Estimation of Coherent Demand Systems with Many Binding Non-Negativity Constraints

Mark M. Pitt

Daniel L. Millimet[1]

Brown University

March 1999

[1]An earlier draft of this paper was presented at the Winter Meeting of the Econometric Society held in New York City, January 1999. We benefitted from discussions with Moshe Buchinsky and Tony Lancaster and from the comments of participants in the Brown Econometrics Workshop. Email: *mp@pittsun.pstc.brown.edu* and *millimet@econ.pstc.brown.edu*

Abstract

Two econometric issues arise in the estimation of complete systems of producer or consumer demands when many non-negativity constraints are binding for a large share of observations, as frequently occurs with micro-level data. The first is computational. The econometric model is essentially an endogenous switching regimes model which requires the evaluation of multivariate probability integrals. The second is the relationship between demand theory and statistical coherency. If the indirect utility or cost function underlying the demand system does not satisfy the regularity conditions at each observation, the likelihood is incoherent in that the sum of the probabilities for all demand regimes is not unity and maximum likelihood estimates are inconsistent. The solution presented is to use the Gibbs Sampling technique and data augmentation algorithm and rejection sampling, to solve both the dimensionality and coherency problem. With rejection sampling one can straightforwardly impose only the necessary conditions for coherency, *coherency at each data point* rather than global coherency. The method is illustrated with a series of simulated demand systems derived from the translog indirect random utility function. The results highlight the importance of imposing regularity when there are many non-consumed goods and the gains from imposing such conditions locally rather than globally.

JEL Classification Codes: C3, C4, D0

Key Words: coherency, Gibbs Sampling, demand systems, translog, data augmentation, Markov Chain Monte Carlo

1 Introduction

An unresolved problem in applied demand analysis is the estimation of complete systems of demand equations from household surveys having highly disaggregated consumption data when there are substantial numbers of non-consumers for many goods. This characterizes demand analysis with individual, household, or firm-level data.

Derivation of an econometric model from the maximization of a random direct utility function subject to the Kuhn-Tucker conditions characterizes the primal solution to the consumer problem (Wales and Woodland (1983)). This method, however, rules out the use of more flexible demand specifications for which no explicit specification of the direct utility function can be given. The dual solution consists of deriving consumer demand systems from indirect cost or utility functions including popular flexible functional forms such as the translog by specifying virtual (or reservation) prices which are dual to the Kuhn-Tucker conditions (Lee and Pitt (1986, 1987)).

Two econometric issues arise in the estimation of such demand systems. The first is the dimensionality problem. The econometric model is essentially an endogenous switching regimes model which requires the evaluation of multivariate probability integrals. The computational burden can be reduced by placing a factor structure on the random components of the indirect utility function (Lee and Pitt (1986)) or by using simulated maximum likelihood methods (Lee (1997)).

The second econometric issue which must be addressed is the coherency problem (Ransom (1987); van Soest and Kooreman (1990); van Soest *et al.* (1993)). Coherency of the likelihood function is not guaranteed over the entire parameter space. In an incoherent likelihood the sum of the probabilities for all demand regimes is not one. Quasi-concavity of the cost or utility function is sufficient (but not necessary) to guarantee coherency. However, current methods to impose quasi-concavity locally have been characterized as "unsatisfactory" by Diewert and Wales (1987) and parametric restrictions required to guarantee global quasi-concavity of the translog and other demand systems destroy their flexibility (Lau (1978); Jorgensen and Fraumani (1981); Gallant and Golub (1984); Diewert and Wales (1987); Terrell (1991)). In addition, Hausman (1985) and MaCurdy *et al.* (1990) have shown that in models of individual labor supply, the restrictions necessary to guarantee global concavity impose *a priori* undesirable limitations on income and wage elasticities.

A number of empirical researchers have been concerned that the flexible functional forms used in various econometric models fail to satisfy the theoretical regularity conditions – strict monotonicity and quasi-concavity. The standard approach in these studies is, at most, to check *ex post* for regularity at each data point and report it as a "statistic" reflecting the reasonableness of the restrictions imposed by demand theory. In fact, the likelihood function for any demand system having corner solutions which is not quasi-concave at every point is incoherent and the maximum likelihood estimates are inconsistent (van Soest *et al.* (1993)). Circumventing this problem through the specification of globally concave functional forms, such as the Linear Expenditure System, imposes unrealistic restrictions on behavior, particularly with disaggregate demands and micro-level data. A recent, notable addition to this literature is Terrell (1996), who employs Bayesian Gibbs Sampling techniques, along with rejection sampling, to impose the theoretical concavity conditions locally to demand systems with only interior solutions.

In this paper, we extend the work of Pitt and Lee (1986, 1987) to the estimation of coherent demand systems in the presence of binding non-negativty constraints. Our solution is to use the Gibbs Sampling technique, as did Terrell, along with the data augmentation algorithm (Tanner and Wong (1987)), to solve both the dimensionality and the coherency problem. With rejection sampling one can straightforwardly impose only the necessary conditions for coherency, *coherency at each data point* rather than global regularity. The remainder of the paper is organized as follows: section 2 presents the dual approach to the consumer maximization problem when there are binding non-negativity constraints; section 3 discusses the coherency problem in the context of the translog demand system; section 4 details the estimation algorithm; section 5 presents the Monte Carlo results; and, section 6 contains some concluding remarks.

2 The Dual Approach to Binding Non-Negativity Constraints

Following Lee and Pitt (1986a), let $H(v;\theta,\epsilon)$ be an indirect utility function defined as

$$H(v;\theta,\epsilon) = \overset{max}{q} \quad \{U(q;\theta,\epsilon) \ \mid \ vq = 1\} \quad , \tag{1}$$

where $U(.;\theta,\epsilon)$ is a strictly quasi-concave utility function defined over k commodities, v is a vector of normalized (by income) market prices, θ is a vector of unknown parameters, and ϵ is a vector of random components. The latent demand equations $q(v;\theta,\epsilon)$ for a set of k goods are derived by applying Roy's Identity:

$$q_i = \frac{\frac{\partial H(v;\theta,\epsilon)}{\partial v_i}}{\sum_{j=1}^{k} v_j \frac{\partial H(v;\theta,\epsilon)}{\partial v_j}} \quad i = 1,...,k \ . \tag{2}$$

These demand equations are latent because they are not restricted to the positive orthont as the maximization problem given by (**??**) does not contain any non-negativity constraints. The latent demands q_i correspond to a vector of non-negative observed demands, x_i as follows: there exists a vector of positive virtual prices π which can exactly support these zero demands (or any other allocation) as long as the preference function is strictly quasi-concave, continuous, and strictly monotonic (Neary and Roberts (1980)). Although analytically deriving virtual price functions when demands are rationed has been shown to be difficult for many popular functional forms (Deaton and Muellbauer (1981)), the problem is enormously simplified when the "ration" is at zero. In this special case, the denominator in (**??**) drops out of the virtual price function. If demands for the first m goods are zero, the virtual prices $\pi_i(v_{m+1}, ..., v_k)$ are solved from the equations

$$0 = \frac{\partial H(\pi_1(\bar{v}), ..., \pi_m(\bar{v}), \bar{v}; \theta, \epsilon)}{\partial v_i} \quad i = 1, ..., m \tag{3}$$

where $\pi_i(\bar{v})$ is the virtual price of good i and $\bar{v}$ is the set of market prices of the positively consumed goods $m+1$ to k.

The market prices $\bar{v}$ are also the virtual prices for the consumed goods as they exactly support the observed positive demands. The remaining (positive) demands are

$$x_i = \frac{\frac{\partial H(\pi_1(\bar{v}), ..., \pi_m(\bar{v}), \bar{v}; \theta, \epsilon)}{\partial v_i}}{\sum_{j=1}^{k} v_j \frac{\partial H(\pi_1(\bar{v}), ..., \pi_m(\bar{v}), \bar{v}; \theta, \epsilon)}{\partial v_j}} \quad i = m+1, ..., k \tag{4}$$

The equations (**??**) are estimable and the parameters of the latent demand equations (**??**) can be identified by estimating this conditional demand system.

Selection among different regimes – defined by the set of positively consumed goods at the optimum – is done by a comparision of the virtual and market prices. The regime in which the first m goods are not consumed is characterized by the conditions

$$\pi_i(\bar{v}) \leq v_i \qquad i = 1, ..., m \tag{5}$$

This characterization follows directly from the Kuhn-Tucker conditions.

3 Coherency of the Translog Demand System

When estimating flexible functional forms of consumer or producer demands, imposing the theoretical curvature restrictions is crucial if the data contains corner solutions, unlike when only interior solutions are observed. Even if the true demand model generating the data is quasi-concave at each observation, failure to impose this condition in the presence of corners may result in inconsistent estimates of the parameters of the model if the iterative estimation process leaves the regular region of the parameter space (van Soest *et al.* (1993)). This inconsistency results from the incoherency of the econometric model. In an incoherent model, the sum of the probabilities of the demand regimes is not one. The conditions required for coherency are, first, that every possible vector ϵ of random components generates a unique set of demands, and second, that every demand regime can be generated by some ϵ vector. Restricting the parameter space to the regular region is sufficient to guarantee coherency. However, failure to impose regularity does not necessarily lead to inconsistent estimates if the iterative process does not exit the regular region of the parameter space. For example, Lee and Pitt (1987) estimate a linear translog demand system without imposing regularity *ex ante*, but confirm the coherency of their estimates *ex post*.

3.1 Non-Linear Translog Demand System

To illustrate the coherency problem, consider the translog indirect utility function of Christensen *et al.* (1975):

$$H(v;\theta,\epsilon) = \sum_{i=1}^{k} \alpha_i v_i + \frac{1}{2}\sum_{i=1}^{k}\sum_{j=1}^{k} \beta_{ij} v_i v_j + \sum_{i=1}^{k} \epsilon_i v_i \quad , \tag{6}$$

where v is a vector of *log* normalized market prices, $\epsilon \sim N_k(0,\Sigma)$, and there are k goods. A convenient normalization is

$$\sum_{i=1}^{k} \alpha_i = 1 \quad ; \quad \sum_{i=1}^{k} \epsilon_i = 0 \quad . \tag{7}$$

The latent budget share equations are

$$v_i q_i = \frac{\alpha_i + \sum_{j=1}^{k} \beta_{ij} v_j + \epsilon_i}{D} \quad , \tag{8}$$

where

$$D = 1 + \sum_{i=1}^{k}\sum_{j=1}^{k} \beta_{ij} v_j > 0 \quad , \tag{9}$$

which is strictly positive as a consequence of utility increasing in income.

For the model to be coherent, the indirect utility function must be locally quasi-concave, which requires that

$$ss' - \hat{s} + \frac{1}{D + y'\beta e}[\beta - s(\beta e)' - \beta e s' + e'\beta e s s'] \tag{10}$$

must be a negative semi-definite matrix at each data point, where β is a k x k matrix of slope parameters, $\hat{s}$ is a k x k diagonal matrix with s_i, $i = 1, ..., k$, the observed shares, along the diagonal,

s is a k-dimensional vector of observed shares, y is a k-dimensional vector of the differences between log market and virtual prices, and e is a k-dimensional vector of ones (van Soest and Kooreman (1990)). In contrast, the model is globally coherent if the indirect utility function is globally quasi-concave, which requires that β is a negative semi-definite matrix. However, as Diewert and Wales (1987) demonstrate, this restriction destroys much of the flexibility that motivates the use of these functional forms.

3.2 Linear Translog Demand System

A common representation for a firm's cost function is the linear (homogeneous) translog indirect cost function. It is a special case of (**??**), obtained by imposing the restrictions that $\sum_{i=1}^{k} \beta_{ij} = 0$. As a result, $\beta e = 0$, D in (**??**) reduces to one, and the local coherency condition in (**??**) simplifies to the requirement that

$$\beta - \hat{s} + ss' \tag{11}$$

is a negative semi-definite matrix for each observation (Diewert and Wales (1987)). As in the non-linear model, global coherency requires β to be a negative semi-definite matrix.

4 Estimation via Gibbs Sampling

Examples of the linear and non-linear translog demand models are estimated using the Markov Chain Monte Carlo (MCMC) technique of Gibbs Sampling. In models with a simple latent structure, the observed data may be augmented in order to make application of the Gibbs Sampler straightforward. This process converts the observed demands into latent demands. However, drawing latent quantities when the demand system is not regular (i.e., augmenting the data conditional on parameters which fail to satisfy the local curvature conditions) restricts the space of possible values that the latent quantities can take. As a result, in an irregular demand system the Gibbs Sampler corresponds to MCMC estimation of the underlying incoherent and inconsistent likelihood. Consequently, imposing the curvature condition sufficient for coherency of the likelihood is also necessary to guarantee convergence of the Gibbs Sampler to the appropriate distribution.

The desired density to be sampled from is $p(\alpha, \beta, \Sigma | s)$, corresponding to the random (stochastic) indirect utility function in (**??**), where s is the vector of observed budget shares. After augmenting the data, the actual density sampled from is $p(\alpha, \beta, \Sigma | s^*)$, where s^* is a vector of latent budget shares. While sampling directly from this density is quite burdensome, sampling from the component conditionals is not. This density may be factored into the following conditional densities:

i. $p(s^*|\alpha, \beta, \Sigma, s)$

ii. $p(\Sigma|s^*, \alpha, \beta)$

iii. $p(\alpha, \beta|s^*, \Sigma)$

The Gibbs Sampler constructs sequences of draws by sampling consecutively from the above conditional distributions. The precise steps are discussed below.

4.1 Augmenting Budget Shares

The first step is to augment the data by simulating the latent budget shares conditional on the observed data and initial values for the parameters of the model. The data augmentation algorithm is more complicated than in the multinomial probit model considered by McCulloch and Rossi (1994). In the case of a demand system with binding non-negativity constraints, one begins by augmenting only the budget shares for the non-consumed goods. The augmented (latent) demands uniquely identify a vector of virtual prices. The remaining random components, ϵ, of the indirect utility function corresponding to the goods for which the observed budget share is strictly positive can be solved for from these virtual prices. Finally, the augmented (latent) demands for the consumed goods are uniquely determined given these random components. Thus, even for the goods not at a corner, the latent and observed budget shares are generally not identical.

4.1.1 Linear Translog Demand System

Below we describe three cases characterized by the number of goods at a corner. A single good at a corner is differentiated from the case of multiple corners only by the ease of exposition.

Case 1. $s_1 = 0$, $s_i > 0 \;\; \forall i \neq 1$. The log virtual price for input 1 is given by

$$\pi_1 = -\frac{1}{\beta_{11}}(\alpha_1 + \sum_{j=2}^{k} \beta_{1j} v_j + \epsilon_1) \quad . \tag{12}$$

Since the observed demands for the consumed inputs depend on π_1 and not v_1, the share equations are

$$s_i \;\; = \;\; \alpha_i - \frac{\beta_{i1}}{\beta_{11}}\alpha_1 + \sum_{j=2}^{k}(\beta_{ij} - \beta_{i1}\frac{\beta_{1j}}{\beta_{11}})v_j + [\epsilon_i - \frac{\beta_{i1}}{\beta_{11}}\epsilon_1]$$

$$\equiv \alpha_i - \frac{\beta_{i1}}{\beta_{11}}\alpha_1 + \sum_{j=2}^{k}(\beta_{ij} - \beta_{i1}\frac{\beta_{1j}}{\beta_{11}})v_j + \tilde{\epsilon}_i \qquad i = 2, ..., k \tag{13}$$

To appropriately augment the data, first derive the components of Ω, the covariance matrix of ϵ_1 and $\tilde{\epsilon}_i$, $i = 2, ..., k-1$, given the initial estimate of Σ. In addition, values for $\tilde{\epsilon}_i$ are obtained conditional on the initial estimates for α and β and the observed data. Next, one can draw $\epsilon_1|\tilde{\epsilon} \sim N(\Omega_{12}\Omega_{22}^{-1}\tilde{\epsilon}, \Omega_{11} - \Omega_{12}\Omega_{22}^{-1}\Omega_{21})$, where $\tilde{\epsilon}$ is a vector of length $k-2$ of observed residuals and ϵ_1 is truncated from above at $-(\alpha_1 + \sum_{i=1}^{k}\beta_{1i}v_i)$. Once an appropriate value has been drawn for ϵ_1, the latent cost shares for all k inputs are uniquely determined. However, even though ϵ_1 is drawn such that s_1^* is non-positive, there is no assurance that the resulting implied values for s_i^*, $i = 2, ..., k$, are consistent with the observed demand regime. As a result, one must check the regime conditions for the consumed inputs. Such conditions were derived in Lee and Pitt (1987) and require $s_i^* - \frac{\beta_{i1}}{\beta_{11}}s_1^* \in (0,1)$, $i = 2, ..., k$. If any of these conditions are violated, ϵ_1 must be re-drawn.

Case 2. $s_k > 0$**,** $s_i = 0 \;\; \forall i \neq k$**.** For demand regimes where more than one input is not consumed, virtual prices may be obtained by matrix inversion. Log virtual prices for inputs $1, ..., k-1$ are

$$\begin{bmatrix} \pi_1 \\ \vdots \\ \pi_{k-1} \end{bmatrix} = -B^{-1}\begin{bmatrix} \alpha_1 + \beta_{1,k}v_k \\ \vdots \\ \alpha_{k-1} + \beta_{k-1,k}v_k \end{bmatrix} - B^{-1}\begin{bmatrix} \epsilon_1 \\ \vdots \\ \epsilon_{k-1} \end{bmatrix} \tag{14}$$

where

$$B = \begin{bmatrix} \beta_{11} & \cdots & \beta_{1,k-1} \\ \vdots & & \vdots \\ \beta_{k-1,1} & \cdots & \beta_{k-1,k-1} \end{bmatrix} .$$

Estimates of the latent cost shares are obtained by drawing a $k-1$ vector of ϵ's unconditionally from a multivariate normal distribution with mean zero and a covariance matrix given by the upper $k-1$ by $k-1$ portion of Σ. Regime conditions for this demand regime require $B^{-1}s^* \geq 0$, where s^* is a vector of length $k-1$ whose elements are the augmented (latent) cost shares. Each element of the vector $B^{-1}s^*$ must be non-negative for virtual prices to be less than or equal to market prices since

$$\begin{bmatrix} \pi_1 \\ \vdots \\ \pi_{k-1} \end{bmatrix} = \begin{bmatrix} v_1 \\ \vdots \\ v_{k-1} \end{bmatrix} - B^{-1}\begin{bmatrix} s_1^* \\ \vdots \\ s_{k-1}^* \end{bmatrix} .$$

If any element of the vector $B^{-1}s^*$ is negative, the entire vector of ϵ's is re-drawn.

Case 3. $s_i > 0\ \forall i$. In this case, the observed cost shares are the latent shares as well since the non-negativity constraints are not binding. There is no need to augment the data.

4.1.2 Non-Linear Translog Demand System

Case 1. $s_1 = 0,\ s_i > 0\ \forall i \neq 1$. When the "ration" is at zero, D drops out of the virtual price equations. Thus, (**??**) still gives the log virtual price for good 1. Define $\pi_1 \equiv \mu_1 - \frac{1}{\beta_{11}}\epsilon_1$. The budget shares for the consumed goods are

$$s_i = \frac{\alpha_i + \beta_{i1}\pi_1 + \sum_{n=2}^{k}\beta_{in}v_n + \epsilon_i}{\tilde{D}} \qquad i = 2, ..., k \tag{15}$$

where

$$\begin{aligned} \tilde{D} &= 1 + \sum_{m=1}^{k}\sum_{n=2}^{k}\beta_{mn}v_n + \sum_{m=1}^{k}\beta_{m1}\pi_1 \\ &= 1 + \sum_{m=1}^{k}\sum_{n=2}^{k}\beta_{mn}v_n + \sum_{m=1}^{k}\beta_{m1}\mu_1 - \frac{1}{\beta_{11}}\sum_{m=1}^{k}\beta_{m1}\epsilon_1 \\ &\equiv \hat{D} - \frac{1}{\beta_{11}}\sum_{m=1}^{k}\beta_{m1}\epsilon_1 \quad . \end{aligned} \tag{16}$$

One can re-write the shares for the consumed goods as

$$\begin{aligned} \hat{D}s_i &= \alpha_i + \beta_{i1}\mu_1 + \sum_{n=2}^{k}\beta_{in}v_n + [\epsilon_i + \frac{1}{\beta_{11}}\epsilon_1(s_i\sum_{m=1}^{k}\beta_{m1} - \beta_{i1})] \\ &= \Delta_i + [\epsilon_i + \frac{1}{\beta_{11}}\epsilon_1[(\frac{\Delta_i + \epsilon_i - \frac{\beta_{i1}}{\beta_{11}}\epsilon_1}{\tilde{D}})\sum_{m=1}^{k}\beta_{m1} - \beta_{i1}]] \\ &\equiv \Delta_i + \tilde{\epsilon}_i \qquad i = 2, ..., k \end{aligned} \tag{17}$$

where

$$\Delta_i = \alpha_i + \beta_{i1}\mu_1 + \sum_{n=2}^{k}\beta_{in}v_n \tag{18}$$

by substituting in equation (**??**) for s_i on the right-hand side of (**??**).

Proceeding according to the same logic as in the linear case, the data theoretically could be augmented by drawing ϵ_1 conditional on $\tilde{\epsilon}_i$, $i = 2, ..., k-1$, where $\tilde{\epsilon}_i$ is defined by equation (**??**). While the mean and variance of $\tilde{\epsilon}_i$ are computable, obtaining its distribution and the conditional distribution $\epsilon_1|\tilde{\epsilon}_i$ is quite complex. As a result, ϵ_1 is drawn unconditionally. Once ϵ_1 is drawn, equation (**??**) is used to solve for ϵ_i, $i = 2, ..., k-1$, given ϵ_1, α,and β as well as the observed data. After the full set of residuals has been computed, the regime conditions must be checked. In the non-linear case, the regime conditions imply that $\pi_1 < p_1$ and $\frac{D}{\tilde{D}}(s_i^* - \frac{\beta_{i1}}{\beta_{11}}s_1^*) \in (0, 1)$,

$i = 2, ..., k-1$. The latter conditions differ slightly from the regime conditions in the linear case since in the linear case $\frac{D}{D}$ reduces to one. If any condition is violated, ϵ_1 must be re-drawn.

Case 2. $s_k > 0$**,** $s_i = 0 \;\; \forall i \neq k$**.** When only 1 good is consumed, the only change from the linear case is the equation for the latent shares. Since D is not restricted *a priori* to be one, once the ϵ's are drawn, the formula for obtaining the latent shares has changed (although this is trivial computationally). The rest of the augmentation procedure is unchanged from the linear case, including the regime conditions.

Case 3. $s_i > 0 \;\; \forall i$**.** As before, there are no binding non-negativity constraints and the observed and latent budget shares are identical.

4.2 Remainder of the Algorithm

4.2.1 Linear Translog Demand System

Once every observation has a full set of latent cost shares, an updated estimate of Σ is obtained first by computing the residuals conditional on the new set of latent shares and current estimates of α and β, and then drawing a new $\Sigma \sim$ Inverse Wishart$(\hat{\Sigma})$ as in (ii), where $\hat{\Sigma}$ is the estimated covariance matrix from these residuals. As shown in (iii), once a new estimate of the covariance matrix has been obtained, α and β are estimated by SUR conditonal on Σ and the latent shares. The values of α and β are then updated by drawing

$$\alpha, \beta \sim N(\begin{bmatrix} \hat{\alpha} \\ \hat{\beta} \end{bmatrix}, \Sigma_{\alpha\beta}) \quad , \tag{19}$$

where $\hat{\alpha}, \hat{\beta}$ are the SUR estimates and $\Sigma_{\alpha\beta}$ is the covariance matrix of the estimated parameters. After drawing new estimates for α and β, coherency is checked at each data point. If coherency is violated at any point, new values are drawn for α and β. This procedure is then cycled through until the iteration of the algorithm converges to the desired density.

4.2.2 Non-Linear Translog Demand System

Once every observation has a full set of latent budget shares, an updated estimate of Σ is constructed as in the linear case. Once a new estimate of the covariance matrix has been obtained, α and β are estimated using non-linear SUR conditonal on Σ and the latent shares.

To obtain new estimates of α and β, rather than minimizing the sum of the squared residuals (conditional on the latent shares), we minimize the sum of $\frac{1}{D^2}\epsilon'\epsilon$ over all observations. Because D^2 is itself a function of a subset of the parameters being estimated (i.e., β), an iterative procedure is employed *within* each loop of the Gibbs Sampler, where $D^2(\beta)$ is computed, estimates of α and β are computed conditional on the prior $D^2(\beta)$, and then $D^2(\beta)$ is updated given the new estimates of the parameters. This is done until the estimates of α and β converge. After this iterative procedure is complete, the covariance matrix is computed for $\hat{\alpha}$ and $\hat{\beta}$ and final values for α and β are drawn from the appropriate multivariate normal distribution. After drawing new estimates for α and β, coherency is checked at each data point. If coherency is violated at any point, new values are drawn for α and β.

5 Monte Carlo Results

5.1 Linear Translog Demand System

To test the algorithm, 1000 data sets were generated, each with 500 observations and three goods, from a homogeneous cost function. Log prices were randomly drawn from a uniform distrbution (with bounds zero and two) and income was drawn from a normal distribution (with mean zero and a standard deviation of 1.25). The model was then estimated for each data set according to the procedure outlined, running 501 loops of the Gibbs Sampler and discarding the initial 500 sets of parameter estimates. This was done twice, once for 1000 data sets with relatively small variances and once for 1000 data sets with relatively large variances. The higher variances lead to many more corners in the observed data. For each cost function, this procedure was performed twice, once imposing global coherency and once imposing only local coherency. The results are presented in Tables 1 and 2. In addition to the median and standard deviation for each parameter, the average number of coherency rejections per loop (counting only those loops which were kept) is reported as well.

TABLE 1. Monte Carlo Results: Linear Demand System (3 Goods) with Small Variances.

PARAMETERS	TRUE VALUE	LOCAL COHERENCY		GLOBAL COHERENCY	
		MEDIAN	STANDARD DEVIATION	MEDIAN	STANDARD DEVIATION
α_1	0.200	0.2035	0.0127	0.2013	0.0130
α_2	0.300	0.2989	0.0129	0.3001	0.0119
β_{11}	-0.020	-0.0313	0.0258	-0.0317	0.0232
β_{12}	0.007	0.0112	0.0195	0.0117	0.0181
β_{22}	-0.025	-0.0315	0.0217	-0.0325	0.0210
σ_1	0.200	0.1965	0.0111	0.1975	0.0108
σ_2	0.150	0.1500	0.0092	0.1507	0.0094
ρ_{12}	-0.500	-0.4950	0.0907	-0.5110	0.0831
Ave. Regularity Rejections Per Loop†		3.236		5.142	
Number of Loops†		1,000		1,000	

†Figures include only those loops which were not discarded.

Table 1 reports the results of the 1,000 draws from the actual density as well as the true parameter values used to generate the data. Columns 3 - 4 report the results imposing coherency locally, while columns 5 - 6 report the results imposing global coherency. Local coherency is imposed by checking that condition (**??**) holds for each draw of α and β. Global coherency, on the other hand, requires that each draw of β is a negative semi-definite matrix. It should be noted that the true β matrix is negative semi-definite. Consequently, one would expect the parameter estimates in the column 5 to be closer on average to the true values.

Comparing the results across the two models, one finds that both models predict the true values extremely well. In addition, one finds little evidence of incoherency, although there is a small difference in the mean number of rejections per loop between the local and global models. Given the lack of a substantial number of non-consumed goods, the fact that coherency is an issue at all is perhaps surprising.

TABLE 2. Monte Carlo Results: Linear Demand System (3 Goods) with Larger Variances.

PARAMETERS	TRUE VALUE	LOCAL COHERENCY		GLOBAL COHERENCY	
		MEDIAN	STANDARD DEVIATION	MEDIAN	STANDARD DEVIATION
α_1	0.200	0.2360	0.0730	0.2403	0.0130
α_2	0.300	0.2868	0.0792	0.2838	0.0119
β_{11}	-0.020	-0.0870	0.0824	-0.0827	0.0232
β_{12}	0.007	0.0358	0.0607	0.0362	0.0181
β_{22}	-0.025	-0.0717	0.0637	-0.0742	0.0210
σ_1	0.600	0.6014	0.1024	0.5960	0.0108
σ_2	0.450	0.4618	0.0623	0.4641	0.0094
ρ_{12}	-0.500	-0.5773	0.2309	-0.5803	0.0831
Ave. Regularity Rejections Per Loop†		41.119		51.357	
Number of Loops†		1,000		1,000	

†Figures include only those loops which were not discarded.

Table 2 reports the results obtained using the data generated from the linear model but with relatively large variances for the stochastic elements. The larger variances substantially increase the prevelance of non-consumed goods. Consequently, the number of coherency rejections rises in both models. When coherency is imposed globally, there were on average nearly 51.4 rejections per loop of the Gibbs Sampler (counting only those loops after which the Gibbs Sampler converges to the appropriate density). Even when coherency is only checked at each data point, there is still over 41.1 rejections per loop on average. However, the fact that there are fewer rejections in the local model illustrates the expansion of the acceptable parameter space (and additional flexibility of the model) which is gained with the MCMC technique. Unlike in the smaller variance case, however, the parameter estimates are not as accurate; although there is little difference between the local and global coherency models.

TABLE 3. Monte Carlo Results: Linear Demand System (6 Goods).

PARAMETERS	TRUE VALUE	LOCAL COHERENCY		GLOBAL COHERENCY	
		MEDIAN	STANDARD DEVIATION	MEDIAN	STANDARD DEVIATION
α_1	0.100	0.0800	0.0291	0.0790	0.0292
α_2	0.200	0.2170	0.0257	0.2160	0.0245
α_3	0.050	0.0410	0.0243	0.0395	0.0244
α_4	0.150	0.1490	0.0296	0.1480	0.0317
α_5	0.070	0.0760	0.0279	0.0790	0.0262
β_{11}	-0.060	-0.0840	0.0134	-0.0835	0.0131
β_{12}	0.010	0.0100	0.0097	0.0100	0.0098
β_{13}	0.020	0.0300	0.0010	0.0290	0.0098
β_{14}	0.015	0.0220	0.0117	0.0210	0.0118
β_{15}	0.020	0.0280	0.0109	0.0280	0.0111
β_{22}	-0.080	-0.0890	0.0114	-0.0890	0.0114
β_{23}	0.030	0.0365	0.0097	0.0370	0.0010
β_{24}	0.010	0.0120	0.0112	0.0100	0.0113
β_{25}	0.015	0.0170	0.0117	0.0170	0.0119
β_{33}	-0.070	-0.0870	0.0124	-0.0870	0.0129
β_{34}	0.010	0.0100	0.0120	0.0120	0.0122
β_{35}	0.005	0.0040	0.0122	0.0030	0.0125
β_{44}	-0.100	-0.1200	0.0154	-0.1195	0.0153
β_{45}	0.025	0.0300	0.0117	0.0310	0.0114
β_{55}	-0.090	-0.1040	0.0135	-0.1040	0.0134
Ave. Regularity Rejections Per Loop†		1.300		5.477	
Number of Loops†		1,000		1,000	

†Figures include only those loops which were not discarded.

To illustrate the usefulness of the model in estimating demand systems with more than three goods, Table 3 presents the results from a six good linear translog demand demand system. As before, coherency is checked both locally and globally, with the model imposing coherency globally rejecting over four times as often per loop.

TABLE 3 (cont.). Monte Carlo Results: Linear Demand System (6 Goods).

PARAMETERS	TRUE VALUE	LOCAL COHERENCY		GLOBAL COHERENCY	
		MEDIAN	STANDARD DEVIATION	MEDIAN	STANDARD DEVIATION
σ_1	0.170	0.2420	0.0394	0.2400	0.0381
σ_2	0.240	0.2640	0.0228	0.2610	0.0221
σ_3	0.180	0.2240	0.0276	0.2220	0.0275
σ_4	0.250	0.3000	0.0320	0.3020	0.0327
σ_5	0.210	0.2440	0.0275	0.2420	0.0266
ρ_{12}	0.500	0.3395	0.1160	0.3270	0.1197
ρ_{13}	0.050	-0.1130	0.1451	-0.1075	0.1473
ρ_{14}	-0.200	-0.2490	0.1467	-0.2255	0.1470
ρ_{15}	0.120	-0.1440	0.1490	-0.1290	0.1504
ρ_{23}	0.000	-0.1580	0.1329	-0.1615	0.1262
ρ_{24}	-0.270	-0.2620	0.1191	-0.2400	0.1224
ρ_{25}	-0.620	-0.5070	0.0989	-0.5055	0.0972
ρ_{34}	0.180	0.0970	0.1368	0.1025	0.1429
ρ_{35}	-0.120	-0.0810	0.1554	-0.0740	0.1530
ρ_{45}	0.330	0.1785	0.1733	0.1620	0.1274
Ave. Regularity Rejections Per Loop†		1.300		5.477	
Number of Loops†		1,000		1,000	

†Figures include only those loops which were not discarded.

5.2 Non-Linear Translog Demand System

The same procedure as detailed in the previous section is used to test the non-linear model. The results are presented in Tables 4 and 5. Unlike in the linear model, however, additional conditions for rejection sampling are imposed in the non-linear model for purely computational reasons. First, in some cases, given the values of the parameters for α and β from the previous iteration of the Gibbs Sampler, the data could not be augmented such that the appropriate regime conditions were satisfied. Because the observed shares are augmented by drawing the stochastic elements ϵ from a (multivariate) normal distribution, with infinite support, augmenting the data is always feasible given a suitably large number of draws. However, given the limitations of the random number generator, extremely low probability events are never sampled.[1] As a result, if after 500,000

[1]For example, 10 million draws from a standard normal were all between 0 and, approximately, 5.5 in absolute value.

attempts at augmenting a particular observation, a suitable ϵ vector could not be found, the results of the previous loop of the Gibbs Sampler are discarded and the previous loop is re-done so that the iterative process may proceed. For example, if given the parameter estimates from iteration 400, appropriate latent data cannot be simulated, the results from loop 400 are discarded and iteration 400 is re-computed given the parameter values from iteration 399. If the problem occurred during the first loop of the Gibbs Sampler, the initial values are altered slightly and then the process is re-started. The limit of 500,000 was chosen because after several tests of the random number generator within the context of the non-linear model, it was found that if the data could not be augmented within the initial 500,000 draws, the probability of simulating suitable latent data with further draws was negligable.

Second, because an iterative procedure is used to perform non-linear SUR estimation within each iteration of the Gibbs Sampler, convergence of the SUR estimates is an issue. If after a set number of iterations and changes in maximization algorithms and initial values the process had not converged, the same action was taken as detailed above; namely, the Gibbs Sampler is "backed up" a loop and then allowed to proceed. This potentially introduces a bias into the estimation if the non-convergence of the SUR estimates is linked to the particular observed data. However, as is reported below, non-convergence of the SUR estimates is quite rare.

A final source of rejection pertains to the coherency criterion. Again, because the β's are drawn from a multivariate normal distribution with infinite support, as long as the space of coherent parameters is non-empty, it is always possible to draw a set of coherent parameter values. However, computationally, this is not the case. As a result, if after 500,000 attempts a coherent β matrix could not be drawn, the Gibbs Sampler was backed up an iteration and allowed to proceed.

TABLE 4. Monte Carlo Results: Non-linear Demand System (3 Goods) with Small Variances.

PARAMETERS	TRUE VALUE	LOCAL COHERENCY		GLOBAL COHERENCY	
		MEDIAN	STANDARD DEVIATION	MEDIAN	STANDARD DEVIATION
α_1	0.050	0.0495	0.0043	0.0498	0.0044
α_2	0.150	0.1505	0.0061	0.1481	0.0060
β_{11}	-0.050	-0.0501	0.0040	-0.0511	0.0041
β_{12}	0.030	0.0303	0.0038	0.0297	0.0038
β_{13}	0.010	0.0098	0.0037	0.0091	0.0034
β_{22}	-0.060	-0.0602	0.0063	-0.0639	0.0056
β_{23}	0.035	0.0357	0.0058	0.0312	0.0043
β_{33}	-0.040	-0.0366	0.0301	-0.0643	0.0170
σ_1	0.060	0.0606	0.0034	0.0616	0.0036
σ_2	0.090	0.0902	0.0046	0.0910	0.0044
ρ_{12}	-0.500	-0.5007	0.0505	-0.5004	0.0517
Ave. Regularity Rejections Per Loop†		0.002		477.681	
Number of Loops†		1,000		1,000	

†Figures include only those loops which were not discarded.

Table 4 presents the results from the non-linear model with small variances. As in the previous tables, column 2 gives the true values of the parameters, columns 3 - 4 contains the results when coherency is imposed locally, and columns 5 - 6 gives the results when global coherency is imposed on the parameter values. While the parameter estimates are virtually identical across the two models, imposing coherency globally significantly constrains the space of acceptable parameter values relative to the space of locally coherent values. Consequently, on average one rejected over 477 parameter draws per iteration of the Gibbs Sampler when imposing global coherency, while rarely rejecting any parameter draws per loop of the Gibbs Sampler when imposing coherency locally.

In terms of the additional conditions checked within each loop of the Gibbs Sampler, seven loops were discarded in the local coherency model because a coherent β matrix was not found after 500,000 draws; 911 iterations were discarded in the model imposing coherency globally.[2] In addition, in the local coherency model, one iteration was thrown out because the non-linear SUR

[2]Note, because these loops are discarded, the 500,000 rejections do not figure in the averages reported in the tables.

estimates failed to converge and 156 loops were discarded because of an inability to augment the observed data after 500,000 attempts. In the global coherency model, 609 loops were discarded due to an inability to augment the data, but the non-linear SUR estimates always converged.

TABLE 5. Monte Carlo Results: Non-linear Demand System (3 Goods) with Large Variances.

PARAMETERS	TRUE VALUE	LOCAL COHERENCY		GLOBAL COHERENCY	
		MEDIAN	STANDARD DEVIATION	MEDIAN	STANDARD DEVIATION
α_1	0.050	0.0534	0.0056	0.0495	0.0055
α_2	0.150	0.1498	0.0081	0.1477	0.0080
β_{11}	-0.050	-0.0502	0.0054	-0.0520	0.0052
β_{12}	0.030	0.0303	0.0052	0.0303	0.0051
β_{13}	0.010	0.0099	0.0050	0.0089	0.0047
β_{22}	-0.060	-0.0611	0.0086	-0.0656	0.0076
β_{23}	0.035	0.0356	0.0075	0.0307	0.0061
β_{33}	-0.040	-0.0378	0.0408	-0.0683	0.0243
σ_1	0.080	0.0812	0.0052	0.0830	0.0052
σ_2	0.120	0.1213	0.0068	0.1226	0.0069
ρ_{12}	-0.500	-0.5011	0.0555	-0.4985	0.0560
Ave. Regularity Rejections Per Loop†		0.031		356.715	
Number of Loops†		1,000		1,000	

†Figures include only those loops which were not discarded.

Table 5 presents the results from the model with slightly larger variances. While the values chosen for the variances are larger than one might find in actual household or individual data, the variances are significantly smaller than the values chosen in the linear model. When the non-linear model was tested with values larger than those actually used, the model became computationally burdensome.

The results are very similar across the two models, however the average number of coherency rejections per iteration of the Gibbs Sampler differs dramatically across the two models. The model imposing coherency globally rejected parameter draws on average over 350 times more often than in the local coherency model. In addition, nearly 1600 loops of the Gibbs Sampler were discarded in the global coherency model because coherent parameter values could not be simulated within the 500,000 draw limit. In the model imposing coherency locally, only 46 loops were thrown out.

As for the remaining conditions checked to ensure that the Gibbs Sampler could always proceed, ten loops were discarded in the local coherency model because the non-linear SUR estimates failed to converge; eight in the global coherency model. Ten iterations had to be thrown out in the local coherency model because of an inability to augment the data; five in the global coherency model.

6 Conclusion

In the estimation of demand systems involving data at the individual, household, or firm level, the issue of how to handle the existance of many non-consumed goods has yet to be resolved empirically. Two problems arise in the estimation of such models. First, the model is essentially an endogenous switching regimes model, involving the evaluation of high-dimensional probability integrals which can be quite burdensome computationally. Second, the problem of statistical coherency – where the sum of the probabilities of the various demand regimes is not one – must be addressed. The curvature condition imposed by demand theory, quasi-concavity of the cost or utility function, is a sufficient condition for coherency.

The use of the Gibbs Sampling algorithm, along with data augmentation, solves both of these issues. Given the simple latent structure of the demand system, augmenting the data removes the need to directly evaluate any probability integrals. In addition, use of rejection sampling within each iteration of the Gibbs Sampler allows one to impose only the minimum restrictions necessary to guarantee coherency: coherency at each data point, *not* global coherency. Previous attempts to impose coherency in models with corners have not been successful. Either more restrictive functional forms have been utilized, such that coherency is guaranteed globally, or global coherency has been imposed on more flexible functional forms, destroying their flexibility. Attempts to impose concavity restriction locally have also been "unsatisfactory", according to Diewert and Wales (1987). A recent exception has been the work of Terrell (1996), who employs Gibbs Sampling to the estimation of demand systems with only interior solutions.

The Gibbs Sampling algorithm along with the data augmentation technique is used in this paper to estimate translog indirect utility and cost functions with simulated data. The results confirm not only the accuracy of the Gibbs Sampler estimates, but also the importance of addressing the problem of coherency. In all of the models estimated, the parameter vector frequently entered the space of incoherent values and rejecting on the basis of global coherency significantly restricts the acceptable parameter space relative to the model imposing coherency locally. Of course, the emprical importance of imposing local coherency requires estimation of demand systems using real firm or individual data. It is in this direction, we plan on moving in the future.

References

[1] Chib, S. and E. Greenberg. 1996. "Markov Chain Monte Carlo Simulation Methods in Econometrics." *Econometric Theory.*

[2] Christensen, L., Jorgensen, D., and L.J. Lau. 1975. "Transcendental Logarithmic Utility Functions." *American Economic Review.* vol. 53.

[3] Diewert, W.E. and T.J. Wales. 1987. "Flexible Functional Forms and Global Curvature Conditions." *Econometrica.* vol. 55, no. 1.

[4] Gallant, A.R. and E.H. Golub. 1984. "Imposing Curvature Restrictions on Flexible Functional Forms." *Journal of Econometrics.* vol. 26.

[5] Hausman, J. 1985. "The Econometrics of Nonlinear Budget Sets." *Econometrica.* vol. 53.

[6] Lancaster, Tony. 1997. "Exact Structural Inference in Optimal Job-Search Models." *Journal of Business and Economic Statistics.* vol. 15, no. 2.

[7] Lau, L.J. 1978. "Testing and Imposing Monotonicity, Convexity, and Quasi-Convexity Constraints" in *Production Economics: A Dual Approach to Theory and Applications.* vol. 1. Ed. by M. Fuss and D. McFadden. Amsterdam: North-Holland.

[8] Lee, L.F. "Simulation Estimation of Dynamic Switching Regression and Dynamic Disequilibrium Models – Some Monte Carlo Results." *Journal of Econometrics.* vol. 78, no. 2.

[9] Lee, L.F. and M.M. Pitt. 1986. "Microeconometric Demand Systems with Binding Nonneagtivity Constraints: The Dual Approach." *Econometrica.* vol. 54, no. 5.

[10] Lee, L.F. and M.M. Pitt. 1987. "Microeconometric Models of Rationing, Imperfect Markets, and Non-Negativity Constraints." *Journal of Econometrics.* vol. 36.

[11] MaCurdy, T., Green, D., and H. Paarsch. 1990. "Assessing Empirical Approaches for Analyzing Taxes and Labor Supply." *Journal of Human Resources.* vol. 25.

[12] McCulloch, R. and P.E. Rossi. 1994. "An Exact Likelihood Analysis of the Multinomial Probit Model." *Journal of Econometrics.* vol. 64.

[13] Ransom, M.R. 1987. "A Comment on Consumer Demand Systems with Binding Non-Negativity Constraints." *Journal of Econometrics.* vol. 34.

[14] Terrell, D. 1996. "Incorporating Monotonicity and Concavity Conditions in Flexible Functional Forms." *Journal of Applied Econometrics.* vol. 11.

[15] van Soest, A., Kapteyn, A. and P. Kooreman. 1993. "Coherency and Regularity of Demand Systems with Equality and Inequality Constraints." *Journal of Econometrics.* vol. 57.

[16] van Soest, A. and P. Kooreman. 1990. "Coherency of the Indirect Translog Demand System with Binding Nonnegativity Constraints." *Journal of Econometrics.* vol. 44.

[17] Wales, T. and A. Woodland. 1983. "Estimation of Consumer Demand Equations with Binding Non-Negativity Constraints." *Journal of Econometrics.* vol. 21.