DEPARTMENT OF AGRICULTURAL AND RESOURCE ECONOMICS
DIVISION OF AGRICULTURE AND NATURAL RESOURCES
UNIVERSITY OF CALIFORNIA

WORKING PAPER NO. 483

EVALUATING PRIOR BELIEFS IN A DEMAND SYSTEM: THE CASE OF MEATS DEMAND IN CANADA

by

James A. Chalfant, Richard S. Gray, and Kenneth J. White

California Agricultural Experiment Station
Giannini Foundation of Agricultural Economics
January, 1989

EVALUATING PRIOR BELIEFS IN A DEMAND SYSTEM:

THE CASE OF MEATS DEMAND IN CANADA

James A. Chalfant, Richard S. Gray, and Kenneth J. White*
Department of Agricultural and Resource Economics
University of California at Berkeley

Abstract

An almost ideal demand system for meats is estimated using Canadian data. We use a Bayesian approach to impose restrictions on substitution elasticities, via Monte Carlo integration and importance sampling, in order to conform with prior beliefs about curvature restrictions and substitution relationships. Results are consistent with demand theory, but not with the prior belief that all meats included are substitutes.

*Assistant Professor of Agricultural and Resource Economics, University of California at Berkeley; Graduate Research Assistant, University of California at Berkeley; Professor, University of British Columbia, respectively.

EVALUATING PRIOR BELIEFS IN A DEMAND SYSTEM: THE CASE OF MEATS DEMAND IN CANADA

1. Introduction

Widespread use of flexible functional forms in demand analysis has given researchers the ability to model consumer preferences with no restrictions on the nature of substitution or complementarity relationships between pairs of goods. Unlike Cobb-Douglas or CES preferences, the arbitrary utility functions approximated by more general demand systems, such as the translog or almost ideal forms, do not impose the restriction that all goods are equally substitutable. Unfortunately, theoretical restrictions automatically met by simpler forms need not hold with flexible forms.

When restrictions can be imposed using equality constraints on the parameters of the demand system, estimation and hypothesis testing are straightforward. This applies to symmetry and homogeneity restrictions. Curvature restrictions are another story. Even with symmetry and homogeneity imposed, a system still may be inconsistent with a well-behaved utility function, if the matrix of substitution elasticities is not negative semi-definite (implying that the expenditure function is not concave). It is also possible to find violations of the monotonicity restriction, in that budget shares predicted for particular combinations of prices and expenditures may not be between 0 and 1. Curvature or monotonicity restrictions require inequality restrictions on parameters, not easily handled in conventional estimation approaches.

In this paper, we illustrate how the Bayesian approach to inference can handle inequality restrictions in demand systems, using a method based on Geweke's work. He shows how to make inferences about or impose inequality restrictions in regression models. Chalfant and

White used his method to impose curvature and monotonicity restrictions on the translog cost function. Here, we apply the same procedure to the estimation of a demand system for meats using Canadian data.

In addition to these restrictions from consumer theory, we suggest a new set of inequality restrictions, not generally imposed on a demand system, but which have a fairly compelling motivation. It seems reasonable to expect that no pair of foods that play essentially the same role in the diet—sources of protein in the present application—should be complements. One thinks of beef and gravy, chicken and dumplings, but not beef and fish, say, as complementary items. Yet, it is common for parameter estimates from a demand system which fits well by other measures—percentage of variation explained, plausible income or own-price elasticities, etc.—to imply that a particular pair of goods which are natural substitutes are instead complements. The problem is that using flexible forms to allow elasticities of substitution between different pairs of goods to have different values also means that they can have different signs. Similarly, the elasticity of substitution between two goods may be positive at one set of prices and negative at another. While restrictions on the signs of elasticities of substitution are suggested by empirical observation, rather than the underlying theory, they seem just as important as compatibility with theory in judging the degree to which an estimated demand system conforms to our beliefs about consumer behavior. For applications such as this one, then, the constraint that all meats are substitutes can be viewed as another requirement that any well-behaved demand system must satisfy, in order to match prior beliefs.

Unfortunately no commonly used flexible form allows arbitrary values for elasticities of substitution without also allowing some values to be positive while some are negative. To restrict all meats to be substitutes thus requires an additional set of inequality constraints. Therefore, we impose not only concavity and monotonicity, but require the elasticity of substitution between any pair of goods to be positive.

One motivation for the approach comes from the recent literature on testing the stability of meats demand. Chalfant and Alston found that data from the U.S. and Australia are consistent with stable preferences, using nonparametric demand analysis. Since that method does not produce elasticities, it is worth asking if the implied elasticities satisfy prior expectations.

We show below how to impose these inequality restrictions using time series data from Canada for prices and per capita consumption of beef, pork, chicken, and fish over the 1960-1984 period. Any of the curvature, monotonicity, or substitutability restrictions can be imposed alone or together. Thus, the procedure can provide parameter estimates and the probability that the restrictions are correct for the cases of theoretical restrictions alone; substitution restrictions alone; and the two combined. We illustrate the method using the almost ideal demand system (Deaton and Muellbauer).

2. The Almost Ideal Demand System

Demand theory suggests that the demand for a good should be a function of its own-price, the prices of closely related goods, and income. In order to estimate demand relationships in a system of a reasonable size, it is common to invoke weak separability—choices concerning the allocation of expenditures among a subset of goods consumed are assumed to be made independently from the prices of goods outside that group. For example, the quantity of beef consumed is likely to be a function of the prices of beef, pork, chicken, fish, and total expenditure on meat, but is not a function of the price of bananas.[1]

[1] Except, of course, to the extent that the prices of goods outside the group under study affect the total group

Whether or not it is appropriate to assume separability for a particular demand system is an empirical question. Theory suggests that any partial demand system representing the separable parts of larger systems should satisfy the conditions of symmetry, homogeneity, monotonicity, and concavity. Indeed, one interpretation of these conditions not holding in an estimated system is that the goods included in the demand system do not make up a separable group—a relevant price has presumably been omitted. Violation of those conditions can also be treated as indicating the presence of structural change, aggregation bias, or some other specification error.

Deaton and Muellbauer suggested the almost ideal demand system as a particular representation of price-independent, generalized logarithmic (PIGLOG) preferences. Such preferences are consistent with the aggregation of individual preferences. In addition, the functional form they chose is locally flexible, in the sense used by Barnett—it can attain arbitrary values for substitution elasticities at a given set of prices. The equations for budget shares take the following form:

$$S_i = \alpha_i + \sum_{j=1}^{n} \gamma_{ij} \ln P_j + \beta_i \ln (x/P),$$

where P_j is the *jth* good's price, x denotes total expenditure on the n goods, and P is Stone's geometric price index.[2] A system of these share equations can be estimated to obtain parameter estimates, and simple formulas convert the parameter estimates to elasticities.

---

expenditure, perhaps in a preliminary stage of allocating expenditures to aggregates such as meats, other foods, shelter, etc.

[2] Use of Stone's index for prices gives rise to an approximate almost ideal demand system. Blanciforti and Green termed it the linear approximate model, since unlike the original case it is linear in parameters. Deaton and Muellbauer or Blanciforti and Green can be consulted for details. In this paper we use only the linear approximate model, and refer to it as the almost ideal demand system for simplicity. However, the method we use could be applied to the nonlinear almost ideal demand system, or indeed, any functional form of interest.

This demand system is easily restricted to satisfy symmetry ($\gamma_{ij} = \gamma_{ji} \;\; \forall\; i,j$) and homogeneity ($\sum_{j=1}^{n} \gamma_{ij} = 0 \;\; \forall\; i$). The adding-up property holds, given these restrictions, provided $\sum_{i=1}^{n} \alpha_i = 1$ and $\sum_{i=1}^{n} \beta_i = 0$. Concavity or monotonicity restrictions are more difficult, as they involve multiple inequality restrictions on the parameters. For concavity, the matrix of second derivatives of the expenditure function, or equivalently, of elasticities of substitution, must be negative semi-definite. For monotonicity, predicted budget shares must all be between 0 and 1, to ensure that predicted quantities consumed are positive. Such restrictions are difficult to impose using most econometric packages and even harder to interpret statistically.

3. The Bayesian Approach to Testing Inequality Restrictions

An alternative approach to imposing inequality restrictions in a demand system is made possible using a Bayesian approach, which permits the formal inclusion of such prior information. Often, prior information can be imposed by choice of functional form. An extreme case is the Cobb-Douglas utility function, which would impose all of the restrictions from consumer theory, plus some less desirable ones, such as additivity of preferences and that elasticities of substitution are each one. Symmetry and homogeneity restrictions from demand theory represent prior information that is often imposed through equality restrictions on the parameters of less restrictive demand systems.

Such restrictions reduce the dimensionality of the parameter space, when demand systems are estimated—the symmetry and homogeneity restrictions, for instance, provide considerable gains in degrees of freedom. Prior information taking the form of an inequality restriction is less informative than such equality restrictions, in the sense that this information serves to truncate the parameter space, rather than reduce the number of free parameters. For instance, a particular parameter θ may be restricted to be positive. Conventional approaches to estimation do not permit the formal inclusion of such information (e.g. Judge et. al.), and most econometric packages do not permit such restrictions to be imposed.

When flexible functional forms for demand systems are estimated, it is quite common to observe results which conflict with prior beliefs. Symmetry or homogeneity restrictions may be violated, when tested, but are generally imposed using equality restrictions on the parameters. The signs of estimated elasticities may be implausible, in which case there are two choices. One is to search over alternative flexible forms, which has obvious undesirable consequences for inferences, once a well-behaved demand system has been obtained. The other is to impose inequality restrictions, through adding constraints to a maximum likelihood procedure. Such constraints are difficult to interpret statistically—the usual likelihood-ratio test does not apply without adjustments, for instance. Even if testing the restrictions is not the goal, imposing inequality constraints in this manner is likely to yield parameter values that lie on one of the constraints. For instance, constraining a demand elasticity to be non-positive may well produce a vertical demand curve, if the constraint is binding. Thus, a constrained maximum-likelihood approach is neither intuitively satisfying nor statistically convenient.

The problem of prior beliefs that take the form of inequality constraints is easily handled in the context of Bayesian inference, however. The Bayesian approach begins with a prior density function, defined over the vector of parameters, θ, call it $p(\theta)$.[3] This prior density summarizes all of the information the researcher has about θ prior to estimation. Specifying $p(\theta)$ permits the formal inclusion of information about the parameters. For instance, if a

[3] See Zellner (1971) or Judge et al. for much more detailed descriptions of the properties of the Bayesian approach.

particular parameter is considered equally likely to be positive or negative, a zero median characterizes the marginal probability density function used to describe prior beliefs about that parameter. If there is no prior information about θ, $p(\theta)$ is simply defined to be proportional to a constant over all real numbers, thus making it an improper density. Alternatively, $p(\theta)$ could be a proper density which reflects various beliefs in the form of probability statements. A very simple case is the prior p.d.f. which says that θ is contained in some region D with probability one:

$$p(\theta) \propto c \qquad \forall \ \theta \in D$$

D may be an open or closed interval, depending on the application. We consider below how such a prior density can be used to represent prior information about the parameters of a demand system.

Bayes' Theorem shows how to combine prior and sample information to obtain a posterior distribution for the parameters in θ given a data set y:

$$f(\theta \mid y) \propto p(\theta)\, L(\theta \mid y).$$

where $\propto$ denotes "is proportional to" and $L(\cdot)$ is the likelihood function based on the observed data. Unlike the sampling-theoretic approach to estimation, the Bayesian approach recognizes that posterior beliefs are conditional on the *observed* data set, rather than emphasizing the performance of estimators in repeated samples.

The posterior distribution $f(\theta \mid y)$ summarizes all information available about θ, both prior and sample information. It can serve as the end result of an investigation, or it can be used to calculate various confidence intervals and probabilities related to hypotheses about θ, or to obtain a point estimate of θ or some related quantity such as a demand elasticity. The optimal point estimate for θ depends on the investigator's objective function. Constrained maximum-likelihood estimation, which yields the mode of the posterior distribution as a point estimate, corresponds to a "zero-one" loss function (e.g. Zellner (1988)). This is an unusual loss function, in the sense that it places the same weight on being wrong for estimates arbitrarily close to the true θ and for choices very far from that value. More plausible loss functions can certainly be imagined, and different point estimates will result. For instance, if the investigator's loss function is quadratic, the mean of the posterior distribution for θ minimizes expected loss (e.g. Judge et al.). All that is needed, then, to find Bayesian point estimates of the parameters of a demand system is a means to describe prior beliefs in the form of inequality restrictions using $p(\theta)$, a way to obtain the posterior density function, and then a way to find its mean.

Below we illustrate how this approach can be applied, using the quadratic loss function. With the inequality restrictions imposed, it is straightforward to obtain the mean of the posterior distribution, call it $\overline{\theta}$. This serves as the optimal point estimate of the parameters of the demand system. Also of interest is some measure of how plausible the restrictions might be, given the data. Suppose prior beliefs are completely uninformative, that is, all parameter values are considered equally likely. In this case, the sample information dominates the posterior density function and an optimal point estimate is the mean of the (unrestricted) posterior distribution. Meanwhile, the probability that the restrictions are correct can be calculated using the unconstrained posterior density. This probability is interpreted as the degree of belief that the restrictions are true, based on observed data, found by obtaining the probability that θ lies within the interval D.

Both restrictions on the signs of substitution elasticities and restrictions on the entire matrix of substitution elasticities, to satisfy curvature restrictions, can be examined by calculating substitution elasticities. The elasticity of substitution between any two goods in the almost ideal demand system is

$$\sigma_{ij} = 1 + \frac{\gamma_{ij}}{S_i S_j} \qquad i \neq j$$

where S_i denotes the *ith* budget share. Own-elasticities of substitution can be found through the homogeneity restriction, or by

$$\sigma_{ii} = 1 + \frac{\gamma_{ii}}{S_i S_i} - \frac{1}{S_i}.$$

To evaluate these inequality restrictions, then, it is necessary to examine the behavior of elasticities of substitution everywhere in the parameter space where the researcher wishes to impose them. Similarly, the monotonicity restriction can be evaluated using predicted budget shares. Each set of restrictions is then imposed by truncating the parameter space so that each restriction holds. To obtain a Bayesian point estimate, (with a quadratic loss function), the researcher must find the mean of the truncated posterior distribution for the parameter vector.

While these calculations are in principle straightforward, requiring that integrals over the posterior density function be evaluated, the analytic solutions cannot be obtained in practice, except for fairly simple models. The dimension of the posterior density is likely to be too great, even if the density function and the region of the parameter space of interest can be described easily. Instead, it is necessary to evaluate the integrals using Monte Carlo integration. This permits estimating the solutions to integrals by random sampling.[4]

To describe the method, we begin by specifying a data-generating process. We assume

[4] Kloek and van Dijk (1978), van Dijk and Kloek (1980), and Geweke (1986) provide the foundations for the Monte Carlo integration and importance sampling, described below. The application to demand systems follows the discussion in Chalfant and White.

that prices and expenditures may be treated as exogenous, so that the parameters of a system of $n-1$ equations of the form

$$S_i = \alpha_i + \sum_{j=1}^{4} \gamma_{ij} \ln(P_j) + \beta_i \ln(x/P) + e_i$$

could be estimated using seemingly-unrelated regressions (SUR). As is well known, the equation for the *nth* budget share cannot be included without implying a singular contemporaneous covariance matrix for the error terms in the n share equations (Barten), but deleting the *nth* share and using restrictions on the parameters allows the complete set of parameter estimates to be obtained. Use of iterated SUR was shown by Barten to lead to estimates that are invariant to the equation chosen for deletion.

We assume that the $n-1$ vector of errors, and therefore the shares themselves, follow the multivariate normal distribution. Strictly speaking, one might prefer a distribution more compatible with the fact that observed shares are bounded by 0 and 1 (e.g. Woodland, Rossi), but we prefer to stick with the more widely used distribution to illustrate our method. The approach we take could easily be adjusted for non-normal errors.

To illustrate Monte Carlo integration, suppose that Σ, the variance-covariance matrix of the errors, is known. Suppose also that $p(\theta)$, the prior information we have about the $\alpha_i's$, $\gamma_{ij}'s$, and $\beta_i's$, indicates that some region D, a proper subset of R^p, contains the true parameter vector, where p denotes the number of free parameters in the model and R^p denotes the p-dimensional real numbers. Finally, suppose that the investigator has a quadratic loss function, and desires a point estimate of θ; as noted earlier, the mean of the posterior density for θ minimizes expected loss.

The steps involved in finding an estimate of the mean, $\bar{\theta}$, are straightforward. With no prior information about θ, the posterior distribution from iterated SUR would be the multivariate normal, centered at $\hat{\theta}$ with variance-covariance matrix $V(\hat{\theta})$, where $\hat{\theta}$ and $V(\hat{\theta})$ are obtained using iterated SUR. Given our prior information, the posterior distribution for θ then becomes the *truncated* multivariate normal, since θ is known to lie in D. Our task becomes finding the mean of a truncated, p-variate normal:

$$E(\theta) = \int_{\theta_1 \in D} \int_{\theta_2 \in D} \cdots \int_{\theta_n \in D} \theta \, N_p[\theta \mid \hat{\theta}, V(\hat{\theta})] \, d\theta_1 d\theta_2 \cdots d\theta_n$$

Needless to say, such a calculation is infeasible for all but trivial examples.

Monte Carlo integration is based on the idea that an expectation such as the one above can be estimated (arbitrarily accurately, given the Law of Large Numbers) using random sampling. One way of estimating the mean of a random variable with p.d.f. $f(\theta)$ is to generate a large number of replications in a random sample from that distribution, and calculate

$$\bar{\theta} = \frac{\sum_{i=1}^{N} \theta_i}{N},$$

where N is the number of replications. $\bar{\theta}$ serves as an estimate of $E(\theta)$, of course. Since N is determined by the investigator, $E(\theta)$ can be obtained with an arbitrarily high degree of accuracy. To apply that approach using the multivariate normal requires 5 steps.

1. Estimate the parameters of the share equations, obtaining $\hat{\theta}$ and $V(\hat{\theta})$.

2. Treat these as parameters of the posterior distribution for θ which would be consistent with no restrictions on θ's range, the p-variate normal density.

3. Use $N_p[\hat{\theta}, V(\hat{\theta})]$ and a random number generator to obtain a random sample from this multivariate normal. Omit those draws θ_i which are not contained in D, leaving a random sample of size n from the truncated multivariate normal.

4. Estimate $E(\theta)$ using the average of the n replications in D:

$$\bar{\theta} = \frac{\sum_{i=1}^{n} \theta_i}{n}.$$

5. A by-product of the procedure is that $\hat{p} = n/N$ estimates the area under the multivariate normal density contained in D, i.e., the probability that the restrictions hold, *given no prior information*. If either p or $E(\theta)$ are estimated with less than the desired precision, increase N and repeat the process.

While computer-intensive, these steps are certainly feasible. They can be performed using the commonly available statistical packages (e.g. SHAZAM, SAS), for any posterior distribution that the researcher specifies. All that is required is a random number generator and some simple calculations.

4. Importance Sampling for Exact Results

The procedure outlined above relies on the asymptotic properties of the estimation procedure, by making use of a normal approximation. While this is comparable to what is done using non-Bayesian approaches, it will not yield results consistent with the exact posterior density function. Unfortunately, the posterior distribution for θ is only of the multivariate normal form if the variance matrix Σ is known, rather than estimated jointly with the parameter vector. Effectively, the procedure outlined above substituted a conditional distribution $f(\theta | y, \hat{\Sigma})$ for the marginal distribution $f(\theta | y)$.

Such a marginal distribution for θ can be obtained from a joint posterior density for θ and the parameters of Σ, by integrating over a posterior density for Σ. Following Zellner (1971, p. 242) and Judge et al. (p. 478), for a diffuse prior density for both Σ and θ, the resulting posterior density for θ is given by

$$f(\theta | y) \propto |A|^{\frac{-T}{2}}.$$

A typical element a_{ij} of the G by G matrix A is given by

$$a_{ij} = [(e_i(\theta))'(e_j(\theta))],$$

where $e_i(\theta)$ is the vector of residuals for equation i evaluated using any value of θ where the posterior density is defined. This density function corresponds to the posterior density for θ with no prior information about likely values. Should one wish to impose the restriction that θ could take on only values consistent with the inequality restrictions, a truncated version of this posterior density must be used. In this framework, imposing the Bayesian restrictions and finding a posterior mean $\overline{\theta}$ requires sampling from the truncated density $f^R(\theta \mid y)$, but this is not a familiar density, and therefore it is difficult to obtain an appropriate random sample. The procedure outlined above, whereby the untruncated density $f(\theta \mid y)$ could be used, also cannot be applied, for the same reason.

Instead, it is necessary to modify the steps outlined above, correcting for the fact that the multivariate normal is at best only approximately the correct posterior density. The technique for doing so is called importance sampling (Kloek and van Dijk (1978), van Dijk and Kloek (1980)). The concept which underlies importance sampling is relatively straightforward. Before returning to the problem at hand, we illustrate its use with a simple example.

Consider estimating the mean of Y, a Beta random variable with α=9 and β=1 and density function $f(y)$. Such a random variable has

$$E(Y) = \int_0^1 y\, f(y)\, dy = \frac{\alpha}{\alpha+\beta} = .9$$

Of course, this is an example where the integral could be evaluated, and where random draws can also be obtained from the correct distribution, but suppose that we had available only a Uniform random number generator, whose density function we denote $g(y)$. How might we calculate the mean of the Beta distribution, using random draws from $U(0, 1)$?

The sample mean of replications drawn from the $U(0, 1)$ will underestimate $E(Y)$, since it will tend toward 1/2. The reason, of course, is that values close to zero for Y occur more often under $U(0, 1)$ and values closer to one will occur less often than under the Beta distribution we have chosen. Importance sampling corrects for this, by adjusting the "importance" given to each replication. It turns out that the appropriate weight for each replication y_i is the ratio of the probability density function of the Beta distribution at y_i, $f(y_i)$, and the density of the uniform at y_i, $g(y_i)$. In this way, those values draw which are closer to one will receive a large weight while those closer to zero will receive a smaller weight.

To see why this works, note that the expected value of Y using the density function given by $f(y)$ can be found by integrating over $g(y)$ instead:

$$E(Y) = \int y\, f(y)\, dy = \int y \frac{f(y)}{g(y)} g(y)\, dy$$

In the first instance, $E(Y)$ is taken with respect to f and in the second, $E[Y\, f(Y)/g(Y)]$ is taken with respect to g. Just as $E(Y)$ could be estimated using a sample mean of replications from $f(y)$, then, so could it be estimated by drawing from $g(y)$ and calculating

$$\frac{\sum_{i=1}^{N} y_i \frac{f(y_i)}{g(y_i)}}{N}.$$

One surprising aspect of this procedure is that *any* density function can be used as $g(y)$, provided it is strictly positive over the range of Y, determined from $f(y)$. Otherwise, division by $g(y)$ within the integral is not allowed, and the implication would be that some values of Y, which do occur when sampling from $f(y)$, would never be drawn using $g(y)$. Naturally, if the weights applied to each y_i are close to one, so that $f(y)$ and $g(y)$ are similar, fewer draws will be required to obtain good estimates of the values of these integrals (Kloek and van Dijk (1978), van Dijk and Kloek (1980); Geweke (1986, 1988)).

This procedure can be applied for estimating the mean of the exact posterior density for θ. That distribution plays the role of the Beta distribution in the case above, in that it is the correct density but difficult to work with. Meanwhile, the multivariate normal is used as was the Uniform, to generate replications for θ. In the case of uninformative priors, draws from the multivariate normal can be adjusted by the ratio of the two density functions to obtain an estimate of the probability that the inequality restrictions hold. Alternatively, the inequality restrictions can be imposed, so that the posterior density is truncated or restricted. To find the posterior mean, find the solution to

$$E(\theta) = \int_D \theta\, f^R(\theta | y)\, d\theta.$$

At the same time, to calculate the probability that the restrictions hold, find

$$p = \int_{\forall\theta} I(\theta)\, f(\theta | y)\, d\theta$$

where $I(\theta) = 1$ if the restrictions hold and $I(\theta) = 0$ otherwise.

Each of these could be accomplished by sampling from the exact posterior density $f(\theta | y)$, if it were of a known form, and the steps outlined earlier could be used. Since these integrands are too complicated to permit analytic solutions, importance sampling must be used. To reiterate, notice that the posterior mean can be found by

$$E(\theta | y) = \int_D \theta\, \frac{f^R(\theta | y)}{g^R(\theta | y)}\, g^R(\theta | y)\, d\theta$$

where $g^R(\theta | y)$ is the truncated multivariate normal (or any other convenient) p.d.f. and D is the region of the parameter space consistent with the concavity and monotonicity restrictions. Again, the density $g^R(\theta | y)$ must have the same range (D) as the posterior density $f^R(\theta | y)$. We use the multivariate normal p.d.f.

$$g(\theta \mid y) \propto \exp[-\tfrac{1}{2}\,\hat{\theta}'\, V(\hat{\theta})^{-1}\,\hat{\theta}].$$

The modified steps now required for the calculations, taken from Chalfant and White, are given below:

1. Estimate the parameters of the demand system using iterated seemingly unrelated regressions, to obtain maximum likelihood estimates $\hat{\theta}$ and the estimated covariance matrix $V(\hat{\theta})$, in this case a 12 element vector and a 12 by 12 matrix.

2. Calculate a matrix H such that $HH' = V(\hat{\theta})$. Draw a random vector of length 12 from the standard normal distribution

$$w \sim N(0, I)$$

where I is the identity matrix of order 12. Replications of θ can be generated using

$$\theta^A = \hat{\theta} + Hw$$

and its "antithetic replication"

$$\theta^B = \hat{\theta} - Hw.$$

The latter step was suggested by Geweke (1988) to improve convergence.

3. Check each replication to see if it violates concavity, monotonicity or substitutability. For each, note whether concavity and monotonicity hold jointly, and whether concavity, monotonicity, and substitutability all hold or whether there is some violation.

To do so, we calculated elasticities of substitution using each replication and the means of observed budget shares for the 4 meats. We checked concavity and substitutability using these elasticities. To check monotonicity, we used each replication to obtain new predicted shares for all 25 data points.

4. Estimate the mean of the posterior distribution using the n draws of θ^A or θ^B which satisfy the restrictions:

$$\bar{\theta} = \frac{\sum_{k=1}^{n} \theta_k \cdot \frac{f^R(\theta_k \mid y)}{g^R(\theta_k \mid y)}}{\sum_{k=1}^{n} \frac{f^R(\theta_k \mid y)}{g^R(\theta_k \mid y)}}.$$

As noted by Chalfant and White, if $f(\theta \mid y)$ and $g(\theta \mid y)$ were proper density functions, a denominator of N would suffice. Otherwise, the denominator serves as a normalizing constant to correct for the fact that we use only the kernels of proper densities.

5. To estimate the probability that the restrictions hold, use *all* replications, letting the first n be those consistent with the restrictions, and calculate

$$\hat{p} = \frac{\sum_{k=1}^{N} I(\theta_k) \frac{f(\theta_k \mid y)}{g(\theta_k \mid y)}}{\sum_{k=1}^{N} \frac{f(\theta_k \mid y)}{g(\theta_k \mid y)}} = \frac{\sum_{k=1}^{n} \frac{f^R(\theta_k \mid y)}{g^R(\theta_k \mid y)}}{\sum_{k=1}^{N} \frac{f(\theta_k \mid y)}{g(\theta_k \mid y)}}.$$

In addition to $\overline{\theta}$ and $\hat{p}$, standard deviations of the posterior distribution are easily calculated. Interval estimates or histograms can also be used to summarize the information about these values that is contained in the posterior distribution. Following Geweke, a numeric standard error for $\hat{p}$ can be calculated using the formula

$$s.e.(\hat{p}) = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}.$$

6. Check to see if the number of replications is large enough to arrive at stable estimates of $\overline{\theta}$ or $\hat{p}$ and of their standard errors. If not, increase N.

We calculated the quantities of interest for several sample sizes, as the number of replications is increased.

5. Application to Aggregate Meat Consumption in Canada

In this section, an almost ideal demand system[5] for meat and fish products is estimated using aggregate Canadian data for the years 1960 to 1984, taken from Van Kooten. The demands for four goods were examined—beef, pork, poultry, and fish—henceforth, the meats group. It was assumed that consumer preferences for the meats group are weakly separable from all other goods. Prices and per capita quantities consumed for each meat are given in Van Kooten. Real expenditures were obtained by deflating total meats expenditures by Stone's geometric price index, as suggested by Deaton and Muellbauer.

[5] Other applications of this system to agricultural data include Blanciforti, Green, and King, Chalfant, and Hayes, Wahl, and Williams. The latter case, which examines Japanese consumption of meat products, also considers the restriction that all meats are substitutes.

A check for consistency with the generalized axiom of revealed preference (Varian) revealed no violations of the axiom, adding support to the notion that meats are indeed weakly separable from other goods. As did Chalfant and Alston for United States and Australian data, we interpret this as evidence consistent with the stability of demands for the meats group over time. As a result, one is justified in fitting a demand system which includes only prices and expenditures, without trends or other "taste shifters", in explaining patterns of consumption through time.

Recall that the expression for the share of the budget allocated to the *ith* meat is

$$S_i = \alpha_i + \sum_{j=1}^{4} \gamma_{ij} \ln(P_j) + \beta_i \ln(x/P).$$

A system of three such equations was estimated using the nonlinear regression (NL) procedure of Version 6 of SHAZAM (White et al.). The fourth equation was deleted due to singularity of the variance matrix for all four equations, and parameters of that equation were obtained through the homogeneity and symmetry restrictions. By iterating over both the parameters and the error variance-covariance matrix, the estimates are obtained are invariant to the equation chosen for deletion (Barten). The parameter estimates are not in themselves of any interest and are used only to calculate elasticities or obtain predicted budget shares, and so are not reported. Elasticities of substitution are reported in Table 1, in the column denoted "Unconstrained". These were calculated at the mean budget shares observed in the sample, and were obtained without the inequality constraints.

The negative own-elasticities are as one would expect. The positive cross-elasticities indicate that the meats tend to be substitutes for one another at the midpoint of the sample. The unrestricted results which are contrary to prior belief are the negative substitution elasticities, indicating complementarity, between fish and beef and between fish and pork, though the magnitudes of these elasticities are rather small. Concavity and monotonicity also hold with these estimates at every point in the sample.

Using the procedure outlined earlier, we estimated the probability that the concavity and substitution restrictions hold for this demand system. By applying the Bayesian approach with this application, we illustrate an important point. When the restriction does not hold "in sample", as is the case with our substitution restriction, the approach is necessary to find parameter estimates consistent with the constraint. However, concavity does hold. That does not mean, however, that the posterior probability that the restriction holds is one, nor is it zero when the maximum likelihood estimate violates concavity. We still must obtain the posterior distribution of the parameter estimates to find the probability that concavity holds, and to find the posterior mean to use as a parameter estimate.

Table 1: Elasticity of Substitution

	Unconstrained	Restricted: Concavity, Normal Approx.	Restricted: Concavity, Importance Sampling	Restricted: Substitutability, Normal Approx.	Restricted: Substitutability, Importance Sampling
$\sigma_{beef\ beef}$	-0.867	-0.870	-0.874	-0.935	-0.909
$\sigma_{beef\ pork}$	1.057	1.056	1.054	1.037	1.016
$\sigma_{beef\ poultry}$	0.711	0.708	0.705	0.623	0.597
$\sigma_{beef\ fish}$	-0.207	-0.197	-0.182	0.051	0.047
$\sigma_{pork\ pork}$	-2.138	-2.140	-2.137	-2.078	-2.106
$\sigma_{pork\ poultry}$	0.839	0.837	0.845	0.583	0.649
$\sigma_{pork\ fish}$	-0.116	-0.110	-0.117	0.078	0.102
$\sigma_{poultry\ poultry}$	-5.789	-5.954	-5.832	-5.094	-5.098
$\sigma_{poultry\ fish}$	2.524	2.537	2.567	2.441	2.411
$\sigma_{fish\ fish}$	-1.676	-1.770	-1.767	-2.438	-2.436

Table 2: Price Elasticities With Concavity Imposed

Quantity i

Price j	Beef	Pork	Poultry	Fish
Beef	-0.350	0.422	0.282	-0.078
Pork	0.267	-0.540	0.214	-0.029
Poultry	0.112	0.139	-0.962	0.423
Fish	-0.033	-0.021	0.462	-0.321

Table 3: Price Elasticities With Concavity and Substitution Imposed

Quantity i

Price j	Beef	Pork	Poultry	Fish
Beef	-0.363	0.407	0.239	0.019
Pork	0.257	-0.532	0.164	0.026
Poultry	0.098	0.107	-0.841	0.398
Fish	0.008	0.018	0.438	-0.442

Table 4: Replications and Parameter Values[7]

Parameter Estimate	2000 Replications	5000 Replications	10000 Replications
θ_1	-.447 (.0046)	-.450 (.0052)	-.450 (.0069)
θ_2	.1001 (.00020)	.101 (.00019)	.102 (.00029)
θ_3	.0063 (.0000066)	.0059 (.0000078)	.0062 (.0000079)
θ_4	-.019 (.000012)	-.019 (.000011)	-.019 (.000019)
θ_5	.206 (.00083)	.206 (.00094)	.206 (.00121)
θ_6	.718 (.010)	.721 (.012)	.726 (.016)
θ_7	.052 (.000059)	.055 (.000060)	.052 (.000056)
θ_8	-.0066 (.000011)	-.0068 (.000018)	-.0072 (.000014)
θ_9	-.094 (.00021)	-.095 (.00025)	-.096 (.00033)
θ_{10}	.541 (.0069)	.535 (.0069)	.530 (.0076)
θ_{11}	-.020 (.000035)	-.020 (.000046)	-.019 (.000040)
θ_{12}	-.093 (.00024)	-.092 (.00024)	-.092 (.00026)
$\hat{p}_{cm}$	.926 (.0000357)	.914 (.0000164)	.909 (.00000865)
$\hat{p}_{subs}$	.00773 (.000767)	.0116 (.000326)	.00847 (.000147)

[7] Standard errors are shown in parentheses. $\hat{p}_{cm}$ is the probability that concavity and monotonicity hold jointly. $\hat{p}_{subs}$ is the probability that concavity, monotonicity and all meats are substitutes hold jointly.

We followed the procedure outlined earlier to obtain a sample of size 10,000 from the multivariate normal, again using SHAZAM. We used our in-sample results for $\hat{\theta}$ and $V(\hat{\theta})$ to specify the parameters of this distribution. We checked concavity for each replication by calculating substitution elasticities using the parameter values given by each replication. For each matrix of elasticities of substitution, we calculated the eigenvalues. We found consistency with the concavity restriction (a substitution matrix without positive eigenvalues) over 95% of the time.[8] This turns out to be slightly larger than our estimate of the probability that the restrictions hold when the exact probability is calculated using importance sampling—the probability that concavity and monotonicity holds falls to .91. These results imply that the demand system is certainly well behaved by this criterion; imposing the restriction by removing the 464 trials which violated this condition is likely to have little effect on the parameter estimates given by the posterior mean, $\overline{\theta}$. Elasticities of substitution calculated using the concavity restriction are given in Table 1; note that there is not much difference between the conditional results from the multivariate normal and those obtained using importance sampling. Price elasticities calculated using $\overline{\theta}$ are shown in Table 2.

The system was not consistent with the restriction that all meats should be substitutes. In our random sample, out of 10,000 draws only 57, or .57 percent, met this restriction, slightly less than the probability of .0085 obtained with importance sampling. This strong rejection of our prior beliefs casts some doubt on the estimated system being a valid representation of preferences. Nonetheless, we calculated posterior means for the remaining elasticities. These numbers are shown in Table 1 for the substitution elasticities and in Table 3 for the price elasticities. In spite of the low posterior probability associated with the substitution restriction, the magnitudes of elasticities are not affected dramatically.

It is important to note that violations of the substitution restriction are not due solely to the cases of complementarity relationships implied by $\hat{\theta}$—beef and fish or pork and fish. Any elasticity could be responsible. Unless the posterior density for any particular elasticity of substitution implies that it is positive with probability one, each can be responsible for any particular replication violating the constraint.

In order to examine whether a sample size of 10,000 was sufficiently large to get an accurate measure of the posterior distribution via importance sampling we examined parameter estimates for 2,000, 5,000, and 10,000 draws. These estimates are show in Table 4. The 10,000 replications seemed to be sufficiently large to estimate the posterior mean of θ. The estimates for $\hat{p}$ seemed to be more sensitive to sample size, although there is only a small difference in the estimate between the sample sizes of 5,000 and 10,000. These results suggest that 10,000 replications gave a reliable estimate of the desired quantities.

6. Summary and Conclusions

An unfortunate by-product of the use of demand systems which do not restrict substitution elasticities is that theoretical restrictions such as symmetry or homogeneity, which are automatically met by the simpler forms, are often violated. More difficult to cope with, we have argued, are those restrictions that involve inequality restrictions. The familiar problem of curvature restrictions is the best example, but we suggested in this paper that the signs of elasticities of substitution between goods were also good examples. In order to determine whether an estimated demand system is entirely consistent with our prior beliefs, at least for cases such as the present application where substitution relationships seem fairly likely, it is important to be able to impose or make inferences about inequality restrictions.

We showed that a Bayesian procedure handled this problem nicely. It produces constrained parameter estimates and also an estimate of the probability that the restrictions are true. For the demands for beef, pork, chicken, and fish in Canada, we found substantial support for the concavity of the consumer's expenditure function underlying an almost ideal demand system. By itself, this is an encouraging finding, and lends support to the nonparametric results we obtained. On the other hand, the sample information is not consistent with prior beliefs about substitution relationships, since it reveals a very low probability that these four goods are all substitutes.

Because the necessary integrals over the exact posterior density were quite complicated, it was necessary to use Monte Carlo integration to estimate parameter values. The p.d.f. of the exact posterior was known but not recognizable, which made sampling from the exact posterior difficult. This problem was overcome by the use of importance sampling. As a by-product of using the multivariate normal to perform the steps involved in importance sampling, we found that there is not a great difference between the exact and conditional posterior results. This finding may not hold in other instances. In any event, the method of importance sampling is not difficult to apply to the estimation of inequality-constrained demand systems.

Our findings are conditional not only on the observed data but on the specification of the almost ideal demand system. All such inferences in demand systems are also conditional on separability and aggregation assumptions (Chalfant and Alston), but if these are valid, the results must be interpreted as questioning either the prior belief that all these goods are substitutes, or the functional form for the demand system. Further research with other functional forms can help to answer this question. In that light, the procedure we have outlined and the restrictions we suggest serve not only as a means to interpret the data, but as a way of evaluating alternative functional forms for demand systems.

REFERENCES

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

Barten, Anton P. "Maximum Likelihood Estimation of a Complete System of Demand Equations" *European Economic Review* 1(1969):7-73.

Blanciforti, L., and R. D. Green. "An Almost Ideal Demand System Incorporating Habits: An Analysis of Expenditures on Food and Aggregate Commodity Groups" *Review of Economics and Statistics* 65(1983):511-515.

Blanciforti, L., and R. D. Green. "The Almost Ideal Demand System: A Comparison and Application to Food Groups" *Agricultural Economics Research* 35(1983):1-10.

Chalfant, James A. "A Globally Flexible, Almost Ideal Demand System" *Journal of Business and Economic Statistics* 5(1987):233-242.

Chalfant, James A., and Julian M. Alston. "Accounting for Changes in Tastes" *Journal of Political Economy* 96(1988):390-410.

Chalfant, James A., and Kenneth J. White. "Estimation of Demand Systems with Concavity and Monotonicity Constraints" University of California, Department of Agricultural and Resource Economics, 1988.

Deaton, Angus, and Muellbauer, John. "An Almost Ideal Demand System" *American Economic Review* 70(1980):312-326.

Geweke, J. "Exact Inference in the Inequality Constrained Normal Linear Regression Model" *Journal of Applied Econometrics* 1(1986):127-141.

Geweke, J. "Antithetic Acceleration of Monte Carlo Integration in Bayesian Inference," *Journal of Econometrics* 38(1988):73-89.

Hayes, Dermot J., Thomas I. Wahl, and Gary W. Williams. "Testing Restrictions on a Model of Japanese Meat Demand" Paper presented at the 1988 American Agricultural Economics Association meetings, Knoxville, TN.

Judge, George. G., W. E. Griffiths, R. Carter Hill, Helmut Lutkepohl, and Tsoung-Chao Lee. *The Theory and Practice of Econometrics,* 2nd. ed. New York: John Wiley and Sons, 1985.

Kloek, T., and H. K. van Dijk. "Bayesian Estimates of Equation System Parameters: An Application of Integration by Monte Carlo," *Econometrica* 46(1978):1-19.

Rossi, P. "Comparison of Alternative Functional Forms in Production" *Journal of Econometrics* 30(1985):345-361.

Stone, J. R. N. *The Measurement of Consumers' Expenditure and Behavior in the United Kingdom, 1920-1938, Vol. 1.* Cambridge: Cambridge University Press, 1954.

White, K. J. "SHAZAM: A General Computer Program for Econometric Methods (Version 5)" *American Statistician* 41(1987):80.

Van Dijk, H. K., and T. Kloek. "Further Experience in Bayesian Analysis Using Monte Carlo Integration" *Journal of Econometrics* 14(1980):307-328.

Van Kooten, G. C. " The Economic Impacts on Consumers of Government Intervention in the Poultry and Egg Sectors: A Comparison of Alternative Welfare Measures" Agriculture Canada Policy Branch Working Paper No. 5/87, 1987.

Varian, Hal R. "The Nonparametric Approach to Demand Analysis" *Econometrica* 50(1982):945-973.

White, Kenneth J., Shirley A. Haun, Nancy G. Horsman, and S. Donna Wong. *SHAZAM Econometrics Computer Program: User's Reference Manual, Version 6.1.* New York: McGraw-Hill Book Company, 1988.

Wohlgenant, Michael K. "Estimating Cross Elasticities of Demand for Beef" *Western Journal of Agricultural Economics* 10(1985):322-329.

Woodland, A. D. "Stochastic Specification and the Estimation of Share Equations" *Journal of Econometrics* 10(1980):361-383.

Zellner, A. "An Efficient Method of Estimating Seemingly Unrelated Regression and Tests for Aggregation Bias" *Journal of the American Statistical Association* 57(1962):348-368.

Zellner, A. *An Introduction to Bayesian Inference in Econometrics* New York: John Wiley and Sons, 1971.

Zellner, A. "Bayesian Analysis in Econometrics" *Journal of Econometrics* 37(1988):27-50.