Bayesian inference of a smooth transition dynamic almost ideal model of food demand in the US

Balcombe, Kelvin and Bailey, Alastair

Imperial College at Wye

2006

Bayesian inference of a smooth transition dynamic almost ideal model of food demand in the US.

Kelvin Balcombe and Alastair Bailey
Imperial College London, Wye Campus.

Summary

A dynamic 'smooth transition' Almost Ideal model is estimated for food consumption in the US. A Metropolis-Hastings algorithm is employed to map the posterior distributions and rejection sampling is used to evaluate and impose curvature restrictions at more than one point in the sample. The findings support the contention of structural change of a 'smooth transition' nature. Notably, the income food elasticity of demand becomes smaller through time, and the own price elasticities for food and non food become more elastic.

1. Introduction

The estimation of Dynamic 'Almost Ideal' demand Systems has been the subject of extensive research (Anderson and Blundell 1984, .Ng 1995; Attfield, 1997, Chambers and Nowman, 1997, Duffy, 2002, Pesaran and Shin, 2002, Balcombe, 2003). Recently, Dechamps (2000, 2003) employed a Bayesian approach to estimating dynamic demand and Griffiths et al. (2000) employed a Bayesian when estimating demand for inputs within agriculture. This paper also considers a Bayesian approach to the estimation of the dynamic demand for US food from 1966 to 2000 that is of the 'Almost Ideal' (Deaton and Meullbauer, 1980) type. However, it extends the existing applied literature by estimating a smooth transition model. This is facilitated using a Metropolis-Hastings (M-H) algorithm (see Albert and Chib 1993; and, Chib and Greenberg 1995a,1995b). Previous work in estimating demand systems along Bayesian line also include Chalfant et al. (1991), and Tiffin and Tiffin (1999) which applied the 'importance sampling' approach outlined in Geweke (1988,1989). The M-H shares common principles with the importance sampling approaches, but is more general.

While consumption models using simple log linear forms continue to be estimated (De Crombrugghe et al., 1997), within the agricultural economics more emphasis has been placed on the microeconomic foundations of demand models along with the power of 'flexible functional forms' to conform to theoretical requirements. Arguably, the operational value of microeconomic foundations are eroded when they are set in a static framework, yet applied in a dynamic setting. Moreover, their application requires strong assumptions concerning aggregation across commodities and individuals. Nevertheless, in common with some other recent papers (e.g. Pesaran and Shin, 2002) we take the view that models should embody basic properties such as homogeneity, symmetry and curvature where the parameters of interest characterize 'long-run' equilibrium relationships. It is these restrictions which enable the parameters to be interpreted as structural demand parameters, rather than quantities which merely summarize the statistical relationships between quantity and price and income. Admittedly, restrictions such as symmetry and homogeneity have been commonly rejected. However, it is also recognized that there are a variety reasons why this might occur. Buse (1994) highlighted problems regarding both misspecification of the AIDs linear approximation, although much of the work in this area continues to use the linear approximation. More generally, it is recognized that most models (including the full AIDs), cannot impose the relevant restrictions globally, or if they can the imposition of these restrictions impose unreasonable compromises on the 'flexibility' of the system. Buse (1998) also outlines problems with divergence between asymptotic and finite sample results leading to the rejection of restrictions such as symmetry and homogeneity, and the classical literature has recently explored this question in some depth (e.g. Balcombe, 2003) from a 'cointegration' perspective.

Unlike the classical approach, the Bayesian approach does not have a 'discontinuity' in its theory regarding the treatment of unit root processes compared with those that are stationary. Moreover, it does not rely as heavily on 'asymptotics' for its justification. However, to many, the main appeal of a Bayesian approach will lie in the ease with which basic theoretical requirements of the inequality type (such as curvature restrictions) can be evaluated and (locally) enforced explicitly and transparently. Classical methods can also be employed to enforce these type of conditions (Lau, 1978). However, unlike the Classical approach, the Bayesian approach generates point estimates that will lie very close to the inequality boundary only when the sample data is highly unsupportive of a restriction. The Bayesian approach is more in tune with the fact that when we set a boundary, this usually reflects our belief that the estimate lies somewhere inside that boundary. While some classical practitioners will cling to the view that Bayesian approaches are tainted by the use of subjective priors. The imposition of virtually any type of restriction reflects prior beliefs regardless of whether one is employing a supposedly Classical or Bayesian methodology. The priors used in this paper are informative only where they are required to identify the structural parameters of interest.

A major leap forward in the implementation of the Bayesian methodology within the last decade has been through the implementation of Gibbs Sampling and/or M-H algorithms (again, readers are referred to Chib and Greenberg, 1995b). These algorithms have been used ostensibly in the Bayesian literature, although they are not exclusively Bayesian in nature. While they appear complex at first, their underlying simplicity enables models to be estimated in a way that solves some of the theoretical and practical difficulties in deriving explicit or numerically simulated posterior distributions. Using the M-H approach, parameters can be sampled from their posterior distributions, without knowing the exact form of that posterior distribution (i.e. Normal, Student t, Inverse Gamma, etc.). Moreover, inequalities may be enforced by building in 'rejection' steps into the algorithms.

In a study using data over 35 years, such as this one, the preferences of consumers are likely to have changed. Therefore, the estimation of demand over time may require not only the modelling of 'dynamics' using lag structures, but change in the parameters also. This may require anything from the introduction of dummy variables, to more complex solutions such as models that allow for random walk parameters (Morrison et al. 2003, Dechamps, 2003). An attractive alternative is to model the parameters using a 'smooth transition' in the parameters and this is the approach employed herein. There are a wide variety of smooth transition models, as outlined in Terasvirta (1994). The type used in this paper are fairly limited, with the transition being a deterministic function of time. A smooth transition model allows the parameters to change in a logistic fashion. It is flexible enough to allow for large sudden changes in the parameters, as well as the slow evolution in the parameters in an almost linear fashion, and can be easily implemented using a Bayesian approach.

The estimation of an AIDS model using time series also raises issues regarding the way that elasticities are calculated, and curvature restrictions are enforced. A general practice has been to calculate elasticities at a mean point in the sample. However, there are compelling arguments against this practice. Griffiths et al. (2000) favor imposing curvature at all points (though in a production context). The alternative which is employed herein, calculates the elasticities at the beginning and end of the sample and curvature is imposed at both these points. However, curvature effectively holds at most or all points within the sample when taking this approach.

We proceed by briefly outlining the AIDS model, and the derivation of elasticities. Section 3 examines how the dynamic AIDS model can be specified within the smooth transition autoregressive distributed lag model. Section 4 briefly covers the use of the Metropolis Hastings Algorithms. Section 5 presents and discusses the empirical results, and Section 6 concludes.

2. The Almost Ideal Demand System

The 'Almost Ideal Demand System' (AIDS) was developed by Deaton and Meullbauer (1980). Allowing for the parameters to change with time, the Almost

Ideal form (ignoring residuals) can be expressed as:

$$s_t = \alpha_t + A_t p_t + \beta_t \Lambda_t (m_t, p_t) \tag{0.1}$$

where s_t is the vector of expenditure shares p_t is a vector of logged prices, $\Lambda_t (m_t, p_t)$ is 'real expenditure'

$$\Lambda_t (m_t, p_t) = m_t - \left(\alpha_t' p_t + \frac{1}{2} p_t' A_t p_t \right) \tag{0.2}$$

where m_t is logged total expenditure. Although the system in [0.1] is dynamic in the sense that the parameters are time dependent, it is not dynamic in the sense that expenditures depend on past prices or expenditures. In this article [0.1] is treated here as an depiction of the (time dependent) equilibrium relationship, from which the parameters α_t, A_t, β_t are calculated as long-run multipliers. This will be discussed in more depth in Section 3. The adding up restrictions require that the column sums of α_t is equal to one and the column sums of A_t and β_t are equal to zero. In addition, theory suggests that the row sums of A_t are zero (price homogeneity), and that A_t is symmetric. Given the adding up restrictions, homogeneity and symmetry are equivalent in dual good system such as the one which is employed here.

2.1 Curvature and Elasticities

The AIDs share equations can be derived as the derivative of a logged indirect expenditure function. The Hessian matrix of the logged cost function should have eigenvalues which are non-positive (e.g. the Hessian is semi negative definite). This restriction cannot be enforced globally within the AIDS model (e.g. for all prices and incomes). Therefore, it is common practice to calculate the elasticities and examine the Hessian matrix around a mean point in the data. This practice has a compelling rationale when the price and expenditure data are stationary, but is less defensible when the data contain trends. Few papers have highlighted this as an issue. Exceptions are Griffiths et al. (2000), who impose curvature in input demand models at multiple points in the sample. The same concerns arise with regard to the results of many AIDS models in the existing literature. While they may yield plausible estimates at their 'mean point' they would yield implausible elasticities were the elasticities calculated at different points within the sample. Consequently, while it does not represent a solution to the problem that the AIDS model does not have all its theoretical properties globally, it is worthwhile examining the elasticities at more than one point in the data set, and to impose curvature restrictions at these points if possible. This view is given added weight in circumstances where the parameters are treated as variable, as in this paper.

Therefore, this paper therefore adopts an alternative approach whereby elasticities are calculated both that the beginning and end of the sample, using the parameters and variable values at those points. Moreover, the Bayesian rejection sampling approach is employed to impose curvature at these points. Defining the matrix $\frac{\partial s_t}{\partial p_t'}|_m = V_t$, it can be observed that:

$$\frac{\partial s_t}{\partial p_t'}|_m = V_t = A_t + \beta_t \beta_t' \Lambda_t \left(m_t, p_t\right) - \beta_t s_t' \tag{0.3}$$

The Uncompensated Price Elasticities are (using the notation $\delta\left(s_t\right)$ to denote the diagonalised matrix with the elements of s_t constituting its diagonal elements):

$$U_t = \delta\left(s_t\right)^{-1} V_t - I \tag{0.4}$$

The Expenditure Elasticities are:

$$\eta_t = \delta\left(s_t\right)^{-1} \beta_t + \mathbf{1}_k \tag{0.5}$$

where $\mathbf{1}$ is a conformable column vector of ones. Using the Slutsky decomposition, the Compensated Price Elasticities are:

$$\xi_t = U_t + \eta_t s_t' = \delta\left(s_t\right)^{-1} V_t - I + \delta\left(s_t\right)^{-1} \beta_t s_t' + \mathbf{1} s_t' \tag{0.6}$$

If the share equation is derived from an indirect cost function, then the matrix

$$K_t = \delta\left(s_t\right)\left(\xi_t\right) = A_t + \beta_t \beta_t' \Lambda_t \left(m_t, p_t\right) - \delta\left(s_t\right) + s_t s_t' \tag{0.7}$$

must have non-positive eigenvalues if the curvature restrictions are obeyed at the point t.

As argued above, there is a compelling case for this enforcing curvature at all points t=1,.... T. However, if the parameters are allowed to evolve in a smooth transition manner, then enforcing the restrictions at the beginning, and end, of the sample (using $K_1$and K_T respectively) is likely to give results that broadly conform to curvature requirements throughout the sample, while significantly decreasing the computational requirements when using the Bayesian methods that are discussed in Section 4. The eigenvalues can subsequently be evaluated at all points in the sample in order to ascertain whether they conformed to the curvature restrictions throughout the sample.

3. Dynamic Smooth Transition AIDS

This paper presents the results for a model estimated using annual data from 1966 to 2000. Both the characteristics of foods, and the tastes and habits of consumers are likely to have altered during this time. Thus, the underlying utility functions and consequentially the parameters of demand equations are not likely to be invariant throughout the sample. Moreover, as already noted, the constant parameter AIDS model itself is known to have good 'local properties' but cannot preserve these properties globally (at all prices and incomes). Allowing the parameters to evolve may therefore mitigate some of these shortcomings. The solution proposed and implemented here is to allow for changes in the parameter using a 'smooth transition regression' (STR) version of the AIDS. This approach is restrictive in the sense that the parameters must evolve monotonically. Random parameter models, such as those which allow for random walk parameters, are flexible in the sense that 'reversals' can take place, but are restrictive in other ways (i.e. cannot model large one-off shifts under a constant innovation variance). Smooth transition models allow the parameters to change in a smooth or sudden way without requiring too many additional parameters.

The variables used in this paper are the share of disposible income spent on food $s_{f,t}$, the logged food price $p_{f,t}$, logged non-food price $p_{n,t}$, real disposable income m_t and the AIDS price index $\Lambda(p_t)$ defined using the price vector $p_t = (p_{f,t}, p_{n,t})$. The construction of the price index is discussed below. Under homogeneity on the relative price is required $x_t = (p_{f,t} - p_{n,t})$, along with real income $z_t = \Lambda_t(m_t, p_t)$ (as in [0.2]). The main specification considered in this paper is the smooth transition form[1]:

$$\begin{aligned} s_{f,t} &= \delta_1(1-f_t) + \delta_T f_t + \lambda s_{f,t-1} + \\ &\quad + (1-f_t)\sum_{i=0}^{1}\theta_{1i}x_{t-i} + f_t\sum_{i=0}^{1}\theta_{T,i}x_{t-i} \\ &\quad + (1-f_t)\sum_{i=0}^{1}\pi_{1i}z_{t-i} + f_t\sum_{i=0}^{1}\pi_{T,i}z_{t-i} \\ &\quad + u_t \end{aligned} \tag{0.8}$$

where $u_t \overset{iid}{\sim} N(0,\sigma^2)$ and

$$f_t = F(t,\gamma) = c_1\left(\gamma_{1,}\gamma_{2,}\right)(\gamma_1 + \exp(-\gamma_2 t))^{-1} - c_2\left(\gamma_{1,}\gamma_{2,}\right) \tag{0.9}$$

[1] A second order model was intitally considered. However, a preliminary examination of the ADL equation without smooth transition suggested that second order lags were not at all significant (using Classical Tests).

where $c_1\left(\gamma_{1,}\gamma_{2,}\right)$ and $c_2\left(\gamma_{1,}\gamma_{2,}\right)$ are 'normalizing constants' calculated so that for any $\gamma=\left(\gamma_1,\gamma_{2,}\right)$, $F(1,\gamma)=0$ and $F(T,\gamma)=1$. The parameter γ_2 governs the speed of the transition, whereby as γ_2 goes towards infinity, the transition becomes a sharp break. As γ_2 goes to one, the transition becomes linear. The parameter $\gamma_{1,}$ allows for variation in the central point at which the transition takes place.

Since λ is held constant, transitions are therefore limited to the independent variable coefficients, rather than the lag coefficient for the share of food. While, in principle, it would be possible to allow for λ to be time dependent, treating this one coefficient as constant simplifies the computational difficulties considerably. This model can be viewed as a time dependent long-run equilibrium model, with the rate of adjustment towards this equilibrium being constant.

A reparameterisation of the autoregressive equation [0.8] into an error correction form gives two sets of the 'long-run' multipliers (under the condition that $\lambda\neq 1$) at the beginning of the sample (t=1) and at the end of the sample (t=T). The long-run multipliers are for:

- The Intercept:

$$\mu_1=\left(\frac{\delta_1}{1-\lambda}\right) \text{ and } \mu_T=\left(\frac{\delta_T}{1-\lambda}\right) \tag{0.10}$$

- Prices:

$$\phi_0=\left(\frac{\theta_{10}+\theta_{11}}{1-\lambda}\right) \text{ and } \phi_T=\left(\frac{\theta_{T0}+\theta_{T1}}{1-\lambda}\right) \tag{0.11}$$

- Real income:

$$\varphi_0=\left(\frac{\pi_{10}+\pi_{11}}{1-\lambda}\right) \text{ and } \varphi_T=\left(\frac{\pi_{T0}+\pi_{T1}}{1-\lambda}\right) \tag{0.12}$$

In each case the time dependent long-run multipliers can be constructed as:

$$\begin{aligned}\mu_t &= (1-f_t)\,\mu_1+f_t\mu_T \\ \phi_t &= (1-f_t)\,\phi_1+f_t\phi_T \\ \varphi_t &= (1-f_t)\,\varphi_1+f_t\varphi_T\end{aligned} \tag{0.13}$$

The time dependent long run equilibrium error can be computed as: $\varepsilon_t=s_{ft}-(\mu_t+\phi_t x_t+\varphi_t z_t)$. However, this requires the construction of $z_t=\Lambda_t(m_t,p_t)$. The approach employed in this paper constructs $\Lambda_t(m_t,p_t)$ from the long-run multipliers, using the adding up and homogeneity conditions. Therefore the parameters of [0.1] are constructed as:

$$\begin{aligned} \alpha_t' &= (\mu_t, 1-\mu_t) \\ A_t &= \begin{pmatrix} \phi_t & -\phi_t \\ -\phi_t & \phi_t \end{pmatrix} \\ \beta_t' &= (\varphi_t, -\varphi_t). \end{aligned} \tag{0.14}$$

Using these values the elasticities and other quantities in equations [0.4] to [0.7] can be computed, using the parameters, and s_t calculated at its long-run equilibrium $\hat{s}_{ft} = (\mu_t + \phi_t x_t + \varphi_t z_t)$ and $\hat{s}_{nt} = 1 - \hat{s}_{ft}$. The implied error correction form is:

$$\Delta s_{ft} = (\lambda - 1)\,\varepsilon_{t-1} + \theta_{1t}\Delta x_t + \pi_{1t}\Delta z_t + u_t \tag{0.15}$$

where $\theta_{1t} = \theta_{11}(1-f_t) + \theta_{1T} f_t$ and $\pi_{1t} = \pi_{11}(1-f_t) + \pi_{1T} f_t$. Therefore, the parameters which require estimation within this model are

$$\Omega = \left(\delta_1, \delta_T, \theta_{11}, \theta_{T1}, \pi_{1t}, \sigma^2, \gamma_1, \gamma_2\right). \tag{0.16}$$

An examination of evidence supporting the homogeneity and curvature restrictions is likely to be of interest to readers. This can be done in a variety of ways, since homogeneity can be evaluated with and without imposing curvature and vice versa. We evaluate the evidence for symmetry by augmenting the equation [0.8] with

$$(1-f_t)\sum_{i=0}^{1} \pi_{1i}^* p_{ft-i} + f_t \sum_{i=0}^{1} \pi_{T,i}^* p_{ft-i}. \tag{0.17}$$

From a classical perspective, $\pi_{10}^* + \pi_{11}^*$ and $\pi_{T0}^* + \pi_{T1}^*$ should be insignificantly different from zero if homogeneity holds,. However, from a Bayesian perspective, there should be a substantial proportion of the posterior mass for each of these quantities either side of zero. Likewise, by examining the posterior mass of the largest eigenvalues of [0.7], the evidence supporting curvature restrictions can be assessed. An estimate of the posterior mass can be made using the techniques outlined in the next section.

4. Estimation Using the Metropolis Hastings Algorithm

A full Bayesian analysis, using the M-H algorithm is relatively straight forward. Moreover, this approach naturally leads to the computation of posterior distributions for the elasticities, and not just for the parameters of the model. A full description of the M-H algorithms could not be done justice in this paper, and readers are referred to Chib and Greenberg (1995b), and Griffiths et al. (2000). However, their usefulness can be outlined in the following way.

Using the 'conditional or approximate approach' (Bauwens et al., 1999; p.135) the data likelihood is treated as normal:

$$L(\Omega) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^T \exp\left(-\frac{\sum_{t=1}^{T} u_t^2}{2\sigma^2}\right) \tag{0.18}$$

with the errors expressed in [0.8]. The classical solution would be to maximise the likelihood function. However, Bayesians would augment this with prior information. If no consideration was given to the ECM or transition parameters, then we could multiply the Likelihood by the 'reference prior'.

$$pr(\Omega) = \frac{1}{\sigma^2} \tag{0.19}$$

This prior is non-informative, in the sense that the prior information will play only a negligible part in the determination of the posterior distributions which are derived subsequently (in the sense that the data will dominate the priors). However, if the parameters of interest are the long-run multipliers [0.10 to 0.12], then these are not identified at $\lambda = 1$. Likewise, as the transition parameter γ_2 tends towards zero, then f_t becomes constant, therefore also inducing non-identification of the model. These issues are discussed in Bauwens *et. al.* (1999). There are at least two solutions to this problem. The first is to use alternative priors which are designed to give higher probabilities to the parameters when they are in the identified region. The second is to use 'rejection sampling' so that only parameters that are within an identified region are deemed acceptable. This can be operationalised within the M-H algorithm. The latter method is used in the example herein. Therefore, the posterior distribution becomes:

$$\begin{aligned} p(\Omega) &= pr(\Omega)\, L(\Omega) \ \text{ if } \Omega \in S_\Omega \\ &= 0 \text{ otherwise} \end{aligned} \tag{0.20}$$

where S_Ω contains only those parameters with $\lambda < .95, \gamma_1 > 0, \gamma_2 > 1$. At $\gamma_1 = 1$ the transition variable f_t becomes approximately linear, whereby as γ_1 increases the transition become 'logistic' and increasingly concentrated over a shorter space of time.

In line with the discussions regarding curvature in the preceding sections, the parameter space S_Ω should also be limited to those producing non-positive eigenvalues of the matrices K_1 and K_T. One type of M-H algorithm would then proceed from a starting point Ω_0 then, by generating parameters using random walk,

$$\Omega_{i+1} = \Omega_i + v_i \tag{0.21}$$

where v_i is a symmetrically distributed iid error term (perhaps normally distributed and with variances that are computed from the data). A step from Ω_i to Ω_{i+1} would be accepted with probability

$$\alpha = \frac{p\left(\Omega_{i+1}\right)}{p\left(\Omega_i\right)} \tag{0.22}$$

which can be performed by generating a uniformly distributed random variable U_i, with acceptance according to the rule $\Omega_{i+1} = \Omega_{i+1}$ if $U_i < \alpha$ and $\Omega_{i+1} = \Omega_i$ otherwise. The general principle behind these algorithms is that after a 'burn in' $i > I$, the values of Ω_{i+1} should be independent of their starting value Ω_0 and should behave as if they were being drawn from $p\left(.\right)$. The values will be highly dependent (obviously, since many of the values will be repeated). However, this is not overly problematic when generating a simulated distribution from a very large number of simulated values, providing the acceptance rate is not too small. Moreover, the dependence can be reduced by sampling every nth value from the simulated values. Care also needs to be taken to ensure that the 'burn in' is sufficiently large to ensure that any dependence on the starting values is negligible, and that the algorithm can be treated as if it is producing simulated values that are in accord with the underlying posterior distributions (convergence of the algorithm). Again, various suggestions are made in Bauwens et al. (1999). The efficiency of the algorithms can be improved by blocking (or grouping) the parameters, and taking steps for each of the parameter blocks sequentially. Likewise, the some of the parameters can be directly simulated from their conditional posterior distributions (such as the Inverse Gamma for the variance) in a manner similar to Gibbs sampling (for which readers are referred to Casalla and George, 1992).

4.1 Testing Restrictions

The literature on demand estimation has devoted considerable energy to the question of 'testing' for the underlying restrictions on the system. The classical approach to this problem is to examine the behavior of certain test statistics under a given pointwise null hypothesis. The Bayesian approach can examine the evidence against point wise restrictions by observing the proportion of the posterior mass that lies above and below certain values.

As covered in Section 3, by adding additional prices to [0.8], a significant proportion of the mass of the posterior distributions of parameters $\pi_{10}^* + \pi_{11}^*$ and

$\pi^*_{T0} + \pi^*_{T1}$ should lie both above and below zero if homogeneity holds. Therefore, this proportion can be calculated using the simulated values M-H algorithm. The closer this proportion is to .5, the less evidence has been accumulated against homogeneity. The curvature restrictions can be tested by accepting parameters which do not conform to the curvature restrictions in another run of the MH algorithm. The simulated maximum eigenvalue (calculated both at the beginning and end of the sample) can be recorded, and the proportion of positively occurring values calculated. The smaller the proportion of eigenvalues above zero, the less evidence has been accumulated against the curvature restrictions. These proportions are presented in Table 1 within Section 5. In each case, the proportion has been calculate having imposed the other restriction (thus the proportion of the posterior mass of the maximum eigenvalue has homogeneity imposed).

Although it is somewhat controversial, the Bayesian approach may have some advantages over the classical approach in this regard. The reason being the 'discontinuity' in the classical theory of estimation and inference in systems containing unit roots. Arguably, the Bayesian approach does not have such a discontinuity. We deal with this problem simply by enforcing the condition that λ is less than one (though we evaluate the evidence supporting this), and make no further analysis of the explanatory variables in our models. A good summary of the 'unit root' controversy is given in Chapter 6, Bauwens et al. (1999). Some of the same issues arise when estimating distributed lag models with potential unit roots in some of the variables. We do not wish stray into this argument any further within this paper since it seems to evoke rather strong perspectives from both Bayesian and Classical statisticians alike. However, we would draw readers attention to the fact that any superiority of the Bayesian approach in dealing with unit root processes is, we concede, disputable.

5 Empirical Section.

5.1 The Data

The data used in this study was taken from the Website of the Economic Research Service, U.S. Department of Agriculture. The data is annual from 1966 to 2000. The share of food is as a proportion of disposable income, and includes all food (consumed at home or away from home). The prices for food and non-food are taken directly from the data set, along with disposable income (per-capita). Strictly speaking, the AIDS framework uses 'total expenditure' rather than disposable income. However, the AIDS framework is static, and as such does not make any provision for saving. In view of this, it is not apparent that using a total expenditure figure is particularly advantageous in a time series setting and, in any case, the two logged series are likely to be closely related. The share of food is around 14.8% of disposable income in 1966, falling (monotonically) to around 10% in 2000. The relative price of food to non-food has generally risen over the period, and real disposable incomes per capita have increased over 70% during the period. Thus, in real terms the consumption of food has increased, even though it is a smaller proportion of the budget. While some of this may be explained by increases in consumption measured in calorie terms, it also suggests that there has been some switch to higher quality items, or at least more expensive food items.

5.2 Results

5.2.1. Convergence

The convergence of the M-H algorithm must be checked prior to analysis of the results. In this paper, the convergence was checked in three ways. First, sequential plots of the sampled parameters and elasticities were examined, along with the acceptance rates. Second, different starting points were used and the results compared. Third, Bauwens et al. suggest that if after N draws, the value of the CUSUM statistic on a given set of parameters lies within $\pm.05$ then the sampler can be considered to have converged after N.

There were several runs of the model, with and without restrictions having (homogeneity and curvature) been imposed. Generally, acceptance rates were extremely small, (around than 2%) consequently to reduce the dependence every 100th value generated by the sampler was recorded. Acceptance rates without homogeneity imposed were even smaller. It is possible that the sampler could have been calibrated to give superior results. However, using this process, the sampler took about an hour which was not unduly problematic. The sequential plots then appeared to fairly 'random' in the sense that they appeared to be consistent with being stationary around a given mean, and without having overly long stretches of repeated values. In each case, then providing a 'burn in' of 10000 was used (that is 10000 collected values with every 100th being sampled, thus 10^6 in total) then the mean and standard errors from the remaining 10000 sampled values were very similar. Plots for the parameters appeared to suggest that the burn in of 10000 was sufficient, although the CUSUM plots for some of the parameters were just inside the boundaries, particularly when homogeneity was not imposed

Turning first to Table 1, the first line gives the proportion of the posterior pertaining to the homogeneity restriction. Upon having approximately 50% of the posterior mass above zero one would conclude that there was little evidence against homogeneity. On the other hand having small values, or very large values would indicate that the data was inconsistent with the homogeneity restriction.

At the beginning of the sample (t=1) it is around 54%, suggesting that there is little evidence against homogeneity. At the end of the sample, the evidence against homogeneity is more pronounced, with a little less than 10% of the posterior mass above zero. The curvature restrictions give approximately the same proportion of the mass of the maximum eigenvalue above zero at both points in the sample. In each case the most of the mass of the maximum eigenvalues is below zero (supporting the curvature restrictions).

Table 1. Restrictions

	t=1	t=T
Homogeneity	$P(\pi^*_{10} + \pi^*_{11} > 0) = .54$	$P(\pi^*_{T0} + \pi^*_{T1} > 0) = .089$
Curvature	$P(MaxEig > 0) = .31$	$P(MaxEig > 0) = .32$

We may conclude from the results in Table 1, that the data is broadly consistent with the homogeneity and curvature restrictions, although the authors acknowledge that a case could be made for not imposing homogeneity given the relatively small proportion of the posterior mass of $\pi^*_{T0} + \pi^*_{T1}$ above zero. However, our view is that removing such theoretical underpinnings undermines the credibility of the elasticities that are subsequently produced. In view of these results the remaining results have these restrictions imposed.

Turning to the parameters in Table 2, these are of limited interest in themselves, with the elasticities arguably being more easily interpretable. Before discussing these parameters it is worth noting that the eigenvalues of the matrix K_t were computed at all points in the sample. This revealed that all the eigenvalues were non-positive, at all points in the sample confirming that the enforcement of curvature at only two points in the sample was sufficient to give estimates that obeyed curvature throughout the whole sample.

Only the mean and the standard deviations of the posterior distributions have been presented in Table 2. The M-H algorithm permits the examination of the full simulated distribution using a histogram or frequency plot. In most cases (though notably not γ_1 and γ_2 which had their distributions truncated for identification purposes) the distributions looked bell shaped symmetric. Therefore, the mean point of the distribution along with the standard deviation should give readers a fairly good idea of the nature of the simulated distributions. A standard deviation of the parameter that is larger (in absolute terms) than the mean will mean that a substantial proportion of the posterior mass of the distribution will lie either side of zero. Most of the coefficients (both short-run and long-run) at the beginning of the sample appear to have a substantial proportion of their posterior mass either side of zero. However, 'end of period' coefficients do not, most notably, the real income variable z_t. This indicates that there is an important 'transition' in the response of food to real income. The consequences of this will be dealt with when the elasticities are subsequently discussed. On the other hand, the parameters γ_1 and γ_2 are not very far away from their lower boundary. As γ_2 becomes close to one, the transition function becomes extremely insensitive to the value of γ_1. Essentially, this result indicates that the transition is approximately linear.

Turning to Table 3 which presents the Uncompensated (Marshallian) and Compensated (Hicksian) elasticities with respect to price. Once again, the mean and the standard deviations for the simulated posterior distributions are being reported. Readers should note that the reported means in this table, vary marginally from the values that would be calculated using the mean of the posterior distribution of the parameters. The Marshallian price elasticities are presented on the left hand side of the table, and the Hicksian on the right hand side. Since the Hicksian elasticities sum to zero across the rows, only the diagonal elements are given so as to avoid repetition. Balcombe et al. (1999) observe that typical income or expenditure elasticities for 'all food' range from .45 to .99 in the majority of the literature on food demand, but with a substantially larger range for individual food items. Typical price elasticities tend to be at least -.5 and lower (more negative). On the other hand developed countries appear to have lower income elasticities.

The estimates of the Marshallian own price elasticities are around unity in 1967, but become elastic throughout time, being highly elastic in 2000. This is also reflected in a substantial increase in cross price elasticities for food and non-food between the beginning and end of the sample. This increase can be seen in as the result of a large growth in the pure substitution effect, as is evident in the decrease in the Hicksian price elasticities with respect to the price of food. It is not the result of an increased income effect. An examination of the elasticities in Table 4, indicate that the income elasticity for food is not only less than unity, but has fallen over the sample period. In 1967, the estimate was .88, falling to .522 at the end of the sample. This decrease has come from two sources. First, there has been a fall in the share of expenditure on food over the period. A constant negative real income coefficient combined with a fall in the share of food consumed will result in a smaller elasticity. However, there has also been a fall in the long-run coefficient ($\varphi_1 > \varphi_T$) of real income. This has worked in the same direction. Likewise, the decrease in the long-run price coefficient ($\phi_1 > \phi_T$) suggests that it is not just changes in prices and income shares that are responsible for the increased substitutability of food for other goods. Visualizing this in terms of an indifference map with food on the horizontal axis, our results would be consistent with a rather flat indifference curves (at the consumption point) at both points at the sample, with this tendency being more pronounced at the end of the sample. This interpretation is counter intuitive if food consumption is thought of in terms of calorie intake, since it is unlikely (we believe) that calorie intake would be highly price sensitive. Consequently, we would interpret this result as implying a preparedness of consumers to reorganize the food consumption bundle, if items within the consumption bundle become more expensive, with consumers spending on less expensive food items and a resulting substitution towards non-food items. Therefore, in summary, the results presented here suggest that food is price elastic, and income inelastic, and that this tendency has increased throughout the period from 1967 to 2000.

Table 2. Parameters

		mean	std dev
int	δ_0	.1120	.0141
x_t	θ_{11}	.0245	.0318
z_t	π_{11}	.0071	.0364
x_{t-1}	θ_{T1}	-.024	.0322
z_{t-1}	π_{T1}	-.021	.0356
$f_t int$	δ_1	-.011	.0079
$f_t x_t$	θ_{12}	-.099	.0612
$f_t z_t$	π_{12}	-.122	.0558
$f_t x_{t-1}$	θ_{T2}	.0157	.0561
$f_t z_{t-2}$	π_{T2}	.0959	.0561
y_{t-1}	λ	.1586	.1120
γ_1		2.277	2.07
γ_2		2.612	2.24
σ^2		1.5×10^{-6}	1.5×10^{-7}
Long-run			
ϕ_1		-0.0004	0.027
φ_1		-0.0156	0.027
ϕ_T		-0.0978	0.041
φ_T		-0.0490	0.019

Table 3: Price Elasticities

	Uncompensated		Compensated	
	Price Food	Price Non-Food	Price Food	Price Non-Food
1967				
Food	-.988 (.213)	-.099 (.206)	-.866 (.200)	
Non-Food	-.0016 (.033)	-1.01 (.032)		-.139 (.0321)
2000				
Food	-1.88 (.400)	1.36 (.404)	-1.835 (.396)	
Non-Food	.102 (.046)	-1.15 (.057)		-0.211 (.0472)

Mean Elasticities are without parentheses

Standard deviations are within parentheses

Table 4: Income Elasticities

	Food	Non-Food
1967	.889	1.02
	(.203)	(.032)
2000	.522	1.054
	(.189)	(.021)

Mean Elasticities are without parentheses
Standard deviations are within parentheses

6. Summary

This paper estimated and presented some results for a dynamic, smooth transition, Almost Ideal model for food consumption in the US using annual data from 1966 to 2000. A Metropolis-Hastings algorithm was employed to map the posterior distributions and rejection sampling was used to evaluate and impose curvature restrictions at more than one point in the sample. The findings supported the contention of structural change of a 'smooth transition' nature. They also broadly supported the curvature and homogeneity restrictions that were placed on the model. The income food elasticity of demand was inelastic and becoming smaller through time, and the own price elasticities for food were elastic and became more elastic through time.

The M-H algorithm and the Bayesian approach to estimation has much to recommend it. An obvious extension of the work in the paper is to extend the analysis beyond the two good case. The drawbacks, we believe, are mainly computational. However, with increasing computing power, and more efficient algorithms, this does not present an insurmountable problem.

References

Albert J.H. and Chib S. (1993) Bayes Inference via Gibbs Sampling of Autoregressive Time Series Subject to Markov Mean and Variance Shifts. Journal of Business & Economic Statistics, January 1993, Vol 11. No 1.

Anderson G. and R. Blundell (1984). Testing Restrictions in a Flexible Dynamic System: An Application to Consumers Expenditure in the U.K. *Economic Journal* 94: 35-44.

Attfield C (1997) Estimating a Cointegrating Demand System. European Economic Review 41:61-73

Balcombe, K. S. Davidova, and Morrison J.A. (1999). Consumer Behaviour in a Country in Transition with a Strongly Contracting Economy, *Journal of Agricultural Economics,* 50, 36-47.

Balcombe K. (2003). Retesting Symmetry and Homogeneity in a Cointegrated Demand System with Bootstrapping: the Case of Meat Demand in Greece. Empirical Economics.

Bauwens, L. Lubrano, M. and J.F. Richard. (1999). Bayesian Inference in Dynamic Econometric Models. Advanced Texts in Econometrics. Oxford University Press.

Buse A (1994) Linearised Almost Ideal Demand System. American Journal of Agricultural Economics 76:4, 781-793

Buse A (1998) Homogeneity in the Almost Ideal Demand System. American Journal of Agricultural Economics 80:1, 208-220

Chalfant, J.A., Gray R.S. and White K.J. (1991). Evaluating Prior Belief in a Demand System: The Case of Meat Demand in Canada, *American Journal of Agricultural Economics*, 73, 476-490

Casella G. and George E.I. (1992) Explaining the Gibbs Sampler. The American Statistician, August 1992, 46, 3, pp.167-174.

Chambers M.J. and Nowman K.B., (1997) Forecasting with the almost ideal demand system: evidence from some alternative dynamic specifications Applied Economics 29, pp.935-943.

Chib S. and Greenberg E. (1995a) Heirarchical Analysis of SUR Models with Extensions to Correlated Serial Errors and Time Varying Parameters, Journal of Econometrics, 68: 339 360.

Chib C. and E. Greenberg, (1995b). Understanding the Metropolis-Hastings Algorithm. The American Statistician, November, 1995, 49. No 4.: 327-335

Deaton, A.S. and J . Muellbauer (1980) An almost ideal demand system. American Economic Review 70, pp312-326.

Deschamps P.J. (2000) Exact small sample interence in stationary fully regular, dynamic demand models. Journal of Econometrics, 97, 51-91.

Deschamps P.J. (2003) Time Varying Intercepts and Equilibrium Analysis: An Extension of the Dynamic Almost Ideal Demand Model, Journal of Applied Econometrics (forthcoming).

De Crombrugge D., Palm F.C. and J.P. Urbain, 'Statistical Demand Function for Food in the USA and the Netherlands, *Journal of Applied Econometrics*, Vol 12. 615-645.

Duffy M., (2002) On the estimation of an advertising-augmented cointegrating demand system. Economic Modelling, 20, pp.181-206.

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

Geweke, J. (1989). Bayesian Inference In econometric Models Using Monte Carlo Integration. Econometrica. 57: 1317-39.

Geweke J. (1997) Using Simulation Methods for Bayesian Econometric Models: Inference, Development and Communication. Paper preparedfor the Australsian meeing so the Econometric Socitey, Melbourne

Griffiths, W.E., O'Donnell, C.J. and A. Tan Cruz (2000) "Imposing regularity conditions on a system of cost and cost-share equations: a Bayesian approach", Australian Journal of Agricultural Economics, 44(1):107-127

Hendry (1993) Econometrics Alchemy of Science? Essays in Econometric Methodology. Blackwell, Oxford UK.

Lau, L.J. (1978), Testing and imposing monotonicity, convexity and quasiconcavity constraints', in Fuss M. and McFadden, D. Production Economics: A Dual Approach to Theory and Applications, Vol 1, North Holland, Amsterdam.

Morrison J.A. K.Balcombe, A.Bailey, S Klonaris, and G.Rapsomanikis (2003). Expenditure on different cataegoris of meat in Greece: the influence of changing tastes, Agricultural Economics, 28: 139-150.

Ng S (1995). Testing for Homogeneity in Demand Systems when the Regressors are Non-Stationary. Journal of Applied Econometrics, 10, pp.147-163.

Terasvirta, T. (1994). Specification, estimation, and evaluation of smooth transition autoregressive models. Journal of Applied Econometrics, 7, S119-S136

Tiffin A. and Tiffin R. (1999). Estimates of Food Demand Elasticities for Great Britain: 1972-1994, *Journal of Agricultural Economics,* 50, 140-147.

Pesaran MH Shin Y (2002) Long-run Structural Modelling. Econometrics Reviews 21:49-87