = \mathbf{X}\beta + \varepsilon$$

where the definitions are obvious. The parametric restrictions implied by equations (4), (9) and (10) and our assumptions concerning the error vector ε can also be written more compactly as:

$$(13)\quad \mathbf{R}\beta = \mathbf{r}$$

$$(14)\quad E\{\varepsilon\} = \mathbf{0}$$

(15) $E(\varepsilon\varepsilon') = \Omega = \Sigma \otimes I_{NT}$

where $\Sigma = [\sigma_{lm}]$ and **r** and **R** are known matrices of order (20×31) and (20×1) respectively. The model given by equations (12) to (15) is a standard restricted SUR model (see Judge *et al*, 1985, p.469-473).

To obtain maximum likelihood estimates we note that the restricted Generalised Least Squares (GLS) estimator for β is

(16) $\bar{\beta} = \hat{\beta} + \mathbf{C}\mathbf{R}'(\mathbf{R}\mathbf{C}\mathbf{R}')^{-1}(\mathbf{r}-\mathbf{R}\hat{\beta})$

where $\mathbf{C} = [\mathbf{X}'(\Sigma^{-1}\otimes I_{NT})\mathbf{X}]^{-1}$ and $\hat{\beta} = \mathbf{C}\mathbf{X}'(\Sigma^{-1}\otimes I_{NT})y$ is the unrestricted GLS estimator. In practice, restricted Estimated Generalised Least Squares (EGLS) estimates can be obtained by replacing Σ in equation (16) with an estimator $\dot{\Sigma}$ constructed using restricted or unrestricted OLS residuals. Of course, another estimate of β can be obtained by replacing Σ in equation (16) with an estimator based on the restricted EGLS residuals (rather than OLS residuals). In fact, we can continue to update our estimates of β and Σ in an iterative way and, if the disturbances are multivariate normal, this iterative process will yield the maximum likelihood estimates.

The iterative process described above can be time-consuming if the number of restrictions to be imposed and parameters to be estimated at each step is large. An alternative but equivalent estimation procedure, which is not only faster but can also be usefully exploited in our Bayesian approach, involves maximum likelihood estimation of the subset of 11 'free' parameters in β. After convergence, the remaining 20 maximum likelihood estimates are derived from the 20 parametric restrictions $\mathbf{R}\beta = \mathbf{r}$. To implement the procedure we rearrange the rows of β and the columns of **X** and **R** in such a way that equations (12) and (13) can be written in the following partitioned form:

(17) $y = \mathbf{X}\beta + \varepsilon = [\mathbf{X}_1 \;\; \mathbf{X}_2]\begin{bmatrix}\eta \\ \gamma\end{bmatrix} + \varepsilon$

(18) $\mathbf{R}\beta = [\mathbf{R}_1 \;\; \mathbf{R}_2]\begin{bmatrix}\eta \\ \gamma\end{bmatrix} = \mathbf{r}$

where X_1, X_2, R_1, R_2, γ and η are (NT×20), (NT×11), (20×20), (20×11), (11×1) and (20×1) respectively. The vector γ contains the subset of 11 'free' parameters to be estimated in the first stage, and η contains the 20 remaining parameters in β which will be estimated using estimates of γ and the following equivalent form of equation (18):

$$(19) \qquad \eta = R_1^{-1}(r - R_2\gamma)$$

Recall that the vector γ of 'free' parameters contains parameters which cannot be obtained from other parameters and restrictions. To estimate γ we use (19) to rewrite (17) in the form:

$$(20) \qquad y = X^*\gamma + \varepsilon$$

where $X^* = X_2 - X_1R_1^{-1}R_2$ and we have exploited the fact that r is a zero vector. The model given by (20), (14) and (15) is an unrestricted SUR model, with (unrestricted) GLS estimator for γ given by

$$(21) \qquad \hat{\gamma} = C^*X^{*\prime}(\Sigma^{-1}\otimes I_{NT})y$$

where $C^* = [X^{*\prime}(\Sigma^{-1}\otimes I_{NT})X^*]^{-1}$. Again, in practice, EGLS estimates can be obtained by replacing Σ with an estimator constructed using unrestricted OLS residuals. Moreover, if the disturbances are multivariate normal then a maximum likelihood estimate for γ can be obtained using the iterative procedure described above.

3.2 Bayesian Estimation

The formulation of our empirical model as an unrestricted SUR model (equations (20), (14) and (15)) is convenient for Bayesian analysis because a number of relevant results already appear in the mainstream econometrics literature (eg. Judge *et al*, p.478-80). The cornerstone of Bayesian analysis is Bayes Theorem which, in the present context, allows us to state that:

(22) $f(\gamma, \Sigma \mid y) \propto L(y \mid \gamma, \Sigma)\, p(\gamma, \Sigma)$

where $\propto$ denotes 'proportional to', $f(\gamma, \Sigma \mid y)$ is the posterior joint density function for γ and Σ given y (the posterior density summarises all the information about γ and Σ after the sample y has been observed), $L(y \mid \gamma, \Sigma)$ is the likelihood function (summarising all the sample information), and $p(\gamma, \Sigma)$ is the prior density function for γ and Σ (summarising the nonsample information about γ and Σ). Our interest lies in the posterior density $f(\gamma, \Sigma \mid y)$, and characteristics (eg. means and variances) of posterior marginal densities which can be derived from it.

We begin with a standard Bayesian treatment of the unrestricted SUR model and assume that ε is multivariate normal. Under this assumption the likelihood function is given by (Judge *et al*, p.478)

(23) $L(y \mid \gamma, \Sigma) \propto |\Sigma|^{-NT/2} \exp[-.5\, tr(A\Sigma^{-1})]$

where A is the 4x4 symmetric matrix with $(i, j)^{th}$ element $a_{ij} = (y_i - X_i^* \gamma)'(y_j - X_j^* \gamma)$. In addition, we use a 'noninformative' joint prior:

(24) $p(\gamma, \Sigma) = p(\gamma)\, p(\Sigma)\, I(\gamma \in \Gamma_s)$ $\qquad s = 1, 2$

where $p(\gamma) \propto$ constant, $p(\Sigma) \propto |\Sigma|^{-(I+1)/2}$ is the limiting form of a Wishart density, the Γ_s are the sets of permissible parameter values when monotonicity and concavity information is (s=1) and is not (s=2) available, and $I(\gamma \in \Gamma_s)$ is an indicator function which takes the value 1 if $\gamma \in \Gamma_s$ and takes the value 0 otherwise. We choose a noninformative prior because it allows us to compare our maximum likelihood results with our Bayesian results, whether or not monotonicity and concavity information is available. Note that the algebraic form of the prior $p(\gamma, \Sigma)$ is unchanged by the availability of monotonicity and concavity information, even though the region over which it is defined is different. The same is true of the joint posterior density which is given by (Judge *et al*, p.479):

$$f(\gamma, \Sigma \mid y) \propto |\Sigma|^{-(NT+I+1)/2} \exp[-.5(y - X^*\gamma)'(\Sigma^{-1}\otimes I_{NT})(y - X^*\gamma)]\, I(\gamma\in\Gamma_s) \qquad s = 1, 2 \tag{25}$$

$$\propto |\Sigma|^{-(NT+I+1)/2} \exp[-.5\, tr(A\Sigma^{-1})]\, I(\gamma\in\Gamma_s) \qquad s = 1, 2 \tag{26}$$

We are particularly interested in the posterior marginal densities of the elements of γ, and the means and standard deviations of those posterior densities. Unfortunately, these results cannot be obtained from equations (25) and (26) analytically. Instead, we must use MCMC methods to draw a sample from $f(\gamma \mid y)$. We then use these sample observations to plot estimates of the marginal densities of the elements of γ, and to estimate their moments. The two MCMC algorithms we use to generate these samples are the Gibbs sampler and Metropolis-Hastings algorithms.

The Gibbs Sampler

Useful illustrations and discussions of the workings of the Gibbs sampler can be found in Casella and George (1992) and Chib and Greenberg (1993). In the present context, the Gibbs sampler is an algorithm which effectively samples from $f(\gamma \mid y)$ by iterating as follows:

Step 1: Specify starting values γ^0, Σ^0. Set i=0.

Step 2: Generate γ^{i+1} from $f(\gamma \mid \Sigma^i, y)$

Step 3: Generate Σ^{i+1} from $f(\Sigma \mid \gamma^{i+1}, y)$

Step 4: Set i=i+1 and go to Step 2.

This iteration scheme produces a Gibbs sequence or 'chain' $\gamma^1, \Sigma^1, \gamma^2, \Sigma^2, \ldots$ with the property that, for large k, γ^k is effectively a sample point from $f(\gamma \mid y)$. Thus, in practice, $\gamma^k, \ldots, \gamma^{k+m}$ can be regarded as a sample from $f(\gamma \mid y)$. In this paper we set k=6,000 (the 'burn-in' period) and draw a sample of size m=24,000. We generate samples of this size because the observations in our Gibbs chain are correlated; smaller sample sizes could have been used if our sample of size m was constructed using the only the last observation in m independent Gibbs chains.

Notice from Steps 2 and 3 that in order to make the Gibbs sampler operational we need the conditional densities $f(\gamma \mid \Sigma, y)$ and $f(\Sigma \mid \gamma, y)$.

To obtain (the kernel of) the conditional posterior pdf $f(\gamma \mid \Sigma, y)$ we use (25) and view Σ as a constant, yielding

$$(27) \qquad f(\gamma \mid \Sigma, y) \propto \exp[-.5(\gamma - \hat{\gamma})'\mathbf{X}^{*\prime}(\Sigma^{-1} \otimes I_{NT})\mathbf{X}^{*}(\gamma - \hat{\gamma})]\, I(\gamma \in \Gamma_1)$$

where $\hat{\gamma}$ is the GLS estimator given by equation (21). Thus, $f(\gamma \mid \Sigma, y)$ is proportional to the density function of a multivariate normal random variable with mean vector $\hat{\gamma}$ and covariance matrix $[\mathbf{X}^{*\prime}(\Sigma^{-1} \otimes I_{NT})\mathbf{X}^{*}]^{-1}$. Accordingly, in Step 2 we generate observations γ^{i} using (see Dhrymes, 1970, p.14)

$$(28) \qquad \gamma^{i} = \hat{\gamma} + \mathbf{P}\mathbf{z}$$

where $\mathbf{z} = (z_1, \ldots, z_{11})'$ is a vector of independent standard normal variates $z_j \sim N(0,1)$ $(j=1, \ldots, 11)$, and $\mathbf{P}$ is the (11×11) matrix such that $\mathbf{PP}' = [\mathbf{X}^{*\prime}(\Sigma^{-1} \otimes I_{NT})\mathbf{X}^{*}]^{-1}$.

Finally, to obtain the kernel of the conditional posterior pdf $f(\Sigma \mid \gamma, y)$ we use (27) and view γ as a constant, yeilding

$$(29) \qquad f(\Sigma \mid \gamma, y) \propto \frac{1}{|\Sigma|^{(NT+I+1)/2}} \exp[-.5\, tr(A\Sigma^{-1})]$$

Thus, $f(\Sigma \mid \gamma, y)$ is proportional to an Inverted Wishart density function with parameters $\mathbf{A}$, NT and I (see Zellner, 1971, p.395). Moreover, in Step 3 we can generate observations Σ^{i} using (see Dhrymes, p.14)

$$(30) \qquad \Sigma^{i} = [\mathbf{HZZ'H'}]^{-1}$$

where $\mathbf{Z} = [z_{ij}]$ is an $(I \times NT)$ matrix of independent standard normal variates $z_{ij} \sim N(0,1)$ $(i=1, \ldots, I;\ j=1, \ldots, NT)$ and $\mathbf{H}$ is the $(I \times I)$ matrix such that $\mathbf{HH}' = \mathbf{A}^{-1}$.

The Metropolis-Hastings Algorithm

A description of the workings of the Metropolis-Hastings algorithm can be found in Chib and Greenberg. In the present context, a Metropolis-Hastings algorithm which allows us to impose monotonicity and concavity at price vectors $w_1, \ldots, w_T$ proceeds iteratively as follows:

Step 1: Specify an arbitrary starting value γ^0 which satisfies the constraints. Set i=0.

Step 2: Given the current value γ^i, use a symmetric transition density $q(\gamma^i, \gamma^c)$ to generate a candidate for the next value in the sequence, γ^c.

Step 3. Use the candidate value γ^c and prices $w_1, \ldots, w_T$ to evaluate the monotonicity and concavity constraints. If any constraints are violated return to Step 2.

Step 4: Calculate $\alpha(\gamma^i, \gamma^c) = \min(f(\gamma^c)/f(\gamma^i)), 1)$

Step 5: Generate an independent uniform random variable U from the interval [0,1].

Step 6: Set $\gamma^{i+1} = \begin{cases} \gamma^c \text{ if } U \leq \alpha(\gamma^i, \gamma^c) \\ \gamma^i \text{ if } U > \alpha(\gamma^i, \gamma^c) \end{cases}$

Step 7: Set i=i+1 and go to Step 2.

Again, this iteration scheme produces a chain $\gamma^1, \gamma^2, \ldots$ with the property that, for large k, γ^k is effectively a sample point from $f(\gamma \mid y)$. Thus, the sequence $\gamma^k, \ldots, \gamma^{k+m}$ can once again be regarded as a sample from $f(\gamma \mid y)$. Importantly, this sequence satisfies the monotonicity and concavity constraints at price vectors $w_1, \ldots, w_T$. In this paper we use a burn-in period of k=45,000, we draw a sample of size m=65,000, and we impose the monotonicity and concavity at the T sets of average input prices $w_t = \Sigma w_{nt}/N$ (t=1, ..., T). In Step 3 the monotonicity constraint is evaluated using the signs of the predicted cost shares, while the concavity constraint is evaluated using the maximum eigenvalue of the estimated Hessian matrix of the cost function.

Notice from Steps 1, 2 and 4 that in order to make the Metropolis-Hastings algorithm operational we need an arbitrary starting value γ^0 which satisfies the constraints. We also need the transition density $q(\gamma^i, \gamma^c)$ and the kernel of $f(\gamma)$.

We use the starting values α_i=0.25 (i=1, ..., 4) and α_{ij}=0 for all i≠j. These starting values satisfy the constraints implied by economic theory but may be some distance from the mean of f(γ | y). Our unusually long burn-in period helps ensure the convergence of our Metropolis-Hastings chain.

The transition density $q(\gamma^i, \gamma^c)$ is assumed to be multivariate normal with mean γ^i and covariance matrix $[X^{*\prime}(\hat{\Sigma}^{-1}\otimes I_{NT})X^*]^{-1}$ (the estimated covariance matrix of the restricted SUR estimator $\hat{\gamma}$). In practice, it is commonplace to multiply the (arbitrarily chosen) covariance matrix by a constant h in order to manipulate the rate at which the candidate γ^c is accepted as the next value in the sequence. In this paper we set h=0.02 in order to obtain an acceptance rate of approximately 0.5. This constant was chosen by trial and error. Given our assumptions concerning the transition density it is possible to generate γ^c using (see Dhrymes, p.14)

(31) $$\gamma^c = \gamma^i + \sqrt{h}\,\mathbf{P}z$$

where **P** has been defined above.

The kernel of the marginal density f(γ) can be obtained by integrating Σ out of the joint posterior (26) (see Judge *et al.* p.479):

(32) $$f(\gamma \mid y) \propto |A|^{-NT/2}\, I(\gamma \in \Gamma_2)$$

In practice, we attempt to reduce numerical problems by calculating $f(\gamma^c)/f(\gamma^i)$ as:

(33) $$f(\gamma^c)/f(\gamma^i) = \exp[\,-(NT/2)\log(|A^c|) + (NT/2)\log(|A^i|)\,]$$

where A^c and A^i represent A evaluated at γ^c and γ^i respectively.

4. DATA

The data was originally collected by the Australian Bureau of Agricultural Economics (ABARE) as part of its Australian Sheep Industry Surveys (ASIS). Our sample consists of 310 time-series and cross-section observations on Australian merino woolgrowers, covering the periods 1952-53 to 1962-63 (t=1, ..., 11) and 1964-65 to 1975-76 (t=13, ..., 24). Each observation in the original data set is a record of the average financial and physical characteristics of a group of firms. These obervations are used to construct observations on output (q), total cost (C), input quantities (**x**) and input prices (**w**). Inputs were grouped into one of four broad categories: land, capital, livestock and other inputs (including labour, equipment, materials and services). A more complete description of the data can be found in O'Donnell and Woodland.

5. RESULTS

The results were generated using SHAZAM (White, 1978). In this section the results are evaluated in terms of estimates of the unknown parameters, predicted cost shares, eigenvalues of the estimated Hessian matrix of second order derivatives of the cost function, and estimates of the own- and cross-price elasticities.

Parameter Estimates

Maximum likelihood estimates of the structural parameters β are presented in Table 1 along with the means of the samples we obtained using the Gibbs and Metropolis-Hastings algorithms. The numbers in parentheses are either the standard errors of the maximum likelihood estimates or the standard deviations of our MCMC samples.

Our maximum likelihood estimates are similar to the maximum likelihood estimates presented by O'Donnell and Woodland. Thus, it appears that our specification of a less complex stochastic structure and our focus on only one woolgrowing sector instead of three has had little or no effect on the signs or magnitudes of the slope coefficients or their standard errors. Note that all coefficients are statistically different from zero at the usual levels of significance. Also note that the coefficient of the time variable in the cost function is a measure of the rate of technical progress, and our maximum likelihood estimates suggest that the annual proportional reduction in unit costs as a result of technical change is in the order of 3.2%, only slightly higher than the estimate of 2.9% reported by O'Donnell and Woodland.

The strong similarity between the maximum likelihood and Gibbs estimates presented in Table 1 reflects our use of a noninformative prior. The use of a noninformative prior implies that the location and shape of the likelihood function $L(y \mid \gamma, \Sigma)$ will govern the location and shape of the posterior density $f(\gamma, \Sigma \mid y)$, and, of course, our maximum likelihood and Gibbs results have been obtained using these two functions. Note, however, that the standard deviations of the Gibbs samples are always higher than the standard errors of the maximum likelihood estimates. These differences arise because the standard errors of the maximum likelihood estimates do not account for the uncertainty associated with the estimation of the variance-covariance matrix Σ. The standard deviations of the Gibbs samples do account for this uncertainty. For this reason, and because the maximum likelihood and Gibbs estimates are very similar, we shall ignore the maximum likelihood estimates in the remainder of this paper.

Finally, there is a reasonable similarity between the Gibbs and Metropolis-Hastings estimates presented in Table 1. In fact, only the first- and second-order coefficients associated with the livestock input appear to be very different: the means of the distributions of α_3, α_{23} and α_{33} appear to change significantly with the imposition of the monotonicity and concavity constraints. Violations of these constraints are assessed below in terms of predicted cost shares and the eigenvalues of the estimated Hessian matrix.

<u>Predicted Cost Shares</u>

Monotonicity requires that the predicted cost shares be positive. The Gibbs samples were used to check this requirement at the 23 sets of average input prices $w_t = \Sigma w_{nt}/N$ ($t=1, \ldots, 11, 13, \ldots, 24$). The distributions of the predicted cost shares were uniformly found to lie between zero and one, indicating that monotonicity was satisfied without the imposition of constraints.

Eigenvalues

For the estimated cost function to be consistent with economic theory it must be concave, requiring that the estimated Hessian matrix of second-order price derivatives be negative semi-definite. A necessary and sufficient condition for negative semi-definiteness is that all the eigenvalues be non-positive. O'Donnell and Woodland used this criterion in their study and found that the concavity condition was violated in 36 out of 69 cases, with many violations attributed to a positive response of the livestock input to a change in its own price. The livestock input seems to play an equally important role in our study: recall that only the (means of the distributions of the) livestock coefficients appear to change significantly with the imposition of the constraints.

Our Gibbs samples were used to check concavity at our 23 sets of average input prices and some of the results are presented in Table 2. Table 2 presents the means and standard deviations of the maximum eigenvalues calculated using the Gibbs samples. 11 out of 23 of these means are positive, a result which is consistent with the results of O'Donnell and Woodland. Table 2 also presents the means of the maximum eigenvalues calculated using the Metropolis-Hastings samples. Of course, the eigenvalues calculated using the Metropolis-Hastings samples have been constrained to be non-positive.

The effects of imposing concavity on our estimated cost function are further illustrated in Figures 1 to 4 where we present the estimated posterior pdf's of the maximum eigenvalues in four representative time periods (t=1, 5, 23 and 24).

Several important observations emerge from Figures 1 to 4. First, in all cases there is a noticeable leftward shift in the posterior pdf after the imposition of the concavity constraint, even when the unconstrained pdf already has a high positive probability over the negative range of the real line (eg. Figure 2). This suggests that the posterior correlations between the maximum eigenvalues in different time periods may be high. Second, most of the constrained posterior pdf's have a regular shape, without severe truncation at zero, even when the unconstrained pdf has a high positive probability over the positive range of the real line (eg. Figure 1). An exception is the constrained posterior pdf in time period t=24 (Figure 4). An explanation may be found in Table 2, where we observe that the posterior mean of the unconstrained maximum eigenvalue in time period t=24 is higher than in any other time period. This suggests that constraining this eigenvalue to be negative may have been sufficient to impose negativity on the remainder. Third, the constrained posterior pdf's appear to have smaller variance than the unconstrained pdf's, an observation which is confirmed by looking at the standard deviations presented in Table 2. This is not surprising: the imposition of restrictions will make the Metropolis-Hastings estimates more precise. Finally, the unconstrained posterior pdf's depicted in Figures 2 and 3 illustrate how the means of the eigenvalues reported in Table 2 may understate the degree to which the concavity constraint is violated. Even when the posterior mean is negative, concavity can be regarded as being violated whenever the posterior pdf of the maximum eigenvalue has positive probability over the positive range of the real line. Using this criterion, the Gibbs estimates violate concavity in 14 out of 23 time periods.

Elasticities

The imposition of regularity conditions on our estimated cost function leads to significant changes in the posterior distributions of a number of own- and cross-price elasticities. To briefly illustrate, Table 3 reports the means and standard deviations of the pdf's of input price elasticities calculated at average prices $w = \Sigma w_{nt}/NT$.

Three features of Table 3 are of particular interest, not least because they tend to confirm our earlier judgement concerning the role of the livestock input. First, (the means of) all own-price elasticities are correctly signed and indicate that all input demands are inelastic with respect to their own prices. Moreover, the only own-price elasticity which seems to be affected by the imposition of the constraints is the own-price elasticity for livestock: the mean of this elasticity decreases from -0.13 to -0.33. Second, the standard deviations of the constrained and unconstrained pdf's are generally similar. Again, the only notable exception is the standard deviation of the own-price elasticity for livestock, which falls dramatically with the imposition of the constraints. Finally, the two cross-price elasticities which measure the relationships between the prices and quantities of livestock and other inputs undergo a sign reversal with the imposition of the constraints. Thus, all pairs of inputs now appear to be substitutes in production.

6. CONCLUSION

In this paper we use Markov Chain Monte Carlo (MCMC) methods to impose regularity conditions on a system of cost and cost-share equations. This Bayesian methodology represents an alternative to conventional sampling theory techniques which typically destroy the flexibility properties of many of the more popular functional forms. The methodology has previously been used by Terell to estimate a cost function using the well-known Berndt and Wood (1975) data set. Our own empirical application has been motivated by the large number of regularity violations reported in the study by O'Donnell and Woodland. Thus, our empirical model is based on the translog model of O'Donnell and Woodland and estimated using (a part of) their data set.

The empirical results we present include parameter estimates, eigenvalue estimates and estimates of input price elasticities for models with and without regularity constraints imposed. Our unconstrained MCMC estimates are almost identical to our maximum likelihood estimates and the maximum likelihood estimates of O'Donnell and Woodland. Our constrained MCMC estimates differ from our unconstrained estimates in several respects: all maximum eigenvalues become negative in accordance with economic theory, standard deviations become smaller, and the signs and magnitudes of coefficients and elasticities associated with the livestock input undergo substantial change. This last result is consistent with the finding of O'Donnell and Woodland that a large number of regularity violations are associated with the livestock input.

Our paper demonstrates how Bayesian techniques can be used to impose a large number of inequality constraints on an estimated cost or profit function. We also demonstrate the use of two important numerical algorithms which can be used to generate samples from (analytically) problematic marginal probability densities. Finally, we show how the Bayesian approach can be used to conveniently summarise important information in the form of posterior pdf's.

Some possible extensions of our study include the specification of a more complex error structure. One possibility is to consider the heteroskedastic error components structure of O'Donnell and Woodland. Another possibility is the truncation of one or more of these error components in line with the stochastic specifications popular in the frontier literature. Other possible extensions include the use of alternative functional forms and relaxation of the assumption of constant returns to scale.

REFERENCES

Barnett, W A. (1983) "New Indices of Money Supply and the Flexible Laurent Demand System", *Journal of Business and Economic Statistics*, 1, 7-23.

Berndt, E.R. and D.O. Wood (1975) "Technology, Prices and the Derived Demand for Energy", *Review of Economics and Statistics*, 57, 259-68.

Berndt, E.R. and M.S. Khaled (1979) "Parametric Productivity Measurement and Choice Among Flexible Functional Forms", *Journal of Political Economy*, 87, 1220-1245.

Casella, G and E.I. George (1992) 'Explaining the Gibbs Sampler', *The American Statistician*, 46(3):167-173.

Chalfant, J.A. and N.E. Wallace (1992) "Bayesian Analysis and Regularity Conditions on Flexible Functional Forms: Application to the US Motor Carrier Industry", in Griffiths, W.E., H. Lutkepohl and M.E. Bock (eds) *Readings in Econometric Theory and Practice: A Volume in Honor of George Judge*, North-Holland, Amsterdam.

Chambers, R.G. (1988) *Applied Production Analysis*, Cambridge University Press, Cambridge.

Chib, S and E. Greenberg (1993) 'Markov Chain Monte Carlo Methods in Econometrics', *mimeo*.

Christensen, L.R., D.W. Jorgensen and L.J. Lau (1971) "Conjugate Duality and the Transcendental Logarithmic Production Function", *Econometrica*, 39, 255-256.

Coelli, T.J. (1996) "Measurement of Total Factor Productivity Growth and Biases in Technological Change in Western Australian Agriculture", *Journal of Applied Econometrics*, 11, 77-91.

Dhrymes, P.J. (1970) *Econometrics: Statistical Foundations and Applications*, Springer-Verlag, NY.

Diewert, W.E. (1971) "An application of the Shephard Duality Theorem: A Generalised Leontief Production Function", *Journal of Political Economy*, 79, 481-507.

Diewert, W.E. and T.J. Wales (1987) "Flexible Functional Forms and Global Curvature Conditions", *Econometrica*, 55(1), 43-68.

Gallant A.R. and E.H. Golub (1984) "Imposing Curvature Restrictions on Flexible Functional Forms", *Journal of Econometrics*, 26, 295-321.

Jorgensen, D.W. and B.M. Fraumeni (1981) "Relative Prices and Technical Change" in Berndt, E. and B. Field (eds) *Modelling and Measuring Natural Resource Substitution*, MIT Press, Cambridge Mass.

Judge, G.G., W.E. Griffiths, R.C. Hill, H. Lutkepohl and T.C. Lee (1985) *The Theory and Practice of Econometrics* (2nd ed.), John Wiley, NY.

Lau, L.J. (1978) "Testing and Imposing Monotonicity, Convexity and Quasiconvexity Constraints", in Fuss, M. and D. McFadden (eds) *Production Economics: A Dual Approach to Theory and Applications, Vol. 1*, North-Holland, Amsterdam.

O'Donnell, C.J. and A.D. Woodland (1995) "Estimation of Australian Wool and Lamb Production Technologies Under Uncertainty. An Error-Components Approach", *American Journal of Agricultural Economics*, 77(August), 552-565.

Talpaz, H., W.P. Alexander and C.R. Shumway (1989) "Estimation of Systems of Equations Subject to Curvature Constraints", *Journal of Statistical Computation and Simulation*, 32(July), 201-214.

Terrell, D. (1996) 'Incorporating Monotonicity and Concavity Conditions in Flexible Functional Forms', *Journal of Applied Econometrics*, 11(2):179-94.

Wiley, D.E., W.H. Schmidt and W.J. Bramble (1973) "Studies of a Class of Covariance Structure Models", *Journal of the American Statistical Association*, 68, 317-323.

White, K.J. (1978) "A General Computer Program for Econometric Methods - SHAZAM", *Econometrica*, January 1978, 239-240.

Zellner. A. (1971) *An Introduction to Bayesian Inference in Econometrics*, John Wiley, NY.

Table 1: Structural Parameters

		ML[a]	Gibbs[b]	Metropolis-Hastings[b]
Constant		-0.5953 (0.0578)	-0.5959 (0.0619)	-0.8351 (0.0491)
α_1	Land	0.2503 (0.0053)	0.2501 (0.0058)	0.2514 (0.0053)
α_2	Capital	0.6742 (0.0188)	0.6743 (0.0203)	0.6658 (0.0180)
α_3	Livestock	0.4400 (0.0126)	0.4401 (0.0132)	0.3445 (0.0071)
α_4	Other	-	-	-
α_{11}	Land/Land	0.0234 (0.0007)	0.0234 (0.0007)	0.0235 (0.0007)
α_{12}	Land/Capital	0.0176 (0.0008)	0.0176 (0.0009)	0.0177 (0.0008)
α_{13}	Land/Livest.	-0.0062 (0.0006)	-0.0062 (0.0007)	-0.0060 (0.0007)
α_{14}	Land/Other	-	-	-
α_{22}	Capital/Capital	0.1151 (0.0059)	0.1151 (0.0064)	0.1151 (0.0059)
α_{23}	Capital/Livest.	-0.0074 (0.0024)	-0.0074 (0.0026)	-0.0064 (0.0022)
α_{24}	Capital/Other	-	-	-
α_{33}	Livest./Livest.	0.0757 (0.0022)	0.0757 (0.0024)	0.0567 (0.0011)
α_{34}	Livest./Other	-	-	-
α_{44}	Other/Other	-	-	-
α_T	Time	-0.0321 (0.0020)	-0.0320 (0.0025)	-0.0333 (0.0023)

[a] Numbers in parentheses are standard errors.

[b] Numbers in parentheses are standard deviations of the MCMC samples.

Table 2: Maximum Eigenvalues[a]

Year	t	Gibbs	Metropolis-Hastings
1952-53	1	0.0248 (0.0046)	-0.0101 (0.0019)
1953-54	2	-0.0149 (0.0055)	-0.0250 (0.0043)
1954-55	3	-0.0099 (0.0047)	-0.0276 (0.0039)
1955-56	4	-0.0149 (0.0047)	-0.0364 (0.0035)
1956-57	5	-0.0120 (0.0046)	-0.0328 (0.0036)
1957-58	6	-0.0220 (0.0056)	-0.0321 (0.0040)
1958-59	7	0.0098 (0.0041)	-0.0213 (0.0016)
1959-60	8	-0.0145 (0.0047)	-0.0345 (0.0036)
1960-61	9	0.0100 (0.0040)	-0.0212 (0.0013)
1961-62	10	0.0131 (0.0041)	-0.0189 (0.0015)
1962-63	11	0.0179 (0.0042)	-0.0153 (0.0015)
1964-65	13	-0.0217 (0.0054)	-0.0371 (0.0037)
1965-66	14	-0.0190 (0.0051)	-0.0379 (0.0037)
1966-67	15	-0.0190 (0.0051)	-0.0347 (0.0037)
1967-68	16	-0.0055 (0.0043)	-0.0329 (0.0022)
1968-69	17	0.0054 (0.0041)	-0.0244 (0.0015)
1969-70	18	-0.0006 (0.0043)	-0.0295 (0.0017)
1970-71	19	0.0006 (0.0043)	-0.0287 (0.0017)

Table 2: continued.

Year	t	Gibbs	Metropolis-Hastings
1971-72	20	0.0229 (0.0042)	-0.0116 (0.0012)
1972-73	21	0.0093 (0.0040)	-0.0215 (0.0012)
1973-74	22	0.0131 (0.0067)	-0.0201 (0.0035)
1974-75	23	-0.0068 (0.0051)	-0.0352 (0.0025)
1975-76	24	0.0357 (0.0043)	-0.0006 (0.0006)

[a] Numbers in parentheses are standard deviations of the MCMC samples.

TABLE 3: Input-Price Elasticities at Average Prices[a]

	Price of Land	Price of Capital	Price of Livestock	Price of Other Inputs
Gibbs				
Qty of Land	-0.6468 (0.0109)	0.4934 (0.0104)	0.0269 (0.0070)	0.1263 (0.0179)
Qty of Capital	0.1483 (0.0033)	-0.3145 (0.0219)	0.0725 (0.0088)	0.0937 (0.0221)
Qty of Livestock	0.0245 (0.0067)	0.2187 (0.0270)	-0.1263 (0.0238)	-0.1168 (0.0338)
Qty of Other Inputs	0.0216 (0.0035)	0.0530 (0.0125)	-0.0219 (0.0063)	-0.0527 (0.0147)
Metropolis-Hastings				
Qty of Land	-0.6440 (0.0111)	0.4948 (0.0098)	0.0304 (0.0076)	0.1188 (0.0167)
Qty of Capital	0.1479 (0.0032)	-0.3299 (0.0201)	0.0765 (0.0074)	0.1055 (0.0199)
Qty of Livestock	0.0271 (0.0073)	0.2276 (0.0224)	-0.3260 (0.0054)	0.0712 (0.0215)
Qty of Other Inputs	0.0200 (0.0032)	0.0593 (0.0111)	0.0135 (0.0041)	-0.0928 (0.0122)

[a] Numbers in parentheses are standard deviations of the MCMC samples.

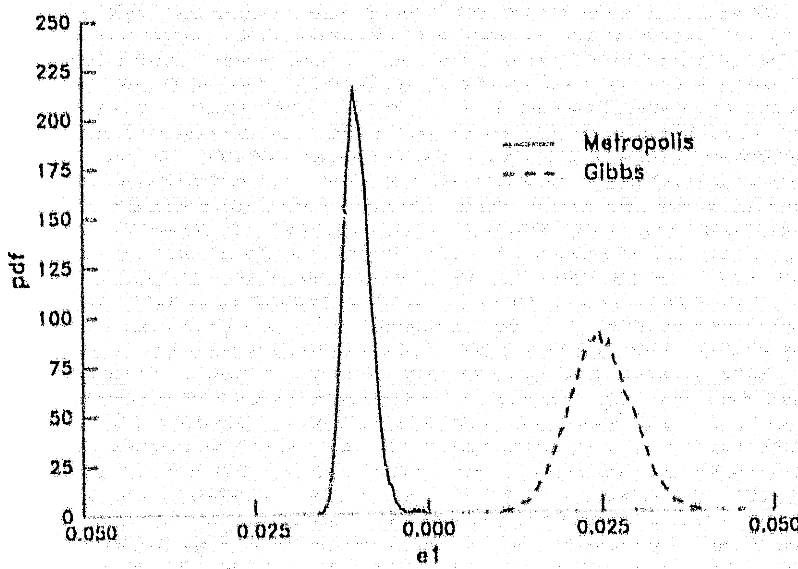

Figure 1: Posterior distribution of the maximum eigenvalue: average 1952-53 prices (t=1)

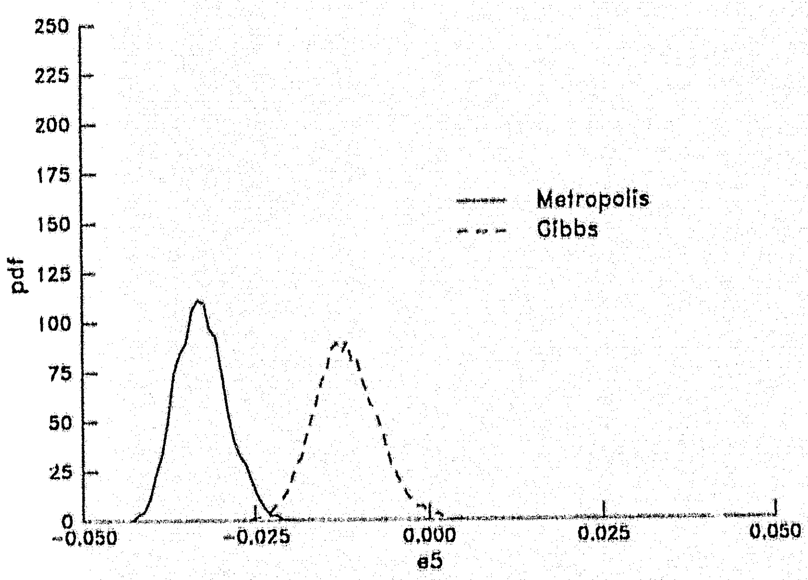

Figure 2: Posterior distribution of the maximum eigenvalue: average 1956-57 prices (t=5)

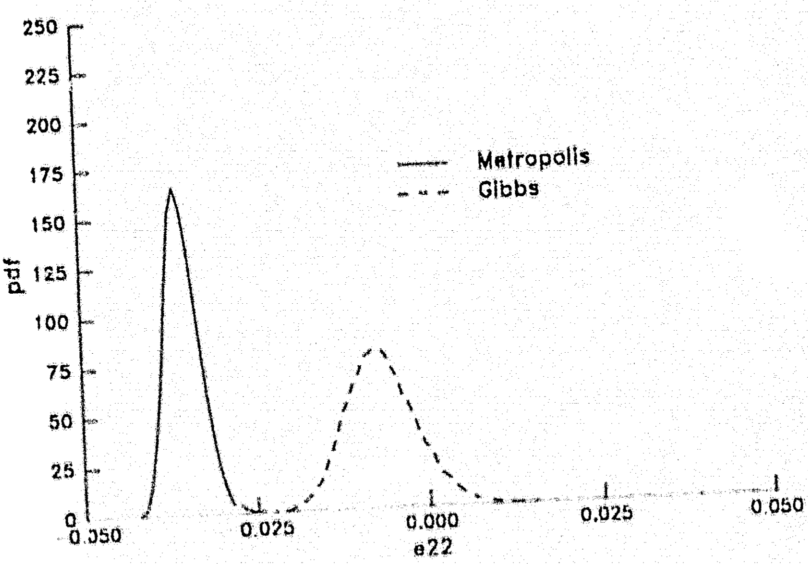

Figure 3. Posterior distribution of the maximum eigenvalue: average 1973-74 prices (t=23)

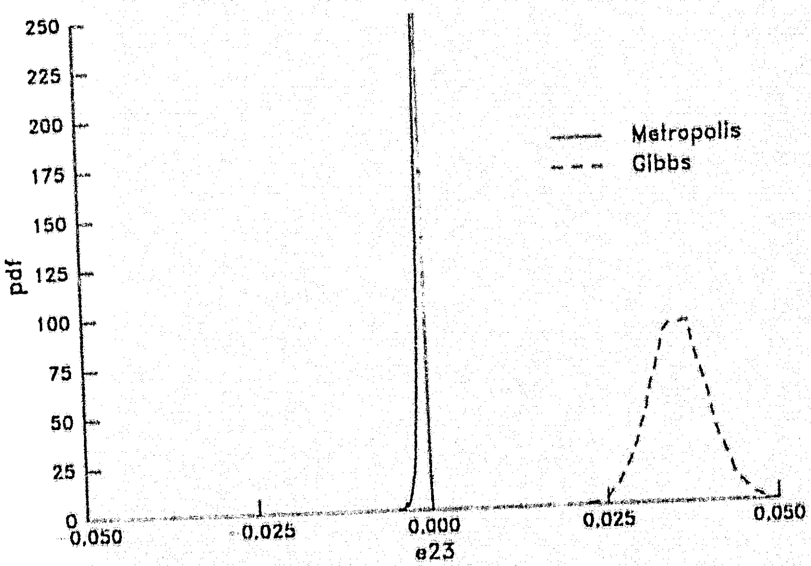

Figure 4: Posterior distribution of the maximum eigenvalue: average 1974-75 prices (t=24)