Bayesian Model Averaging in Consumer Demand Systems with Inequality Constraints

C.L Chua
mchua@unimelb.edu.au
Melbourne Institute of Applied Economic and Social Research
University of Melbourne
Melbourne, Australia

W.E. Griffiths
b.griffiths@unimelb.edu.au
Department of Economics
University of Melbourne
Melbourne, Australia

C.J O'Donnell
codonnel@metz.une.edu.au
School of Economic Studies
University of New England
Armidale, Australia

June, 2001

Paper for presentation at 2001 Australasian Meeting of The Econometric Society, July 6- 8, Auckland, New Zealand

* The paper produced here is part of my PHD work done in University of New England. I would like to thank all the staff in School of Economic Studies, UNE especially those in Econometrics discipline

Abstract

Share equations for the translog and almost ideal demand systems are estimated using Markov Chain Monte Carlo. A common prior on the elasticities and budget shares evaluated at average prices and income is used for both models. It includes equality restrictions (homogeneity, adding up and symmetry) and inequality restrictions (monotonicity and concavity). Posterior densities on the elasticities and shares are obtained; the problem of choosing between the results from the two alternative functional forms is resolved by using Bayesian model averaging. The application is to USDA data for beef, pork and poultry.

Keywords: Marginal likelihood; Metropolis Hastings Approximation; Modified Harmonic Mean; Kleibergen Procedure.

JEL classification: C11, C32, E21.

Bayesian Model Averaging in Consumer Demand Systems with Inequality Constraints

C.L. Chua, W.E. Griffiths and C.J. O'Donnell

1. Introduction

Consumer demand systems, a highly researched area in applied economics, has seen several advances over recent decades. The introduction of flexible functional forms such as the almost ideal demand system (Deaton and Muellbauer, 1980), the translog demand system (Jorgenson, Lau and Stoker, 1982) and the normalised quadratic demand system (Diewert and Wales, 1988) have provided a rich source of consumer demand models. The main attractiveness of a flexible functional form is that the demand model has enough free parameters to provide a second order approximation to an arbitrary twice differentiable function at a particular set of prices. In other words, parameters can mimic cross, own price elasticities and budget shares at a particular set of prices and quantities. However, being able to mimic the shares and elasticities comes at the expense of introducing additional parameters or implies other restrictions relative to non flexible specifications[1]. In addition, unlike other functional forms, a flexible function does not automatically satisfy inequality constraints implied by theory, like concavity and monotonicity. Instead, the estimated demand system is often tested for satisfaction of concavity and monotonicity constraints by checking whether the Slutsky matrix is negative semi definite. Alternatively, the inequality constraints are sometimes imposed on the boundary of the parameter space, a strategy that is neither intuitively satisfying nor statistically attractive.

Two forms of concavity have evolved in empirical demand systems: local and global concavity. Local concavity is a weaker restriction than global concavity where imposed constraints in a demand system lead to concavity at each data point. Which is more suitable depends on the objectives of the research. If the aim is to provide estimates (e.g. cross and own price elasticities) over the observed region, then imposing local concavity is sufficient for the study. On the other hand, global concavity is imposed if the aim is to provide estimates over a wide region that lies outside the observed region.

[1] The number of parameters in the flexible functional forms increases with the square of the number of goods.

Although, global concavity is more attractive than local concavity, Diewert and Wales (1988) found that imposing global concavity destroyed the flexibility of their normalised quadratic reciprocal demand system. More recently, Ryan and Wales (1998)[2] proposed a technique for imposing inequality constraints locally on the almost ideal demand system, the linear translog demand system and the normalised quadratic demand system. Moschini (1999) further developed Ryan and Wales (1998) methods to the basic translog (BTL) and log translog systems (log TL) for which Ryan and Wales believed their methods to be unworkable.

In this paper we follow the Bayesian approach to incorporating inequality constraints as prior beliefs to the almost ideal system and the log translog demand system. Geweke (1986) and Griffiths (1988) showed the feasibility of imposing inequality constraints via Bayesian inference in the normal linear regression model. Their imposition in systems of linear equations has been considered within a demand context by Chalfant, Gray and White (1981) who imposed inequality constraints on linearised almost ideal demand systems (LAIDS) using importance sampling. Along similar lines, Hasegawa, Kozumi, and Hashimoto (1999) imposed the constraints on a Rotterdam Model using Gibbs sampling. The demand systems that we use are systems of non linear equations and the inequality constraints are imposed at every data point. The Markov Chain Monte Carlo algorithm that we employ for the imposition of inequality constraints in demand systems is a Metropolis algorithm similar to that used for expenditure functions by Griffiths and Chotikapanich (2000) and for production systems by Griffiths, O'Donnell and Tan-Cruz (2000).

Given the numerous empirical demand systems suggested in the literature, researchers are faced with choice of a particular model or group of models for their research. In both classical and Bayesian econometrics, different tests are put forward to select the most desirable model. Such tests, still yield doubts about the chosen and unchosen models. Has information been lost by discarding some models? Is uncertainty reflected in measures of precision such as standard errors? Bayesian model averaging (BMA) is a means for resolving these issues. The contribution of each model to estimation is weighted by the probability it is correct; posterior standard deviations reflect the existence of model uncertainty.

[2] Ryan and Wales (2000) have extended their procedure to translog and generalised Leontief cost functions in a production context.

Instead of choosing a suitable demand system from existing demand systems, the demand systems are averaged.

It is within this framework that this paper applies model averaging and methods for imposing inequality constraints to empirical demand systems; a Metropolis-Hastings Markov-chain Monte Carlo algorithm is our main sampling technique. The paper is organised as follows: Section 2 shows how inequality constraints are imposed in a general Bayesian framework. Section 3 discusses how BMA is conducted in the field of economics where, in most cases, we want to ensure that, before averaging, priors are specified for each of the model's parameters. Section 4 applies the methods to the AIDS and Log TL models. An empirical example is given in Section 5, where BMA is applied to the AIDS and Log TL models using quarterly aggregated USDA data for beef, pork and poultry.

2. Imposing Inequality Constraints in Bayesian Statistics

In this section, we consider the imposition of inequality constraints in a general Bayesian framework. The motivation comes from Geweke (1986) who demonstrated the plausibility of including inequality constraints as prior information in a general linear regression framework.

According to Bayes theorem, posterior information is formulated on the basis of sample information and prior information.

Let γ denote a vector of unknown parameters for a given model, and let y denote a vector of sample observations. Then, posterior information is expressed in terms of a posterior probability density function (pdf), $p(\gamma \mid y)$, which is given by

$$p(\gamma \mid y) = \frac{p(y \mid \gamma)\, p(\gamma)}{p(y)} \tag{1}$$

where $p(\gamma)$ is the prior for model parameters, $p(y \mid \gamma)$ is the *likelihood function* which is the conditional density for y given γ, and $p(y)$ is known as the marginal likelihood.

Generally, equation (1) can be rewritten as

$$p(\gamma \mid y) \propto p(y \mid \gamma)\, p(\gamma) \tag{2}$$

Equation (2) recognises that $p(y)$ is a normalizing constant that ensures $p(\gamma \mid y)$ integrates to unity over the posterior space. Also, $p(y)$ measures how well the given model predicts the observed sample y; it is simply the integral of $p(y \mid \gamma)p(\gamma)$ with respect to γ, and is the basis for model comparison and averaging.

To impose the inequality constraints, we attach an indicator function, $I(\Xi)$, to the prior pdf where $I(\Xi)$ takes a value of 1 if the inequality constraints are satisfied and a value of 0 otherwise. That is,

$$p(\gamma) = kg(\gamma)I(\Xi) \tag{3}$$

where $g(\gamma)$ is the kernel of the prior density, and k is the normalising constant. Because of truncation of the parameter space by the inequality constraints, k is generally unknown and needs to be computed or estimated. Also, $g(\gamma)$ is such that $p(\gamma)$ is proper. The inequality constraints can range from simple inequality constraints that confine single parameters to ranges to more complicated inequality constraints that are function of parameters and data. For model comparison or model averaging k can be directly computed for simple inequality constraints and estimated using sampling techniques for complex inequality constraints.

3. Bayesian Model Averaging

Instead of selecting a single "best" model from the pool of uncertain models, BMA uses information from all considered models to draw inferences on what we call “economic quantities of interest” (EQI). Let θ denote the EQI, and M_j be the j^{th} model, $j = 1,...,M$. The posterior distribution of θ obtained by model averaging is as follows:

$$p(\theta \mid y) = \sum_{j=1}^{M} p(\theta \mid y, M_j)P(M_j \mid y) \tag{4}$$

where $p(\theta \mid y, M_j)$ is the posterior distribution for the EQI when the model is M_j, and $P(M_j \mid y)$ is the posterior model probability. Thus, BMA yields a posterior pdf for the EQI that is a weighted average of the posterior distributions for each model.

The posterior model probability is obtained using the discrete version of Bayes rule.

$$P(M_j \mid y) = \frac{P(M_j)\, p(y \mid M_j)}{\sum_{j=1}^{M} P(M_j)\, p(y \mid M_j)} \tag{5}$$

where the $P(M_j)$ are the prior probabilities for each of the models $j = 1,\dots,M$ such that $\sum_{j=1}^{M} P(M_j) = 1$, and $p(y \mid M_j)$ is the marginal likelihood of model j. The essence of model averaging lies with a model's marginal likelihood which measures how well each model predicts the observed data.

Although the computation of a posterior model probability is straightforward, the hindrance in most practical problems lies with the marginal likelihood. In most situations, the marginal likelihood is intractable. One obvious, but inefficient way to estimate it is to simulate parameters from the prior pdf and average the likelihood function. Estimation from the posterior simulation is more efficient. In this regard, several approaches are available for approximation of the marginal likelihood such as the: modified harmonic mean (MHM) method of Gelfand and Dey (1994), and the Metropolis-Hastings approximation (MHA) suggested by Geweke (1999). The Laplace approximation (LA) described by Tierney and Kadane (1986)[3] is another alternative. An advantage of simulating from the prior pdf is that it only requires the normalizing constant from the likelihood. The latter approaches have to account for the normalising constants in both the likelihood function and the prior. Thus, when dealing with inequality constraints, the normalising constant created by the inequality constraints has to be factored into the estimation of the marginal likelihood via MHA, MHM and LA. Depending on the type of inequality constraints, sometimes it is straightforward to compute normalising constants directly and sometimes they are estimated through sampling from the prior.

Instead of dealing with the posterior density of EQI, in practice, we deal with expectations of functions which cover parameters means, second moments around zero, variances, and the proportion of observations in an interval for a histogram.

$$E[g(\theta) \mid y] = \sum_{j=1}^{M} E[g(\theta) \mid y, M_j] P(M_j \mid y) \tag{6}$$

where $g(\theta)$ is a function of θ.

BMA[4] has been widely applied in social statistics and biostatistics. The models used in these fields are mostly linear hierarchical equations with model averaging being done over different combinations of independent variables. Fernandez, Ley and Steel (1997) model average using a linear hierarchical model for fishing activities in the North Atlantic. In this instance, the number of models can be enormous due to the number of independent variables. Madigan and Raftery (1994) proposed Occam's window algorithm for reducing the number of possible models.

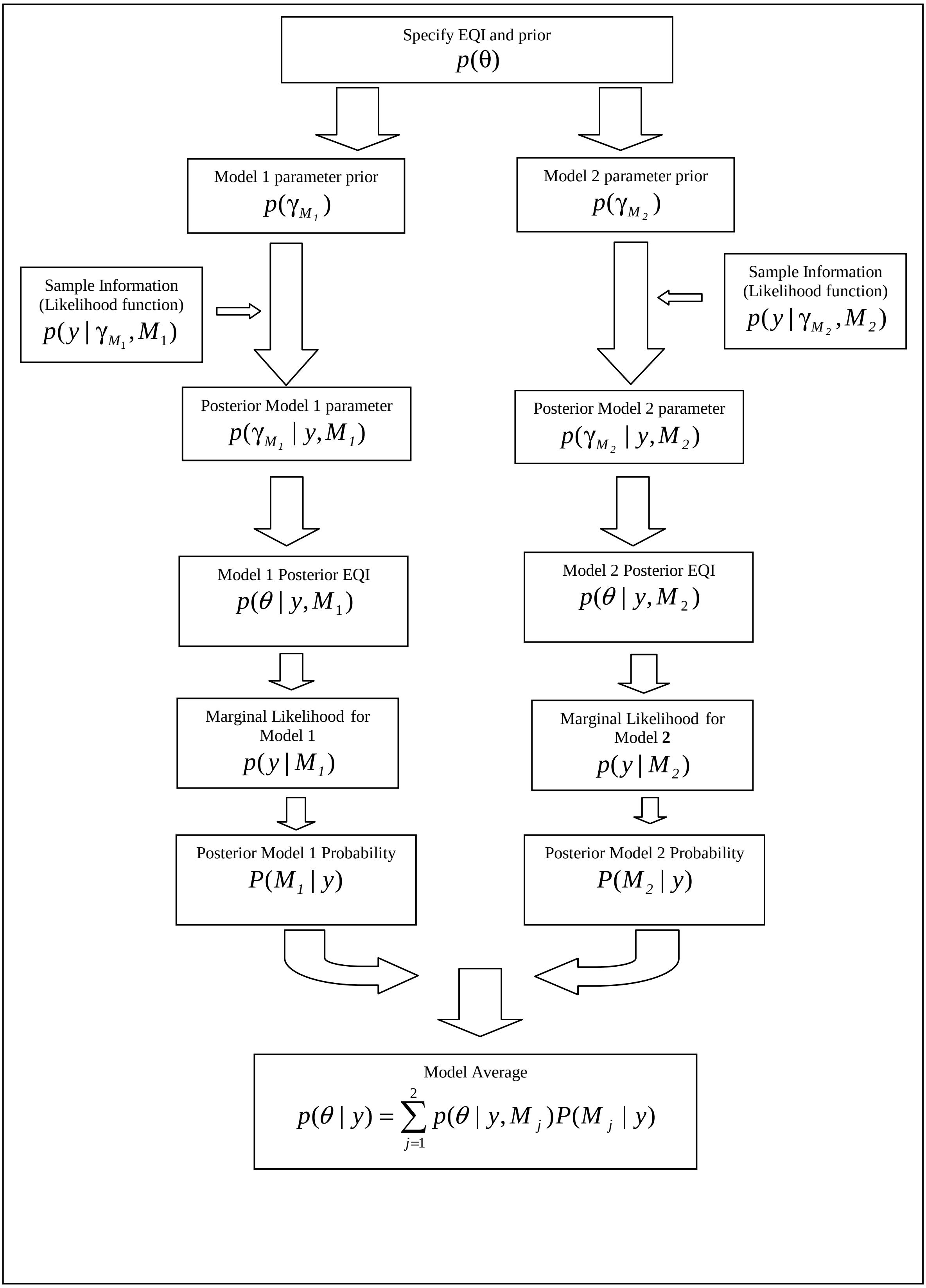
For economic models, model averaging is significantly different in several respects. Firstly, the models used range from simple single equations to systems of non linear equations. The models to be averaged are not normally nested within one another. Secondly, the prior pdfs for each of the model's parameters are derived from a prior on the EQI such that the priors on parameters in different models have approximately the same "content" of information before introducing the likelihood and model averaging. Model averaging in this context involves significantly fewer models and can be more complicated because of the presence of non linear equations and transformations from a prior on the EQI to priors on the parameters for each of the models. The flowchart in Figure 1 shows the basic idea of how model averaging is conducted on two economic models.

---

[3] Raftery (1996) comments that the Laplace method is often not applicable because the derivatives that it requires are not easily available, particular true for complex models that require MCMC.

[4] There is a website for BMA, http://www.research.att.com/~volinsky/bma.html; most of the work that appears on it is based on linear hierarchical equations in the areas of biostatistics and social statistics.

Figure 1: Flowchart on Model Averaging of Two Models.



4. BMA between AIDS and Log TL

We illustrate BMA with two alternative demand models AIDS and Log TL, with the imposition of local concavity and monotonicity in each model. Discussion will mostly focus on the AIDS model; similar methodology is applied to the Log TL for deriving priors on the model parameters, and computing posterior pdfs and marginal likelihoods.

4.1 Budget Shares of AIDS and Log TL

The budget shares for the AIDS model take the following form

$$s_i = \alpha_{Ai} + \sum_j \gamma_{Aij} \log p_j + \beta_{Ai} \left(\log y - \log g(p) \right) \tag{7}$$

where $\log g(p) = \alpha_{A0} + \sum_k \alpha_{Ak} \log p_k + \frac{1}{2} \sum_k \sum_j \gamma_{Akj} \log p_k \log p_j$, s_i is the budget share of the i^{th} good, p_k is the price of the k^{th} good, y is total expenditure and $\alpha_{A0}, \alpha_{Ai}, \beta_{Ai}$ and γ_{Aij} are unknown parameters, $i, j, k = 1,2,\cdots,n$. To ensure the theoretical properties of adding up, homogeneity and symmetry are satisfied, the following equality constraints are imposed on the parameters:

$$\sum_i \alpha_{Ai} = 1, \qquad \sum_i \beta_{Ai} = 0, \qquad \sum_j \gamma_{Aij} = 0, \qquad \gamma_{Akj} = \gamma_{Ajk}$$

The parameter α_0 has been explicitly set to 0 indicating the minimum outlay of each commodity is none.

The Log TL model, in budget share form, is expressed as

$$s_i = \frac{\alpha_{Bi} + \sum_k \gamma_{Bik} \log p_j + \log y \sum_k \gamma_{Bik}}{1 - \sum_j \sum_k \gamma_{Bkj} \log p_j} \tag{8}$$

where s_i is the budget share of the i^{th} good, p_k is the price of the k^{th} good, y is the expenditure and α_{Bi} and γ_{Bij} are unknown parameters, $i, j, k = 1,2,\cdots,n$. The Log TL is always homogenous of degree zero. This means that one less set of equality constraints is placed to ensure homogeneity (see

Christensen, Jorgenson and Lau (1975) for a proof). Thus, the following equality constraints are imposed for adding up and symmetry.

$$\sum_i \alpha_{Bi} = 1, \qquad \sum_i \sum_j \gamma_{Bij} = 0, \qquad \gamma_{Bkj} = \gamma_{Bjk}$$

We denote the AIDS and Log TL models as A and B, respectively. Their parameters have been distinguished in this way in equations (7) and (8). Let γ_A be a vector of unknown parameters for model A after substituting out the equality restrictions; γ_B is similarly defined for model B. A reduced set of non-redundant parameters for a 3 commodity demand system $(n = 3)$, and the one we employ in our analysis, is

$$\gamma_A = [\alpha_{A1} \;\; \alpha_{A2} \;\; \gamma_{A11} \;\; \gamma_{A12} \;\; \gamma_{A22} \;\; \beta_{A1} \;\; \beta_{A2}]' \tag{9}$$

$$\gamma_B = [\alpha_{B1} \;\; \alpha_{B2} \;\; \gamma_{B11} \;\; \gamma_{B12} \;\; \gamma_{B13} \;\; \gamma_{B22} \;\; \gamma_{B23}]' \tag{10}$$

4.2 EQI Prior and Reduced EQI Prior

In applied demand systems, researchers are typically interested in drawing inferences about the marginal effects on consumption when prices or income change. Henceforth, we consider own-price, cross-price and income elasticities and budget shares at mean prices and income as our EQI. As a consequence the EQI vector θ is of dimension (15×1). Using obvious notation, we write it as:

$$\begin{aligned}\theta &= [\theta_1 \;\; \theta_2 \;\; \theta_3 \;\; \theta_4 \;\; \theta_5 \;\; \theta_6 \;\; \theta_7 \;\; \theta_8 \;\; \theta_9 \;\; \theta_{10} \;\; \theta_{11} \;\; \theta_{12} \;\; \theta_{13} \;\; \theta_{14} \;\; \theta_{15}]' \\ &= [\bar{\eta}_{11} \;\; \bar{\eta}_{12} \;\; \bar{\eta}_{13} \;\; \bar{s}_1 \;\; \bar{s}_2 \;\; \bar{\eta}_{21} \;\; \bar{\eta}_{32} \;\; \bar{\eta}_{1x} \;\; \bar{\eta}_{2x} \;\; \bar{s}_3 \;\; \bar{\eta}_{31} \;\; \bar{\eta}_{3x} \;\; \bar{\eta}_{23} \;\; \bar{\eta}_{33} \;\; \bar{\eta}_{22}]'\end{aligned} \tag{11}$$

The data are rescaled such that prices and expenditure at the mean are equal to 1, $\bar{p}_1 = \bar{p}_2 = \bar{p}_3 = \bar{y} = 1$. By doing so, we can simplify the expressions for the budget shares and elasticities, which in general, are function of the parameters as well as prices and expenditure. Rescaling of prices and expenditure has no effect on the elastcities and budget shares because it induces corresponding changes in the parameter values. To allow the data to be the main determinant of the outcomes, the EQI priors were chosen to be relatively uninformative, but proper. We assumed the EQI priors to be uniformly and independently distributed as follows:

(a) $\bar{s}_1, \bar{s}_2, \bar{s}_3 \sim U(0,1)$

(b) $\bar{\eta}_{11}, \bar{\eta}_{22}, \bar{\eta}_{33} \sim U(-3,0)$

(c) $\bar{\eta}_{12}, \bar{\eta}_{13}, \bar{\eta}_{21}, \bar{\eta}_{23}, \bar{\eta}_{31}, \bar{\eta}_{32} \sim U(-3,3)$

(d) $\bar{\eta}_{1x}, \bar{\eta}_{2x}, \bar{\eta}_{3x} \sim U(-3,3)$

Thus,

$$p(\theta) = \left(\frac{1}{3}\right)^3 \left(\frac{1}{6}\right)^9 I(\Gamma_\theta) \qquad (12)$$

where $I(\Gamma_\theta)$ is an indicator function which takes a value of 1 if all EQI are within their respective parameter supports and 0 otherwise. The inequality restrictions on the elasticities are not imposed at this point. Concavity and monotonicity are to be imposed locally in the sense that they are imposed at every data point in the sample. Because the dependence of the elasticities on the data is different for each of the models, the nature of the restrictions will be different for each of the models. Accordingly, these restrictions are introduced later, on the vector of coefficients for each model.

Having specified the EQI prior, the next task is to translate the EQI information into priors for each of the model's parameters. To do so, we first need to reduce the dimension of the prior on the EQI to equal the number of parameters in each model. Then the transformation variable method can be employed to map the reduced EQI prior to the prior for the model's parameters. An obvious solution is to have 7 EQI instead of 15 EQI. However, which 7 to choose is an arbitrary decision which is likely to favour some EQI over others. A more attractive alternative solution is to employ a procedure that was suggested by Kleibergen (1996) and applied by Tan-Cruz (1998) for model averaging the translog and normalised quadratic cost functions in a production system. Kleibergen (1996) relies on the existence of a set of equality relationship among the EQI, which permits a reduction in the number of EQI. In our case, 8 independent equality relationships are required to reduce the number of EQI from 15 to 7.

<u>Kleibergen (1996) procedure</u>

The Kleibergen procedure involves the following steps:

Step 1 : Define $\xi = f(\theta)$ and $\lambda = g(\theta)$, where ξ is a (7×1) one to one transformation on a non redundant subset of θ, and λ is a (8×1) vector such that the equality relationships on elasticities and budget shares hold when $\lambda = 0$.

Step 2 : Transform $p(\theta)$ to $p(\xi,\lambda)$ through a transformation of variables .

$$p(\xi,\lambda) = p(\theta)\left|\frac{\partial\theta}{\partial\xi'} \quad \frac{\partial\theta}{\partial\lambda'}\right| \tag{13}$$

Step 3 : To reduce the 15 EQI to 7 EQI, we evaluate $p(\xi \mid \lambda)$ evaluated at $\lambda = 0$, where a suitable prior $p(\xi)$ is given by

$$p(\xi \mid \lambda) = \frac{p(\xi,\lambda)}{p(\lambda)} \tag{14}$$

and, when $\lambda = 0$,

$$\begin{aligned} p(\xi \mid \lambda)\big|_{\lambda=0} &= \frac{p(\xi,\lambda)\big|_{\lambda=0}}{p(\lambda)\big|_{\lambda=0}} \\ &\propto p(\xi,\lambda)\big|_{\lambda=0} \end{aligned} \tag{15}$$

Thus we have $p(\xi) = p(\xi \mid \lambda)\big|_{\lambda=0}$ which implies that

$$p(\xi) \propto p(\theta)\left\|\frac{\partial\theta}{\partial\xi'} \quad \frac{\partial\theta}{\partial\lambda'}\right\|_{\lambda=0} \tag{16}$$

A suitable choice for ξ is

$$\xi = \left[\bar{\eta}_{11} \;\; \bar{\eta}_{12} \;\; \bar{\eta}_{13} \;\; \bar{s}_1 \;\; \bar{s}_2 \;\; \bar{\eta}_{21} \;\; \bar{\eta}_{32}\right]' \tag{17}$$

and the 8 equality restrictions suggested by demand theory are given by

$$\begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \\ \lambda_5 \\ \lambda_6 \\ \lambda_7 \\ \lambda_8 \end{bmatrix} = \begin{bmatrix} \bar{\eta}_{11} + \bar{\eta}_{12} + \bar{\eta}_{13} + \bar{\eta}_{1x} \\ \bar{s}_1\bar{\eta}_{12} + \bar{s}_1\bar{s}_2\bar{\eta}_{1x} - \bar{s}_2\bar{\eta}_{21} - \bar{s}_1\bar{s}_2\bar{\eta}_{2x} \\ 1 - \bar{s}_1 - \bar{s}_2 - \bar{s}_3 \\ \bar{s}_1 + \bar{s}_1\bar{\eta}_{11} + \bar{s}_2\bar{\eta}_{21} + \bar{s}_3\bar{\eta}_{31} \\ 1 - \bar{s}_1\bar{\eta}_{1x} - \bar{s}_2\bar{\eta}_{2x} - \bar{s}_3\bar{\eta}_{3x} \\ \bar{s}_2\bar{\eta}_{23} + \bar{s}_2\bar{s}_3\bar{\eta}_{2x} - \bar{s}_3\bar{\eta}_{32} - \bar{s}_2\bar{s}_3\bar{\eta}_{3x} \\ \bar{s}_3 + \bar{s}_1\bar{\eta}_{13} + \bar{s}_2\bar{\eta}_{23} + \bar{s}_3\bar{\eta}_{33} \\ \bar{\eta}_{21} + \bar{\eta}_{22} + \bar{\eta}_{23} + \bar{\eta}_{2x} \end{bmatrix} = \begin{bmatrix} \theta_1 + \theta_2 + \theta_3 + \theta_8 \\ \theta_4\theta_2 + \theta_4\theta_5\theta_8 - \theta_5\theta_6 - \theta_4\theta_5\theta_9 \\ 1 - \theta_4 - \theta_5 - \theta_{10} \\ \theta_4 + \theta_4\theta_1 + \theta_5\theta_6 + \theta_{10}\theta_{11} \\ 1 - \theta_4\theta_8 - \theta_5\theta_9 - \theta_{10}\theta_{12} \\ \theta_5\theta_{13} + \theta_5\theta_{10}\theta_9 - \theta_{10}\theta_7 - \theta_5\theta_{10}\theta_{12} \\ \theta_{10} + \theta_4\theta_3 + \theta_5\theta_{13} + \theta_{10}\theta_{14} \\ \theta_6 + \theta_{15} + \theta_{13} + \theta_9 \end{bmatrix} \tag{18}$$

It can be shown that the reduced EQI prior is consequently given by

$$p(\xi) \propto \left(\frac{1}{3}\right)^3 \left(\frac{1}{6}\right)^9 \left| \frac{1}{\theta_4\theta_5^2(1-\theta_4-\theta_5)^3} \right| I(\Gamma_\theta) \tag{19}$$

where $\left| \frac{\partial\theta}{\partial\xi'} \quad \frac{\partial\theta}{\partial\lambda'} \right|_{\lambda=0} = \left| \frac{1}{\theta_4\theta_5^2(1-\theta_4-\theta_5)^3} \right|$

4.3 Prior for Σ_A and γ_A

In terms of model A, we need to specify or derive priors for the coefficient vector γ_A and the error covariance matrix Σ_A. We treat these quantities as priori independent such that

$$p(\gamma_A, \Sigma_A) = p(\gamma_A)\,p(\Sigma_A) \tag{20}$$

The prior for model B is treated in a similar manner. Also, since the error covariance matrix Σ_A is also the covariance matrix for the expenditure shares, which represent the dependent variables in both models A and B, we use the same prior for both models, $p(\Sigma_A) = p(\Sigma_B)$. As we shall see, $p(\gamma_A)$ and $p(\gamma_B)$ are obtained via a transformation of $p(\xi)$.

Priors for Σ_A

For a prior for Σ_A an inverted Wishart pdf is used; one which is relatively uninformative but proper was chosen. Specifically,

$$p(\Sigma_A \mid v) = h|C|^{v/2}|\Sigma_A|^{-(v+N+1)/2} \exp\left\{-\frac{1}{2}tr\left(C\Sigma_A^{-1}\right)\right\} \qquad v \geq N \tag{21}$$

where N is the number of equations in the model, C is an $(N \times N)$ matrix of prior parameters and

$h^{-1} = 2^{vN/2} \pi^{N(N-1)/4} \prod_{i=1}^{N} \Gamma\left(\frac{v+1-i}{2}\right)$.

For a proper but relatively uninformative prior, v is set to its smallest possible value, in this case, $v = N = 2$. The matrix C is specified as a diagonal matrix. This specification implies the budget shares are a priori independent, which is unlikely. However, it does overcome the problem of setting an a priori correlation and it does let the data be the main determinant of a posterior correlation between budget shares. For assigning values to c_{11} and c_{22}, we use the marginal distribution of $\left(c_{ii} / \sigma_{ii} \right)$ which is $\chi^2_{(1)}$; the value of c_{ii} should reflect that the budget shares lie between 0 and 1. Hence, it is unlikely that

$$\sqrt{\sigma_{ii}} > 0.2 \qquad\qquad i = 1,2 \tag{22}$$

Setting the probability of $\sqrt{\sigma_{ii}}$ being greater than 0.2 equal to 0.05 yields a c_{ii} = 0.00016.

Prior for γ_A

Having obtained $p(\xi)$, $p(\gamma_A)$ is obtained through

$$p(\gamma_A) = p(\xi) \left| \frac{\partial \xi}{\partial \gamma'_A} \right| \tag{23}$$

In addition, at this point we impose the inequality restrictions which are model dependent. To do so, we modify equation (23) to yield

$$p(\gamma_A) = k_A p(\xi) \left| \frac{\partial \xi}{\partial \gamma'_A} \right| I(\Xi_A) \tag{24}$$

where $I(\Xi_A)$ is indicator function for the inequality constraints; it takes the value 1 when monotonicity and concavity are satisfied, and 0 otherwise. Imposition of the inequality constraints truncates the parameter space. Thus, an unknown constant k_A is added to ensure that the probability under $p(\gamma_A)$ is unity. It can be shown that

$$p(\gamma_A) \propto k_A \left(\frac{1}{3}\right)^3 \left(\frac{1}{6}\right)^9 \left| \frac{1}{\alpha_{A1}^3 \alpha_{A2}^3 (1-\alpha_{A1}-\alpha_{A2})^4} \right| I(\Xi_A) I(\Gamma_\theta) \qquad (25)$$

4.4 Posterior pdf for γ_A

Having specified the priors, they are then combined with the likelihood function to give the joint posterior density for the model parameters, $p(\gamma_A, \Sigma_A \mid s)$.

$$p(\gamma_A, \Sigma_A \mid s) \propto p(\gamma_A, \Sigma_A)\, p(s \mid \gamma_A, \Sigma_A) \qquad (26)$$

The likelihood function is

$$\begin{aligned} p(s \mid \gamma_A, \Sigma_A) &= (2\pi)^{-TN/2} |\Sigma_A|^{-T/2} \exp\left\{ -\frac{1}{2}(s - f(X, \gamma_A))' (\Sigma_A^{-1} \otimes I)(s - f(X, \gamma_A)) \right\} \\ &= (2\pi)^{-TN/2} |\Sigma_A|^{-T/2} \exp\left(-\frac{1}{2} tr\left(A \Sigma_A^{-1}\right) \right) \end{aligned} \qquad (27)$$

where $(A)_{ij} = (s_i - f_i(X, \gamma_A))' (s_j - f_j(X, \gamma_A))$ and f_i and f_j represent the AIDS function for shares i and j. Combining the various pdfs into equation (26) and integrating out Σ_A yields

$$\begin{aligned} p(\gamma_A \mid s) &= \int_0^\infty p(\gamma_A, \Sigma_A \mid s) d\Sigma_A \\ &\propto p(\gamma_A)\, p^*(s \mid \gamma_A) \end{aligned} \qquad (28)$$

where $p^*(s \mid \gamma_A) = k_1^{-1} h (2\pi)^{-TN/2} |C|^{V/2} |A + C|^{-(v+T)/2}$ and $k_1^{-1} = 2^{(v+T)N/2} \pi^{N(N-1)/4} \prod_{i=1}^{N} \Gamma\left(\frac{v + T + 1 - i}{2} \right)$

$p^*(s \mid \gamma_A)$ can be regarded as the data density with the covariance matrix factored out. As such the marginal posterior pdf for the model's parameters is made up of the model parameter prior, $p(\gamma_A)$ and a data density $p^*(s \mid \gamma_A)$.

Substituting equation (25) into (28) and ignoring all constants gives

$$p(\gamma_A \mid s) \propto |A + C|^{-(v+T)/2} \left| \frac{1}{\alpha_{A1}^3 \alpha_{A2}^3 (1-\alpha_{A1}-\alpha_{A2})^4} \right| I(\Xi_A) I(\Gamma_\theta) \qquad (29)$$

Similarly, it can be shown that the prior and posterior pdfs for the parameters of the Log TL model are as follows:

$$p(\gamma_B) \propto k_B \left(\frac{1}{3}\right)^3 \left(\frac{1}{6}\right)^9 \left| \frac{1}{\alpha_{B1}^3 \alpha_{B2}^3 (1-\alpha_{B1}-\alpha_{B2})^4} \right| I(\Xi_B) I(\Gamma_\theta) \tag{30}$$

$$p(\gamma_B \mid s) \propto \left| B + C \right|^{-(v+T)/2} \left| \frac{1}{\alpha_{B1}^3 \alpha_{B2}^3 (1-\alpha_{B1}-\alpha_{B2})^4} \right| I(\Xi_B) I(\Gamma_\theta) \tag{31}$$

The similarity of equations (25) and (30) suggests that both priors have the same information content.

4.5 Model Averaging

Model averaging centres on the evaluation of the marginal likelihood of each model

$$\begin{aligned} p(s \mid A) &= \int\int p(\gamma_A, \Sigma_A) p(s \mid \gamma_A, \Sigma_A) d\gamma_A d\Sigma_A \\ &= \int p(\gamma_A) p^*(s \mid \gamma_A) d\gamma_A \end{aligned} \tag{32}$$

$$p(s \mid B) = \int p(\gamma_B) p^*(s \mid \gamma_B) d\gamma_B \tag{33}$$

In this case $p(s \mid A)$ and $p(s \mid B)$ are intractable; MHM and MHA approaches are utilised for the estimating of $p(s \mid A)$ and $p(s \mid B)$. In addition, MHA is employed to estimate k_A and k_B which have to be included in the respective estimation of marginal likelihoods based on MHM and MHA. k_B is the normalising constant created by the inequality constraints in the Log TL model.

The posterior model probabilities are then

$$P(A \mid s) = \frac{p(s \mid A) P(A)}{p(s \mid A) P(A) + p(s \mid B) P(B)} \tag{34}$$

$$P(B \mid s) = \frac{p(s \mid B) P(B)}{p(s \mid A) P(A) + p(s \mid B) P(B)} \tag{35}$$

where $P(A)$ and $P(B)$ are our model prior beliefs. Subsequently, the posterior pdf for the EQI obtained by model averaging AIDS and Log TL is

$$p(\theta \mid s) = p(\theta \mid s, A) P(A \mid s) + p(\theta \mid s, B) P(B \mid s) \tag{36}$$

or

$$E[g(\theta)\mid s] = E[g(\theta)\mid s, A]P(A\mid s) + E[g(\theta)\mid s, B]P(B\mid s) \quad (37)$$

where $g(\theta)$ is a function of θ.

4.6 Metropolis Hastings Algorithm, Metropolis Hastings Approximation, Modified Harmonic Mean

Due to the complicated nature of equations (29) and (31), the marginal densities of individual parameters are analytically intractable. With the recent development of Markov Chain Monte Carlo (MCMC), these densities can be estimated. The Metropolis-Hastings algorithm is employed here.

Metropolis Hastings Algorithm

The Metropolis Hastings algorithm was first developed by Metropolis et al (1953) and first employed in a statistical context by Hasting (1970). This method draws samples of observations from the marginal posterior model parameter densities, $p(\gamma_A \mid s)$ and $p(\gamma_B \mid s)$. Thus, the posterior means and variances of individual parameters can be estimated from the simulated sample of observations.

The process of drawing observation from $p(\gamma_A \mid s)$ is as follows.

Step 1: Specify an arbitrary starting value $\gamma_A^{(i)}$ which satisfies the constraint. Set $i = 0$.

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

Step 3: Use the candidate value $\gamma_A^{*(i)}$ to evaluate the monotonicity and concavity constraints. If any constraints are violated set

$\alpha(\gamma_A^{(i)}, \gamma_A^{*(i)}) = 0$ and go to step 5

Step 4: Calculate $\alpha(\gamma_A^{(i)}, \gamma_A^{*(i)}) = \min\left(\dfrac{p(\gamma_A^{*(i)} \mid s)}{p(\gamma_A^{(i)} \mid s)}, 1\right)$

Step 5: Generate an independent uniform random variable.

U from the interval $[0,1]$

Step 6: Set $\gamma_A^{(i+1)} = \begin{cases} \gamma_A^{*(i)} & if \quad u < \alpha(\gamma_A^{(i)}, \gamma_A^{*(i)}) \\ \gamma_A^{(i)} & if \quad u \geq \alpha(\gamma_A^{(i)}, \gamma_A^{*(i)}) \end{cases}$

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

This iterative scheme produces a sequence of draws, $\gamma_A^{(1)}, \gamma_A^{(2)}, \cdots, \gamma_A^{(m)}$. For a large number of draws, says s onwards, $\gamma_A^{(s)}, \gamma_A^{(s+1)} \cdots \text{and}\, \gamma_A^{(m)}$, are draws from $p(\gamma \mid s)$ [5]. Thus, the posterior mean (a typical point estimate) can be estimated as

$$\tilde{\gamma}_A = \frac{1}{m-s} \sum_{i=s}^{m} \gamma_A^{(i)} \tag{38}$$

MH requires a transition density, $q(\gamma_A^{(i)}, \gamma_A^{*(i)})$, to move from one stage to another. The choice of transition density is arbitrary. A random walk random transition density is selected and is assumed to have a multivariate normal distribution

$$q(\gamma_A^{(i)}, \gamma_A^{*(i)}) \sim N\left(\gamma_A^{(i)}, r_A \left[\frac{\partial e'}{\partial \gamma_A} (\hat{\Sigma}_A^{-1} \otimes I_T) \frac{\partial e}{\partial \gamma_A} \right]^{-1} \Bigg|_{\bar{\gamma}_A} \right) \tag{39}$$

where $\gamma_A^{(i)}$ is the mean of the transition density and $r_A \left[\frac{\partial e'}{\partial \gamma_A} (\hat{\Sigma}_A^{-1} \otimes I_T) \frac{\partial e}{\partial \gamma_A} \right]^{-1} \Bigg|_{\bar{\gamma}_A}$ is the covariance matrix. The term $\left[\frac{\partial e'}{\partial \gamma_A} (\hat{\Sigma}_A^{-1} \otimes I_T) \frac{\partial e}{\partial \gamma_A} \right]^{-1} \Bigg|_{\bar{\gamma}_A}$ is the maximum likelihood covariance matrix of γ_A and r_A is a constant used to manipulate the rate at which the candidate $\gamma_A^{(i)}$ is accepted in the next sequence. By doing so, we control the rate at which the MH algorithm explores the permissible parameter space.

Metropolis Hastings Approximation

MHA computes a weight for each iteration using the MH algorithm. These weights are given by the following ratio, with all densities including their proper normalising constants

$$w(\gamma_A^{(i)}, \gamma_A^{*(i)}) = \frac{p(\gamma_A^{*(i)}) \, p^*(s \mid \gamma_A^{*(i)})}{q(\gamma_A^{(i)}, \gamma_A^{*(i)})} \tag{40}$$

[5] In such case, the MCMC sequence is said to have 'converged'. Plotting γ against the number of draws helps establish a suitable s.

Since the expectation of $w(\gamma_A^{(i)}, \gamma_A^{*(i)})$ is equal to $p(s \mid A)$, an estimate of $p(s \mid A)$ is given by an average of the weights. See Geweke (1999, p44) for proof.

Thus,

$$\hat{p}(s \mid A) = \frac{1}{m}\sum_{i=1}^{m} w(\gamma_A^{(i)}, \gamma_A^{*(i)}) \tag{41}$$

Similarly, the marginal likelihood estimate for Log TL is

$$\hat{p}(s \mid B) = \frac{1}{m}\sum_{i=1}^{m} w(\gamma_B^{(i)}, \gamma_B^{*(i)}) \tag{42}$$

Modified Harmonic Mean

MHM, unlike MHA, requires only the posterior output for estimating the marginal likelihood. For an arbitrary function $f(\gamma_A)$ that integrates to one, and is such that $\frac{f(\gamma_A)}{p(\gamma_A)p^*(s \mid \gamma_A)}$ is bound above, taking the expectation of $\frac{f(\gamma_A)}{p(\gamma_A)p^*(s \mid \gamma_A)}$ with respect to the posterior density $p(\gamma_A \mid s)$ yields the inverse of the marginal likelihood. Specifically,

$$\begin{aligned} E\left[\frac{f(\gamma_A)}{p(\gamma_A)p^*(s \mid \gamma_A)}\middle| s, A\right] &= \int_{\gamma_A} \frac{f(\gamma_A)}{p(\gamma_A)p^*(s \mid \gamma_A)}\, p(\gamma_A \mid s)\, d\gamma_A \\ &= \int_{\gamma_A} \frac{f(\gamma_A)}{p(\gamma_A)p^*(s \mid \gamma_A)} \frac{p(\gamma_A)p^*(s \mid \gamma_A)}{\int_{\gamma_A} p(\gamma_A)p^*(s \mid \gamma_A)\, d\gamma_A}\, d\gamma_A \\ &= \frac{\int_{\gamma_A} f(\gamma_A)\, d\gamma_A}{\int_{\gamma_A} p(\gamma_A)p^*(s \mid \gamma_A)\, d\gamma_A} \\ &= p(s \mid A)^{-1} \end{aligned} \tag{43}$$

Following suggestions from Geweke (1992, p46) and Gelfand and Dey (1994, p511), we choose $f(\gamma_A)$ to be a truncated normal distribution with the amount of truncation varying with $p \in (0,1)$, as given in equation (44).

$$f(\gamma_A) = p^{-1}(2\pi)^{-k/2}\left|\hat{\Sigma}_{\gamma_A}\right|^{-1/2} \exp\left[-\frac{1}{2}(\gamma_A - \hat{\gamma}_A)'\hat{\Sigma}_{\gamma_A}^{-1}(\gamma_A - \hat{\gamma}_A)\right] I(\Gamma_A) \tag{44}$$

where $\hat{\gamma}_A = \frac{1}{m}\sum_{i=1}^{m}\gamma_A^{(i)}$, $\hat{\Sigma}_{\gamma_A} = \frac{1}{m}\sum_{i=1}^{m}(\gamma_A^{(i)} - \hat{\gamma}_A)(\gamma_A^{(i)} - \hat{\gamma}_A)'$ and $I(\Gamma_A)$ is an indicator function that equals 1 for γ_A satisfying $(\gamma_A - \hat{\gamma}_A)'\hat{\Sigma}_{\gamma_A}^{-1}(\gamma_A - \hat{\gamma}_A) \le q$ and q is such that $P(\chi^2_{(k)} \le q) = p$.

$\hat{\gamma}_A$ and $\hat{\Sigma}_{\gamma_A}$ are estimated from the posterior output, l is the number of model A parameters[6]. Thus, the marginal likelihood estimates can be obtained through

$$\hat{p}(s \mid A)^{-1} = \frac{1}{m}\sum_{i=1}^{m}\frac{f(\gamma_A^{(i)})}{p(\gamma_A^{(i)})\,p^*(s \mid \gamma_A^{(i)})} \tag{45}$$

Similarly

$$\hat{p}(s \mid B)^{-1} = \frac{1}{m}\sum_{i=1}^{m}\frac{f(\gamma_B^{(i)})}{p(\gamma_B^{(i)})\,p^*(s \mid \gamma_B^{(i)})} \tag{46}$$

5. Empirical Results

The BMA techniques discussed in the previous sections were applied to quarterly US aggregated data for meat: beef, pork and poultry for the period 1979:1 - 1995:2. These data was first applied by Alston, Chalfant and Piggot (1998) in a nested PIGLOG model.

Before estimation, we treat the two models as equally likely and hence set $P(A) = P(B)$. To prevent singularity in the demand system, the equation for poultry is dropped during estimation.

Table 1 reports the log marginal likelihood estimates and posterior model probability estimates for AIDS and Log TL. In the parentheses in Table 1 are their associated batch standard errors (BSE) that are used to assess the reliability of the marginal likelihood estimates. The BSE is employed instead of a normal standard error because the Metropolis-Hastings generated observations are correlated. The BSE is calculated as

$$\text{BSE} = \frac{\text{Sample Standard Deviation of batch means of size } n}{\sqrt{\text{number of batches}}}$$

[6] Both models are having the same number of model parameters.

In Table 2, the EQI point estimates are reported for estimates obtained under MLE, individual model Bayesian estimation and Bayesian model averaging.

From Table 1 the model probabilities indicate that the AIDS model dominates the Log TL model, with the MHA result suggesting greater dominance than that suggested by MHM. Comparing the BSE from both approaches, MHA seem to be a less efficient approach than MHM.

In Table 2, the EQI estimates from the various methods of estimation are relatively similar, except for $\overline{\eta}_{33}$, which is the own price elasticity for poultry. In both models, the values for $\overline{\eta}_{33}$ under maximum likelihood estimation are theoretically implausible, since own price elasticities should lie in the negative region. Interestingly, the posterior pdfs for the EQI, graphed in Figure 2, are almost identical irrespective of whether Log TL, AIDS or model averaging is used. This suggests that without going through BMA to remove model uncertainties, both models are equally suited for making inferences on the EQI. On the other hand, the posterior model probabilities suggest some grounds for conducting BMA.

The posterior odds between AIDS and Log TL (which is also the Bayes factor[7] in this case) shows a value of 4 from MHA, and 2.4 from MHM, in favour of AIDS. According to Jeffreys (1961), if the Bayes factor is between 1 and 3, it is not worthwhile selecting a particular model; the odds are not sufficiently strong. Thus, in this case, BMA can account for the indecisiveness of selecting a best model out of a pool of models. In cases where the posterior odds are strongly in favour of one model (greater than 100), BMA has no impact on model uncertainties. The averaged results are the same as those from the best model.

[7] Bayes factor is the weight of evidence provided by the data for one model against another model. Whereas posterior odds is the weight of evidence provided by data and prior beliefs.

Table 1 Log Marginal Likelihood and Posterior Model Probabilities for AIDS and Log TL

	Log Marginal Likelihood		Posterior Model Probability	
	Aids	Log TL	Aids	Log TL
MHA	332.5651 (0.342768)**	331.1853 (0.402838)**	0.79896209 (0.084959)**	0.2010379 (0.084959)**
*MHM	333.0057 (0.143137)**	332.13398 (0.307929)**	0.705103468 (0.070608)**	0.2948965 (0.070608)**

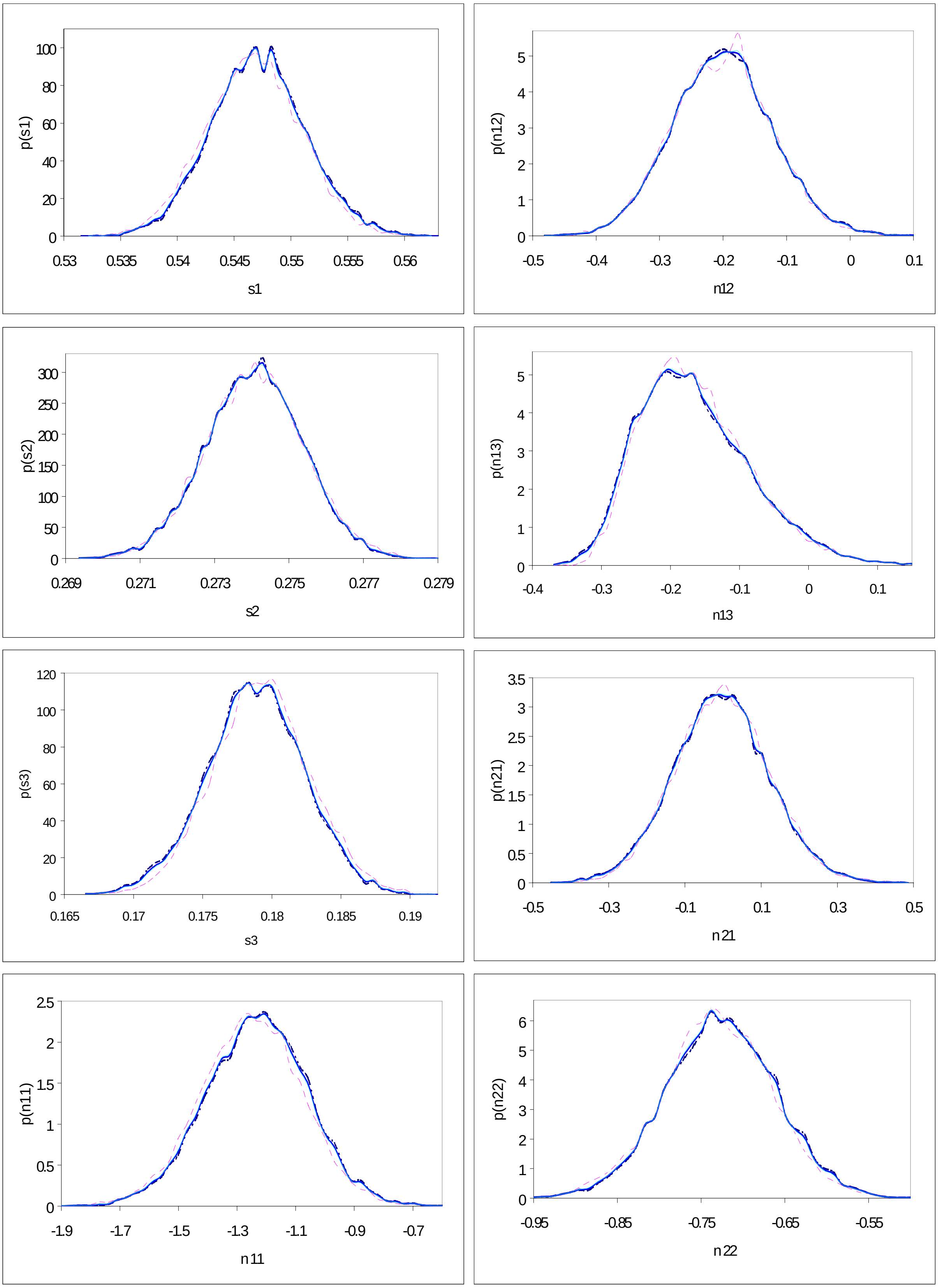
** BSE, * p=0.9

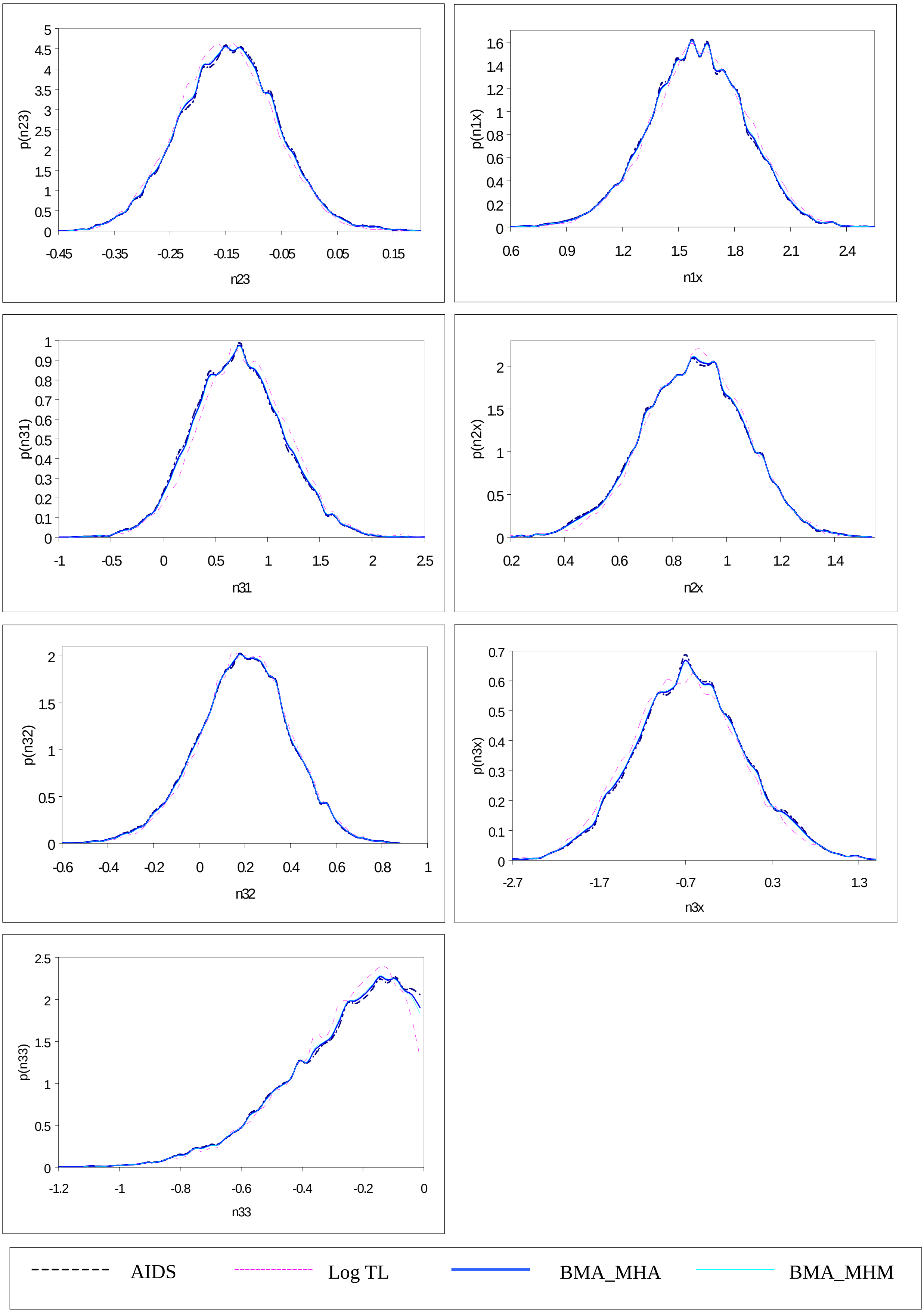
Table 2 Posterior EQI point estimates from AIDS , Log TL and Model Averaging

EQI	MLE		Bayesian		BMA	
	AIDS	Log TL	AIDS	Log TL	MHA	*MHM
$\bar{s}_1$	0.54693 (0.0042789)	0.54666 (0.0042201)	0.54720 (0.0041795)	0.54667 (0.0041943)	0.5470886 (0.0041878)	0.5470388 (0.0041908)
$\bar{s}_2$	0.27414 (0.0012957)	0.27413 (0.0013429)	0.27406 (0.0012993)	0.27409 (0.0013302)	0.2740680 (0.0013056)	0.2740704 (0.0013086)
$\bar{s}_3$	0.17893 (0.0035694)	0.17921 (0.0034534)	0.17874 (0.0035241)	0.17925 (0.0035107)	0.1788434 (0.0035272)	0.1788908 (0.0035276)
$\bar{\eta}_{11}$	-1.22501 (0.17448)	-1.22482 (0.17857)	-1.2291 (0.17288)	-1.2519 (0.17384)	-1.233694 (0.173312)	-1.235829 (0.1734724)
$\bar{\eta}_{22}$	-0.71765 (0.066803)	-0.72022 (0.064132)	-0.72601 (0.066016)	-0.73167 (0.065262)	-0.7271474 (0.0659036)	-0.7276792 (0.06584457)
$\bar{\eta}_{33}$	0.021175 (0.32030)	0.043352 (0.32968)	-0.27439 (0.19928)	-0.27706 (0.19003)	-0.2749253 (0.197457)	-0.2751765 (0.1965996)
$\bar{\eta}_{12}$	-0.24101 (0.088363)	-0.24443 (0.090328)	-0.20010 (0.078659)	-0.20124 (0.078303)	-0.2003299 (0.07858783)	-0.2004362 (0.07855481)
$\bar{\eta}_{13}$	-0.26550 (0.10681)	-0.27334 (0.12264)	-0.16541 (0.082144)	-0.16175 (0.078746)	-0.1646736 (0.08148421)	-0.1643297 (0.08117272)
$\bar{\eta}_{21}$	-0.0025549 (0.11722)	-0.0075905 (0.11832)	-0.0057510 (0.12577)	-0.0018810 (0.12370)	-0.00497295 (0.1253653)	-0.004609724 (0.1251744)
$\bar{\eta}_{23}$	-0.13683 (0.084694)	-0.13701 (0.086279)	-0.14322 (0.088548)	-0.15058 (0.085280)	-0.1446993 (0.08794873)	-0.1453896 (0.08765969)
$\bar{\eta}_{31}$	0.69166 (0.43649)	0.69738 (0.44730)	0.70967 (0.42654)	0.77071 (0.43209)	0.7219383 (0.4283546)	0.7276676 (0.4290821)
$\bar{\eta}_{32}$	0.30408 (0.22319)	0.31763 (0.23025)	0.19186 (0.20237)	0.20286 (0.19804)	0.1940720 (0.2015566)	0.1951042 (0.2011668)
$\bar{\eta}_{1x}$	1.731517 (0.28539)	1.74260 (0.29392)	1.5946 (0.26029)	1.6148 (0.25933)	1.598698 (0.2602178)	1.600595 (0.2601651)
$\bar{\eta}_{2x}$	0.85704 (0.18295)	0.86482 (0.14380)	0.87498 (0.19287)	0.88413 (0.18685)	0.8768197 (0.1917095)	0.8776785 (0.1911589)
$\bar{\eta}_{3x}$	-1.01692 (0.72923)	-1.05836 (0.74876)	-0.62714 (0.64564)	-0.69650 (0.65042)	-0.641085 (0.6471925)	-0.6475953 (0.6478168)

* The averaged EQI are estimated based on $p = 0.9$

Figure 2 Posterior EQI Pdf s from AIDS, Log TL and Model Averaging based on MHA and MHM for with Inequality Constraints




p(n23)
5 4.5 4 3.5 3 2.5 2 1.5 1 0.5 0
-0.45 -0.35 -0.25 -0.15 -0.05 0.05 0.15
n23
p(n1x)
1.6 1.4 1.2 1 0.8 0.6 0.4 0.2 0
0.6 0.9 1.2 1.5 1.8 2.1 2.4
n1x
p(n31)
1 0.9 0.8 0.7 0.6 0.5 0.4 0.3 0.2 0.1 0
-1 -0.5 0 0.5 1 1.5 2 2.5
n31
p(n2x)
2 1.5 1 0.5 0
0.2 0.4 0.6 0.8 1 1.2 1.4
n2x
p(n32)
2 1.5 1 0.5 0
-0.6 -0.4 -0.2 0 0.2 0.4 0.6 0.8 1
n32
p(n3x)
0.7 0.6 0.5 0.4 0.3 0.2 0.1 0
-2.7 -1.7 -0.7 0.3 1.3
n3x
p(n33)
2.5 2 1.5 1 0.5 0
-1.2 -1 -0.8 -0.6 -0.4 -0.2 0
n33
AIDS
Log TL
BMA_MHA
BMA_MHM

6. Conclusion

This paper has shown how model averaging is conducted along with imposition of inequality constraints in the context of a consumer demand framework. The use of BMA centres on the evaluation of marginal likelihoods that are generally intractable here. The advent of sampling techniques has not only allowed us to compute posterior estimates for AIDS and Log TL, with complex inequality constraints, but has also permitted estimation of the marginal likelihood.

Also, the posterior model probability can be used as a basis for choice between model selection and model averaging. On a final note, the methodology discussed could be extended to other areas of study, such as production systems, stochastic frontiers and modelling of the Philips curve.

References

Alston, J.M., J.A. Chalfant and N.E. Piggott (1998), "A Globally Flexible Model of the Effects of Generic Advertising of Beef and Pork on U.S. Meat Demand," Selected paper, AAEA meetings, August 1998, Salt Lake City, Utah

Chalfant, J.A, R.S. Gray and K.J. White (1981), "Evaluating Prior Beliefs in a Demand System: The Case of Meat Demand in Canada," *American Journal of Agriculture Economic, 73*, p466-489

Christensen, L.R, D.W. Jorgenson and L.J. Lau (1975), "Transcendental Logarithmic Utility Functions," *American Economic Review*, *65*, p367-383

Deaton, A. and J. Muellbauer (1980), "An Almost Ideal Demand System," *American Economic Review, 70*, p312-326

Diewert, W.E. and T.J. Wales (1988), "A Normalized Quadratic Systems of Consumer Demand Functions," *Journal of Business and Economic Statistics, 6*, p303-312

Fernandez, C., E. Ley and M.F.J, Steel (1997), "Statistical Modelling of fishing Activities in the North Atlantic," *CentER discussion papers 9708, University of Tilburg*

Gallant, R.A. and G.H. Golub (1984), "Imposing Curvature Restrictions on Flexible Functional Forms," *Journal of Econometrics, 26*, p295-321

Gelfand, A.E. and D.K. Dey (1994), "Bayesian Model Choice: Asymptotics and Exact Calculations," *Journal of the Royal Statistical Society Series B, 56*, p501-514

Geweke, J. (1986), "Exact Inference in the Inequality Constrained Normal Linear Regression Model," *Journal of Applied Econometrics, 1*, p127-141

Geweke, J. (1999), "Using Simulation Methods For Bayesian Econometric Models: Inference, Development, and Communciation," *Econometric Review, 18*, p1-73

Griffiths, W.E. (1988), "Bayesian Econometrics and How to get Rid of Those Wrong Signs," *Review of Marketing and Agricultural Economics, 56,* p36-56

Griffiths, W.E. and D. Chotikapanich (2000), "Posterior Distribution for the Gini Coefficient Using Grouped Data," *Australia and New Zealand Journal of Statistics, 42*, p383-392

Griffiths, W.E., C.J. O'Donnell and A.T. Cruz (2000), "Imposing regularity conditiond on a system of cost and factor share equations," *The Australian Journal of Agricultura land Resource Economics*, *44*, p107-127

Hasegawa, H, H. Kozumi and N. Hashimoto (1999), "Testing for negativity in a demand system: A Bayesian approach," *Empirical Economics*, *24*, p211-223

Hastings, W.K. (1970), "Monte Carlo Sampling Methods using Markov Chains and their Applications," *Biometrika*, *57*, p97-109

Jeffreys, H.S. (1939), *Theory of Probability*, Oxford: university Press [2nd edn (1948), 3rd edn (1961)]

Jorgenson, D.W., L.J. Lau and T.M. Stoker (1982), "The Transcendental Logarithmic Model of Aggregate Consumer Behaviour," *Advances in Economics (Vol 1)*, eds. R.L Basmann and G. Rhodes, Greenwich, CT: JAI Press, p 97-238

Kleibergen, F. (1996), "Equality Restricted Random Variables: Densities and Sampling Algorithms," *Econometrica Institute Reports*, No 9662/A

Madigan, D. and A.E. Raftery (1994), “Model Selection and Accounting for Model Certainty in Graphical Models Using Occam’s Window,” *Journal of the American Statistical Association, 89*, p1535-1546

Metropolis, N., A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller and E. Teller (1953), “Equations of State Calculations by Fast Computing Machines,” *Journal of Chemical Physics, 21*, p1087-1092

Moschini, G. (1999), “Imposing Local Curvature conditions in flexible Demand system”, *Journal of Business and Economic Statistics, 17*, p487-490

Ryan, D.L., and T.J. Wales (1998), “A Simple Method for Imposing Local Curvature in Some Flexible Consumer -Demand Systems,” *Journal of Business and Economic Statistics*, 16, 331-338.

Tierney, L and Kadane, J.B. (1986), “Accurate Approximations for Posterior Moments and Marginal Densities,” *Journal of American Statistical Association, 81*, 501-514

Tan-Cruz, A.B. (1998), “Estimating Production Technologies with Inequality Constraints,” *PhD Thesis*, University of New England