Chapter 5

Maximum Likelihood Estimation of Household Equivalence Scales

The preceding chapter has been concerned with the estimation of equivalence scales using the Engel approach which showed the overall effect of a change in demographic composition on the total consumption of the household. In this chapter, commodity-specific scales are estimated using the scaling-procedure of Barten to incorporate demographic variables into a demand system. An iterative procedure based on maximum likelihood techniques is developed and used in the estimation of the scales. The results update both the commodity-specific and general scales previously obtained for Australia using the extended linear expenditure system.

5.1 The Model

The demand system employed here is the extended linear expenditure system (ELES) of Lluch (1973) that was later modified by Kakwani (1977) for equivalence scale estimation. A feature of this system is linearity, an assumption which is often questioned. Also, the utility function from which it is derived is directly additive and, as shown in Deaton and Muellbauer (1980), this is a restrictive assumption, particularly in studies using a detailed disaggregation of commodities. Inspite of these disadvantages, the ELES was chosen for this study for a number of important reasons. First, the ELES is a convenient vehicle for carrying out relatively sophisticated research on consumer behaviour even when available data on private consumption are limited. Because time series data on private consumption are not disaggregated over various commodity groups, one can only use cross-section information for estimating demand parameters. Since purely cross-section data generally give no price variation, inference about price effects requires strong theoretical specifications. Second, the ELES is chosen for its historical significance in equivalence scale research. ELES-based equivalence scales have been repeatedly estimated in the past and are particularly popular among researchers using Australian data. See, for example, Kakwani (1980), Binh and Whiteford (1990), and Bradbury (1994). Using the ELES in this study facilitates comparison of results to these earlier ones. Thirdly, the ELES is used because of its simplicity. In this chapter (and succeeding chapters) new methods of estimation of equivalence scales are derived. For these purposes, the ELES is ideal because the system remains mathematically tractable but is still sufficiently complicated to warrant a number of econometric innovations. Once these innovations have been developed, they can be more readily applied to more complicated models at a later time. Special care was taken to split the sample into groups where there were few parametric restrictions on the scales and estimation was restricted to just eleven broad commodity groups, thereby mitigating the assumption of additive utility. The exercise is an natural starting point for the demonstration of the derivation of a new iterative procedure using maximum likelihood techniques (in this chapter) and the application of Bayesian methodologies (in Chapters 6 and 7) for equivalence scale estimation. The study as a whole provides a useful addition to the available empirical evidence on equivalence scales in Australia.

5.1.1 The Extended Linear Expenditure System

To describe the model, consider n commodity groups indexed by $i = 1, 2, ..., n$ and H types of households indexed by $h = 1, 2, ..., H$ where household types are defined according to the number of adults and the number of children in the household. Define q_{ih} as the quantity of the i^{th} commodity consumed by the h-type household and s_{ih} is the i^{th} commodity-specific equivalence scale for the h-type household.

The s_{ih} are factors used to adjust q_{ih} values in utility functions to show the effect of a change in the household's demographic composition on household utility and on specific commodity expenditures. On a per unit basis, a given q_{ih} provides less utility if it is shared with more people. How q_{ih} should be deflated to give the same per unit utility will depend on the commodity i and on the household type h. Thus the scale is subscripted by i and h. The unit for which the utility function is specified is the household type where $s_{ih} = 1$. For example, if the reference unit is a household with two adults without children, then the q_{ih} in the utility function is scaled by s_{ih} to give a comparable two-adult-zero-children utility function for all households.

Given this background, consider now the Klein-Rubin utility function where the consumption quantities q_{ih} have been Barten-scaled as follows:

$$u_h = \sum_{i=1}^{n} b_i \ln \left\{ \frac{q_{ih}}{s_{ih}} - c_i \right\} \tag{5.1}$$

where

b_i = is the marginal contribution to utility of the i^{th} commodity and satisfies the constraints $0 < b_i < 1$ and $\sum_{i=1}^{n} b_i = 1$;

c_i = is a parameter which, if interpreted as the subsistence quantity of the i^{th} commodity, satisfies the constraint $c_i > 0$; Pollak and Wales (1978) prefer not to give c_i a strict subsistence interpretation letting negative values be a possibility.

Let p_i be the price of the i^{th} commodity and v_h be the total expenditure for the h-type household. The Barten-scaled prices and consumption quantities are denoted by $p^*_{ih} = p_i s_{ih}$ and $q^*_{ih} = \frac{q_{ih}}{s_{ih}}$. Maximising the utility function (5.1) subject to the budget constraint $\sum_{i=1}^{n} p_i q_{ih} = v_h$ is equivalent to maximising the utility function

$$u(q^*_{ih}) = \sum_{i=1}^{n} b_i \ln(q^*_{ih} - c_i) \tag{5.2}$$

such that $\sum_{i=1}^{n} p^*_i q^*_{ih} = v_h$. The Lagrangean equation for the problem above is written as

$$L = \sum_{i=1}^{n} b_i \ln\,(q^*_{ih} - c_i) + \lambda \left(v_h - \sum_{i=1}^{n} p^*_i q^*_{ih} \right) \tag{5.3}$$

where λ is the Lagrange multiplier. Differentiating L with respect to q^*_{ih} and λ yields the following first order conditions

$$\frac{b_i}{q^*_{ih} - c_i} = \lambda p^*_i \tag{5.4}$$

$$v_h = \sum_{i=1}^{n} p^*_i q^*_{ih} \tag{5.5}$$

Solving for q^*_{ih} and λ, we get from (5.4)

$$b_i = \lambda p^*_i (q^*_{ih} - c_i) = \lambda (p^*_i q^*_{ih} - p^*_i c_i) \tag{5.6}$$

Summing over all i and recognising that $\sum_{i=1}^{n} b_i = 1$ and that $v_h = \sum_{i=1}^{n} p^*_i q^*_{ih}$, the following expression for λ is obtained

$$\lambda = \frac{1}{\left(v_h - \sum_{i=1}^{n} p^*_i c_i \right)} \tag{5.7}$$

Substituting (5.7) back into (5.4) yields

$$\frac{b_i}{q^*_{ih} - c_i} = \frac{p_i}{\left(v_h - \sum_{j=1}^{n} p^*_j c_j \right)} \tag{5.8}$$

Simplification leads to

$$p^*_i q^*_{ih} = p^*_i c_i + b_i \left(v_h - \sum_{j=1}^{n} p^*_j c_j \right) \tag{5.9}$$

or, equivalently,

$$p_i q_{ih} = p_i s_{ih} c_i + b_i \left(v_h - \sum_{j=1}^{n} p_j s_{jh} c_j \right) \tag{5.10}$$

A household whose demand system is an LES is often described as first purchasing "necessary", "subsistence" or "committed" quantities of each good $(s_{1h}c_1, ..., s_{nh}c_n)$ and then dividing its remaining or "supernumerary" expenditure $(v_h - \sum_{i=1}^{n} p_i s_{ih} c_i)$, among the goods in fixed proportions, $(b_1, ..., b_n)$.

The system in (5.10) can be more compactly expressed as

$$v_{ih} = a_{ih} + b_i(v_h - a_h) \tag{5.11}$$

where

v_{ih} = $p_i q_{ih}$ is expenditure on the i^{th} commodity by the h-type household, and

a_{ih} = $p_i c_i s_{ih}$ is subsistence expenditure for the i^{th} commodity and h-type household.

5.1.2 Identification of Parameters

The objective is to estimate a_{ih} and b_i, with these estimates later being used to estimate the scales s_{ih}. Specifically, if $s_{ir} = 1$ denotes the scale for the reference household type, then

$$s_{ih} = \frac{p_i c_i s_{ih}}{p_i c_i s_{ir}} = \frac{a_{ih}}{a_{ir}} \tag{5.12}$$

However, without further information, not all of the a_{ih} are identified. The identification problem arises because for a given household type, one of the n equations in (5.11) is redundant, redundancy being illustrated by summing both sides. Summing (5.11) over all commodities yields

$$\sum_{ih}^{n} v_{ih} = \sum_{i=1}^{n} a_{ih} + \sum_{i=1}^{n} b_i(v_h - a_h) \tag{5.13}$$

or

$$v_h = a_h + (v_h - a_h) \tag{5.14}$$

where $a_h = \sum_{i=1}^{n} a_{ih}$ is total subsistence expenditure for a h-type household. The redundancy of one equation means that separate information is only available from $n-1$ equations. Our problem is to estimate n intercept terms with only $n-1$ equations.

One solution to this identification problem is to include in the linear expenditure system in (5.11) a micro-consumption function given by

$$v_h = a_h + b(x_h - a_h) \tag{5.15}$$

where v_h is the total expenditure, x_h is net income b is a common marginal propensity to consume for all households. This function shows that total expenditure v_h is composed of "committed" or "subsistence" expenditure a_h and a proportion b of "uncommitted" expenditure $(x_h - a_h)$. The extended linear expenditure system or ELES is thus comprised of equations (5.11) and (5.15).

To estimate the parameters in the ELES, Kakwani (1977) appended errors to these equations, and assumed the error variances can be different for each household type and for each commodity. He suggested first estimating a_h and b from (5.15), and then replacing a_h in each of the commodity equations in (5.11) with its estimate from (5.15). Then, to estimate a_{ih} and b_i in (5.11), weighted least squares which allows for heteroskedasticity across different household types was applied to each of these equations. Using an external estimate of a_h identifies the remaining parameters.

The estimation procedure that is developed in this chapter attempts to improve on Kakwani's procedure in two ways. First, because Kakwani estimated each of the commodity equations separately, he ignored any correlation that might exist between the errors that correspond to different commodity equations for a given household. Second, the '2-step' nature of the procedure ignored the effect of using estimates from one equation on the properties of the estimates from a second equation. An estimator which allows for error correlation across different commodity equations and which estimates all parameters simultaneously would seem more desirable.

To investigate how all the parameters might be jointly estimated, (5.15) is substituted into (5.11) to obtain

$$\begin{aligned} v_{ih} &= a_{ih} + b_i[(a_h + b(x_h - a_h)) - a_h] \\ &= a_{ih} + b_i b(x_h - a_h) \\ &= a_{ih} - b_i b a_h + b_i b x_h \\ &= \theta_{ih} + \eta_i x_h \qquad (5.16) \end{aligned}$$

where

$$\theta_{ih} = a_{ih} - b_i b a_h$$

$$\eta_i = b_i b$$

Consider now the estimation of the θ_{ih} and the η_i and how estimates of the structural parameters, a_{ih}, b_i, b and a_h can be retrieved from these estimates. Given estimates of θ_{ih} and η_i, estimates of the structural parameters b, a_h, b and a_{ih} can be obtained using the expressions

$$b = \sum_{i=1}^{n} \eta_i \tag{5.17}$$

$$a_h = \frac{\sum_{i=1}^{n} \theta_{ih}}{1 - \sum_{i=1}^{n} \eta_i} \tag{5.18}$$

$$b_i = \frac{\eta_i}{\sum_{j=1}^{n} \eta_j} \tag{5.19}$$

$$a_{ih} = \theta_{ih} + \frac{\eta_i \sum_{j=1}^{n} \theta_{jh}}{1 - \sum_{j=1}^{n} \eta_j} \tag{5.20}$$

The system in (5.16) does not suffer from an identification problem because there are no redundant equations. All the n commodity equations for a given household type can be utilised.

5.2 Expressions for the Commodity-Specific and General Scales

By definition, the equivalence scale s_{ih} is the ratio of the subsistence expenditure for the i^{th} commodity of the h-type household to that of the reference household. It is given by

$$s_{ih} = \frac{a_{ih}}{a_{ir}} \tag{5.21}$$

where r refers to the reference household. This shows the amount by which the commodity expenditures must be deflated so that they are expressed in per household equivalent terms.

To obtain an expression for the general scales, we first consider the demand equations in (5.10). Dividing through by $p_i s_{ih}$, we get

$$\frac{q_{ih}}{s_{ih}} = c_i + \frac{b_i b}{p_i s_{ih}} \left(x_h - \sum_{j=1}^{n} p_j s_{jh} c_j \right) \tag{5.22}$$

Equation (5.22) is then substituted into the direct utility function in (5.1) and noting that $\sum_{i=1}^{n} b_i = 1$, we get

$$\begin{aligned} u_h &= \sum_{i=1}^{n} b_i \ln \left[\frac{b_i b}{p_i s_{ih}} \left(x_h - \sum_{j=1}^{n} p_j s_{jh} c_j \right) \right] \\ &= \sum_{i=1}^{n} b_i \left[\ln b_i - \ln p_i - \ln s_{ih} + \ln b + \ln \left(x_h - \sum_{j=1}^{n} p_j s_{jh} c_j \right) \right] \\ &= \ln b + \ln \left(x_h - \sum_{i=1}^{n} p_i s_{ih} c_i \right) + \sum_{i=1}^{n} b_i \ln b_i - \sum_{i=1}^{n} b_i \ln p_i - \sum_{i=1}^{n} b_i \ln s_{ih} \end{aligned} \tag{5.23}$$

For the standard reference household where $s_{ir} = 1$, the indirect utility function is thus expressed as

$$u_r = \ln b + \ln(x_r - \sum_{i=1}^{n} p_i c_i) + \sum_{i=1}^{n} b_i \ln b_i - \sum_{i=1}^{n} b_i \ln p_i \tag{5.24}$$

The general scale for the h-type household is given by the ratio of incomes,

$$s_h = x_h / x_r$$

that equates the two indirect utility functions. Working in this direction, we set $u_r = u_h$ to obtain

$$\begin{aligned} &\ln b + \ln(x_r - \sum_{i=1}^{n} p_i c_i) + \sum_{i=1}^{n} b_i \ln b_i - \sum_{i=1}^{n} b_i \ln p_i \\ &= \ln b + \ln \left(x_h - \sum_{i=1}^{n} p_i s_{ih} c_i \right) + \sum_{i=1}^{n} b_i \ln b_i - \sum_{i=1}^{n} b_i \ln p_i - \sum_{i=1}^{n} b_i \ln s_{ih} \end{aligned} \tag{5.25}$$

Simplifying (5.25), we have

$$\ln \left(x_r - \sum_{i=1}^{n} p_i c_i \right) = \ln \left(x_h - \sum_{i=1}^{n} p_i s_{ih} c_i \right) - \sum_{i=1}^{n} b_i \ln s_{ih}$$

$$
\begin{aligned}
\ln\left(x_h - \sum_{i=1}^{n} p_i s_{ih} c_i\right) &= \ln\left(x_r - \sum_{i=1}^{n} p_i c_i\right) + \sum_{i=1}^{n} b_i \ln s_{ih} \\
\left(x_h - \sum_{i=1}^{n} p_i s_{ih} c_i\right) &= \left(x_r - \sum_{i=1}^{n} p_i c_i\right)\left(\prod_{i=1}^{n} s_{ih}^{b_i}\right) \\
&= x_r \prod_{i=1}^{n} s_{ih}^{b_i} - \left(\prod_{i=1}^{n} s_{ih}^{b_i}\right)\left(\sum_{i=1}^{n} p_i c_i\right) \\
x_h &= x_r \prod_{i=1}^{n} s_{ih}^{b_i} + \sum_{i=1}^{n} p_i s_{ih} c_i - \left(\prod_{i=1}^{n} s_{ih}^{b_i}\right)\left(\sum_{i=1}^{n} p_i c_i\right) \\
\frac{x_h}{x_r} &= \prod_{i=1}^{n} s_{ih}^{b_i} + \frac{1}{x_r}\left[\sum_{i=1}^{n} p_i s_{ih} c_i - \left(\prod_{i=1}^{n} s_{ih}^{b_i}\right)\left(\sum_{i=1}^{n} p_i c_i\right)\right] \quad (5.26)
\end{aligned}
$$

Noting that $a_{ih} = p_i s_{ih} c_i$ and $a_h = \sum_{i=1}^{n} a_{ih}$, (5.26) can be equivalently written as

$$
s_h = \frac{x_h}{x_r} = \prod_{i=1}^{n} s_{ih}^{b_i} + \frac{1}{x_r}\left[\sum_{i=1}^{n} a_{ih} - \prod_{i=1}^{n} s_{ih}^{b_i} a_r\right] \quad (5.27)
$$

$$
= \frac{a_h}{x_r} + \prod_{i=1}^{n} (s_{ih})^{b_i}\left[1 - \frac{a_r}{x_r}\right] \quad (5.28)
$$

These general scales s_h are the final quantities of interest. They capture the overall effect of a change in demographic composition on the total expenditure of the household. From (5.28), they are shown to be functions of the commodity-specific scales s_{ih}'s and are calculated based on a chosen reference income level of the reference household. If we rewrite the first term in (5.28) as

$$
\frac{a_h}{a_r}\frac{a_r}{x_r}
$$

we can see that a general scale is weighted average of the $\prod_{i=1}^{n} (s_{ih})^{b_i}$ term and a_h/a_r, where the former is a geometric mean of the s_{ih}'s and the latter is a ratio of relative subsistence costs.

5.3 Stochastic Assumptions and ML Estimation

Suppose now that there are M_h observations (households) with demographic composition type h. In the notation that follows, the symbols v_{ih} and x_h which previously represented scalar quantities for a given household, will become $(M_h \times 1)$

vectors containing all observations on households of type h. Returning to equation (5.16), and adding stochastic terms, the system we wish to estimate can be written as

$$\begin{bmatrix} v_{1h} \\ v_{2h} \\ \vdots \\ v_{nh} \end{bmatrix} = \begin{bmatrix} z_h & & & \\ & z_h & & \\ & & \ddots & \\ & & & z_h \end{bmatrix} \begin{bmatrix} \theta_{1h} \\ \theta_{2h} \\ \vdots \\ \theta_{nh} \end{bmatrix} + \begin{bmatrix} x_h & & & \\ & x_h & & \\ & & \ddots & \\ & & & x_h \end{bmatrix} \begin{bmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_n \end{bmatrix} + \begin{bmatrix} e_{1h} \\ e_{2h} \\ \vdots \\ e_{nh} \end{bmatrix} \tag{5.29}$$

or

$$\mathbf{V}_h = \mathbf{Z}_h \boldsymbol{\Theta}_h + \mathbf{X}_h \boldsymbol{\eta} + \mathbf{E}_h \tag{5.30}$$

where

h	$= 1, 2, ..., H$ refers to household composition type h;
n	refers to the number of commodity groups;
v_{ih}	is an $(M_h \times 1)$ vector of observations on expenditure for the i^{th} commodity and the h-type household;
z_h	is an $(M_h \times 1)$ vector of ones;
x_h	is an $(M_h \times 1)$ vector of observations on income for households of type h;
e_{ih}	is an $(M_h \times 1)$ vector of errors;
$\mathbf{V}_h$	is of dimension $(nM_h \times 1)$;
$\mathbf{Z}_h$	$= I_n \otimes z_h$ is an $(nM_h \times n)$ matrix of dummy variables;
$\mathbf{X}_h$	$= I_n \otimes x_h$ is an $(nM_h \times n)$ matrix of household incomes;
$\boldsymbol{\Theta}_h, \boldsymbol{\eta}$	are $(n \times 1)$ vectors of unknown parameters;
$\mathbf{E}_h$	is an $(nM_h \times 1)$ vector of errors which is assumed to be distributed as

$$\mathbf{E}_h \sim N[0, \boldsymbol{\Omega}_h \otimes I_{M_h}] \tag{5.31}$$

Thus the error covariance matrix $\boldsymbol{\Omega}_h$ is allowed to be different for different household types. Because $\boldsymbol{\Omega}_h$ is not diagonal, correlation between errors from equations for different commodities, and the same household, is permitted. Zero error correlation is assumed across different households. (The sample is assumed to be random). Thus, in addition to (5.31), $E(\mathbf{E}_h\mathbf{E}'_k) = 0$ for $h \neq k$.

The task is to derive expressions for the maximum likelihood estimators of $\mathbf{\Theta}_h, \mathbf{\Omega}_h$ and $\boldsymbol{\eta}$, as well as asymptotic covariance matrices for these estimates and asymptotic covariance matrices for the consequent maximum likelihood estimates for the parameters in equations (5.17) - (5.20).

5.3.1 Derivation of Maximum Likelihood Estimators

Noting that $\mathbf{V}_1, \mathbf{V}_2, ..., \mathbf{V}_H$ in the system of equations defined in (5.30) are independent, the log-likelihood for all parameters, given data on all household types, can be written as

$$\log \mathrm{L} = \frac{nM}{2}\log\,(2\pi) - \frac{1}{2}\sum_{h=1}^{H} M_h \log \mid \mathbf{\Omega}_h \mid$$
$$- \frac{1}{2}\sum_{h=1}^{H}(\mathbf{V}_h - \mathbf{Z}_h\mathbf{\Theta}_h - \mathbf{X}_h\boldsymbol{\eta})'(\mathbf{\Omega}_h^{-1} \otimes I_{M_h})(\mathbf{V}_h - \mathbf{Z}_h\mathbf{\Theta}_h - \mathbf{X}_h\boldsymbol{\eta}) \quad (5.32)$$

where $M = \sum_{h=1}^{H} M_h$. To maximise this function, the possibility of concentrating out the $\mathbf{\Theta}_h$ is first investigated. Working in this direction, the last term can be written (without the summation) as

$$\begin{aligned} \mathbf{Q}_h &= (\mathbf{V}_h - \mathbf{Z}_h\mathbf{\Theta}_h - \mathbf{X}_h\boldsymbol{\eta})'(\mathbf{\Omega}_h^{-1} \otimes I_{M_h})(\mathbf{V}_h - \mathbf{Z}_h\mathbf{\Theta}_h - \mathbf{X}_h\boldsymbol{\eta}) \\ &= (\mathbf{V}_h - \mathbf{X}_h\boldsymbol{\eta})'(\mathbf{\Omega}_h^{-1} \otimes I_{M_h})(\mathbf{V}_h - \mathbf{X}_h\boldsymbol{\eta}) + \mathbf{\Theta}'_h\mathbf{Z}'_h(\mathbf{\Omega}_h^{-1} \otimes I_{M_h})\mathbf{Z}_h\mathbf{\Theta}_h \\ &\quad -2\mathbf{\Theta}'_h\mathbf{Z}'_h(\mathbf{\Omega}_h^{-1} \otimes I_{M_h})(\mathbf{V}_h - \mathbf{X}_h\boldsymbol{\eta}) \end{aligned} \quad (5.33)$$

Now,

$$\begin{aligned} \frac{\partial \log \mathrm{L}}{\partial \mathbf{\Theta}_h} &= -\frac{1}{2}\frac{\partial \mathbf{Q}_h}{\partial \mathbf{\Theta}_h} \\ &= -\frac{1}{2}\left[2\mathbf{Z}'_h(\mathbf{\Omega}_h^{-1} \otimes I_{M_h})\mathbf{Z}_h\mathbf{\Theta}_h - 2\mathbf{Z}'_h(\mathbf{\Omega}_k^{-1} \otimes I_{M_h})(\mathbf{V}_h - \mathbf{X}_h\boldsymbol{\eta})\right] \end{aligned} \quad (5.34)$$

Setting this derivative to zero and solving for the maximising value $\hat{\mathbf{\Theta}}_h$ gives

$$\hat{\mathbf{\Theta}}_h = \left[\mathbf{Z}'_h(\mathbf{\Omega}_h^{-1} \otimes I_{M_h})\mathbf{Z}_h\right]^{-1}\mathbf{Z}'_h(\mathbf{\Omega}_h^{-1} \otimes I_{M_h})(\mathbf{V}_h - \mathbf{X}_h\boldsymbol{\eta}) \quad (5.35)$$

Now,

$$\begin{aligned}\mathbf{Z}_h'(\mathbf{\Omega}_h^{-1}\otimes I_{M_h})\mathbf{Z}_h &= (I_n\otimes z_h')(\mathbf{\Omega}_h^{-1}\otimes I_{M_h})(I_n\otimes z_h)\\ &= \mathbf{\Omega}_h^{-1}\otimes z_h'z_h\\ &= \mathbf{\Omega}_h^{-1}\otimes M_h\\ &= M_h\mathbf{\Omega}_h^{-1} \qquad (5.36)\end{aligned}$$

Also

$$\mathbf{Z}_h'(\mathbf{\Omega}_h^{-1}\otimes I_{M_h}) = \mathbf{\Omega}_h^{-1}\otimes z_h' \qquad (5.37)$$

Using (5.36) and (5.37) in (5.35) gives

$$\begin{aligned}\hat{\mathbf{\Theta}}_h &= [\mathbf{\Omega}_h^{-1}\otimes M_h]^{-1}(\mathbf{\Omega}_h^{-1}\otimes z_h')(\mathbf{V}_h-\mathbf{X}_h\boldsymbol{\eta})\\ &= (\mathbf{\Omega}_h\otimes M_h^{-1})(\mathbf{\Omega}_h^{-1}\otimes z_h')(\mathbf{V}_h-\mathbf{X}_h\boldsymbol{\eta})\\ &= \left(I_n\otimes\frac{1}{M_h}z_h'\right)(\mathbf{V}_h-\mathbf{X}_h\boldsymbol{\eta}) \qquad (5.38)\end{aligned}$$

Considering the i^{th} row in equation (5.38), we obtain

$$\begin{aligned}\hat{\theta}_{ih} &= \frac{1}{M_h}z_h'(v_{ih}-x_h\eta_i)\\ &= \overline{v}_{ih}-\overline{x}_h\eta_i \qquad (5.39)\end{aligned}$$

where $\overline{v}_{ih} = \frac{1}{M_h}z_h'v_{ih}$ is the average expenditure on commodity i for all households of type h and $\overline{x}_h = \frac{1}{M_h}z_h'x_h$ is the average income of all h-type households. The result in (5.39) is an important one. It means that the $\hat{\theta}_{ih}$'s do not depend on $\mathbf{\Omega}_h$ and can be computed at the end of the maximum likelihood algorithm, after we have estimated $\mathbf{\Omega}_h$ and η_i.

Let $\overline{\mathbf{V}}_h' = (\overline{v}_{1h},\overline{v}_{2h},...,\overline{v}_{nh})$ and $\overline{X}_h = I_n\otimes\overline{x}_h$. Then,

$$\hat{\mathbf{\Theta}}_h = \overline{\mathbf{V}}_h - \overline{\mathbf{X}}_h\boldsymbol{\eta} \qquad (5.40)$$

Substituting (5.40) into (5.33) yields

$$\begin{aligned}\mathbf{Q} &= \left[(\mathbf{V}_h-\mathbf{Z}_h\overline{\mathbf{V}}_h)-(\mathbf{X}_h-\mathbf{Z}_h\overline{\mathbf{X}}_h)\boldsymbol{\eta}\right]'(\mathbf{\Omega}_h^{-1}\otimes I_{M_h})\left[(\mathbf{V}_h-\mathbf{Z}_h\overline{\mathbf{V}}_h)-(\mathbf{X}_h-\mathbf{Z}_h\overline{\mathbf{X}}_h)\boldsymbol{\eta}\right]\\ &= (\mathbf{V}_h^*-\mathbf{X}_h^*\boldsymbol{\eta})'(\mathbf{\Omega}_h^{-1}\otimes I)(\mathbf{V}_h^*-\mathbf{X}_h^*\boldsymbol{\eta}) \qquad (5.41)\end{aligned}$$

where $\mathbf{V}_h^* = \mathbf{V}_h - \mathbf{Z}_h\overline{\mathbf{V}}_h$ is a vector of expenditures expressed in terms of deviations from the mean expenditures for each commodity and household type, and $\mathbf{X}_h^* = \mathbf{X}_h - \mathbf{Z}_h\overline{\mathbf{X}}_h$ is a vector of incomes expressed in terms of deviations from the mean incomes for each household type. The concentrated log-likelihood function can now be written as

$$\begin{aligned}\log \mathrm{L}^* &= -\frac{nM}{2}\log\,(2\pi) - \frac{1}{2}\sum_{h=1}^{H} M_h \log \mid \boldsymbol{\Omega}_h \mid \\ &\quad -\frac{1}{2}\sum_{h=1}^{H}(\mathbf{V}_h^* - \mathbf{X}_h^*\boldsymbol{\eta})'(\boldsymbol{\Omega}_h^{-1} \otimes I_{M_h})(\mathbf{V}_h^* - \mathbf{X}_h^*\boldsymbol{\eta}) \qquad (5.42)\\ &= -\frac{nM}{2}\log\,(2\pi) + \frac{1}{2}\sum_{h=1}^{H} M_h \log \mid \boldsymbol{\Omega}_h^{-1} \mid -\frac{1}{2}\sum_{h=1}^{H} tr\left[\mathbf{W}_h\boldsymbol{\Omega}_h^{-1}\right] \qquad (5.43)\end{aligned}$$

where $\mathbf{W}_h$ is an $(n \times n)$ matrix of $(i,j)^{th}$ element given by

$$[\mathbf{W}_h]_{ij} = (v_{ih}^* - x_h^*\eta_i)'(v_{jh}^* - x_h^*\eta_j) \qquad (5.44)$$

See Judge, et.al. (1988, p.553) for details of the two alternative specifications in (5.42) and (5.43). Judge, et.al. also gives details on how to differentiate (5.42) with respect to $\boldsymbol{\Omega}_h^{-1}$. This differentiation yields

$$\frac{\partial \log \mathrm{L}^*}{\partial \boldsymbol{\Omega}_h^{-1}} = \frac{M_h}{2}\boldsymbol{\Omega}_h - \frac{1}{2}\mathbf{W}_h \qquad (5.45)$$

Setting this derivative equal to zero yields the maximum likelihood estimator for $\boldsymbol{\Omega}_h$ given $\boldsymbol{\eta}$ as follows:

$$\hat{\boldsymbol{\Omega}}_h = \frac{1}{M_h}\mathbf{W}_h \qquad (5.46)$$

To find an expression for the maximum likelihood estimator for $\boldsymbol{\eta}$, given $\boldsymbol{\Omega}_h$, we return to the last term in (5.42) and rewrite it as

$$\begin{aligned}\sum_{h=1}^{H}\mathbf{Q}_h^* &= \sum_{h=1}^{H}(\mathbf{V}_h^* - \mathbf{X}_h^*\boldsymbol{\eta})'(\boldsymbol{\Omega}_h^{-1} \otimes I_{M_h})(\mathbf{V}_h^* - \mathbf{X}_h^*\boldsymbol{\eta})\\ &= \sum_{h=1}^{H}\Big[\mathbf{V}_h^{*\prime}(\boldsymbol{\Omega}_h^{-1} \otimes I_{M_h})\mathbf{V}_h^* + \boldsymbol{\eta}'\mathbf{X}_h^{*\prime}(\boldsymbol{\Omega}_h^{-1} \otimes I_{M_h})\mathbf{X}_h^*\boldsymbol{\eta}\\ &\quad -2\boldsymbol{\eta}'\mathbf{X}_h^{*\prime}(\boldsymbol{\Omega}_h^{-1} \otimes I_{M_h})\mathbf{V}_h^*\Big] \qquad (5.47)\end{aligned}$$

Now,

$$\begin{aligned}\frac{\partial \log \mathrm{L}^*}{\partial \boldsymbol{\eta}} &= -\frac{1}{2}\sum_{h=1}^{H}\frac{\partial \mathbf{Q}_h^*}{\partial \boldsymbol{\eta}} \\ &= -\frac{1}{2}\sum_{h=1}^{H}\left[2\mathbf{X}_h^{*\prime}(\boldsymbol{\Omega}_h^{-1}\otimes I_{M_h})\mathbf{X}_h^*\boldsymbol{\eta} - 2\mathbf{X}_h^{*\prime}(\boldsymbol{\Omega}_h^{-1}\otimes I_{M_h})\mathbf{V}_h^*\right] \quad (5.48)\end{aligned}$$

Setting this quantity equal to zero yields

$$\left[\sum_{h=1}^{H}\mathbf{X}_h^{*\prime}(\boldsymbol{\Omega}_h^{-1}\otimes I_{M_h})\mathbf{X}_h^*\right]\hat{\boldsymbol{\eta}} = \sum_{h=1}^{H}\mathbf{X}_h^{*\prime}(\boldsymbol{\Omega}_h^{-1}\otimes I_{M_h})\mathbf{V}_h^* \quad (5.49)$$

Now,

$$\begin{aligned}\mathbf{X}_h^{*\prime}(\boldsymbol{\Omega}_h^{-1}\otimes I_{M_h})\mathbf{X}_h^* &= (I_n\otimes x_h^{*\prime})(\boldsymbol{\Omega}_h^{-1}\otimes I_{M_h})(I_n\otimes x_h^*) \\ &= \boldsymbol{\Omega}_h^{-1}\otimes x_h^{*\prime}x_h^* \\ &= x_h^{*\prime}x_h^*\boldsymbol{\Omega}_h^{-1} \quad (5.50)\end{aligned}$$

Also,

$$\begin{aligned}\mathbf{X}_h^{*\prime}(\boldsymbol{\Omega}_h^{-1}\otimes I_{M_h})\mathbf{V}_h^* &= (I_n\otimes x_h^{*\prime})(\boldsymbol{\Omega}_h^{-1}\otimes I_{M_h})\mathbf{V}_h^* \\ &= (\boldsymbol{\Omega}_h^{-1}\otimes x_h^{*\prime})\mathbf{V}_h^* \quad (5.51)\end{aligned}$$

In the light of the 2nd line in (5.50), equation (5.51) can be written as

$$\begin{aligned}\mathbf{X}_h^{*\prime}(\boldsymbol{\Omega}_h^{-1}\otimes I_{M_h})\mathbf{V}_h^* &= \left[\boldsymbol{\Omega}_h^{-1}\otimes x_h^{*\prime}x_h^*\right]\left[\boldsymbol{\Omega}_h^{-1}\otimes x_h^{*\prime}x_h^*\right]^{-1}\left[\boldsymbol{\Omega}_h^{-1}\otimes x_h^{*\prime}\right]\mathbf{V}_h^* \\ &= \left(x_h^{*\prime}x_h^*\boldsymbol{\Omega}_h^{-1}\right)\left(\boldsymbol{\Omega}_h\otimes(x_h^{*\prime}x_h^*)^{-1}\right)(\boldsymbol{\Omega}_h^{-1}\otimes x_h^{*\prime})\mathbf{V}_h^* \\ &= x_h^{*\prime}x_h^*\boldsymbol{\Omega}_h^{-1}[I_n\otimes(x_h^{*\prime}x_h^*)^{-1}x_h^{*\prime}]\mathbf{V}_h^* \\ &= x_h^{*\prime}x_h^*\boldsymbol{\Omega}_h^{-1}\hat{\boldsymbol{\eta}}^h \quad (5.52)\end{aligned}$$

where

$$\hat{\boldsymbol{\eta}}^h = [I_n\otimes(x_h^{*\prime}x_h^*)^{-1}x_h^{*\prime}]\mathbf{V}_h^* \quad (5.53)$$

is the OLS estimator for $\boldsymbol{\eta}$ from observations corresponding only to the h-type household. The i^{th} element in $\hat{\boldsymbol{\eta}}$ is given by

$$\hat{\eta}_i^h = (x_h^{*\prime}x_h^*)^{-1}x_h^{*\prime}v_{ih}^* \quad (5.54)$$

Substituting (5.50) and (5.52) into (5.49) yields

$$\left[\sum_{h=1}^{H}(x_h^{*\prime}x_h^*)\boldsymbol{\Omega}_h^{-1}\right]\hat{\boldsymbol{\eta}} = \sum_{h=1}^{H}(x_h^{*\prime}x_h^*)\boldsymbol{\Omega}_h^{-1}\hat{\boldsymbol{\eta}}^h \tag{5.55}$$

or

$$\hat{\boldsymbol{\eta}} = \left[\sum_{h=1}^{H}(x_h^{*\prime}x_h^*)\boldsymbol{\Omega}_h^{-1}\right]^{-1}\sum_{h=1}^{H}(x_h^{*\prime}x_h^*)\boldsymbol{\Omega}_h^{-1}\hat{\boldsymbol{\eta}}^h \tag{5.56}$$

Conditional on $\boldsymbol{\Omega}_h$ and $\boldsymbol{\Theta}_h$, the maximum likelihood estimator for $\boldsymbol{\eta}$ is given by a matrix-weighted average of the h-type household OLS estimators $\hat{\boldsymbol{\eta}}^h$ with weights given by $(x_h^{*\prime}x_h^*)\boldsymbol{\Omega}_h^{-1}$. Maximum likelihood estimators for all the $\boldsymbol{\theta}_h, \boldsymbol{\eta}$ and $\boldsymbol{\Omega}_h$ are given by the simultaneous solution of equations (5.39), (5.46) and (5.56).

5.3.2 An Iterative Estimation Procedure

The results obtained in the last section lead to the following convenient iterative procedure for computing these estimates:

1. Express v_{ih} and x_h in terms of deviations from their household type means. That is, compute $v_{ih}^* = v_{ih} - \overline{v}_{ih}z_h$ and $x_h^* = x_h - \overline{x}_h z_h$, where $\overline{v}_{ih} = M_h^{-1}z_h' v_{ih}$ and $\overline{x}_h = M_h^{-1}z_h' x_h$.

2. Find the least squares estimates $\hat{\eta}_i^h = (x_h^{*\prime}x_h^*)^{-1} x_h^{*\prime}v_{ih}^*$.

3. Find an initial estimate of $\boldsymbol{\Omega}_h$ as[1]

$$\left[\hat{\boldsymbol{\Omega}}_h\right]_{ij} = (v_{ih}^* - x_h^*\hat{\eta}_i^h)'(v_{jh}^* - x_h^*\hat{\eta}_j^h)/M_h \tag{5.57}$$

4. Compute a pooled estimate for $\boldsymbol{\eta}$ as

$$\hat{\boldsymbol{\eta}} = \left[\sum_{h=1}^{H}(x_h^{*\prime}x_h^*)\hat{\boldsymbol{\Omega}}^{-1}\right]^{-1}\sum_{h=1}^{H}(x_h^{*\prime}x_h^*)\hat{\boldsymbol{\Omega}}^{-1}\hat{\boldsymbol{\eta}}^h \tag{5.58}$$

   where $\hat{\boldsymbol{\eta}}^h = (\hat{\eta}_1^h, \hat{\eta}_2^h, ..., \hat{\eta}_n^h)'$.

5. Repeat step (3) with $\hat{\eta}_i^h$ replaced by $\hat{\eta}_i$ that is computed from (5.58).

[1]Note that steps (2) and (3) can be computed at the same time with a seemingly unrelated regression of each of $(v_{ih}^*, v_{2h}^*, ..., v_{nh}^*)$ on x_h^*, with no constant.

6. Repeat steps (4) and (5) until convergence.

7. Compute estimates of the θ_{ih} from

$$\hat{\theta}_{ih} = \overline{v}_{ih} - \overline{x}_h \hat{\eta}_i \tag{5.59}$$

5.4 Derivation of the Asymptotic Covariance Matrices

This section outlines the derivation of the asymptotic covariance matrices for the parameters $\boldsymbol{\theta}_h$ and $\boldsymbol{\eta}$. A second section details the derivation of the asymptotic covariance matrices for b, a_h, b_i and a_{ih}. The last section derives asymptotic variances for the s_{ih} estimators.

5.4.1 Variance Matrices For $\boldsymbol{\theta}_h$ and $\boldsymbol{\eta}$

To specify the asymptotic covariance matrix for the maximum likelihood estimator, the second derivatives of the log-likelihood function specified in (5.32) are required. From (5.34), (5.36) and (5.37), these second derivatives are obtained as follows:

$$\begin{aligned} \frac{\partial^2 \log \mathrm{L}}{\partial \boldsymbol{\Theta}_h \partial \boldsymbol{\Theta}_h'} &= -\mathbf{Z}_h'(\boldsymbol{\Omega}_h^{-1} \otimes I_{M_h})\mathbf{Z}_h \\ &= -M_h \boldsymbol{\Omega}_h^{-1} \end{aligned} \tag{5.60}$$

$$\frac{\partial^2 \log \mathrm{L}}{\partial \boldsymbol{\Theta}_h \partial \boldsymbol{\Theta}_k} = 0 \qquad (h \neq k) \tag{5.61}$$

$$\begin{aligned} \frac{\partial^2 \log \mathrm{L}}{\partial \boldsymbol{\Theta}_h \partial \boldsymbol{\eta}'} &= -\mathbf{Z}_h'(\boldsymbol{\Omega}_h^{-1} \otimes I_{M_h})\mathbf{X}_h \\ &= -M_h \overline{x}_h \boldsymbol{\Omega}_h^{-1} \end{aligned} \tag{5.62}$$

From (5.32) and (5.50), we obtain

$$\begin{aligned}
\frac{\partial^2 \log \mathrm{L}}{\partial \boldsymbol{\eta} \partial \boldsymbol{\eta}'} &= -\sum_{h=1}^{H} \mathbf{X}_h' (\boldsymbol{\Omega}_h^{-1} \otimes I_{M_h}) \mathbf{X}_h \\
&= -\sum_{h=1}^{H} x_h' x_h \boldsymbol{\Omega}_h^{-1} \qquad (5.63)
\end{aligned}$$

It can be shown that the expectation of the cross partial derivatives with respect to θ_{ih} or η_i and the elements of $\boldsymbol{\Omega}_h$, is zero. Thus the information matrix is block diagonal, and, providing the interest is not on the standard errors of the maximum estimator of $\boldsymbol{\Omega}_h$, concern may be confined to the derivatives (5.60)-(5.62).

Specifically, let $\boldsymbol{\Theta}' = (\boldsymbol{\Theta}_1', \boldsymbol{\Theta}_2', ..., \boldsymbol{\Theta}_H', \boldsymbol{\eta}')$, then

$$-\frac{\partial^2 \log \mathrm{L}}{\partial \boldsymbol{\Theta} \partial \boldsymbol{\Theta}'} = \begin{bmatrix} M_1 \boldsymbol{\Omega}_1^{-1} & 0 & \dots & 0 & M_1 \overline{x}_1 \boldsymbol{\Omega}_1^{-1} \\ 0 & M_2 \boldsymbol{\Omega}_2^{-1} & \dots & 0 & M_2 \overline{x}_2 \boldsymbol{\Omega}_2^{-1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & M_H \boldsymbol{\Omega}_H^{-1} & M_H \overline{x}_H \boldsymbol{\Omega}_H^{-1} \\ M_1 \overline{x}_1 \boldsymbol{\Omega}_1^{-1} & M_2 \overline{x}_2 \boldsymbol{\Omega}_2^{-1} & \dots & M_H \overline{x}_H \boldsymbol{\Omega}_H^{-1} & \sum_{h=1}^{H} x_h' x_h \boldsymbol{\Omega}_h^{-1} \end{bmatrix} \qquad (5.64)$$

Since this matrix does not contain any stochastic elements, the information matrix obtained by taking expectations of (5.64) is the same as (5.64). Let

$$\begin{aligned}
D &= \sum_{h=1}^{H} \left[(x_h' x_h - M_h \overline{x}_h^2) \boldsymbol{\Omega}_h^{-1} \right] \qquad (5.65) \\
\overline{x} &= (\overline{x}_1, \overline{x}_2, ..., \overline{x}_H)' \qquad (5.66)
\end{aligned}$$

$$A = \begin{bmatrix} M_1^{-1} \boldsymbol{\Omega}_1 & & & \\ & M_2^{-1} \boldsymbol{\Omega}_2 & & \\ & & \ddots & \\ & & & M_H^{-1} \boldsymbol{\Omega}_H \end{bmatrix} \qquad (5.67)$$

Using results on the partitioned inverse of a matrix and using $V(.)$ to denote the asymptotic covariance matrix, it can be shown that

$$V(\hat{\boldsymbol{\Theta}}) = \left[E - \frac{\partial^2 \log \mathrm{L}}{\partial \boldsymbol{\Theta} \partial \boldsymbol{\Theta}'} \right]^{-1}$$

$$= \begin{bmatrix} A^{-1} & A^{-1}(\overline{x} \otimes I) \\ (\overline{x}' \otimes I)A^{-1} & \sum x_h x_h' \Omega_h^{-1} \end{bmatrix}^{-1} \tag{5.68}$$

$$= \begin{bmatrix} A + \overline{x}\overline{x}' \otimes D^{-1} & -\overline{x} \otimes D^{-1} \\ -\overline{x}' \otimes D^{-1} & D^{-1} \end{bmatrix} \tag{5.69}$$

The relevant variance components from (5.68) are

$$V(\hat{\Theta}_h) = M_h^{-1}\Omega_h + \overline{x}_h^2 D^{-1} \tag{5.70}$$

and

$$V(\hat{\eta}) = D^{-1}. \tag{5.71}$$

5.4.2 Variance Matrices for b, a_h, b_i and a_{ih}

This section provides details of the derivation of the asymptotic covariance matrix of the estimators for the parameters b, a_h, b_i and a_{ih} defined in equations (5.17)-(5.20). From (5.17),

$$V(\hat{b}) = z'D^{-1}z \tag{5.72}$$

where $z = (1, 1, ..., 1)'$. Let $\mathbf{B} = (b_1, b_2, ..., b_n)'$. Then,

$$V(\hat{\mathbf{B}}) = \left(\frac{\partial \mathbf{B}}{\partial \boldsymbol{\eta}'}\right) D^{-1} \left(\frac{\partial \mathbf{B}}{\partial \boldsymbol{\eta}'}\right)' \tag{5.73}$$

Now,

$$\frac{\partial b_i}{\partial \eta_j} = \begin{cases} \frac{1}{b}\left(1 - \frac{\eta_i}{b}\right) & \text{for } i = j \\ \frac{1}{b}\left(-\frac{\eta_i}{b}\right) & \text{for } i \neq j \end{cases} \tag{5.74}$$

Let

$$C = I_n - \frac{1}{b}\boldsymbol{\eta} z' \tag{5.75}$$

From (5.73) and (5.74) it follows that

$$\frac{\partial \mathbf{B}}{\partial \boldsymbol{\eta}'} = \frac{1}{b}C \tag{5.76}$$

and

$$V(\hat{\mathbf{B}}) = \frac{1}{b^2} C D^{-1} C' \tag{5.77}$$

Consider now the covariance matrix for the $\hat{a}_{ih}$. Let $\boldsymbol{\alpha}_h = (a_{1h}, a_{2h}, ..., a_{nh})'$. By definition, we have

$$V(\hat{\boldsymbol{\alpha}}_h) = \begin{bmatrix} \frac{\partial \boldsymbol{\alpha}_h}{\partial \boldsymbol{\Theta}'_h} & \frac{\partial \boldsymbol{\alpha}_h}{\partial \boldsymbol{\eta}'} \end{bmatrix} V \begin{pmatrix} \hat{\boldsymbol{\Theta}}_h \\ \hat{\boldsymbol{\eta}} \end{pmatrix} \begin{bmatrix} \left(\frac{\partial \boldsymbol{\alpha}_h}{\partial \boldsymbol{\Theta}'_h}\right)' \\ \left(\frac{\partial \boldsymbol{\alpha}_h}{\partial \boldsymbol{\eta}'}\right)' \end{bmatrix} \tag{5.78}$$

Now, $\frac{\partial \boldsymbol{\alpha}_h}{\partial \boldsymbol{\Theta}'_h} = C^*$ and $\frac{\partial \boldsymbol{\alpha}_h}{\partial \boldsymbol{\eta}'} = a_h C^*$ where $C^* = I_n + \frac{1}{1-b} \boldsymbol{\eta} z'$. Thus

$$\begin{aligned} V(\hat{\boldsymbol{\alpha}}_h) &= \begin{bmatrix} C^* & a_h C^* \end{bmatrix} \begin{bmatrix} M_h^{-1} \boldsymbol{\Omega}_h + \overline{x}_h^2 D^{-1} & -\overline{x}_h D^{-1} \\ -\overline{x}_h D^{-1} & D^{-1} \end{bmatrix} \begin{bmatrix} C^{*\prime} \\ a_h C^{*\prime} \end{bmatrix} \\ &= C^* [M_h^{-1} \boldsymbol{\Omega}_h + (\overline{x}_h - a_h)^2 D^{-1}] C^{*\prime} \end{aligned} \tag{5.79}$$

Noting that $\hat{a}_h = z' \hat{a}_{ih}$, we also have

$$V(\hat{a}_h) = z' C^* [M_h^{-1} \boldsymbol{\Omega}_h + (\overline{x}_h - a_h)^2 D^{-1}] C^{*\prime} z \tag{5.80}$$

5.4.3 Variance Expressions for the $\hat{s}_{ih}$'s

By definition, the commodity specific scales are ratios of the subsistence expenditures of some household h and the reference household r. Thus,

$$\hat{s}_{ih} = \frac{\hat{a}_{ih}}{\hat{a}_{ir}} \tag{5.81}$$

In this regard, the following expression for the variance of the commodity-specific scales $V(\hat{s}_{ih})$ is obtained

$$\begin{aligned} V(\hat{s}_{ih}) &= \left(\frac{\partial s_{ih}}{\partial a_{ih}}\right)^2 V(\hat{a}_{ih}) + \left(\frac{\partial s_{ih}}{\partial a_{ir}}\right)^2 V(\hat{a}_{ir}) + 2 \left(\frac{\partial s_{ih}}{\partial a_{ih}}\right) \left(\frac{\partial s_{ih}}{\partial a_{ir}}\right) \operatorname{cov}(\hat{a}_{ih}, \hat{a}_{ir}) \\ &= \frac{1}{a_{ih}^2} V(\hat{a}_{ih}) + \frac{a_{ih}^2}{a_{ir}^4} V(\hat{a}_{ir}) - \frac{2a_{ih}}{a_{ir}^3} \operatorname{cov}(\hat{a}_{ih}, \hat{a}_{ir}) \end{aligned} \tag{5.82}$$

The elements $V(\hat{a}_{ih})$ and $V(\hat{a}_{ir})$ are given by the appropriately selected diagonal element in (5.78) and its counterpart for the reference household. The elements $\text{cov}(\hat{a}_{ih}, \hat{a}_{ir})$ are given by the diagonal elements of

$$
\begin{aligned}
\text{cov}(\hat{\boldsymbol{\alpha}}_h, \hat{\boldsymbol{\alpha}}_r) &= \begin{bmatrix} \frac{\partial \boldsymbol{\alpha}_r}{\partial \boldsymbol{\Theta}_h'} & \frac{\partial \boldsymbol{\alpha}_r}{\partial \boldsymbol{\Theta}_r'} & \frac{\partial \boldsymbol{\alpha}_r}{\partial \boldsymbol{\eta}'} \end{bmatrix} V \begin{pmatrix} \hat{\boldsymbol{\Theta}}_h \\ \hat{\boldsymbol{\Theta}}_r \\ \hat{\boldsymbol{\eta}} \end{pmatrix} \begin{bmatrix} \left(\frac{\partial \boldsymbol{\alpha}_h}{\partial \boldsymbol{\Theta}_h'}\right)' \\ \left(\frac{\partial \boldsymbol{\alpha}_h}{\partial \boldsymbol{\Theta}_r'}\right)' \\ \left(\frac{\partial \boldsymbol{\alpha}_h}{\partial \boldsymbol{\eta}'}\right)' \end{bmatrix} \\
&= \begin{bmatrix} 0 & C^* & a_r C^* \end{bmatrix} V \begin{pmatrix} \hat{\boldsymbol{\Theta}}_h \\ \hat{\boldsymbol{\Theta}}_r \\ \hat{\boldsymbol{\eta}} \end{pmatrix} \begin{bmatrix} C^{*\prime} \\ 0 \\ a_h C^{*\prime} \end{bmatrix} \\
&= \begin{bmatrix} 0 & C^* & a_h C^* \end{bmatrix} \begin{bmatrix} M_h^{-1}\boldsymbol{\Omega}_h + \bar{x}_h^2 D^{-1} & \bar{x}_h \bar{x}_r D^{-1} & -\bar{x}_h D^{-1} \\ \bar{x}_h \bar{x}_r D^{-1} & M_h^{-1}\boldsymbol{\Omega}_r + \bar{x}_r^2 D^{-1} & -\bar{x}_r D^{-1} \\ -\bar{x}_h D^{-1} & \bar{x}_r D^{-1} & D^{-1} \end{bmatrix} \begin{bmatrix} C^{*\prime} \\ 0 \\ a_h C^{*\prime} \end{bmatrix} \\
&= (\bar{x}_h - a_h)(\bar{x}_r - a_r) C^* D^{-1} C^{*\prime} \qquad (5.83)
\end{aligned}
$$

5.5 Empirical Application

The data used in this study are derived from the 1988-89 Household Expenditure Survey (HES) conducted by the Australian Bureau of Statistics between the period July 1988 to July 1989. This is described in detail in Chapter 3 of this study. The iterative procedure described in Section 5.3.2 was applied to the 5532 sample households grouped into eight household types (see Table 3.2) and using all the eleven expenditure categories described in Chapter 3.

Table 5.1 presents the parameter estimates of the extended linear expenditure system. The iterative process (Steps 4 and 5) converged on the 5th iteration and yielded the η_i estimates found in the 2nd column. The other columns present the estimates of θ_{ih} corresponding to each commodity group and household type. The table also provides estimates of the asymptotic standard errors of these parameters from equations (5.69) and (5.71). These estimates of η_i and θ_{ih} do not carry a direct economic interpretation but are important to the procedure as they lead to the estimation of the marginal propensities and subsistence expenditures which are presented in Table 5.2. The standard errors are all relatively small, except perhaps for household type (1,3) where the number of households of that type is small.

In Table 5.2, the 2nd column provides the marginal budget shares b_i and the 3rd through to the 10th columns give the estimates of subsistence expenditures a_{ih} for each expenditure category. In general, the subsistence expenditures increase with household size, with wider differentials occuring across two-adult households compared to one-adult households. For all household types, expenditure on food was on top of the shopping list, followed closely by housing, then transport, then household furnishings. Together, these items make up between 62 to 67 percent of subsistence expenditures of a typical Australian household.

Table 5.3 presents the estimates of commodity-specific scales defined in equation (5.21). A two-adult household with no children is chosen to be the reference household for which s_{ih} is set to 1. For most commodities, the scales increased with the increase in the household size. These increases are observed to occur at a decreasing rate indicating economies of scale for additional children. After the first child there exists strong economies of scale for additional children, particularly for expenditures towards housing, food and household furnishings. There are some exceptions to these observations. The magnitude of the scales Alcohol and Tobacco, for instance, declines as the number of children in the household increases. Also, the scales for Medical and Health Care and Others commodity groups exhibit no defined trend for one-adult households. A more thorough investigation of expenditure patterns of households may be required for us to provide definitive explanations for such deviations but one possibility is that the presence of children in the household tends to influence expenses away from 'adult goods' under which alcohol, tobacco and many other miscellaneous goods are classified.

Has there been significant changes in the scale relativities over time? Information from Tables 5.4 and 5.5 provide some answers. In Table 5.4, scale estimates calculated from the Binh and Whiteford (1990) results, that used the 1984 data, are presented and compared with our results that used the 1988 data. Also, since it could be argued that a difference in results may be attributable to the new maximum likelihood estimation procedure, rather than a change in consumption patterns, estimates from the 1988 data set obtained using Kakwani's estimation procedure (the procedure used by Binh and Whiteford (1990)) are also presented. The two sets of 1988 scales are very similar with no one method exhibiting consistently higher or lower values. The estimated standard errors for both sets (not shown) show more divergence, but again do not display any consistent over or underestimation.

There are noticeable changes between the 1984 and 1988 scales. For the one-adult households with children, the estimated scales appear to have mostly decreased over time. In contrast, the scales for the (1,0) household type have increased mostly. For the two-adult households with children, results were mixed. The direction for the change were the same for the (2,1) and (2,2) household types while change in the scale values for the (2,3) type tend to be in the other direction. Interestingly, the only consistent (direction of) change for all household types was observed for Alcohol and Tobacco. For this commodity, scales decreased significantly from 1984 to 1988. The 1988 trend seem to imply that the presence of children has a deterring effect on the consumption of alcohol and tobacco which (from the results above) was not the case in 1984. Many would argue against this interpretation but it seems plausible.

Also, while relative costs for Alcohol and Tobacco increased with the increase in household size in 1984, this trend is reversed in 1988. It is also noted that there are fewer economies of scale in housing in the later data set, but greater economies of scale in food. The largest differences in the scale estimates occurred in the one-adult, three-children household groups. Since the number of households in this group is relatively small, and the standard errors of the estimated scales are relatively high, these differences may reflect sampling error.

The interpretations made above could be misleading. It is recognised that a four-year period may be too short for any significant changes in behavioural patterns to occur with such regularity. It is possible that price level changes may have accounted for a change in the results. Without price movements being incorporated in the analyses, a comparison of estimated scales at two different time points in time must be made with caution.

The general scales computed from equation (5.28) are presented in Table 5.5.

Because these scales depend on income x_r, they are computed for three income levels, the same levels as utilised by Binh and Whiteford (1990). Also presented in Table 5.5 are three estimates of each scale - the Binh and Whiteford 1984 estimates, the 1988 estimates using Kakwani's estimation procedure and the 1988 estimates using maximum likelihood estimation. There is virtually no difference between the two sets of 1988 estimates. Also, no sensitivity to the level of income is detected. Comparing the 1984 estimates with the 1988 estimates, we find the results for the two-adult families are quite similar, although Binh and Whiteford's conclusion that "there is strong evidence of economies of scale in the second child but adding the third child increases these households' needs considerably" no longer holds. For the 1988 data, adding the third child was only slightly more expensive than adding the second child. For one-adult families, the noticeable differences are an increase in the relative cost when there are no children, and a decrease in the relative cost when children are present.

Table 5.1 Parameters Estimates of the Extended Linear Expenditure System

Commodity Type	η_i	b_i	θ_{ih} *Household Type (no. of adults, no. of children)* (1,0)	(1,1)	(1,2)	(1,3)	(2,0)	(2,1)	(2,2)	(2,3)
Housing	0.0577 (0.0024)	0.1721 (0.0041)	31.4384 (1.5904)	39.2799 (3.4719)	44.0618 (4.0361)	48.5558 (8.1841)	31.6131 (2.0571)	52.8594 (3.9800)	52.1784 (3.4128)	56.3839 (4.1472)
Fuel & Power	0.0033 (0.0003)	0.0097 (0.0004)	6.8267 (0.1876)	9.4220 (0.5191)	10.9337 (0.6032)	11.3064 (1.2418)	9.9670 (0.2480)	12.0888 (0.4113)	13.4061 (0.3652)	14.3140 (0.4972)
Food	0.0347 (0.0013)	0.1035 (0.0023)	32.4681 (0.8173)	45.5670 (2.5353)	61.0741 (2.9826)	68.2695 (6.7713)	62.6246 (1.1152)	77.4861 (2.0088)	90.4438 (1.9529)	100.2526 (2.6030)
Alcohol & Tobacco	0.0115 (0.0091)	0.0342 (0.0014)	10.0246 (0.6507)	6.9719 (1.0621)	4.9411 (1.2183)	3.4686 (1.3293)	17.5799 (0.8760)	15.4021 (1.3746)	13.0945 (1.3600)	10.2715 (1.1911)
Clothing & Footwear	0.0235 (0.0013)	0.0701 (0.0022)	5.2716 (0.8410)	11.4805 (2.5512)	10.9911 (2.5906)	19.6190 (5.0014)	10.7708 (1.1941)	14.1915 (1.9730)	15.7439 (1.8881)	19.4025 (2.4064)
Household Furnishings & Equipment	0.0443 (0.0028)	0.1322 (0.0047)	16.8705 (1.6025)	20.1199 (3.2014)	24.0804 (4.6645)	24.1203 (8.6138)	31.9543 (2.5944)	50.1229 (5.7405)	33.7267 (4.0231)	39.9537 (4.5020)
Medical & Health Care	0.0126 (0.0007)	0.0376 (0.0012)	6.5825 (0.7026)	4.6592 (1.2638)	7.9275 (1.9084)	4.4214 (2.2170)	12.6566 (0.6276)	16.1081 (1.3418)	15.9796 (0.8806)	15.8738 (1.1131)
Transport	0.0455 (0.0031)	0.1358 (0.0051)	23.1214 (2.1326)	23.5812 (4.0892)	25.4170 (5.1102)	33.4503 (13.1716)	46.0910 (2.8519)	42.9850 (4.3130)	52.5040 (4.2381)	61.3162 (7.1711)
Recreation & Entertainment	0.0585 (0.0027)	0.1745 (0.0047)	12.5837 (1.5089)	11.2043 (3.7386)	6.2985 (2.8249)	18.6823 (7.6083)	25.3231 (2.6473)	21.2988 (4.1107)	31.4789 (4.4664)	32.5454 (5.5580)
Personal Care	0.0052 (0.0004)	0.0156 (0.0006)	2.9486 (0.2490)	4.5084 (0.6603)	5.7927 (1.0832)	3.7391 (0.9797)	5.6180 (0.3719)	6.6150 (0.6447)	7.1597 (0.5324)	6.0180 (0.6224)
Others	0.0385 (0.0018)	0.1150 (0.0031)	4.0790 (1.2281)	12.3671 (3.1177)	8.0946 (2.0378)	4.4807 (2.5730)	7.6374 (1.6269)	13.1176 (2.5048)	20.7703 (3.4253)	25.2940 (5.0650)

Note: The estimated standard errors are in parentheses.

Table 5.2 Parameter Estimates of Marginal Propensities and Subsistence Expenditures

		Subsistence Expenditures (a_{ih}'s) Household Type (no. of adults, no. of children)							
Commodity Type	b_i	*(1,0)*	*(1,1)*	*(1,2)*	*(1,3)*	*(2,0)*	*(2,1)*	*(2,2)*	*(2,3)*
Housing	0.1721	44.6461	55.6935	62.2498	69.3904	54.3326	80.8232	82.4249	89.4976
	(0.0041)	(1.6292)	(3.9026)	(4.4800)	(9.9430)	(1.8154)	(4.1615)	(3.4164)	(4.4639)
Fuel & Power	0.0097	7.5720	10.3483	11.9601	12.4822	11.2492	13.6669	15.1027	16.1827
	(0.0004)	(0.1717)	(0.5242)	(0.6003)	(1.2544)	(0.1994)	(0.3794)	(0.3524)	(0.4757)
Food	0.1035	40.4086	55.4349	72.0089	80.7954	76.2837	94.2910	108.5187	120.1607
	(0.0023)	(0.8475)	(2.8826)	(3.3039)	(7.6824)	(1.0129)	(2.1782)	(2.0471)	(2.8829)
Alcohol & Tobacco	0.0342	12.6534	10.2387	8.5610	7.6153	22.1017	20.9678	19.0782	16.8621
	(0.0014)	(0.6188)	(1.0512)	(1.2132)	(1.4246)	(0.7419)	(1.2963)	(1.2491)	(1.0574)
Clothing & Footwear	0.0701	10.6531	18.1682	18.4018	28.1081	20.0279	25.5854	27.9937	32.8947
	(0.0022)	(0.8219)	(2.7734)	(2.7713)	(5.7923)	(1.0271)	(1.9432)	(1.7799)	(2.5240)
Household Furnishings	0.1322	27.0147	32.7264	38.0498	40.1224	49.4040	71.6005	56.8178	65.3866
& Equipment	(0.0047)	(1.5508)	(3.5949)	(5.0889)	(9.7593)	(2.2763)	(5.9201)	(3.9197)	(4.5354)
Medical & Health Care	0.0376	9.4661	8.2428	11.8985	8.9702	17.6169	22.2135	22.5436	23.1035
	(0.0012)	(0.6868)	(1.3382)	(1.9893)	(2.4079)	(0.5132)	(1.3274)	(0.7944)	(1.0619)
Transport	0.1358	33.5432	36.5327	39.7687	49.8903	64.0183	65.0505	76.2271	87.4452
	(0.0051)	(2.1086)	(4.4341)	(5.5586)	(14.7220)	(2.4982)	(4.2747)	(3.9894)	(7.4524)
Recreation	0.1745	25.9738	27.8446	24.7377	39.8047	48.3563	49.6488	61.9587	66.1164
& Entertainment	(0.0047)	(1.4963)	(4.3098)	(3.1839)	(9.7877)	(2.4763)	(4.3307)	(4.6053)	(6.0698)
Personal Care	0.0156	4.1486	5.9926	7.4373	5.6231	7.6723	9.1436	9.8783	9.0123
	(0.0006)	(0.2291)	(0.6766)	(1.0986)	(1.0526)	(0.3116)	(0.6180)	(0.4776)	(0.5997)
Others	0.1150	12.9057	23.3362	20.2497	18.4044	22.8208	31.8058	40.8624	47.4238
	(0.0031)	(1.2465)	(3.4371)	(2.2069)	(3.1771)	(1.4203)	(2.5751)	(3.5113)	(5.3230)
Total	1.0001	228.9851	284.5588	315.3233	361.2065	393.8838	484.7969	521.4062	574.0855

Note: The estimated standard errors are in parentheses.

Table 5.3 Estimates of Commodity-Specific Scales

	Commodity Specific Scales							
	Household Type (no. of adults, no. of children)							
Commodity Type	*(1,0)*	*(1,1)*	*(1,2)*	*(1,3)*	*(2,0)*	*(2,1)*	*(2,2)*	*(2,3)*
Housing	0.82	1.03	1.15	1.28	1.00	1.49	1.52	1.65
	(0.04)	(0.08)	(0.09)	(0.19)	(0.00)	(0.09)	(0.08)	(0.10)
Fuel & Power	0.67	0.92	1.06	1.11	1.00	1.21	1.34	1.44
	(0.02)	(0.05)	(0.06)	(0.11)	(0.00)	(0.04)	(0.04)	(0.05)
Food	0.53	0.73	0.94	1.06	1.00	1.24	1.42	1.58
	(0.01)	(0.04)	(0.05)	(0.10)	(0.00)	(0.03)	(0.03)	(0.04)
Alcohol & Tobacco	0.57	0.46	0.39	0.34	1.00	0.95	0.86	0.76
	(0.03)	(0.05)	(0.06)	(0.07)	(0.00)	(0.07)	(0.06)	(0.05)
Clothing & Footwear	0.53	0.91	0.92	1.40	1.00	1.28	1.40	1.64
	(0.05)	(0.15)	(0.15)	(0.30)	(0.00)	(0.11)	(0.11)	(0.15)
Household Furnishings	0.55	0.66	0.77	0.81	1.00	1.45	1.15	1.32
& Equipment	(0.04)	(0.08)	(0.11)	(0.20)	(0.00)	(0.14)	(0.09)	(0.11)
Medical & Health Care	0.54	0.47	0.68	0.51	1.00	1.26	1.28	1.31
	(0.04)	(0.08)	(0.11)	(0.14)	(0.00)	(0.08)	(0.06)	(0.07)
Transport	0.52	0.57	0.62	0.78	1.00	1.02	1.19	1.37
	(0.04)	(0.07)	(0.09)	(0.23)	(0.00)	(0.08)	(0.08)	(0.13)
Recreation	0.54	0.58	0.51	0.82	1.00	1.03	1.28	1.37
& Entertainment	(0.04)	(0.09)	(0.07)	(0.21)	(0.00)	(0.10)	(0.11)	(0.14)
Personal Care	0.54	0.78	0.97	0.73	1.00	1.19	1.29	1.17
	(0.04)	(0.09)	(0.15)	(0.14)	(0.00)	(0.09)	(0.08)	(0.09)
Others	0.57	1.02	0.89	0.81	1.00	1.39	1.79	2.08
	(0.06)	(0.16)	(0.11)	(0.15)	(0.00)	(0.14)	(0.19)	(0.27)

Note: The estimated standard errors are in parentheses.

Table 5.4 A Comparison of Commodity-Specific Scales

		Commodity-Specific Scales (S_{ih}) Household Type (no. of adults, no. of children)							
Commodity Type	*Year**	*(1,0)*	*(1,1)*	*(1,2)*	*(1,3)*	*(2,0)*	*(2,1)*	*(2,2)*	*(2,3)*
Housing	1984	0.80	1.22	1.23	1.24	1.00	1.51	1.52	1.51
	1988a	0.80	0.97	1.09	1.20	1.00	1.47	1.50	1.60
	1988b	0.82	1.03	1.15	1.28	1.00	1.49	1.52	1.65
Fuel & Power	1984	0.61	0.99	1.02	1.38	1.00	1.23	1.32	1.50
	1988a	0.67	0.90	1.04	1.08	1.00	1.19	1.34	1.42
	1988b	0.67	0.92	1.06	1.11	1.00	1.21	1.34	1.44
Food	1984	0.51	0.65	0.95	1.26	1.00	1.17	1.40	1.61
	1988a	0.53	0.71	0.93	1.04	1.00	1.23	1.42	1.56
	1988b	0.53	0.73	0.94	1.06	1.00	1.24	1.42	1.58
Alcohol & Tobacco	1984	0.58	0.64	0.74	0.80	1.00	1.15	1.11	1.15
	1988a	0.57	0.46	0.39	0.34	1.00	0.95	0.87	0.76
	1988b	0.57	0.46	0.39	0.34	1.00	0.95	0.86	0.76
Clothing & Footwear	1984	0.38	0.94	1.46	2.18	1.00	1.15	1.32	1.68
	1988a	0.53	0.89	0.90	1.38	1.00	1.27	1.39	1.63
	1988b	0.53	0.91	0.92	1.40	1.00	1.28	1.40	1.64
Household Furnishings	1984	0.46	0.85	1.05	1.24	1.00	1.14	1.20	1.37
& Equipment	1988a	0.56	0.71	0.83	0.89	1.00	1.48	1.14	1.37
	1988b	0.55	0.66	0.77	0.81	1.00	1.45	1.15	1.32
Medical &	1984	0.47	0.33	0.43	0.63	1.00	1.15	1.20	1.36
Health Care	1988a	0.53	0.44	0.64	0.47	1.00	1.25	1.28	1.28
	1988b	0.54	0.47	0.68	0.51	1.00	1.26	1.28	1.31
Transport	1984	0.46	0.62	0.75	1.19	1.00	1.34	1.16	1.32
	1988a	0.53	0.61	0.66	0.85	1.00	1.01	1.19	1.41
	1988b	0.52	0.57	0.62	0.78	1.00	1.02	1.19	1.37
Recreation	1984	0.53	0.62	0.68	1.01	1.00	0.92	1.15	1.19
& Entertainment	1988a	0.54	0.61	0.54	0.88	1.00	1.03	1.28	1.40
	1988b	0.54	0.58	0.51	0.82	1.00	1.03	1.28	1.37
Personal Care	1984	0.62	0.69	1.33	1.32	1.00	1.19	1.13	1.22
	1988a	0.54	0.76	0.95	0.71	1.00	1.19	1.29	1.17
	1988b	0.54	0.78	0.97	0.73	1.00	1.19	1.29	1.17
Others	1984	0.45	1.34	0.99	1.98	1.00	1.19	1.26	1.86
	1988a	0.57	1.04	0.90	0.82	1.00	1.40	1.80	2.10
	1988b	0.57	1.02	0.89	0.81	1.00	1.39	1.79	2.08

*1984 Scales derived using 1984 unit record data from Binh and Whiteford (1990).
1988a Scales derived using the Kakwani procedure applied to 1988 HES survey data (own calculations).
1988b Scales derived using the proposed MLE procedure applied to 1988 Household Expenditure Survey (own calculations).

Table 5.5 Estimates of General Scales

*Reference Income***	*Year**	*General Scales (S_h)*							
		Household Type (no. of adults, no. of children)							
		(1,0)	*(1,1)*	*(1,2)*	*(1,3)*	*(2,0)*	*(2,1)*	*(2,2)*	*(2,3)*
Low Income ($325 p.w.)	1984	0.53	0.80	0.95	1.27	1.00	1.20	1.28	1.44
	1988a	0.59	0.73	0.82	0.94	1.00	1.23	1.33	1.46
	1988b	0.58	0.72	0.81	0.92	1.00	1.23	1.32	1.45
Medium Income ($450 p.w.)	1984	0.52	0.81	0.94	1.28	1.00	1.20	1.27	1.44
	1988a	0.58	0.73	0.80	0.92	1.00	1.23	1.32	1.47
	1988b	0.58	0.72	0.80	0.91	1.00	1.23	1.33	1.46
High Income ($700 p.w.)	1984	0.52	0.81	0.94	1.29	1.00	1.19	1.26	1.45
	1988a	0.58	0.73	0.78	0.92	1.00	1.23	1.32	1.47
	1988b	0.58	0.72	0.79	0.91	1.00	1.23	1.33	1.47

* 1984 Scales reprinted from Binh and Whiteford (1990) which used 1984 Household Expenditure Survey data.
1988 a Scales derived using the Kakwani procedure applied to data from 1988 Household Expenditure Survey (own calculations).
1988b Scales derive using the proposed MLE procedure applied to data from 1988 Household Expenditure Surve (own calculations).
**The scales have been evaluated using the listed incomes as reference levels.

Chapter 6

A Bayesian Approach to the Estimation of Equivalence Scales

In the last chapter, the procedure used in the estimation of parameters was derived within the sampling theory framework of analysis. In this section, an alternative approach based on Bayes Theorem is proposed and used. Certain features of the Bayesian approach to inference make it attractive for the model considered here. First, the approach is known for its ability to provide finite sample inference procedures in many instances where only asymptotic inference is available from a sampling theory framework. As seen in Chapter 5, the model considered here has the original demand parameters (a_{ih} and b_i) expressed as non-linear functions of the estimated parameters (θ_{ih} and η_i). The commodity-specific scales s_{ih} are likewise non-linear functions of the a_{ih}; and the general scales s_h are complicated functions of a number of the parameters. These nonlinearities meant that considerable effort had to be put into deriving asymptotic covariance matrices for the various sampling theory estimators in Chapter 5. In this chapter, a Bayesian procedure that describes parameter uncertainty in a more straightforward and less cumbersome manner is derived. For the general scales, Bayesian estimators are shown to allow for more complete inference than was possible within the sampling theory framework.

Another advantage of the Bayesian approach is its ability to include pre-sample or "prior" information about parameters by way of a formal framework. In this chapter, non-informative priors are used and, in that sense, it can be argued that this work does not avail of the advantage that Bayesian analysis has in terms of including prior information. This is not because no prior information is available. Indeed, subsistence parameters should be less than the corresponding minimum consumption in the sample, and prior information about the relative magnitudes of the various scales exists. Specifying non-informative priors does, however, provide a first step towards Bayesian analysis of the model, from which more sophisticated future research involving informative priors (inequality restrictions) could emanate. Research within this context, but for a different specification, has been carried out by Griffiths and Chotikapanich (1996, 1997).

A further motivation for the Bayesian approach in this chapter is that it provides a first step towards the development of a methodology to cover the infrequency of purchase problem (Chapter 7) that is too difficult to handle within the sampling theory framework. Finally, for some, presenting results in terms of a subjective (post-sample) distribution is more natural and meaningful than presenting values from estimators that have good pre-sample but not necessarily good post-sample properties (as in the sampling theory approach).

6.1 Bayesian Inference and Bayes Theorem

There is a lively debate in the literature about the relative merits of the sampling theory and Bayesian approaches to inference. See, for example, Zellner (1988). Griffiths, Hill and Judge (1993, Chp. 24) explain in detail the distinction between these two approaches. In conventional sampling theory, inferences are made based on parameter estimates which are evaluated in terms of their performance in repeated samples. Accordingly, an estimate that is unbiased implies that in a large number of hypothetical samples, the average value of the estimates produced by that estimator converges to the true unknown parameter value. A minimum variance unbiased (mvu) estimator is so desired because, relative to estimates provided by any other unbiased estimator, mvu estimates vary less around the true unknown parameter value. To a sampling theorist, the concept of probability of an event is defined in terms of its frequency of occurrence in repeated identical experiments.

Consequently, the precision of a statistical estimation technique used within this framework is evaluated based on its long-term accuracy over repeated samples. Such is the case for the methods of least squares and maximum likelihood estimation. Also estimation techniques used within this framework do not normally involve the formal consideration of non-sample information. The variation in the observed variables is, for the most part, assumed to contain all the information about the unknown parameters of interest.

In the Bayesian framework, probability is a concept that expresses an individual's belief on how likely or unlikely a particular event is to occur. This belief may depend on quantitative or qualitative information or both but it does not necessarily depend on the relative frequency of the event in a large number of future hypothetical experiments. As such, different individuals are allowed to assign different probabilities to the same event, and hence, the association of this notion of 'subjective' probability with Bayesian procedures. A further consequence is that the uncertainty about the value of an unknown parameter can be expressed in terms of a probability distribution; this is one of the main features of Bayesian analysis.

The initial step to presenting and analysing information about a set of unknown parameters $\boldsymbol{\theta}= (\theta_1, \theta_2, ..., \theta_k)$ using the Bayesian approach is the construction of a prior probability density function (pdf), $g(\boldsymbol{\theta})$. This prior pdf summarises all the researcher's information about $\boldsymbol{\theta}$ before the sample is observed. Such prior information would typically be based on economic theory, past studies, or both. Some data relevant to $\boldsymbol{\theta}$ is then collected. Subsequently, consider now the sample information vector

$$y = (y_1, y_2, ..., y_n)$$

which has a probability density function that depends on $\boldsymbol{\theta}$ in some known manner denoted by

$$f(y \mid \boldsymbol{\theta}).$$

What is desired at this point is to combine prior knowledge contained in $g(\boldsymbol{\theta})$ with that of information from the sample $f(y \mid \boldsymbol{\theta})$ to form an expression for $\boldsymbol{\theta}$ known as the posterior probability density function (posterior pdf for short). A posterior pdf summarises all the sample information about $\boldsymbol{\theta}$ after the sample y is taken.

Bayes' theorem for random variables (and vectors) is a mathematical rule that enables us to do just that. Bayes' theorem states that

$$g(\boldsymbol{\theta} \mid y) = \frac{f(y \mid \boldsymbol{\theta})g(\boldsymbol{\theta})}{f(y)} \tag{6.1}$$

where $g(\boldsymbol{\theta} \mid y)$ is the posterior pdf for $\boldsymbol{\theta}$ and $g(\boldsymbol{\theta})$ is the prior density for $\boldsymbol{\theta}$. Here, it is noted that the function $f(y \mid \boldsymbol{\theta})$ is algebraically identical to the likelihood function for $\boldsymbol{\theta}$. Accordingly, if we regard $f(y)$ as a constant, and let $f(y \mid \boldsymbol{\theta})$ be written as the likelihood function $l(\boldsymbol{\theta} \mid y)$, then (6.1) becomes

$$g(\boldsymbol{\theta} \mid y) \propto l(\boldsymbol{\theta} \mid y)g(\boldsymbol{\theta}) \tag{6.2}$$

where $\propto$ denotes 'proportional to'. In simpler form, Bayes theorem in (6.2) can be written as

$$\text{posterior} \propto \text{sample} \times \text{priori}$$

This relationship summarises the way in which Bayesians modify their beliefs in order to take into account the available non-sample information. This feature of the Bayesian approach to estimation is particularly appealing to economists who frequently have prior information about the parameters in their economic relationships.

One of the major obstacles to the application of Bayesian econometrics has been the perceived need for numerical integration. A common way to present results from a Bayesian investigation is to provide graphs of marginal posterior pdfs for each of the unknown parameters of interest. Instead of presenting a point estimate and its corresponding standard error as traditionally done in the sampling theory approach, Bayesians summarise information about an unknown parameter in terms of a pdf that describes how likely (in a subjective probabilistic sense) different values of the parameter are. See Griffiths, et.al. (1993) for an introductory exposition. In order to present information in this way, it is necessary to derive marginal posterior pdfs for each of the unknown parameters from the joint posterior pdf for all the parameters. This process means that unwanted parameters must be integrated out of the joint posterior pdf. If analytical integration is not possible, or the dimension of the joint posterior pdf is greater than three, making numerical integration impractical, some other solution must be found. One such solution, which has led to an enormous explosion in Bayesian literature over the last five years, is the use of Markov Chain Monte Carlo (MCMC) techniques. These techniques provide a way of drawing observations from the joint posterior pdf. Once some observations have been drawn, they can be used to construct histograms as estimates of the marginal posterior pdfs. Since observations are drawn artificially using computer software, we can make estimated marginal posterior pdfs as accurate as we like, by drawing as many observations as are required. Marginal posterior pdfs are not the only way of presenting information that utilises MCMC-estimated integrals. Posterior means and standard deviations which are the Bayesian counterparts of sampling theory point estimates and standard errors often take the form of intractable integrals. These quantities can be readily estimated using the sample means and standard deviations of the MCMC-generated observations. There are two main MCMC techniques: Gibbs sampling and the Metropolis-Hastings (M-H) algorithm. Introductory expositions to these numerical algorithms are found in Casella and George (1992) and Chib and Greenberg (1995a). And for an appreciation of the wide variety of applications of these algorithms, see Albert and Chib (1996), Chib and Greenberg (1996), Tanner (1993) and Gelfand and Sfiridis (1996).

In this chapter and the next, the availability of the Gibbs sampling and M-H algorithm paves the way for the application of Bayesian econometrics to our equivalence scale estimation problem. Gibbs sampling is employed in this chapter while the M-H algorithm is combined with the Gibbs sampling procedure in the next chapter. In the meantime, we now focus on the application of Bayesian estimation techniques to the n-equation linear seemingly unrelated regression system which was analysed in Chapter 5. The MLE-based results from that chapter will be compared with the results derived from this chapter using Bayesian principles.

6.2 The Model and Bayesian Estimation

The model used in this chapter is the extended linear expenditure system which was discussed previously in section 5.1.1 of Chapter 5. The following detailed presentation of the Bayesian solution to the same equivalence scale estimation problem starts off by considering the following set of n-equation linear seemingly unrelated regression (SUR) system (first presented in Chapter 5 as equations (5.29) and (5.30)):

$$\begin{bmatrix} v_{1h} \\ v_{2h} \\ \vdots \\ v_{nh} \end{bmatrix} = \begin{bmatrix} z_h & & & \\ & z_h & & \\ & & \ddots & \\ & & & z_h \end{bmatrix} \begin{bmatrix} \theta_{1h} \\ \theta_{2h} \\ \vdots \\ \theta_{nh} \end{bmatrix} + \begin{bmatrix} x_h & & & \\ & x_h & & \\ & & \ddots & \\ & & & x_h \end{bmatrix} \begin{bmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_n \end{bmatrix} + \begin{bmatrix} e_{1h} \\ e_{2h} \\ \vdots \\ e_{nh} \end{bmatrix}$$

or

$$\mathbf{V}_h = \mathbf{Z}_h \mathbf{\Theta}_h + \mathbf{X}_h \boldsymbol{\eta} + \mathbf{E}_h \tag{6.3}$$

where

h	$= 1, 2, ..., H$ refers to household composition type h;
n	refers to the number of commodity groups;
v_{ih}	is an $(M_h \times 1)$ vector of observations on expenditure for the i^{th} commodity and the h-type household;
z_h	is an $(M_h \times 1)$ vector of ones;
x_h	is an $(M_h \times 1)$ vector of observations on income for households of type h;
e_{ih}	is an $(M_h \times 1)$ vector of errors;
$\boldsymbol{V}_h$	is of dimension $(nM_h \times 1)$;
$\boldsymbol{Z}_h$	$= I_n \otimes z_h$ is an $(nM_h \times n)$ matrix of dummy variables;
$\boldsymbol{X}_h$	$= I_n \otimes x_h$ is an $(nM_h \times n)$ vector of household incomes;
$\mathbf{\Theta}_h, \boldsymbol{\eta}$	are $(n \times 1)$ vectors of unknown parameters;
$\boldsymbol{E}_h$	is an $(nM_h \times 1)$ vector of errors which is assumed

to be distributed as

$$\boldsymbol{E}_h \sim N[0, \boldsymbol{\Omega}_h \otimes I_{M_h}] \tag{6.4}$$

where $\boldsymbol{\Omega}_h$ is a $(n \times 1)$ error covariance matrix.

Application of the Bayesian approach to estimation begins with specification of a joint prior distribution for the unknown parameters. Traditional non-informative priors for the SUR model are specified. Non-informative priors ignore the existence of what could be considerable prior information, but they do have the advantage of objective data-based reporting of posterior information.

The following generic notation will be useful:

$$\begin{aligned}
\boldsymbol{\Omega} &= \{\boldsymbol{\Omega}_h \mid \text{ for all } h\} \\
\boldsymbol{\Theta} &= \{\boldsymbol{\Theta}_h \mid \text{ for all } h\ \} \\
\mathbf{V} &= \{\mathbf{V}_h \mid \text{ for all } h\ \}
\end{aligned}$$

The traditional non-informative or diffuse prior (Judge, et.al. 1985, p.478) for $\boldsymbol{\Omega}_h$ is

$$g(\boldsymbol{\Omega}_h) \propto \mid \boldsymbol{\Omega}_h \mid^{-\frac{n+1}{2}} \tag{6.5}$$

Treating all the $\boldsymbol{\Omega}_h$ as *a priori* independent, the combined prior pdf for all error covariance matrices is

$$g(\boldsymbol{\Omega}) \propto \prod_{h=1}^{H} \boldsymbol{\Omega}_h^{-\frac{n+1}{2}} \tag{6.6}$$

For the location parameters $\boldsymbol{\Theta}_h$ and $\boldsymbol{\eta}$ which can take on any value on the real line, it is customary to assign priors that are proportional to a constant. Therefore, we have

$$g(\boldsymbol{\eta}) \propto \text{ constant} \tag{6.7}$$

$$g(\boldsymbol{\Theta}_h) \propto \text{ constant} \tag{6.8}$$

Assuming *a priori* independence implies

$$g(\Theta) = \prod_{h=1}^{H} g(\Theta_h) \propto \text{ constant} \tag{6.9}$$

Finally, assuming prior independence of Θ, η and Ω, the joint prior pdf will simply be the product of the priors specified in (6.6) (6.7) and (6.9). Thus, we have

$$g(\Theta, \eta, \Omega) = g(\Theta)g(\eta)g(\Omega) \tag{6.10}$$

$$\propto \prod_{h=1}^{H} \mid \Omega_h \mid^{-\frac{n+1}{2}} \tag{6.11}$$

as the joint prior density function.

For specification of the likelihood function, it is first noted that the pdf of the expenditure vector for all households of type h is given by

$$f(V_h \mid \Theta_h, \eta, \Omega_h) \propto |\Omega_h|^{-\frac{M_h}{2}} \exp\left\{-\frac{1}{2}(V_h - Z_h\Theta_h - X_h\eta)'(\Omega_h^{-1} \otimes I_{M_h})(V_h - Z_h\Theta_h - X_h\eta)\right\} \tag{6.12}$$

Then, the complete likelihood function is

$$\begin{aligned} f(V \mid \Theta, \eta, \Omega) &= \prod_{h=1}^{H} f(V_h \mid \Theta_h, \eta, \Omega_h) \\ &\propto \prod_{h=1}^{H} |\Omega_h|^{-\frac{M_h}{2}} \\ &\exp\left\{-\frac{1}{2}\sum_{h=1}^{H}(V_h - Z_h\Theta_h - X_h\eta)'(\Omega_h^{-1} \otimes I_{M_h})(V_h - Z_h\Theta_h - X_h\eta)\right\} \end{aligned} \tag{6.13}$$

Combining (6.11) and (6.13) via Bayes' Theorem yields the joint posterior pdf

$$f(\Theta, \boldsymbol{\eta}, \boldsymbol{\Omega}|\mathbf{V}) \propto \prod_{h=1}^{H} |\boldsymbol{\Omega}_h|^{-\frac{M_h+n+1}{2}} \exp\left\{-\frac{1}{2}\sum_{h=1}^{H}(\mathbf{V}_h - \mathbf{Z}_h\Theta_h - \mathbf{X}_h\boldsymbol{\eta})'(\boldsymbol{\Omega}_h^{-1}\otimes I_{M_h})(\mathbf{V}_h - \mathbf{Z}_h\Theta_h - \mathbf{X}_h\boldsymbol{\eta})\right\} \quad (6.14)$$

This pdf is the source of all our post-sample information on the parameters Θ, $\boldsymbol{\eta}$ and $\boldsymbol{\Omega}$. What is of particular interest is the information on the equivalence scales s_{ih} and s_h that can be derived from the information on Θ and $\boldsymbol{\eta}$. To tackle this problem analytically, the traditional approach is to derive the marginal posterior pdfs $f(s_{ih}|\mathbf{V})$ and $f(s_h|\mathbf{V})$ for all i and h through variable transformation and by integrating unwanted parameters out of the joint posterior pdf. This task is a daunting one. Properties of the inverted Wishart distribution can be used to integrate $\boldsymbol{\Omega}_h$ from (6.14), but the resulting marginal posterior pdf for $(\Theta,\boldsymbol{\eta})$ is not of a recognisable form, and to then employ a transformation to (s_{ih}, s_h) would likewise not be a rewarding experience. A numerical approach is much more promising. If a sample of values of the parameters can be drawn from the posterior pdf in (6.14), corresponding values of the s_{ih} and s_h can be computed. This sample of values is equivalent to sampling from the marginal posterior pdfs $f(s_{ih}|\mathbf{V})$ and $f(s_h|\mathbf{V})$ for each commodity parameter and household type. Thus, the sample values can be used to estimate posterior means and standard deviations for the various equivalence scales, as well as provide information for approximate plots of the posterior pdfs.

Drawing values from the joint posterior pdf in (6.14) is achieved conveniently using Gibbs sampling. The Gibbs sampler is an algorithm for generating observations on random variables from their marginal distributions indirectly without having to derive their pdfs[1]. Through information from the conditional posterior pdfs for each of the parameters in the joint distribution, the procedure enables the generation of observations by sampling iteratively from these pdfs. After a suitable

[1] Appendix A gives a brief description of how the Gibbs sampler is applied in general terms.

'burn-in' period, the draws obtained represent draws from the joint posterior pdf. The application of Gibbs sampling to SUR systems of different kinds has been considered by Percy (1992), Chib and Greenberg (1995b) and Griffiths, Thomson and Coelli (1996). The current model does not fit neatly into any of these earlier studies, and so, the direction taken now is to derive the required conditional posterior pdfs of $\boldsymbol{\theta}_h$, $\boldsymbol{\eta}$ and $\boldsymbol{\Omega}_h$. The details of the application of the Gibbs sampler in this chapter is outlined in section 6.4.

6.3 Conditional Posterior Pdfs

Note that the joint posterior pdf in (6.14) can be conveniently factored as

$$
\begin{aligned}
f(\boldsymbol{\Theta},\boldsymbol{\eta},\boldsymbol{\Omega} \mid \mathbf{V}) \quad & \propto \\
& |\boldsymbol{\Omega}_h|^{-\frac{M_h+n+1}{2}} \quad \exp\left\{-\frac{1}{2}(\mathbf{V}_h-\mathbf{Z}_h\boldsymbol{\Theta}_h-\mathbf{X}_h\boldsymbol{\eta})'(\boldsymbol{\Omega}_h^{-1}\otimes I_{M_h})(\mathbf{V}_h-\mathbf{Z}_h\boldsymbol{\Theta}_h-\mathbf{X}_h\boldsymbol{\eta})\right\} \\
\prod_{k\neq h} & |\boldsymbol{\Omega}_k|^{-\frac{M_k+n+1}{2}} \quad \exp\left\{-\frac{1}{2}\sum_{k\neq h}(\mathbf{V}_k-\mathbf{Z}_k\boldsymbol{\Theta}_k-\mathbf{X}_k\boldsymbol{\eta})'(\boldsymbol{\Omega}_k^{-1}\otimes I_{M_k})(\mathbf{V}_k-\mathbf{Z}_k\boldsymbol{\Theta}_k-\mathbf{X}_k\boldsymbol{\eta})\right\}
\end{aligned}
\tag{6.15}
$$

This factorization is a convenient one for obtaining the conditional posterior pdfs for $\boldsymbol{\Omega}_h$ and $\boldsymbol{\Theta}_h$. Let

$$
\begin{aligned}
\boldsymbol{\Omega}_h^* &= \{\boldsymbol{\Omega}_k \mid \text{ for all } k \neq h\} \\
\boldsymbol{\Theta}_h^* &= \{\boldsymbol{\Theta}_k \mid \text{ for all } k \neq h\ \}
\end{aligned}
$$

To obtain the conditional posterior pdf for $\boldsymbol{\Omega}_h$, (6.15) is viewed as a function of $\boldsymbol{\Omega}_h$ only, with other parameters held constant. Then from (6.15) we have (Judge, et.al. 1985, p.479)

$$
f(\boldsymbol{\Omega}_h \mid \boldsymbol{\Theta},\boldsymbol{\eta},\boldsymbol{\Omega}_h^*,\mathbf{V}) \propto \mid \boldsymbol{\Omega}_h \mid^{-\frac{M_h+n+1}{2}} \exp\left\{-\frac{1}{2}\ \mathrm{tr}\ \left(\mathbf{A}_h\boldsymbol{\Omega}_h^{-1}\right)\right\} \tag{6.16}
$$

where $\mathbf{A}_h$ is an $(n \times n)$ matrix with $(i,j)^{th}$ element equal to

$$a_{ij} = (v_{ih} - z_h\theta_{ih} - x_h\eta_i)'(v_{jh} - z_h\theta_{jh} - x_h\eta_j)$$

The conditional density in (6.16) has an inverted Wishart distribution from which random generation of observations is straightforward (Anderson, 1984, p.238).

Also, from (6.15) the conditional posterior pdf for $\mathbf{\Theta}_h$ can be written as

$$f(\mathbf{\Theta}_h \mid \mathbf{\Theta}_h^*, \boldsymbol{\eta}, \mathbf{\Omega}, \mathbf{V}) \propto \exp\left\{-\frac{1}{2}(\mathbf{V}_h - \mathbf{Z}_h\mathbf{\Theta}_h - \mathbf{X}_h\boldsymbol{\eta})'\,(\mathbf{\Omega}_h^{-1} \otimes I_{M_h})(\mathbf{V}_h - \mathbf{Z}_h\mathbf{\Theta}_h - \mathbf{X}_h\boldsymbol{\eta})\right\} \tag{6.17}$$

Now, let

$$\begin{aligned}
\hat{\mathbf{\Theta}}_h &= \left[\mathbf{Z}_h'(\mathbf{\Omega}_h^{-1} \otimes I_{M_h})\mathbf{Z}_h\right]^{-1} \mathbf{Z}_h'(\mathbf{\Omega}_h^{-1} \otimes I_{M_h})(\mathbf{V}_h - \mathbf{X}_h\boldsymbol{\eta}) \\
&= (I_n \otimes M_h^{-1} z_h')(\mathbf{V}_h - \mathbf{X}_h\boldsymbol{\eta})
\end{aligned} \tag{6.18}$$

By adding and subtracting $\mathbf{Z}_h\hat{\mathbf{\Theta}}_h$ to the term $(\mathbf{V}_h - \mathbf{Z}_h\mathbf{\Theta}_h - \mathbf{X}_h\boldsymbol{\eta})$ in (6.17) and expanding, it can be shown that

$$\begin{aligned}
(\mathbf{V}_h - \mathbf{Z}_h\mathbf{\Theta}_h - \mathbf{X}_h\boldsymbol{\eta})'\,(\mathbf{\Omega}_h^{-1} &\otimes I_{M_h})(\mathbf{V}_h - \mathbf{Z}_h\mathbf{\Theta}_h - \mathbf{X}_h\boldsymbol{\eta}) \\
&= (\mathbf{\Theta}_h - \hat{\mathbf{\Theta}}_h)'\mathbf{Z}_h'(\mathbf{\Omega}_h^{-1} \otimes I_{M_h})\mathbf{Z}_h(\mathbf{\Theta}_h - \hat{\mathbf{\Theta}}_h) \\
&+ \left(\mathbf{V}_h - \mathbf{Z}_h\hat{\mathbf{\Theta}}_h - \mathbf{X}_h\boldsymbol{\eta}\right)'\,(\mathbf{\Omega}_h^{-1} \otimes I_{M_h})\left(\mathbf{V}_h - \mathbf{Z}_h\hat{\mathbf{\Theta}}_h - \mathbf{X}_h\boldsymbol{\eta}\right)
\end{aligned} \tag{6.19}$$

Note that the second term in (6.19) does not include $\mathbf{\Theta}_h$. Therefore, we can write (6.17) as

$$f(\mathbf{\Theta}_h \mid \mathbf{\Theta}_h^*, \boldsymbol{\eta}, \mathbf{\Omega}, \mathbf{V}) \propto \exp\left\{-\frac{1}{2}(\mathbf{\Theta}_h - \hat{\mathbf{\Theta}}_h)'\mathbf{Z}_h'(\mathbf{\Omega}_h^{-1} \otimes I_{M_h})\mathbf{Z}_h(\mathbf{\Theta}_h - \hat{\mathbf{\Theta}}_h)\right\} \tag{6.20}$$

Equation (6.20) suggests that the conditional posterior density for $\mathbf{\Theta}_h$ is a multivariate normal density function with mean $\hat{\mathbf{\Theta}}_h$ and covariance matrix $(M_h)^{-1}\mathbf{\Omega}_h$.

That is,

$$f(\Theta_h \mid \Theta_h^*, \eta, \Omega, \mathbf{V}) \sim N\left[\hat{\Theta}_h, \left[\mathbf{Z}_h'(\Omega_h^{-1} \otimes I_{M_h})\mathbf{Z}_h\right]^{-1}\right]$$

$$\sim N\left[\hat{\Theta}_h, \frac{1}{M_h}\Omega_h\right] \qquad (6.21)$$

To derive the conditional posterior for η consider again the joint posterior density in (6.14). Conditioning on all parameters except η implies that

$$f(\eta \mid \Theta, \Omega, \mathbf{V}) \propto \exp\left\{-\frac{1}{2}\sum_{h=1}^{H}(\mathbf{V}_h - \mathbf{Z}_h\Theta_h - \mathbf{X}_h\eta)'(\Omega_h^{-1} \otimes I_{M_h})(\mathbf{V}_h - \mathbf{Z}_h\Theta_h - \mathbf{X}_h\eta)\right\} \qquad (6.22)$$

If we let $\mathbf{V}_h^o = \mathbf{V}_h - \mathbf{Z}_h\Theta_h$, the summation term in (6.22) can be written as

$$\sum_{h=1}^{H}\mathbf{V}_h^{o\prime}(\Omega_h^{-1} \otimes I_{M_h})\mathbf{V}_h^o + \eta'\left[\sum_{h=1}^{H}\mathbf{X}_h'(\Omega_h^{-1} \otimes I_{M_h})\mathbf{X}_h\right]\eta - 2\eta'\left[\sum_{h=1}^{H}\mathbf{X}_h'(\Omega_h^{-1} \otimes I_{M_h})\mathbf{V}_h^o\right] \qquad (6.23)$$

To enable further simplification, let

$$W = \sum_{h=1}^{H}\mathbf{X}_h'(\Omega_h^{-1} \otimes I_{M_h})\mathbf{X}_h = \sum_{h=1}^{H}(x_h'x_h)\Omega_h^{-1} \qquad (6.24)$$

$$Q = \sum_{h=1}^{H}\mathbf{X}_h'(\Omega_h^{-1} \otimes I_{M_h})\mathbf{V}_h^o = \sum_{h=1}^{H}(\Omega_h^{-1} \otimes x_h')\mathbf{V}_h^o \qquad (6.25)$$

and define

$$\hat{\eta} = \left[\sum_{h=1}^{H}\mathbf{X}_h'(\Omega_h^{-1} \otimes I_{M_h})\mathbf{X}_h\right]^{-1}\sum_{h=1}^{H}\mathbf{X}_h'(\Omega_h^{-1} \otimes I_{M_h})\mathbf{V}_h^o$$

$$= W^{-1}Q \qquad (6.26)$$

It then follows that

$$f(\boldsymbol{\eta} \mid \boldsymbol{\Theta}, \boldsymbol{\Omega}, \mathbf{V}) \propto \exp\left\{-\frac{1}{2}(\boldsymbol{\eta} - \hat{\boldsymbol{\eta}})' W (\boldsymbol{\eta} - \hat{\boldsymbol{\eta}})\right\}$$
$$\propto \exp\left\{-\frac{1}{2}(\boldsymbol{\eta} - \hat{\boldsymbol{\eta}})' \left(\sum_{h=1}^{H} (x_h' x_h) \boldsymbol{\Omega}_h^{-1}\right) (\boldsymbol{\eta} - \hat{\boldsymbol{\eta}})\right\} \quad (6.27)$$

which suggests that $f(\boldsymbol{\eta} \mid \boldsymbol{\Theta}, \boldsymbol{\Omega}, \mathbf{V})$ is a multivariate normal distribution with mean $\hat{\boldsymbol{\eta}}$ and covariance matrix W^{-1}. That is,

$$f(\boldsymbol{\eta} \mid \boldsymbol{\Theta}, \boldsymbol{\Omega}, \mathbf{V}) \sim N \left[\hat{\boldsymbol{\eta}}, \left(\sum_{h=1}^{H} (x_h' x_h) \boldsymbol{\Omega}_h^{-1}\right)^{-1}\right] \quad (6.28)$$

6.4 Applying the Gibbs Sampler

In this section, the Gibbs sampling procedure applied to our model is described. The object of the exercise is to be able to draw observations or "sample" from the Bayesian posterior pdf in (6.14) using information from the conditional posterior pdfs in (6.16), (6.20) and (6.28). Specifically, the Gibbs sampler is employed to generate observations on $\boldsymbol{\Omega}_1$, $\boldsymbol{\Omega}_2$, ..., $\boldsymbol{\Omega}_H$, $\boldsymbol{\Theta}_1$, $\boldsymbol{\Theta}_2$, ..., $\boldsymbol{\Theta}_H$, and $\boldsymbol{\eta}$ using the following steps:

1. Given some initial values for θ_{ih} and η_i, compute for each h,

$$[\mathbf{A}_h]_{ij} = (v_{ih} - z_h \theta_{ih} - x_h \eta_i)'(v_{jh} - z_h \theta_{jh} - x_h \eta_j)$$

2. Draw values $\boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, ..., \boldsymbol{\Omega}_H$ from respective inverted Wishart distributions with parameter matrices $\mathbf{A}_1, \mathbf{A}_2, ..., \mathbf{A}_H$ and degrees of freedom $M_1, M_2, ..., M_H$.

3. Compute $\hat{\boldsymbol{\Theta}}_1, \hat{\boldsymbol{\Theta}}_2, ..., \hat{\boldsymbol{\Theta}}_H$ as defined in (6.18) and, given the $\boldsymbol{\Omega}_h$ drawn in step (2), draw values $\boldsymbol{\Theta}_h$, $h = 1, 2, ..., H$ from $N(\hat{\boldsymbol{\Theta}}_h, M_h^{-1}\boldsymbol{\Omega}_h)$ distributions.

4. Using the values for $\boldsymbol{\Omega}_h$ and $\boldsymbol{\Theta}_h$ drawn in steps (2) and (3), respectively, compute W and $\hat{\boldsymbol{\eta}}$ as defined in (6.24) and (6.26).

5. Draw a value for $\boldsymbol{\eta}$ from a N$(\hat{\boldsymbol{\eta}}, W^{-1})$ distribution.

6. Return to step (1) using the θ_{ij} and η_i drawn in steps (3) and (5), respectively, and continue to proceed iteratively through all the steps, until a large sample has been generated.

Markov Chain theory guarantees that, after a particular point, the observations from this large sample represent observations from the marginal (or joint) posterior pdfs (Geman and Geman, 1984). The point at which they represent points from the marginal pdfs is the point at which the Markov chain (created by the Gibbs sampling procedure) has converged. Because observations at the beginning of the iterative procedure will not necessarily be from the marginal posterior pdfs, it is conventional to drop a number of these, treating them as initial observations in a "burn-in" period.

Estimates for θ_{ij} and η_i, once obtained, enable the calculation of parameter estimates for a_{ih} and b_i from equations (5.17)-(5.20). These in turn lead to the estimation of equivalence scales s_{ih} and s_h from expressions (5.21) and (5.28) derived earlier.

6.5 Empirical Application

The Bayesian procedure described in the previous section is applied here using data from the 1988-89 Australian Household Expenditure Survey. This microunit data set is described in detail in Chapter 3. The Bayesian procedure developed in this chapter is applied to the 5532 sample households grouped into eight household types (see Table 3.2) and using all the eleven expenditure categories described in Chapter 3.

For Bayesian estimation, 18,000 sets of observations were generated for each parameter element in $\boldsymbol{\Omega}_h$, $\boldsymbol{\eta}$ and $\boldsymbol{\Theta}_h$ for all the 8 household types. Observations from the first 3,000 runs in the iterations provided the 'burn-in' period of the Gibbs sampler and hence were discarded, leaving 15,000 observations in the final estimation sample. These observations were used to estimate the posterior means and standard deviations of all the commodity-specific and general scales as well as to provide observations for graphing marginal posterior pdfs for some selected scales. All calculations were carried out using the econometric package SHAZAM.

Checks of convergence of the generated Gibbs sequence were conducted through diagrams. The estimated scale values for the first and last 1,000 observations of that part of the generated series that was retained were plotted adjacent to each other. The plots for the generated values of the food, clothing and housing scales capture the essence of the convergence for the entire set of commodity groupings and are shown here in Figures 6.1-6.3. For all commodity types, the generated values of the scales are stable suggesting that convergence has taken place. Across commodities, the series for the food scales exhibit least variability while the series for the clothing scales exhibit the most variability. The series for the rest of the commodity groups including housing show levels of variability in between these two extremes. Plotted in Figure 6.4 are the generated values for the general scales. It is likewise evident here that convergence of the generated series has taken place.

Table 6.1 shows the posterior means and standard deviations of the commodity-specific scales from the Bayesian estimation. Taking the posterior means as point estimates, the scales for most of the commodities exhibited the expected increase in the per household equivalent expenditure with the increase in household size. Further, the increasing graduation of scales occurred at a decreasing rate indicating economies of scale for additional children. There are some exceptions to these observations. The estimated scales for Alcohol and Tobacco are observed to decline as the number of children in the household increases. Also, the scales for Medical and Health Care and Others commodity groups exhibit no defined trend for one-adult households.

To facilitate comparison of the Bayesian scales with those obtained through the iterative maximum likelihood procedure developed in Chapter 5, Table 5.3 is reproduced here as Table 6.2. It is most obvious, from only a quick inspection of both tables, that the point estimates from both methods are strikingly similar. In most cases, the point estimates differ only from the third decimal point. The two sets of estimates appear to be highly reliable in that they possess small posterior standard deviations and small standard errors, although the Bayesian standard deviations are, in general, higher than the ML standard errors. Across commodities, the largest variances are associated with the estimated scales for Clothing and

Footwear, Recreation and Entertainment and Others commodity types while the smallest variances are associated with the estimated scales for Food and Alcohol and Tobacco. Across households, the largest variances are exhibited by household type (1,3) while the smallest variances belong to those scales for single member households i.e. household type (1,0). This result can be attributed to the relative sizes of the samples: there are more than a thousand households of type (1,0) while there are only 42 households of type (1,3) (see Table 3.2).

The estimates for the general scales are presented in Table 6.3. It may be recalled that general scales depend on a chosen level of income of the reference household x_r. The ML estimates are based on x_r= 450 which is the median income level. The Bayesian estimates are marginalised with respect to the empirical distribution of the x_r, with the inequality restriction $x_r > a_r$ imposed. Symbolically, the posterior pdf for the general scale is

$$f(s_h) = \int_{x_r > a_r} f(s_h \mid x_r, \mathbf{V}) f(x_r) dx_r \tag{6.29}$$

In practice, this means that, in each Gibbs sampling iteration, a value x_r is drawn from the empirical distribution of the x_r and the corresponding value for s_h is calculated from it and the current drawings of the other parameters. If drawings of x_r and a_r were such that $x_r < a_r$ (i.e. income is less than subsistence expenditure), another x_r was drawn until the required inequality was satisfied. The Bayesian posterior means and ML estimates are again very close to each other. Compared to the ML point estimates, the Bayesian scales for the 2-adult households are smaller in magnitude while those for the 1-adult households are slightly larger. The standard errors for the ML estimates are conspicuously absent from the table. Why? While it is theoretically possible to derive standard errors for the ML estimates, doing so would be extremely cumbersome because of the large number of parameters involved and their various interrelationships. With the Bayesian estimates, posterior standard deviations are straightforwardly estimated.

The posterior pdfs of some of the estimated scales are presented in Figures 6.5-6.8. Normal conditional posterior pdfs implied by the ML estimates and standard errors are plotted alongside the posterior pdfs for the Bayesian scales. It is immediately evident that the ML-based posteriors approximate those of the Bayesian posterior pdfs, exhibiting no significant differences for all curves. The figures clearly show that posterior pdfs for the single member households stand out for their sharp steepness relative to the others. In contrast, the posterior pdfs associated with household type (1,3) are the flattest of the lot, reflecting the added imprecision from a small number of households. Also, a comparison of the variability in precision of information across commodities is more readily made. For example, foodscales are observed to have less variability compared to those of the clothing or housing scales.

The pdfs are observed to shift to the right with the addition of children in the household. The magnitudes of the shifts however differ across the commodities. For example, with food, the effects of increasing the number of adults and children in the household has a clear distinct effect; overlapping of the posterior pdfs is minimal. With clothing, changes in the demographic make-up of the household also shift the pdfs to the right but with considerable overlappings of the pdfs. This reflects the fact that there exists lesser gains in economies of scale for food compared to housing or clothing.

The posterior pdfs for the general scales in Fig. 6.8 yield patterns that are consistent with the observations for the posterior pdfs of the three commodity groups analysed earlier. That is, the pdfs shift to the right at a diminishing rate with the addition of children in the household. The least variance was observed for household type (1,0) while the largest variance was observed for household type (1,3). Standard deviations for the general scales are noted to be larger than those for the commodity-specific scales.

6.6 Concluding Remarks

In this chapter, the Bayesian approach was applied to our problem of estimating equivalence scales. It was shown that Bayesian techniques and related numerical methods are relatively straightforward to use to obtain posterior pdfs and summary measures such as the posterior means and standard deviations. Unlike ML estimation, posterior pdfs permit finite sample inferences about non-linear functions of the original parameters like the equivalence scales. Also, Bayesian estimates take account of the uncertainty associated with the error covariance matrix estimation in that they are not conditional on point estimates of the error covariances, as ML estimates are. It is also convenient to be able to present information diagramatically through plots of posterior pdfs. Through such diagrams, it is easy to show the relationships between the different scales in terms of both the magnitudes and reliability of the derived information; such relationships may not be so obvious when analysing point estimates alone.

In this chapter, finite sample inference for all parameters is possible from their marginal posterior probability density functions that were estimated using Gibbs sampling. As it turns out, if one views the ML results through Bayesian eyes, and constructs "posterior" normal distributions from the ML results with the point estimate as the mean of the distribution, and the standard error as the "posterior standard deviation", the asymptotic ML results from Chapter 5 are virtually identical to the Bayesian results obtained in this chapter. This result is perhaps surprising, but is likely due to the fact that the sample size is large and the asymptotic results for the ML estimators are good finite sample approximations. For the general scales, no attempt at an asymptotic variance was made because of its complicated structure. The Bayesian estimators, however, provide a complete posterior distribution for the general scales which permit more complete inference than was possible within the sampling theory framework. A useful future extension of this study would be the imposition of *prior* information in the form of obvious inequalities between scales and on subsistence parameters.

With the conclusion of this chapter, the basic steps are set for the development of a methodology to cover zero expenditures in the next chapter. The Bayesian approach is shown there to facilitate the development of an estimation procedure that is too difficult to handle within the sampling theory framework.

Table 6.1 Bayesian Posterior Means and Standard Deviations for Commodity-Specific Scales

	Commodity Specific Scales							
	Household Type (no. of adults, no. of children)							
Commodity Type	*(1,0)*	*(1,1)*	*(1,2)*	*(1,3)*	*(2,0)*	*(2,1)*	*(2,2)*	*(2,3)*
Housing	0.82269	1.02559	1.14785	1.27995	1.00000	1.48807	1.51560	1.64963
	(0.04040)	(0.08350)	(0.09622)	(0.22548)	-	(0.09238)	(0.07938)	(0.09952)
Fuel & Power	0.67342	0.92021	1.06296	1.11025	1.00000	1.21552	1.34328	1.43859
	(0.01893)	(0.05229)	(0.06003)	(0.13720)	-	(0.04019)	(0.03851)	(0.04961)
Food	0.52986	0.72651	0.94421	1.05883	1.00000	1.23641	1.42314	1.57516
	(0.01302)	(0.04096)	(0.04761)	(0.12146)	-	(0.03288)	(0.03215)	(0.04369)
Alcohol & Tobacco	0.57326	0.46419	0.38810	0.34489	1.00000	0.94899	0.86359	0.76252
	(0.03384)	(0.05340)	(0.05993)	(0.07912)	-	(0.06648)	(0.06332)	(0.05411)
Clothing & Footwear	0.53373	0.90953	0.92086	1.41036	1.00000	1.27952	1.40232	1.64687
	(0.04929)	(0.15584)	(0.15551)	(0.35590)	-	(0.11570)	(0.11224)	(0.15340)
Household Furnishings	0.54832	0.66360	0.77333	0.81383	1.00000	1.45087	1.15187	1.32548
& Equipment	(0.04031)	(0.08347)	(0.11627)	(0.24227)	-	(0.13807)	(0.09439)	(0.11088)
Medical & Health Care	0.53766	0.46764	0.67554	0.50716	1.00000	1.26139	1.28029	1.31282
	(0.04199)	(0.08200)	(0.12328)	(0.16573)	-	(0.08378)	(0.05732)	(0.07144)
Transport	0.52496	0.57064	0.62282	0.78055	1.00000	1.01755	1.19239	1.36633
	(0.03858)	(0.07627)	(0.09542)	(0.28149)	-	(0.07744)	(0.07681)	(0.12984)
Recreation	0.53822	0.57570	0.51331	0.82662	1.00000	1.02866	1.28465	1.36945
& Entertainment	(0.04111)	(0.09893)	(0.07593)	(0.24889)	-	(0.10475)	(0.11507)	(0.14589)
Personal Care	0.54155	0.78131	0.97237	0.73424	1.00000	1.19372	1.29070	1.17633
	(0.03697)	(0.09988)	(0.15972)	(0.16776)	-	(0.09478)	(0.08032)	(0.09259)
Others	0.56738	1.02376	0.88952	0.80893	1.00000	1.39732	1.79408	2.08332
	(0.06528)	(0.17297)	(0.11805)	(0.17583)	-	(0.14236)	(0.19043)	(0.27199)

Note: The posterior standard deviations are in parentheses.

Table 6.2 ML Estimates of Commodity-Specific Scales

	Commodity Specific Scales Household Type (no. of adults, no. of children)							
Commodity Type	*(1,0)*	*(1,1)*	*(1,2)*	*(1,3)*	*(2,0)*	*(2,1)*	*(2,2)*	*(2,3)*
Housing	0.82172	1.02505	1.14572	1.27714	1.00000	1.48756	1.51704	1.64722
	(0.03987)	(0.07964)	(0.09091)	(0.18807)	-	(0.08956)	(0.07835)	(0.09766)
Fuel & Power	0.67312	0.91992	1.06320	1.10961	1.00000	1.21492	1.34256	1.43857
	(0.01902)	(0.04940)	(0.05659)	(0.11329)	-	(0.03918)	(0.03824)	(0.04888)
Food	0.52972	0.72669	0.94396	1.05914	1.00000	1.23606	1.42257	1.57518
	(0.01290)	(0.03902)	(0.04509)	(0.10172)	-	(0.03230)	(0.03196)	(0.04277)
Alcohol & Tobacco	0.57251	0.46325	0.38735	0.34456	1.00000	0.94869	0.86320	0.76293
	(0.03346)	(0.05007)	(0.05641)	(0.06558)	-	(0.06559)	(0.06223)	(0.05348)
Clothing & Footwear	0.53191	0.90715	0.91881	1.40345	1.00000	1.27749	1.39773	1.64244
	(0.04840)	(0.14615)	(0.14618)	(0.29826)	-	(0.11465)	(0.11101)	(0.14993)
Household Furnishings & Equipment	0.54681	0.66242	0.77018	0.81213	1.00000	1.44928	1.15006	1.32351
	(0.03940)	(0.07897)	(0.10895)	(0.20121)	-	(0.13538)	(0.09300)	(0.10880)
Medical & Health Care	0.53733	0.46789	0.67540	0.50918	1.00000	1.26092	1.27966	1.31144
	(0.04163)	(0.07720)	(0.11462)	(0.13755)	-	(0.08260)	(0.05641)	(0.07033)
Transport	0.52396	0.57066	0.62121	0.77931	1.00000	1.01612	1.19071	1.36594
	(0.03817)	(0.07280)	(0.09015)	(0.23206)	-	(0.07606)	(0.07555)	(0.12714)
Recreation & Entertainment	0.53713	0.57582	0.51157	0.82315	1.00000	1.02673	1.28129	1.36727
	(0.04153)	(0.09393)	(0.07086)	(0.20691)	-	(0.10205)	(0.11331)	(0.14254)
Personal Care	0.54072	0.78107	0.96937	0.73291	1.00000	1.19177	1.28752	1.17465
	(0.03637)	(0.09376)	(0.14850)	(0.14053)	-	(0.09237)	(0.07896)	(0.09046)
Others	0.56552	1.02258	0.88733	0.80647	1.00000	1.39372	1.79058	2.07809
	(0.06391)	(0.16369)	(0.11145)	(0.14848)	-	(0.13901)	(0.18630)	(0.26501)

Note: The estimated standard errors are in parentheses.

Table 6.3 Bayesian and ML Estimates of General Scales

Household Type (no. of adults, no. of children)	*Bayesian Estimates*	*ML Estimates*
(2,0)	1.00000	1.00000
	-	-
(2,1)	1.23380	1.23164
	(0.05705)	
(2,2)	1.33410	1.32578
	(0.05458)	
(2,3)	1.47000	1.46043
	(0.07253)	
(1,0)	0.58189	0.58143
	(0.02198)	
(1,1)	0.72262	0.72295
	(0.05323)	
(1,2)	0.78616	0.79635
	(0.05633)	
(1,3)	0.89943	0.91396
	(0.15416)	

Note: The posterior means are treated as the Bayesian point estimates and the values in parentheses are the posterior standard deviations.

Figure 6.1
Plots of the first and last 1000 sample points from the generated series for the food scales *(after discarding the burn-in observations).*

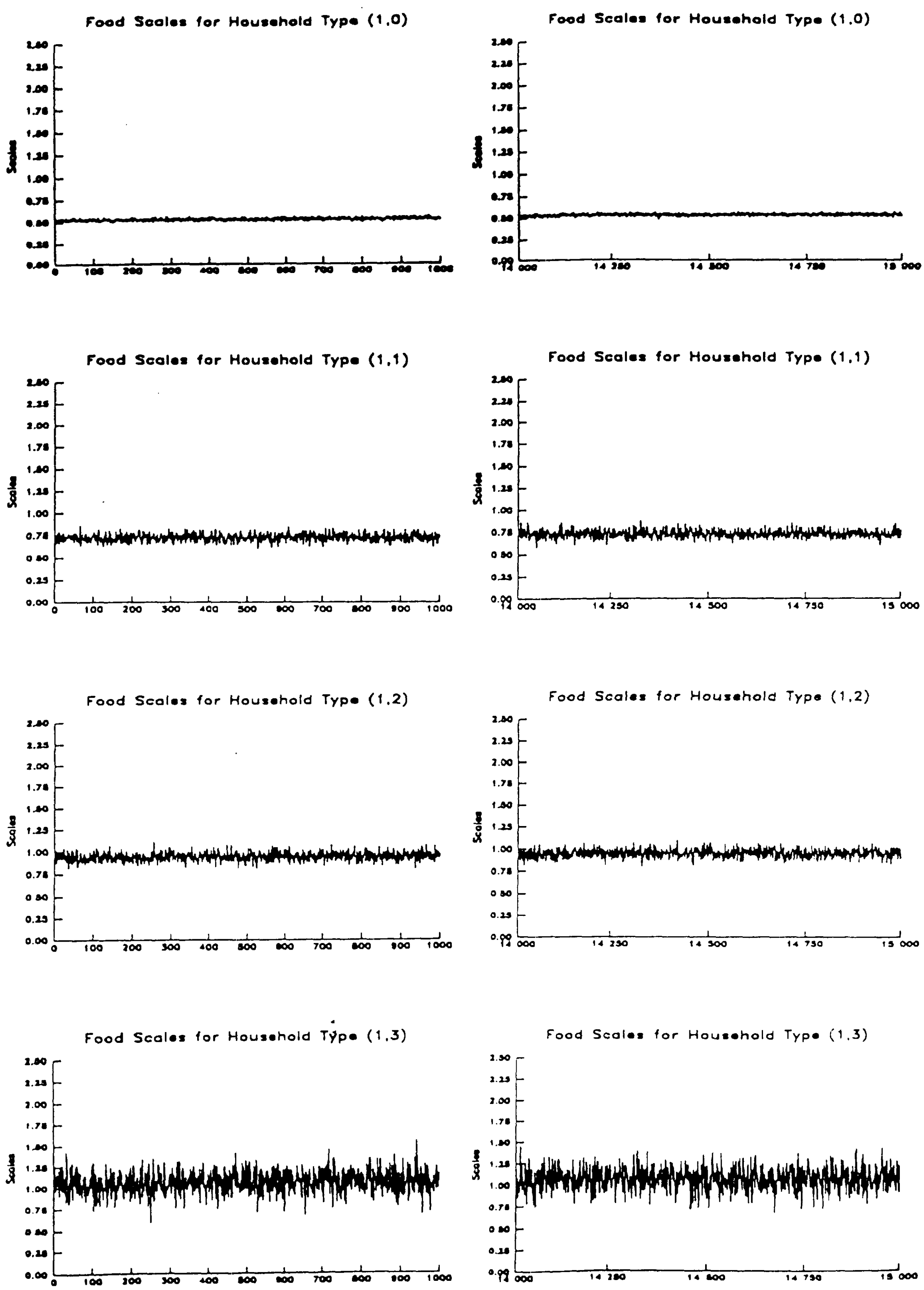

Figure 6.1 (cont.)
Plots of the first and last 1000 sample points from the generated series for the food scales *(after discarding the burn-in observations).*

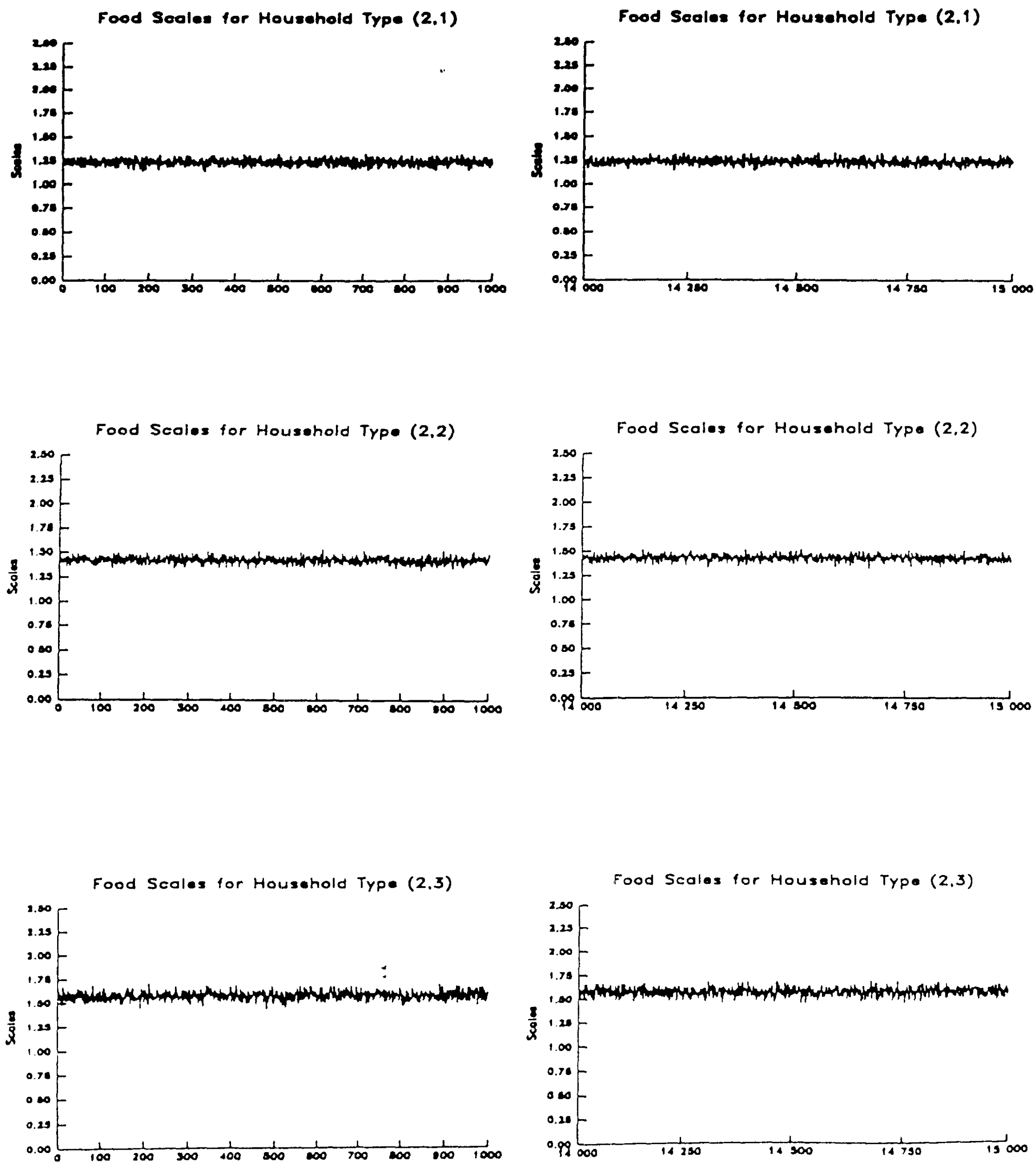

Figure 6.2

Plots of the first and last 1000 sample points from the generated series for the clothing scales *(after discarding the burn-in observations).*

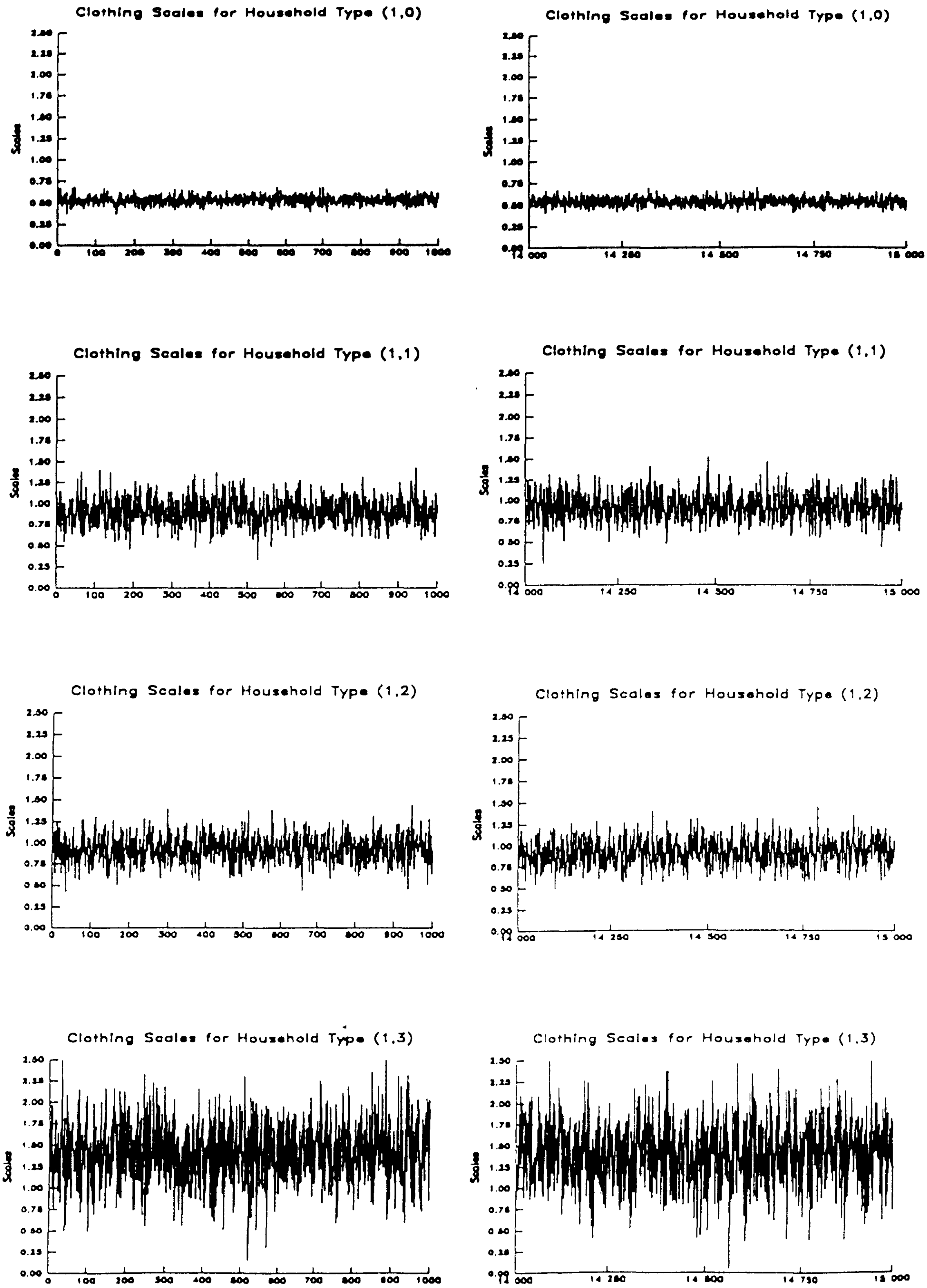

Figure 6.2 (cont.)
Plots of the first and last 1000 sample points from the generated series for the clothing scales *(after discarding the burn-in observations).*

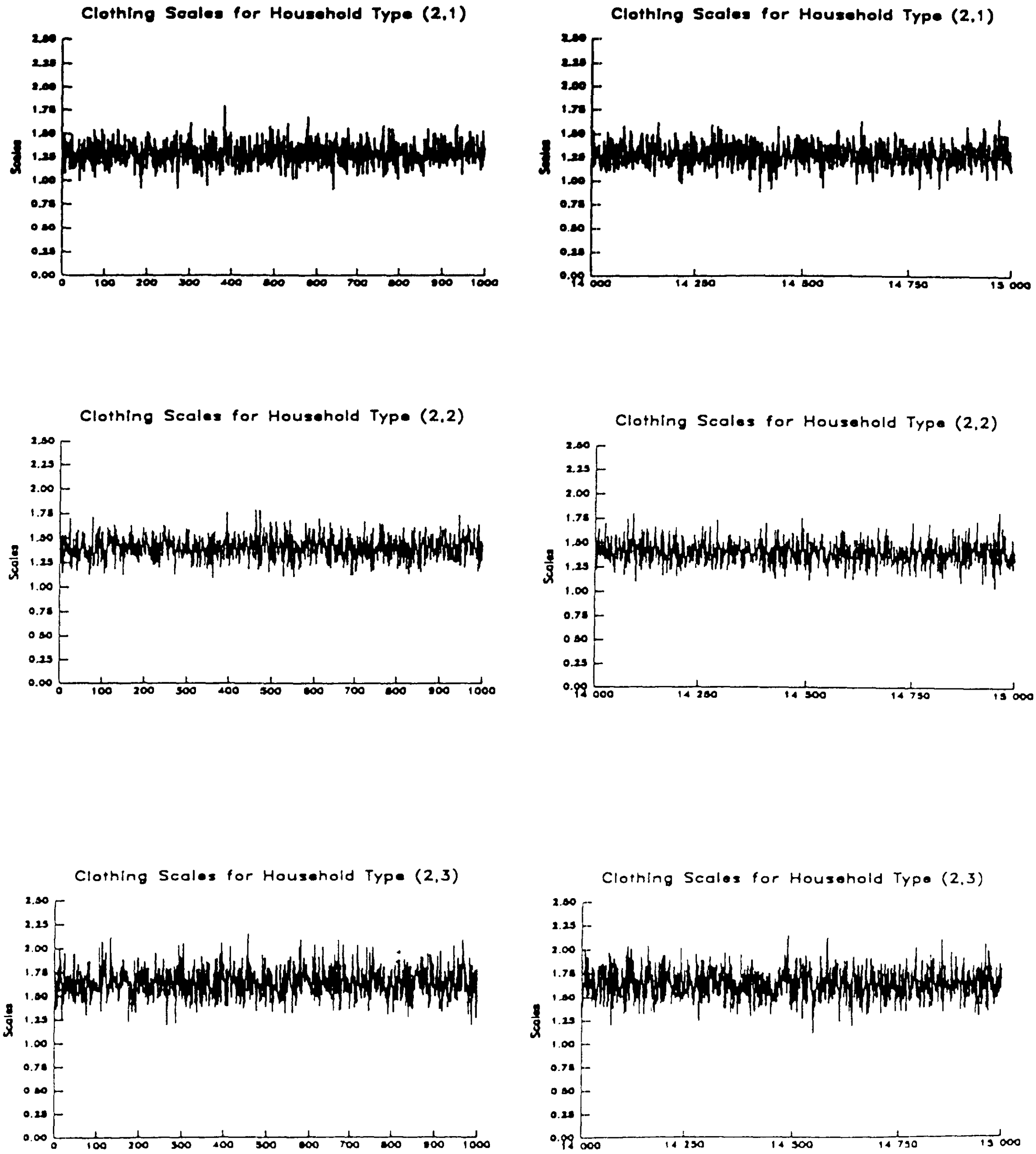

Figure 6.3
Plots of the first and last 1000 sample points from the generated series for the housing scales *(after discarding the burn-in observations).*

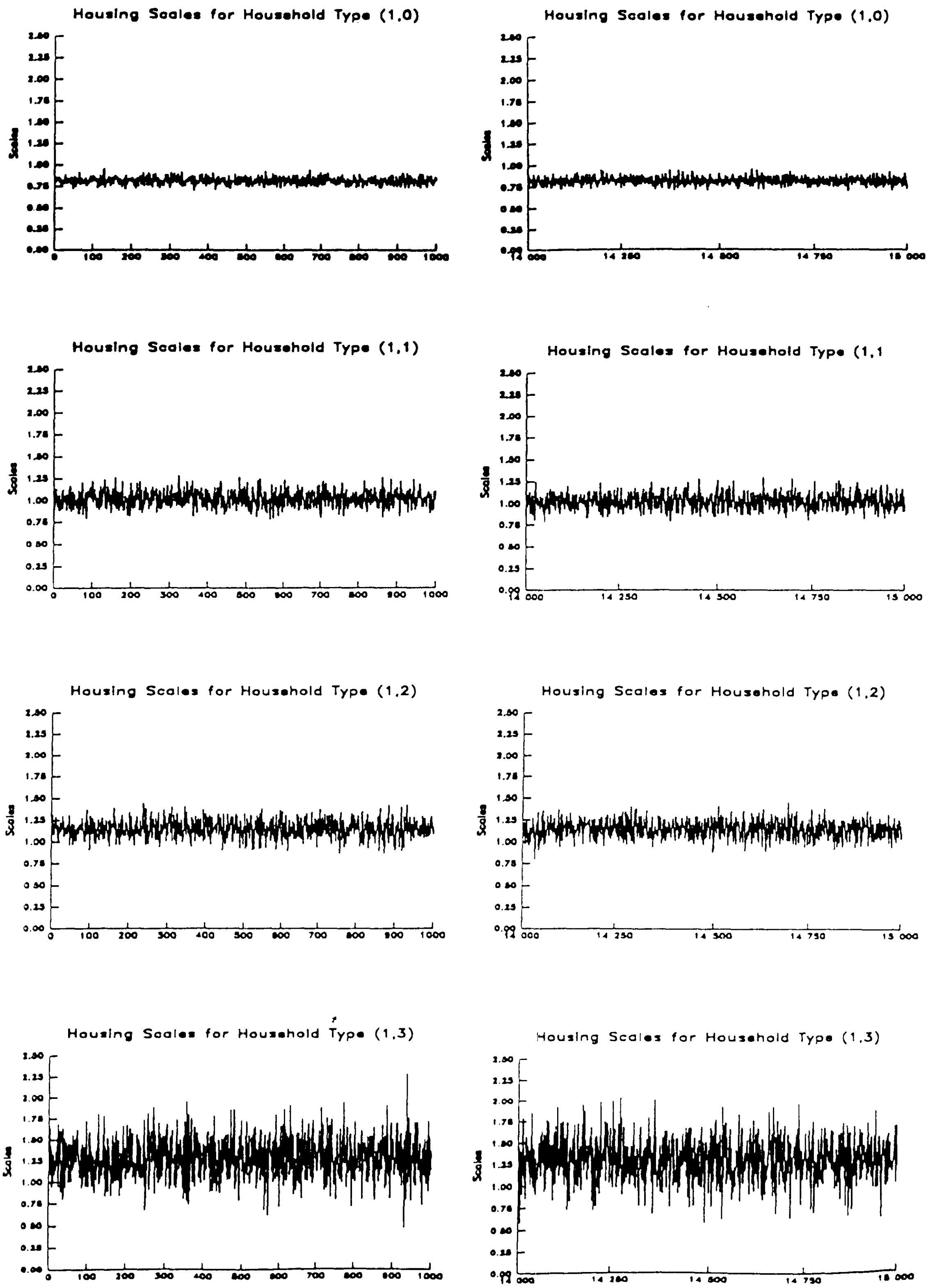

Figure 6.3 (cont.)
Plots of the first and last 1000 sample points from the generated series for the housing scales *(after discarding the burn-in observations).*

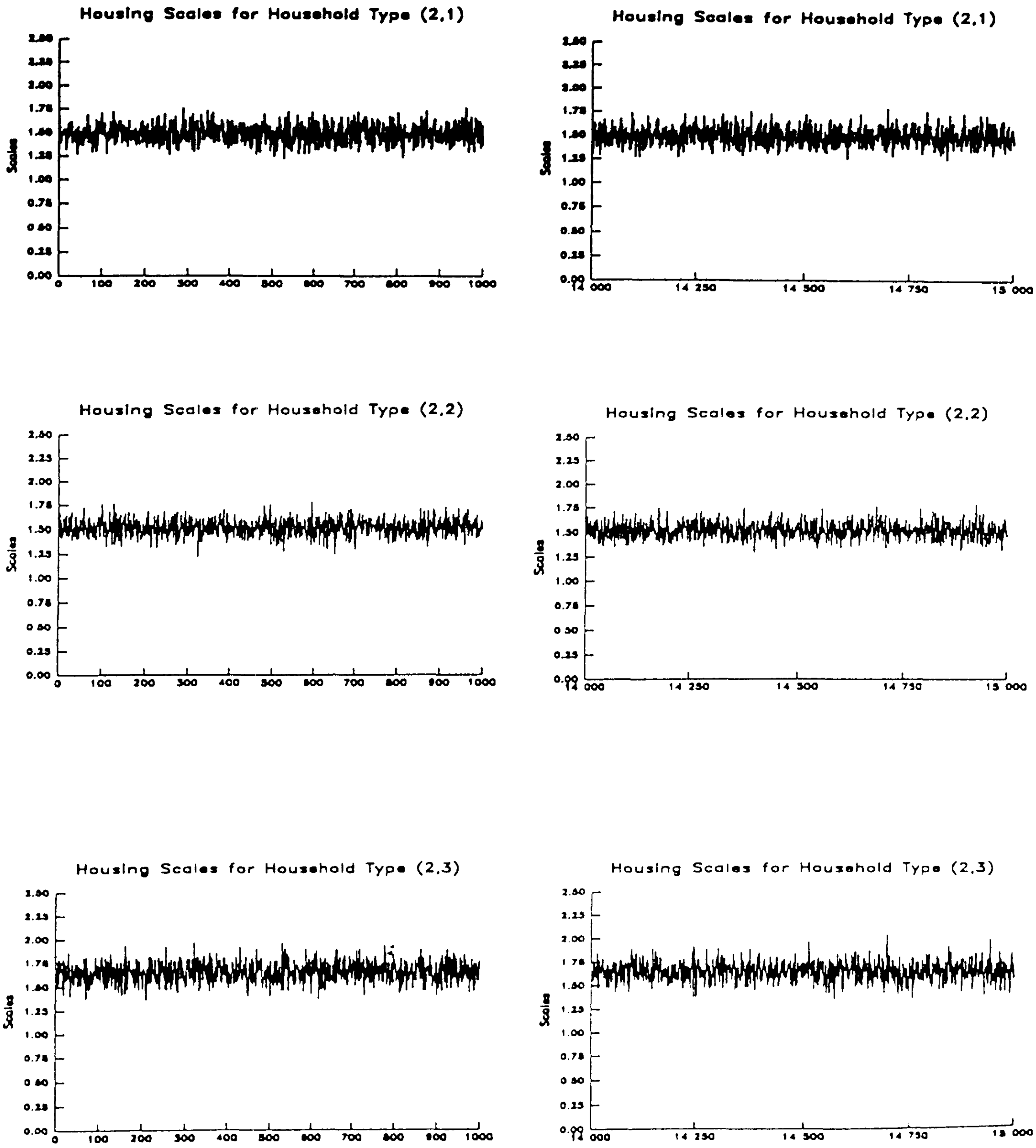

Figure 6.4
Plots of the first and last 1000 sample points from the generated series for the general scales *(after discarding the burn-in observations).*

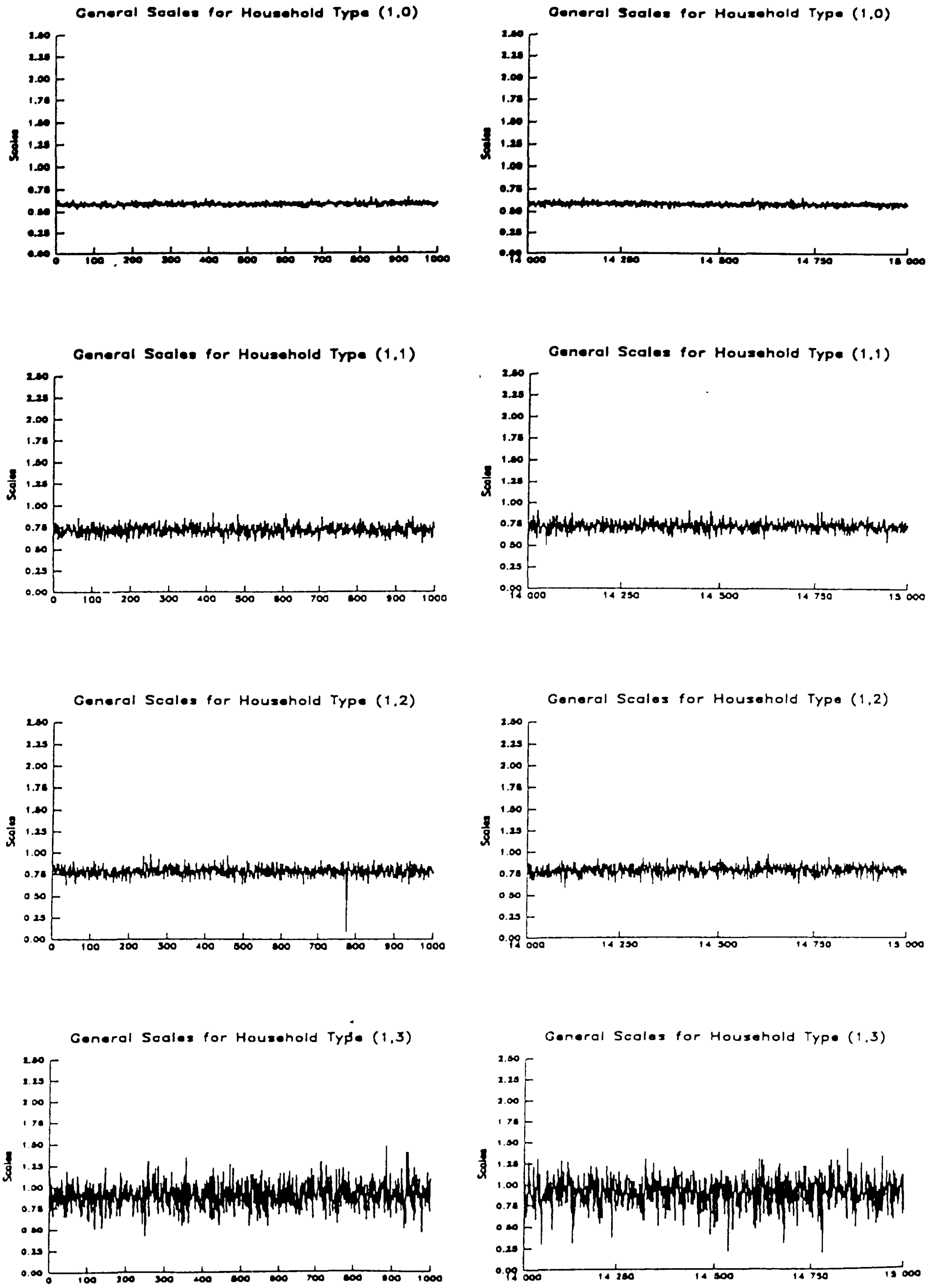

Figure 6.4 (cont.)
Plots of the first and last 1000 sample points from the generated series for the general scales *(after discarding the burn-in observations).*

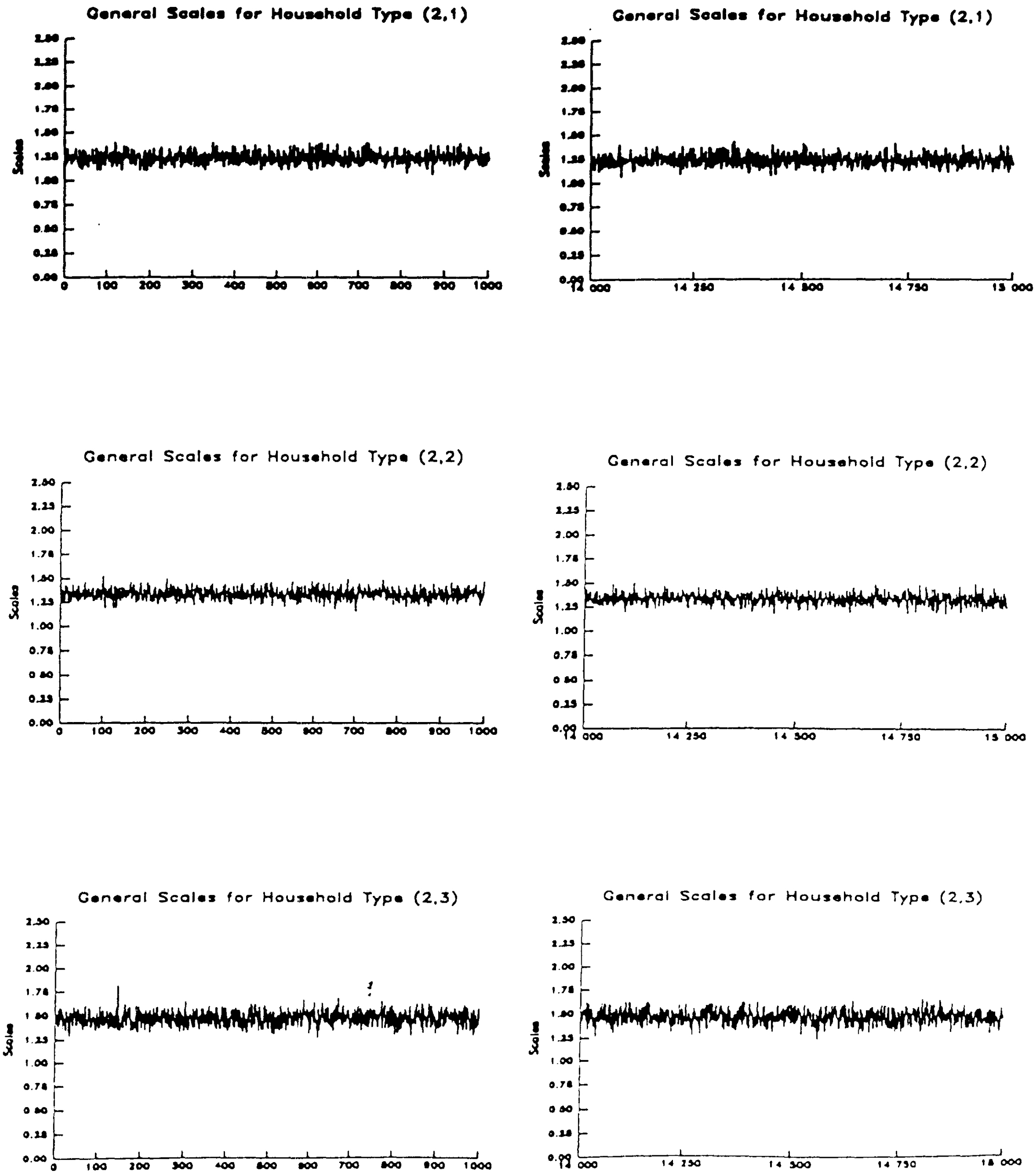

Figure 6.5
Posterior Distributions of Food Scales for each Household Type *(no. of adults, no. of children).*

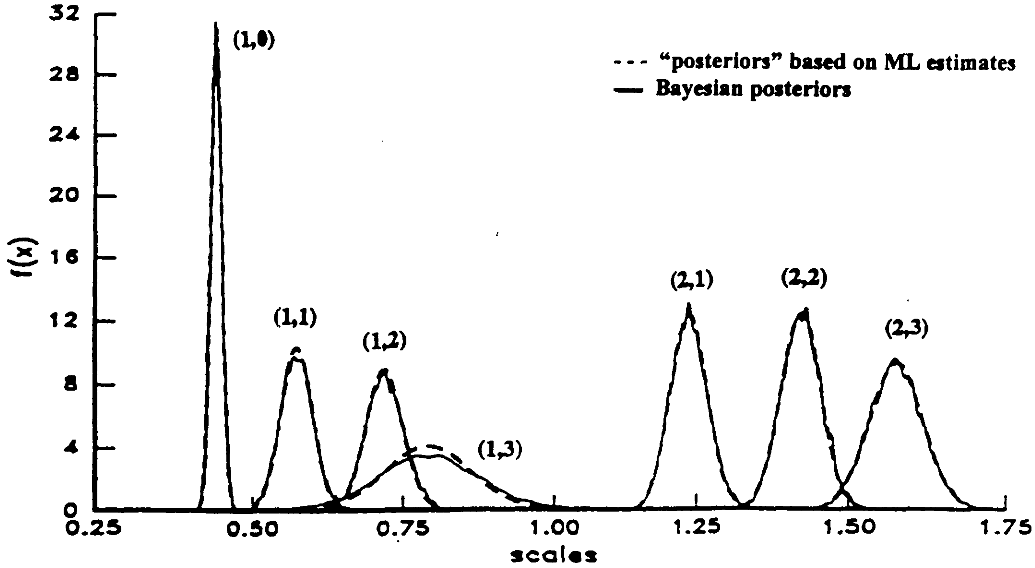

Figure 6.6
Posterior Distributions of Clothing Scales for each Household Type *(no. of adults, no. of children).*

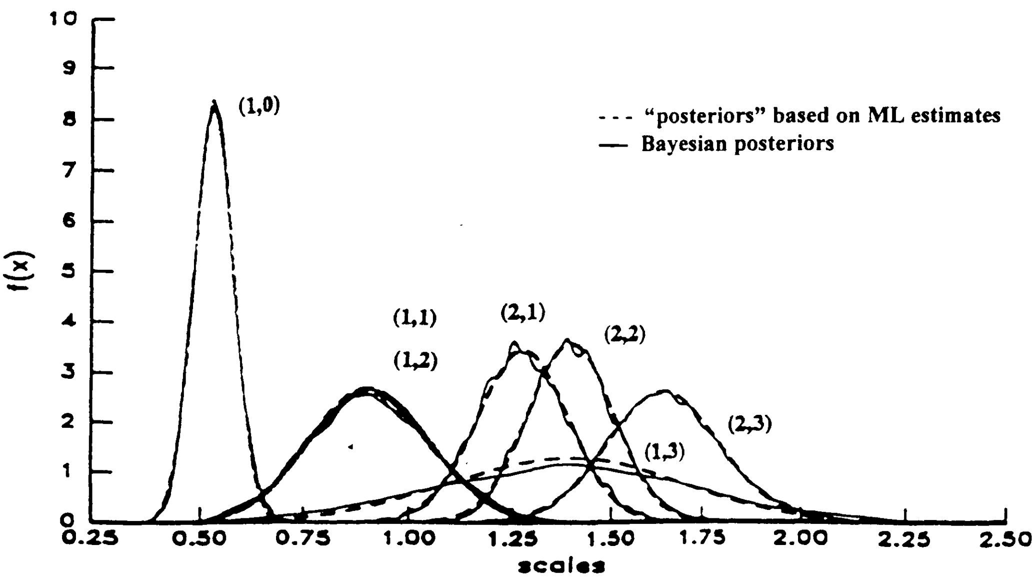

Figure 6.7
Posterior Distributions of Housing Scales for each Household Type *(no. of adults, no. of children).*

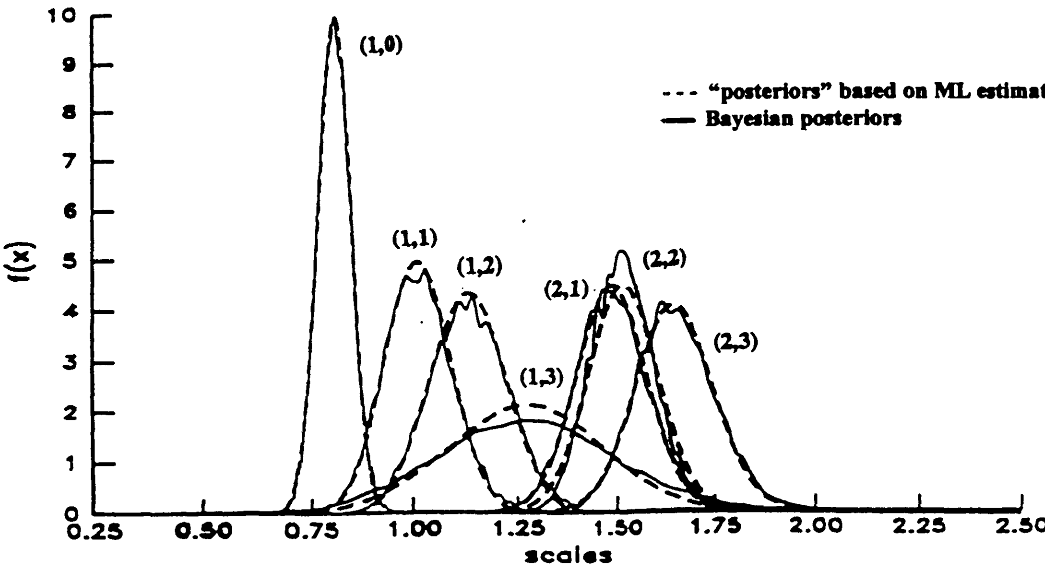

Figure 6.8
Posterior Distributions of General Scales for each Household Type *(no. of adults, no. of children).*